CHAPTER 3

Multistability

3.1 Introduction

As I argued in Chaps. 1 and 2, perceiving is hard. Fundamentally, the proximal sensations at the receptors are underdetermined. They do not specify a unique thing (the *distal* object). There always is ambiguity; there is so much visual, auditory, and tactual information that any scene can be understood in more than one way. For example, a ball seen increasing in size might be approaching or it may be inflating. What rescues us from this confusion is that there are neurological and cognitive processes that capitalize on the physical constraints on the shapes, movements, and sounds of objects in our physical world to yield a single percept. In addition, objects are not presented in isolation and that context may act to bring about one percept. I am certain that part of the match between the physical and perceptual world is partly due to evolution and partly to experience as discussed in Chap. 2.

Normally, one possible percept will be *stronger due the above factors*, and it dominates. But, in other instances two or more possible percepts are *equally strong*. In those instances, given enough time, the resulting percept can bounce around the various possibilities. The physical stimulation is unchanging, but conscious awareness fluctuates and the perception of one of the incompatible percepts suppresses all of the others. Transitions may be marked by very brief periods of indeterminate, mixed, or intermediate appearances. But each percept eventually becomes unified and complete; it is **not** an amalgam of the two or three possible competing percepts.

Such scenes, drawings, or sounds have been termed *reversing* figures or *multistable* figures. Each percept may not occur equally often, but each one is a plausible interpretation of the stimulus. I do not think that reversing or

Electronic Supplementary Material: The online version of this chapter (https://doi.org/10.1007/978-3-319-96337-2_3) contains supplementary material, which is available to authorized users.

S. Handel, *Perceptual Organization*,
https://doi.org/10.1007/978-3-319-96337-2_3

multistable pictures should be considered merely curiosities, since all perception allows for alternative potential outcomes. That said, there are ambiguous stimuli that never undergo alternations. A color on the green/yellow border does not reverse.

The original explanation for the perceptual reversals was that they were due shifts of attention, for example, focusing on different faces of the Necker cube (see Fig. 3.2). But the current conceptualization is that the reversals can be due to many factors in addition to shifts of attention. These range from intrinsic properties of the stimulus, to knowledge of the alternatives, to random fluctuations of neural firing, and to neural fatigue resulting in shifts in firing rates. In the following sections we will describe the various types of visual, auditory, and tactual reversing figures, particularly those in which parts of one stimulus are reconfigured to produce a different percept. Then, we will show how neural processes and higher-level cortical processes affect the alternation.

3.2 Visual Multistability

3.2.1 Multistable Static Figures

In vision, the classic methodology is to use ambiguous static pictures, as illustrated in Figs. 3.1 and 3.2. In some cases participants are given prior information about the multiple percepts, while in others they are naïve to the possibilities. Typically, the participant is presented with a single stimulus and asked to indicate whenever the percept shifts. The rivalry is between two alternative interpretations of the same stimulus, is it two faces or one vase.

One type of reversing figure occurs when the two alternatives seem to be at the same depth. In Fig. 3.1, these two reversing percepts are A, D, E, F, G, H, J, K, L, and I. For A, F, and J the alternatives arise from two different orientations of the same perceived object. For I, the triangles can be seen pointing in three alternative directions so that it is a tristable figure. For D, E, G, H, and K, the alternatives arise from grouping the component parts in different ways to bring forth two distinct objects (e.g., a duck versus a rabbit in E).

The second type are figure-ground reversals; two alternate objects appear at different depths, one in front of the other, and the reversal brings the behind object in front and thereby changes the perception. The reversals in B, C, and L occur when the black and white segments reverse in depth. For B, when the white segment is seen in front it appears as a vase and the black segment is just unimportant background. If the black segment is seen in front, it appears to be the silhouettes of two men facing each other and the white segment is background. The border goes with the part in front, it shapes the object as discussed in Chap. 2.

Consider the classic case of a reversing figure, the Necker cube shown in Fig. 3.2. The six lines can be interpreted as being on the two-dimensional surface, or projected into three dimensions. In (A), we have the basic configurations (also see Fig. 3.1A) with six transparent sides. The basic Necker cube reverses in three dimensions easily; in one configuration one of the square faces

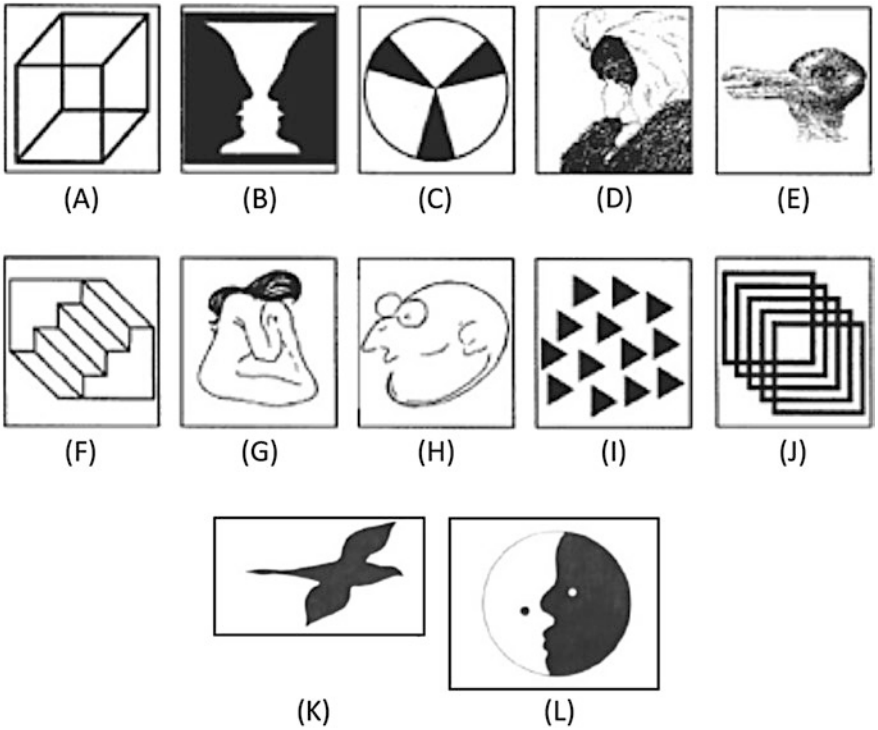

Fig. 3.1 A set of well-known reversing figures. (A) Necker Cube; (B) Face/Vase; (C) Maltese Cross; (D) Wife/Mother-in-Law; (E) Duck/Rabbit; (F) Staircase; (G) Man/Girl; (H) Rat/Man; (I) Triangles; (J) Overlapping Squares; (K) Goose or crow; (L) Gog and Magoog. (Adapted from Long & Toppino, 2004; Fisher, 1968)

appears to point down to the left, in the second configuration the other square face appears to point up to the right. If one of the two square faces is made darker but transparent in 3.2A1, it is still seen in three dimensions and reversals occur easily. If the darker face is opaque, it is still three-dimensional but few reversals occur probably due to the occlusion of one of the corners and one of the connecting edges. The perception of 3.2A as three-dimensional is probably due to two factors. First, all of the corners consist of 90° angles, which, combined with two obtuse angles greater than 90°, signify depth. Second, there are several places where lines cross, which would not be a probable outcome for a two-dimensional object. The more important factor, however, is probably the corner angles because the figure still looks three-dimensional if the crossing lines are eliminated, though now the cube does not reverse (A2).

In 3.2B, the faces meet at a point, four edges still connecting the vertices of the faces. The transparent Necker cube in 3.2B can reverse between two or three dimensions (or reverse in three dimensions). If one face is made transparent so that the connecting edges remain visible (3.2B1), it can still be seen in two or three dimensions. However, removing the hidden edges makes the cube three-dimensional (3.2B2).

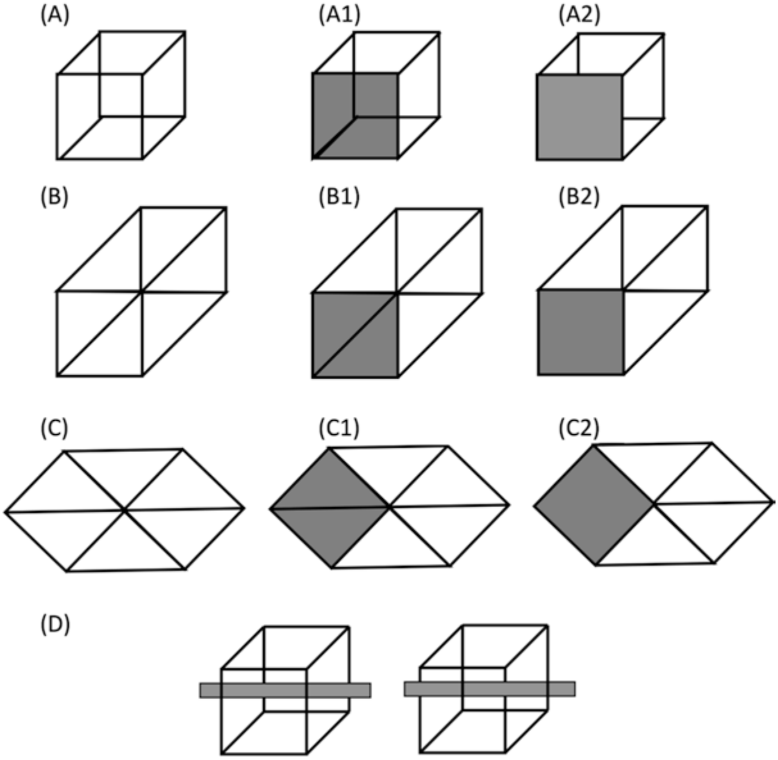

Fig. 3.2 A Necker cube illustrated at different orientations. In (A) through (C), all six faces are transparent. If one face is shaded, different orientations and occlusions give rise to quite different perceptions. In (D), an intersecting rod makes one orientation predominant

If the cube is rotated 45°, it is invariably seen in two dimensions (3.2C) even if one face is made opaque (3.2C1) and the connecting edge is removed (3.2C2). The explanation for the two-dimensional percept for 3.2C and 3.2D is usually based on the "most probable" principle multiplying the prior probabilities by the likelihood that each hypothesis would yield the proximal sensations, that is, Bayesian calculations. Namely, if the distal object was three-dimensional the proximal stimulus would occur only in one improbable orientation and therefore the best bet is to perceive the distal stimulus as a flat two-dimensional object. Such cases, that is, three-dimensional objects oriented to appear two-dimensional, have been termed "accidental coincidences" and thus are unlikely events (Rock, 1983). We can also use the same principle to explain why the preferred orientation of 3.2A and 3.2B is downward to the left. Typically, we would look at a cube at a downward angle on a flat surface so that the front surface would face down and that prevalent viewpoint would predict the preferred orientation of 3.1F and 3.1J. Finally, in 3.2D a rod fixes the orientation of the cube. Adding context stabilizes our perceptions.

YouTube Video
Archival Gibson: There are several videos in the category "Archival Gibson" that illustrate J.J. Gibson's seminal contributions. The videos seem primitive; they were produced before computers and were constructed using single frame photography. But, they illustrate important aspects of visual perception

Bob Shaw-1974. Symmetry and Event Perception: Shows how the perception of the Necker cube varies as a function of orientation

3.2.2 Multistable Dynamic Figures

The simplest multistable case comes about if two lights, tones, or vibrators are flashed alternatively at a fixed distance apart with a constant blank temporal interval between the stimuli. (Fig. 3.3A). The fundamental perceptual issue is determining how the first stimulus relates to the second. The perceptual systems must construct the most plausible correspondence. If the temporal interval is very short, then both stimuli seem to flash at the same time so that they probably represent different objects. If the temporal interval is very long, then the two stimuli flash intermittently and there is no reason to connect the two. However, if the spatial distance and the temporal interval are chosen appropriately, the light seems to move smoothly between the two positions even though the motion is not seen until after the second stimulus is presented. At that timing between the onsets, the illusion of a stimuli moving back and forth alternates with the accurate perception of the alternation. Korte's third law (Korte, 1915) formalized the relationship that to achieve smooth motion the onset interval must increase as the distance increases. This holds true for touch and vision at the best timing for apparent motion and at threshold for audition (Lakatos & Shepard, 1997). However, Gepshtein, and Kubovy (2007) varied the time interval between the lights and found that while Korte's third law was true for higher speeds (i.e., shorter onset intervals) between the lights, a tradeoff between speed and distance occurred at slower speeds. Here, as the distance interval increases, the onset interval must decrease.

If the interval between the first stimulus (left light) and the second stimulus (right light) is shorter than the reverse interval between the right light back to the left, then the perceived motion is left to right (Fig. 3.3B). The shorter interval gives rise to the dominant motion. Freeman and Driver (2008) made use of this asymmetry to show that a tone can affect the direction of the apparent motion, another example of temporal ventriloquism. Specifically, the intervals between the two lights were identical. Then they placed one tone so that it lagged the onset of the first tone slightly and a second tone so that it led the onset of the second light by the same amount. A lag-lead sequence had the effect of making the perceived interval duration shorter and that led to motion from the lag light to the lead light. The effect of the tone offsets was equal to that caused by actually shifting the timing of the lights (without the tones).

The perception of *apparent* motion can be very strong and will occur even if the two lights are different colors or sizes. I have even seen apparent motion between a slide of a mouse and a slide of an elephant, where a trunk grows near

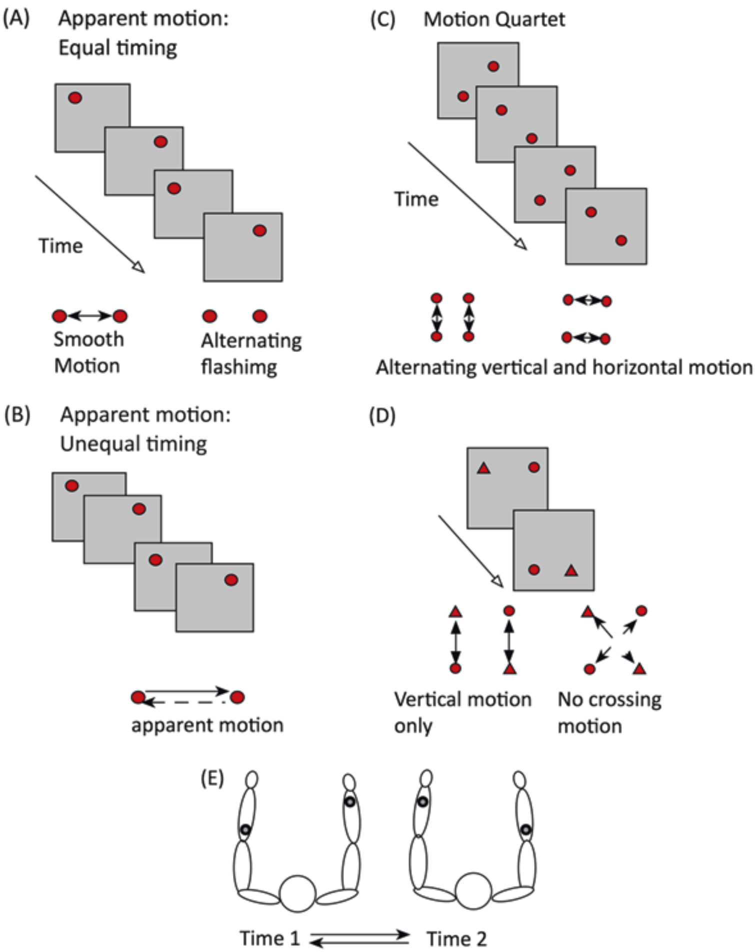

Fig. 3.3 (A) If the same object is flashed alternatively, two perceptions appear: either (1) a smooth movement between the objects or (2) the alternating flashing objects. If the leftward and rightward timing is the same, the speed of the leftward and rightward movement is identical. (B) If the rightward timing is shorter than the leftward timing, the light seems to go quickly to the right and then slowly back to the left. (C) If two pairs of identical diagonal objects are flashed alternatively, either (1) horizontal or (2) vertical motion occurs. Even if the objects are different as in (D), only vertical motion takes place. Crossing motion that connects the same kind of shape does not happen (Web designers beware). In (A), (B), (C), and (D) the different percepts switch back and forth. (E) Apparent motion also occurs between two limbs. The perception crosses the body midline

the elephant and shrinks near the mouse. (I am certain that if the second slide were that of a zebra, then I would have seen stripes appear and disappear in the motion). The visual system solves the correspondence problem by assuming that the two stimuli represent the same object in spite of the physical dissimilarity.

The perception of motion can be so strong that magicians trick audiences into believing that a coin has been pitched between two hands by making the appropriate tossing and catching movements at the right timing as described by Macnik, Martinez-Conde, and Blakeslee (2010).

Furthermore, it is possible to create apparent motion between two vibratory actuators or a light and a tactual actuator on a fingertip (Harrar, Winter, & Harris, 2008). In contrast to apparent motion for two lights, the apparent motion between the vibrators occurred over a shorter range of distances between the stimuli. Apparent motion also occurred between a light and vibrator although the distance between the light and vibrator did not change the strength of the perceived motion. This could suggest a different mechanism for cross-modal apparent movement.

A more complex dynamic visual configuration is based on four lights fixed on the corners of an imaginary square (Fig. 3.3C). In one of the two alternating frames, two diagonal dots are presented and in the other frame the two other diagonal dots are presented. Three percepts occur: (1) the dots are perceived to move back and forth horizontally; or (2) move up and down vertically; or (3) flash periodically. The three percepts can oscillate particularly when the vertical and horizontal distances between the lights are equal. The individual motions are nearly always identical to minimize changes in the configuration. Transforming the square into a tall vertical or a long horizontal rectangle can change the strength of each percept so that strongly horizontal and vertical movement will occur respectively. The motion occurs between the two closest dots and that results in the slowest movement corresponding to gradual movements in the environment. As previously argued, these percepts can be understood in terms of the strong predisposition to perceive a three-dimensional world. It is interesting to note that several apparent motion percepts do not occur: the dots rarely rotate around the square, and the dots do not split apart. Moreover, objects do not appear to move on paths that cross one another (avoiding collisions) even if that results in motions between dissimilar shapes (Fig. 3.3D).

It is possible to create an analogous array using four vibrators. The four vibrators could be placed on one forefinger (Carter, Konkle, Wang, Haywood, & Moore, 2008), or two vibrators could be placed on each forearm (Liaci, Bach, van Elst, Heinrich, & Kommeler, 2016). The latter is illustrated in Fig. 3.3E. The tactual results parallel those for vision with some interesting differences. The kinds of motions are identical: there are parallel vertical or horizontal movements. This occurs even though for the array illustrated in Fig. 3.3E, the vertical movements are within each forearm, while the horizontal movements are between the forearms, crossing the body midline. But, variation in the vertical and horizontal distances among the vibrators has much less effect on the type of movement than for the visual displays. It is possible to create 100% vertical or horizontal movement visually by the placement of the lights, but that does not occur for tactual arrays. The reason for this difference is unclear.

Given that apparent motion is similar for all three modalities, it is possible to explore the interactions when apparent motion occurs in two or three modalities simultaneously. In the basic configuration, there are two lights, two tones, and two vibrators, one placed to the participants left and one placed to the right. The spatial arrangement is identical: the light is placed in front of the audio speaker and the participant grasps the vibrators in front of the speaker. Moreover, the timing of the stimuli in each modality is identical. The stimuli in each modality are excited left then right or the reverse and the task is simply to judge whether the stimuli moved left or right. (There is only one cycle). In some instances only one stimulus was presented, but in other instances two or three stimuli were presented simultaneously. The two or three stimuli could move *congruently* left to right (or right to left) or the stimuli could move *incongruently* so that one stimulus moved left to right but the other stimulus (or stimuli) moved right to left. (This stimulus presentation does not give rise to alternating perceptions because there is only one cycle, but it does make use of the appearance of motion). These options are illustrated in Fig. 3.4.

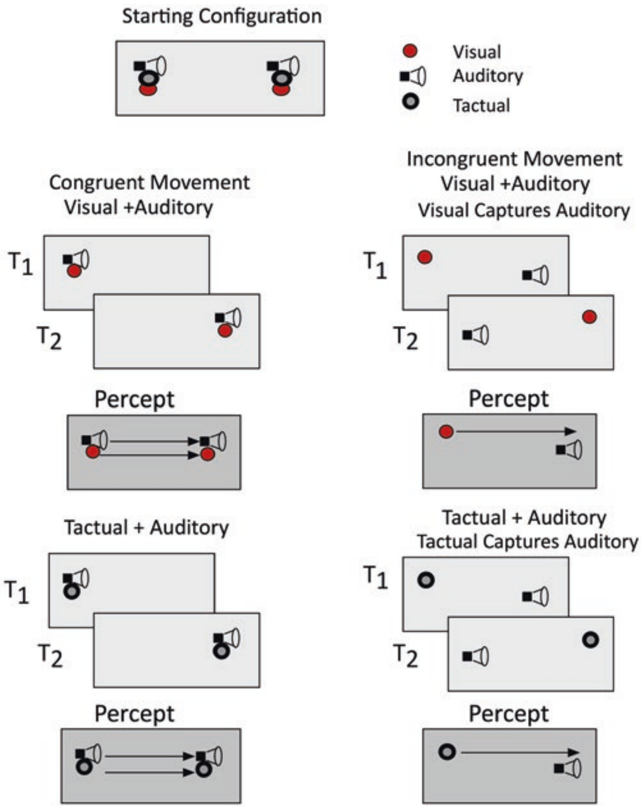

Fig. 3.4 The visual, auditory, and tactual stimuli are stacked on both sides of the display. If the stimuli from two modalities move congruently (either left to right or right to left), the percept is integrated and participants can judge the direction of either modality. If the stimuli move incongruently in opposite directions, the visual and tactual stimuli capture the auditory one. The auditory stimuli either are misjudged to move in the same direction as the visual or tactual stimuli or their direction cannot be judged

If the stimuli in one, two, or three modalities were congruent, then the perception of the direction was essentially perfect. If the stimuli were incongruent, the perception of direction was asymmetric and differentiated the modalities. Vision is the dominant modality. If the visual light and auditory tone or vibrator motions were incongruent, participants were able to perceive the direction of the light perfectly in spite of the tone or vibrators, but were unable to report the direction of the tone or vibrators. The direction of the light captured the directions of the tone or vibrator. In similar fashion, the direction of the touch stimuli captured the direction of the tones although the direction of the tone did affect the direction of the touch somewhat. At least for tones and vibrators, increasing the salience or discriminability of either modality did increase the strength of the capture (Occelli, Spence, & Zampini, 2010). If the direction of the tones was incongruent to the bimodal presentation of lights and touches, the capture of the tones was greater than either lights or touch alone (Sanabria, Soto-Faraco, & Spence, 2005; Soto-Faraco, Spence, & Kingstone, 2004).

Another way to illustrate the influence of a second sense on the perception of multistable figures is by means of binocular rivalry. If different scenes are presented to each eye, say a series of vertical lines to the left eye and a series of horizontal lines to the right eye, the percept oscillates between the two orientations. Rarely do the two scenes fuse into a crisscross pattern. If the observer places one hand on a tactual grid scored with either vertical or horizontal lines, that orientation dominates visually. Switching the orientation of the hand on the tactual surface switches the orientation of the dominant visual percept (Klink, van Wezel, & van Ee, 2012).

3.3 Auditory Multistability

One common approach to study auditory multistability is to make use of the stream segregation paradigm. In the simplest case, two sounds of different frequencies are recycled. If the frequencies are similar or the time interval between the offset of one and the onset of the next is long, the percept is of one oscillating pattern. As the frequency difference is increased, and/or the time interval is shortened, the continuous pattern breaks up and is perceived as two independent sequences or streams; one stream of the low-pitch tones, one stream of the high-pitch tones **(Examples are found in Sound File 2.9)**. For these sequences, sound similarity dominates. In between, there are frequency and time intervals in which the percept switches between the two. In general, the initial percept is that of a single integrated cycling of the tones, another instance of the preference to perceive one source or object. The two interleaved streams percept appears later, as an alternative. After the initial swap, further swaps between the one- and two-stream percept occur after random time intervals. Clearly, the multistability of streaming and apparent visual movement (i.e., movement between the dots versus independent flashing dots, Fig. 3.3A) is analogous both in stimulus configuration and outcome.

The streaming paradigm can be expanded to include three frequencies (Thossen & Bendixen, 2017). The three tones 400 Hz, 635 Hz, and 1008 Hz cycle at the rate of five tones/sec. There are many possible ways of organizing the three tones. For example, the 400 Hz tone could form one stream and the 635 Hz + 1008 Hz tones could form the second. Moreover, the single-tone stream could be heard in the forefront or the two-tone stream could be heard in the forefront. In general, the streams were based on frequency, the 400 Hz tone versus the two higher pitches or the 1008 Hz tone against the two lower pitches. What is most relevant here is that the organization is multistable, and the average time between reversals was roughly 16 sec. The number and pattern of reversals differed greatly among the listeners, which make sense given the number of alternative organizations.

Sound File 3.4: Multistability for three tone sequences

A more complex example of binaural rivalry occurs with sequences of alternating pitches (Brancucci & Tommasi, 2011). Originally discovered by Deutsch (1974), a pictorial representation of the stimulus sequence is shown in Fig. 3.5. The identical alternations of tones are presented to each ear such that the sequences are out of phase with each other. The very surprising outcome for the vast majority of listeners is that one ear hears only the low tones and the other ear hears only the high tones. The perceived rate of each tone is one-half the actual rate. The tones switch ears at random, although they seem to switch ears more frequently at the faster presentation rate. There is no agreed upon explanation (Fig. 3.5).

A second type of auditory multistability occurs if a sound or word is repeated continuously. For example, Warren and Gregory (1958) have constructed a sequence that simply repeats the word “farewell.” As one listens, the word becomes “welfare” and there then is a continual switch between the two words.

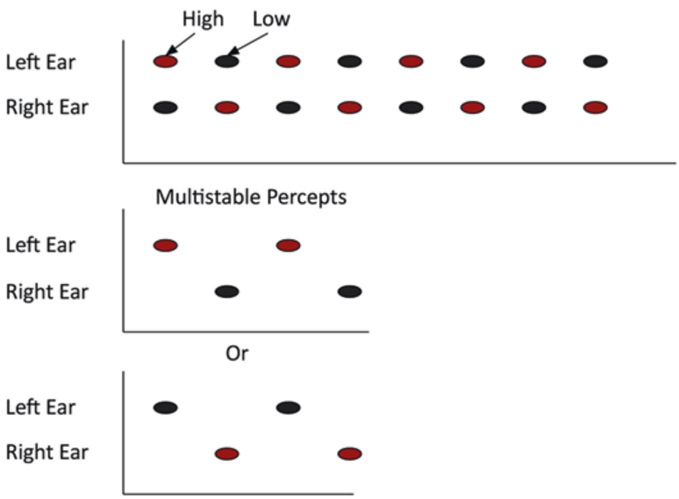

Fig. 3.5 A representation of the out-of-phase low/high tone sequences. The typical perception is either the high tone in the left ear and the low tone is the right ear or vice versa. The two possibilities switch back and forth

Sound Files 3.5: Binaural rivalry creates multistability as illustrated in Fig. 3.5

This is an interesting twist because the two possibilities are due to a timing swap; there is a different starting point for each word, it resembles focusing on a different point in a visual figure. Moreover, as can be seen in Fig. 3.6, the shift to "welfare" involves linking the two syllables across a relatively silent gap. A second example occurs when repeating the word "ace." The perception shifts back and forth between "ace" and "say." This sort of multistability seems analogous to perspective shifts in vision.

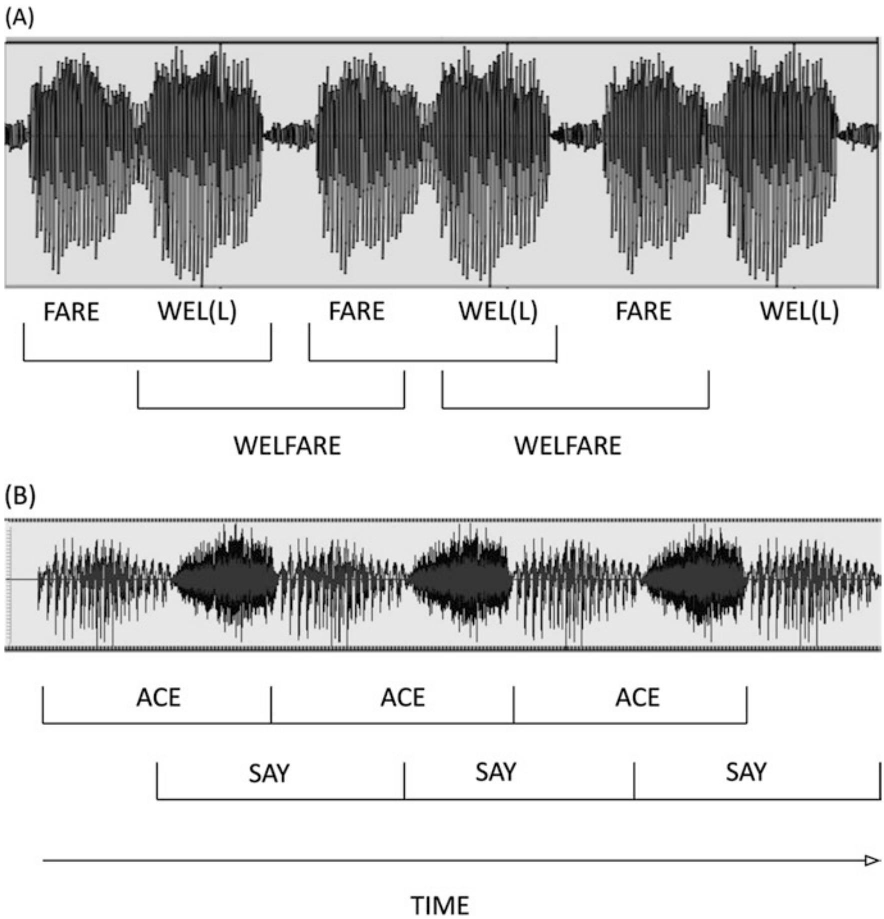

Fig. 3.6 (A) An acoustical representation of three repetitions of the word "welfare." After listening for a short period, the percept changes from "farewell" to "welfare" and then oscillates back and forth. (B) An acoustical representation of two representations of the word "ace." The percept changes from "ace" to "say" and then switches back and forth. www4.uwm.edu/APL/demonstrations.html. These demonstrations are derived from (Warren, 1999)

Sound Files 3.6: Multistability due to repetition of one and two syllable words as diagrammed in Fig. 3.6

3.4 The Nature of the Reversals: No Single Explanation

Taken together, all of this implies that there are many processes underlying these reversals. There are competition mechanisms based on adaptation as well as mutual inhibition at multiple neural-processing levels. Some occur at the initial stages of visual and auditory processing based on neural fatigue and binocular rivalry. Others occur at later cortical areas involved with attention and recognition, the initial local processes embedded in later more global ones. There are always interactions among the levels so that feedback from higher levels prompts reanalysis of lower level percepts. The remainder of this chapter will attempt to tease apart these components. Much of this follows from the reviews by Alais and Blake (2015) and Long and Toppino (2004).

3.4.1 "Bottom-Up" Passive and Automatic Peripheral Processing

As mentioned previously, the original explanation for reversing figures centered on the idea that eye movements aimed at different areas of the figure led to the reversal. It is clearly true that eye movements can expedite and cause reversals, but reversals readily occur without any eye movements at all. For example, if one stares at a reversing picture for a long time and then looks at a blank wall, there will be an afterimage of that picture which will also reverse even though eye movements could not have caused it.

After World War II, there was renewed interest in reversing figures because they seemed to be explained by the electrical brain forces postulated by Gestalt theorists as causing the perception of organized wholes. According to the Gestalt ideas of electrical flow, continuous viewing leads to the satiation and fatigue of those circuits. The fatigue increases the resistance, inhibiting the current in these circuits so that the electrical flow moves to different areas, generating the perceptual switch. Over time, the satiation in the new area increases, so the theory goes, while the satiation in the previous area dissipates, and the electrical flow reverts to its original position and the percept similarly reverts back to the original.

Explanations based on brain currents shifted to those based on inhibition among receptors at the eye or ear because recordings of single cells in cortical regions revealed that neural coding at the receptors created localized responses to particular stimulus features, that is, a horizontal bar versus a vertical one. Such neurons would mutually inhibit each other and the dynamics of the inhibition would create the reversals. Each set of neurons would encode one of the incompatible possible percepts and fatigue or satiation of one set would lead to a reversal.

Among the evidence for a peripheral explanation is that the number of switches increases over time. The recovery from satiation would be incomplete after each cycle and diminish over time. Each swap from percept A to B would occur at shorter intervals because the satiation point for A would be reached faster having begun at a higher recovery level due to less recovery (from the previous reversal). Percept B would then reappear more quickly, as shown in

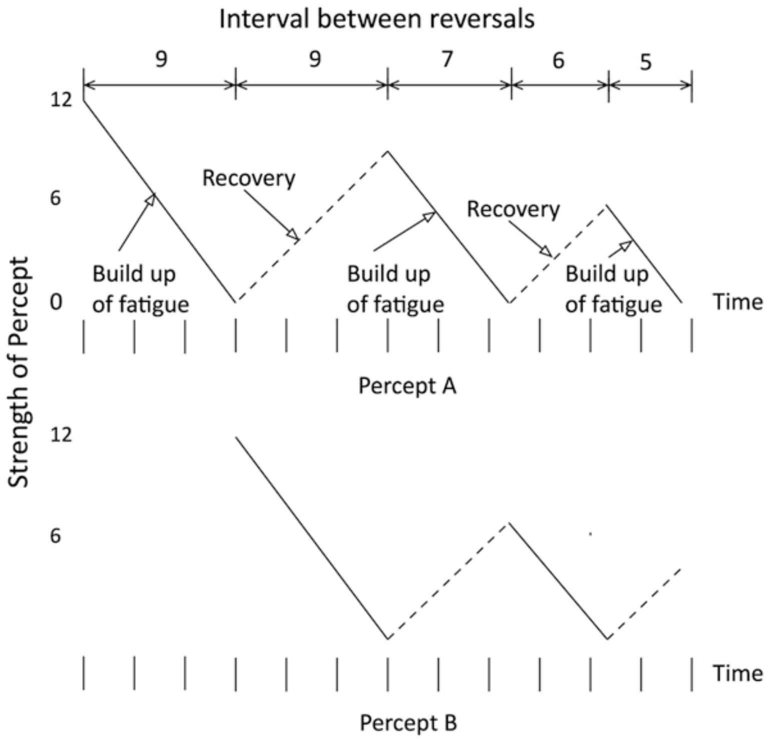

Fig. 3.7 If the fatigue rate and recovery rate are different, then the rate of alternation between the two percepts increases. In the figure, fatigue increases at four steps/time-unit while recovery reduces fatigue at the rate of three steps/time-unit. Percept A builds up to the satiation level over nine time-units where the strength of the percept drops to zero. The percept then switches to B that appears for nine time-units until its strength goes to zero. After Percept B satiates, Percept A reappears but has not recovered fully its former strength so that when it reaches the satiation level after only seven time-units and then Percept B reoccurs. This alternation continues and as the recovery level increases the interval between reversals decreases

Fig. 3.7. If one percept is made more prominent, the duration of the dominant percept does not decrease, but the duration of the weaker percept does.

A related phenomenon is that after a change in retinal location of the reversing figures the original longer intervals between reversals reappear. As stated above, continuous viewing leads to more rapid reversals, but if the same figure is moved to a new retinal position, the timing among reversals reverts to the original timing, as if the figure had never been seen. Fatigue and satiation appears to be limited to a restricted area of the retina.

A second phenomenon that supports peripheral processing (but also top-down processes) occurs when multiple figures are presented at the same time. If two Necker cubes are presented adjacent to each other, they reverse independently of one another (Figure 3.8A). This should not be a surprise as each cube

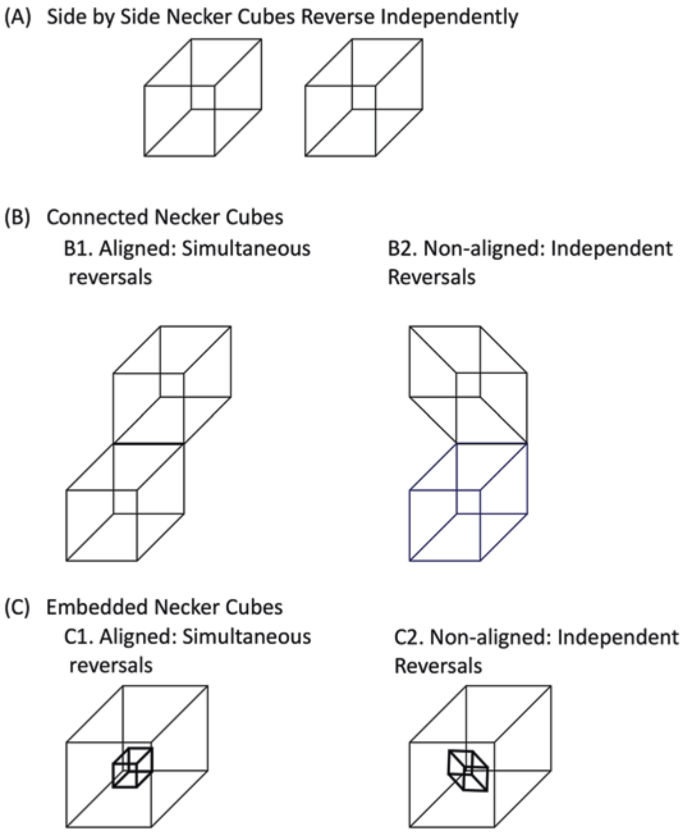

Fig. 3.8 The spatial position and alignment of two Necker cubes determines whether the reversals are independent. (Adapted from Adams & Haire, 1958, 1959)

is localized at a different retinal position and there is no connection between the two cubes. But, if the two cubes are connected, the cubes reverse together if they are aligned as in B1, yet reverse independently if they are not aligned, as in B2. In similar fashion, if an embedded cube is aligned with the larger one, reversals occur simultaneously (C1), but if the embedded cube is not aligned, then the smaller and large cube reverse independently (C2). The independent reversals in B2 and particularly C2 preclude an explanation based only on eye movements, but the simultaneous reversals in B1 and C1 suggest that parts of the visual field that appear grouped undergo the same reversals.

In spite of the large differences in procedures, stimuli, and the nature of the rivalry, the rate and pattern of visual and auditory alternations is identical, suggesting the same underlying neural processes (Brascamp, van Ee, Pestman, & van den Berg, 2005; Pressnitzer & Hupe, 2006). Analyses of the timing of the reversals suggest that the switches occur independently. What this means is that whatever the interval between one switch, the next switch interval could be longer or shorter with equal probability. There are no sequences of fast or slow reversals except by chance. If the frequency of the intervals between switches is tabulated, the best-fitting probability distribution to model the shifts are based on the accumulation of small independent "micro-switches" between A and B until a threshold is reached and then the reversal occurs. Furthermore, if there were simultaneous visual and auditory presentation, for example, visual apparent movement and auditory stream segregation, reversals were independent. The visual reversals occurred at different times than the auditory reversals, with no central mechanism linking the two. These outcomes support the conclusions from Chap. 2 that multisensory cross-modality organization is not as strong as within modality organization.

A third phenomenon illustrating peripheral effects concerns tristable visual figures with three alternative percepts, such as the triangles in Fig. 3.1I that can point in three directions, The perceptual question is whether the three perceptual directions cycle, A-B-C-A-B-C, or undergo alternative unbalanced transitions such as A-B-A-C-A-B-A-C in which one percept serves as a "hub." Wallis and Ringelhan (2013) used a more complex display and found that cycling happens almost exclusively. They argued that the cycling maximizes the interval between the repetitions of any orientation thereby minimizing the fatigue of all the possible percepts

In sum, there is strong evidence that localized peripheral effects can bring about the reversals. Continuous viewing or listening to the same proximal stimulus fatigues a set of cells, so that the percept shifts to other non-fatigued cells that were previously firing at low rates. This shift in active cells gives rise to a different percept; however, the new cells become fatigued in turn so that the percept shifts once again to the firings of the set of now "recovered" original cells. There is reciprocal inhibition. But the peripheral processes do not explain all of the factors involved in the switches. Below we will describe some of the evidence that illustrates the influence of higher-level cognitive processes.

3.4.2 "Top-Down" Active Cognitive Control

The first phenomenon illustrating cognitive control is the ability to control the rate of reversals. If participants are asked to hold on to one percept, then their rate of reversals is roughly half that of passive attending. Conversely, if they are instructed to maximize the rate of alternation, the reversals occur five times the rate under passive attending, from one switch every five seconds to one switch every second. It is important to note that even when participants are asked to hold one percept, they are unable to do so; the alternative always appears. In some experiments (Kornmeier, Hein, & Bach, 2009) participants are asked to hold on to one of the two possible percepts, rather than either percept as above. In following these instructions, participants increase the duration of the primary percept, and tend to decrease the duration of the secondary percept.

A second phenomenon that illustrates top-down processing is that reversals occur only if the participants realize that there is an alternative percept. The participant's prior experience with the alternatives and their intention to reverse the figure are critical. For example, Fisher (1967) has created a set of ambiguous drawings that range from a clearly visible gypsy head with a barely visible girl with a mirror

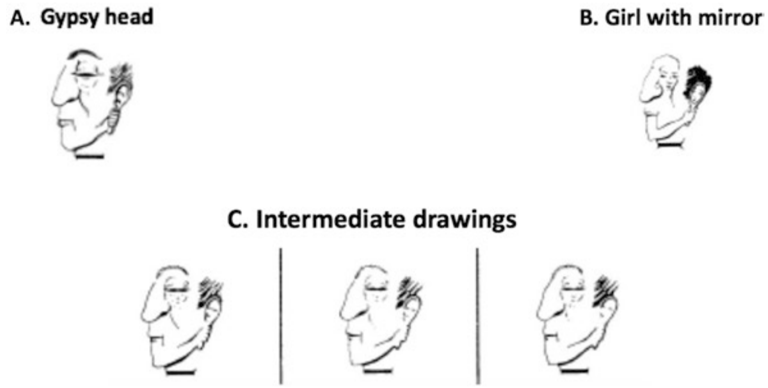

Fig. 3.9 The gypsy head (A) and girl with mirror (B) show the end points of the transformation. It is very difficult to perceive the alternative drawing without first seeing the other end point. But, after seeing both endpoints, it is easy to alternate between the two in the intermediate drawings in (C). (Adapted from Fisher, 1967)

at one end, to a clearly visible girl with mirror and barely visible gypsy head at the other end (Fig. 3.9A, B, & C). I found myself unable to reverse either end drawing without previously viewing a clear representation of the other end drawing.

A third phenomenon is that of set or expectancy. As described in the previous chapter, expectancy and prior learning heavily influence all aspects of perceiving, which the use of reversing figures can point out explicitly. After participants preview an unambiguous version of the figure for short time periods, they are far more likely to see that possibility when viewing an ambiguous version. However, longer viewing periods decrease the probability of seeing the previewed alternative, illustrating the effect of peripheral satiation.

Pastukhov, Vonau, and Braun (2012) have investigated reversals in the perception of rotating donuts. If the outline of the circumference of a donut is rotated around its vertical axis, a striking impression of the donut rotating in depth occurs. The direction of rotation often spontaneously reverses, from clockwise to counterclockwise and vice versa. These possibilities are illustrated in Fig. 3.10.

Pastukhov et al. (2012) were interested in the points at which the donut swapped the direction of rotation since logically the swap should occur at any angle with equal probability. However the majority of alternations occurred at the two orientations where the donut appeared depth symmetric, either flat or edge-on so that only a single surface was visible. The authors argue that the choice of those orientations reflect prior experience with the transitions between objects in the environment. To go from one object representation to another requires prior experience with the validity of that transition.

Finally, Ward and Scholl (2015) investigated whether fleeting cues, possibly unconscious, can bring about switches in the percept. The silhouetted spinning dancer can be seen rotating in either direction, but it is often difficult to reverse and for some individuals it never reverses at all. In the experiment, briefly presented short, white lines indicated which leg was in front as shown in Fig. 3.11, and following those cues there was an increase in the number of reversals.

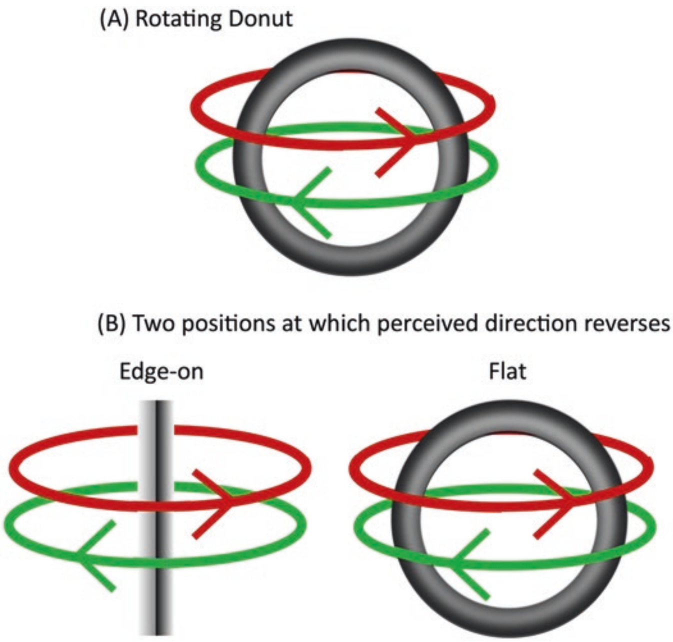

Fig. 3.10 The direction of rotation of the donut reverses spontaneously. There were two places at which the majority of reversals occurred: if the donut was edge-on and when the donut was flat

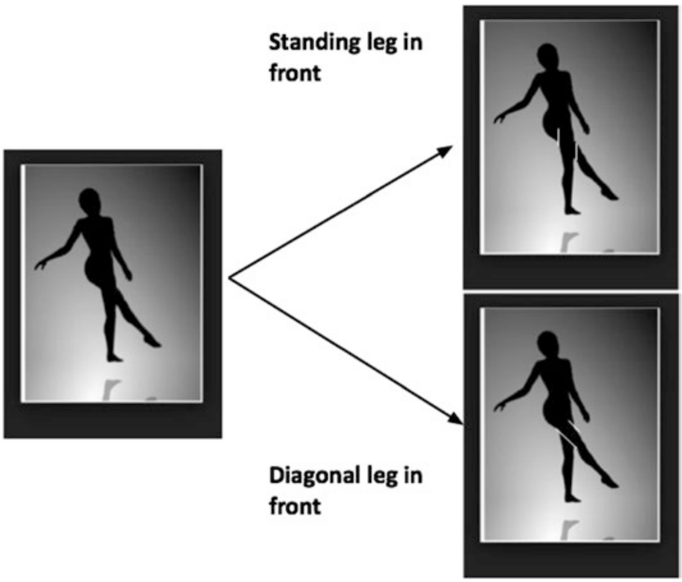

Fig. 3.11 Briefly flashing the white lines indicating which leg is in front can bring about a rotation switch even if the observer is unaware of the lines. (Adapted from Ward & Scholl, 2015)

Spinning Dancer
Wikipedia: http://en.wikipedia.org/wiki/File:Spinning_Dancer.gif
http://www.yale.edu/perception/dancer/

YouTube Videos
Improved 'solution' for spinning dancer girl illusion-alternating directions by Peter Wassink
Bistable perception by Mariushart

3.5 Summary

To return to the basic issue, many different distal objects could generate the same proximal image at the eye, ear, or hand and any distal object can generate many different proximal images. The major difference among perceptual theories lies in how this ambiguity is resolved. As described in the previous chapter, two theoretical positions have emerged. Gestalt theory emphasizes that percepts tend to be the simplest given the actual stimulation. Underlying this notion of *prägnanz* is the belief that the percept follows the operation of a uniform nervous system. Statistical theory on the other hand emphasizes that percepts tend to be the ones most likely to occur in specific situations and that people attend to the most reliable sensations, the Bayesian assumption. Perceptions that originally required conscious calculations become rapid, accurate, and unconscious with practice. As is clear, it is often difficult to distinguish between these two alternative explanations, difficult to define simplest or coherent, and difficult to derive a catalog of all the possible objects in the environment. It is extremely rare for the exact same sensation to occur twice.

As in also clear, no single mechanism is responsible for shifting perceptions, and the interaction among neural and cognitive processes, which of course are also neural, has been a constant theme. Fluctuations in neural responses at both anatomically early and late processing are strongly correlated with the recurrent percepts (Sterzer, Kleinschmidt, & Rees, 2009). But it is unclear whether these fluctuations actually cause the reversals. The dominant stimulus at any time point flows through the nervous system linking with all the connotations of the stimulus. The alternative but suppressed stimulus is disrupted at several points in the pathway although patterns of activity continue to carry information about that percept. There is constant feedback; the present perceptions guide actions that alter perceptions that guide further actions.

We can imagine a simple model for multistability. Suppose there are two alternative percepts, each of which can be represented as a depression in a perceptual energy module, as in Fig. 3.12. There is a barrier between the percepts, and energy is necessary to shift to the other depression and thereby change the percept. This energy comes from a series of small random inputs from different parts of the cortical tracts such as neural satiation, switches in viewing orientation, expectancies, attention shifts, inputs from other senses, and so on. It is a mixed bag of inputs, bottom-up and top-down, some of which seem to lower

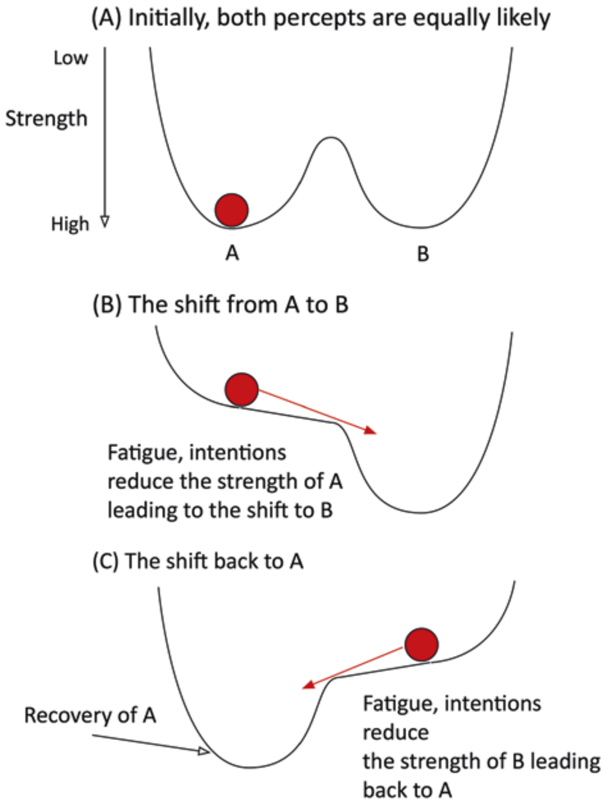

Fig. 3.12 Initially, both percepts could be equally likely based on the depth of the depression. The barrier between the two maintains the initial perception. If the inputs to Percept A reduce the depth of the depression so that it is higher than the barrier, the percept will shift to B. Then, inputs will reduce the depth for Percept B, and at the same time Percept A will recover its strength. The height becomes greater than the barrier and that results in a reversal back to Percept A

the barrier and others which seem to lift the percept. On top of this we should expect each to have different timings between inputs. The percept shifts when the sum of these random inputs reaches a critical value so that the depression becomes higher (i.e., becomes weaker) than the barrier. Fluctuations in the inputs can thus account for the overall randomness of the oscillation between the percepts (Braun & Mattia, 2010).

Models like this are found in several domains. For example, ecologists use such models to explain the shift from one state (e.g., algae-infested turbid water) to another state (clear water), that is, tipping points, (Popkin, 2014). This explanation reinforces the argument from Chap. 1 that perceptions result from the interaction of many neural processes. There is no "single actor" who is in charge (Fig. 3.12).

References

Adams, P. A. A., & Haire, M. (1958). Structural and conceptual factors in the perception of double-cube figures. *American Journal of Psychology, 71*(3), 548–886. https://doi.org/10.2307/1420250

Adams, P. A. A., & Haire, M. (1959). The effect of orientation on the reversal of one cube inscribed in another. *American Journal of Psychology, 72*(2), 296–299. https://doi.org/10.2307/1419384

Alais, D., & Blake, R. (2015). Binocular rivalry and perceptual ambiguity. In J. Wagemans (Ed.), *The Oxford handbook of perceptual organization* (pp. 775–798). Oxford, UK: Oxford University Press.

Brancucci, A., & Tommasi, L. (2011). "Binaural rivalry": Dichotic listening as a tool for the investigation of the neural correlates of consciousness. *Brain and Cognition, 76*, 218–224. https://doi.org/10.1016/jbandc.2011.02.007

Brascamp, J. W., van Ee, R., Pestman, W. R., & van den Berg, A. V. (2005). Distributions of alternation rates in various forms of bistable perception. *Journal of Vision, 5*, 287–298. https://doi.org/10.1167/5.4.1

Braun, J., & Mattia, M. (2010). Attractors and noise: Twin drivers of decsions and multistability. *NeuroImage, 52*, 740. https://doi.org/10.1016/j.neuroimage.2009.12.126

Carter, O., Konkle, T., Wang, Q., Haywood, V., & Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Current Biology, 18*, 1050–1054. https://doi.org/10.1016/j.cub.2008.06.027

Deutsch, D. (1974). An auditory illusion. *Nature, 251*, 307–309.

Fisher, G. H. (1967). Measuring ambiguity. *American Journal of Psychology, 80*(4), 541–557.

Fisher, G. H. (1968). Ambiguity of form: Old and new. *Perception & Psychophysics, 4*(3), 189–192.

Freeman, E., & Driver, J. (2008). Direction of visual apparent motion driven solely by timing of a static sound. *Current Biology, 18*, 1262–1266. https://doi.org/10.1016/j.cub.2008.07.086

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision, 7*(8), 9, 1–15. https://doi.org/10.1167/7.8.9

Harrar, V., Winter, R., & Harris, L. R. (2008). Visuotactile apparent motion. *Perception & Psychophysics, 70*, 807–817. https://doi.org/10.3758/PP.70.5.807

Klink, P. C., van Wezel, R. J. A., & van Ee, R. (2012). United we sense, divided we fail: Context-driven perception of ambiguous visual stimuli. *Philosophical Transactions of the Royal Society B, 367*, 932–941. https://doi.org/10.1098/rstb.2011.0358

Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: When bottom-up and top-down coincide. *Brain and Cognition, 69*, 138–147. https://doi.org/10.1016/j.bandc.2008.06.005

Korte, A. (1915). Kinematoskopische Untersuchungen (Cinematoscopic investigations). *Zeitschrift fur Psychologie, 72*, 193–216.

Lakatos, S., & Shepard, R. N. (1997). Constraints common to apparent motion in visual, tactile and auditory space. *Journal of Exerimental Psychology: Human Perception and Performance, 23*, 1050–1060.

Liaci, E., Bach, M., van Elst, L. T., Heinrich, S. P., & Kommeler, J. (2016). Ambiguity in tactile apparent motion perception. *PLoS One, 11*(5), e0152736. https://doi.org/10.1371/journal.pone.0152736

Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin, 130*, 748–768. https://doi.org/10.1037/0033-2909.130.5.748
Macnik, S. L., Martinez-Conde, S., & Blakeslee, S. (2010). *Sleights of mind*. New York, NY: Henry Holt.
Occelli, V., Spence, C., & Zampini, M. (2010). Assessing the effect of sound complexity on the audiotactile cross-modal dynamic capture task. *The Quarterly Journal of Experimental Psychology, 63*, 694–704. https://doi.org/10.1080/17470210903068989
Pastukhov, A., Vonau, V., & Braun, J. (2012). Believable change: Bistable reversals are governed physical plausibility. *Journal of Vision, 12*, 1–16. https://doi.org/10.1167/12.1.17
Popkin, G. (2014, September 26). On the edge. *Science, 345*, 1552–1554.
Pressnitzer, D., & Hupe, J. M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology, 16*(13), 1351–1357. https://doi.org/10.1016/j.cub.2006.05.054
Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
Sanabria, D., Soto-Faraco, S., & Spence, C. (2005). Assessing the effect of visual and tactile distractors on the perception of auditory apparent motion. *Experimental Brain Research, 166*, 548–558. https://doi.org/10.1007/s00221-005-2395-6
Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Congruency effects between auditory and tactile motion: Extending the phenomenon of cross-modality dynamic capture. *Cognitive, Affective, & Behavioral Neuroscience, 4*, 208–217.
Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Science, 13*, 310–318. https://doi.org/10.1016/j.tics.2009.04.006
Thossen, S., & Bendixen, A. (2017). Subjective perceptual organization of a complex auditory scene. *Journal of the Acoustical Society of America, 141*, 265–276. https://doi.org/10.1121/1.4973806
Wallis, G., & Ringelhan, S. (2013). The dynamics of perceptual rivalry in bistable and tristable perception. *Journal of Vision, 13*, 1–21. https://doi.org/10.1167/13.2.24
Ward, E. J., & Scholl, B. J. (2015). Stochastic or systematic? Seemingly random perceptual switching in bistable events triggered by transient unconscious cues. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 929–939. https://doi.org/10.1037/a0038709
Warren, R. M. (1999). *Auditory perception: A new analysis and synthesis*. Cambridge, UK: Cambridge University Press.
Warren, R. M., & Gregory, R. L. (1958). An auditory analogue of the visual reversible figure. *American Journal of Psychology, 71*, 467–473. https://doi.org/10.2307/1420267