Weakening the Robustness of Perspective: Evidence for a Modified Theory of Compensation in Picture Perception

Tyrone Yang
University of Virginia

Michael Kubovy
University of Virginia

Viewed from the center of projection, a perspective picture presents the pictorial depth information of a scene. Knowing the center of projection, one can reconstruct the depicted scene. Assuming another viewpoint is the center of projection will cause one to reconstruct a transformed scene. Despite these transformations, we appreciate pictures from other viewpoints. The compensation hypothesis states that the visible picture surface allows observers to compensate for transformations by locating the center of projection and experiencing pictorial space from there. We show that observers neither completely compensate nor do they experience transformations of space as geometry would predict. We propose a modified compensation hypothesis according to which different degrees of visibility of the picture surface invoke different degrees of compensation.

Pictures are flat surfaces that show scenes in depth. For the visual system pictures present the problem of integrating conflicting flatness and depth information. Practically, understanding how picture perception resembles or differs from perceiving real space will allow us to create and use spatial displays more effectively. In this paper, we address two problems: (a) how we perceive pictures from different viewpoints even though pictures are geometrically correct for only one viewpoint and (b) how depth information and flatness information interact in perceiving depth in pictures.

A picture mimics the light from a scene to one viewpoint, called the *center of projection*. To all other viewpoints, the picture presents a geometrically transformed pictorial space. However, experience suggests that we can appreciate pictures from many viewpoints. Kubovy (1986) has called this phenomenon the *robustness of perspective*. Some researchers (e.g., Pirenne, 1970; Kubovy, 1986; Goldstein, 1987; Rosinski & Farber, 1980) have proposed that robustness results because seeing the picture surface allows observers to compensate for these transformations of pictorial space. That is, observers perceive the layout of pictorial space as if they were viewing the picture from the center of projection. We call this explanation the *compensation theory of perspective robustness*.

In this paper, we present two versions of the compensation theory—a weak version and a strong one—and show that the strong version is inconsistent with the data of two experiments. In our modified (weak) version of the compensation hypothesis, we suggest that observers neither completely compensate for changes in viewpoint nor does their perception of transformed pictorial space undergo as much transformation as geometry would predict. Increasing the visibility of the picture surface increases our ability to compensate for changes of viewpoint.

Pictures contain some of the spatial information available in natural scenes. A picture acts like a window into a virtual world;[1] it is a frozen cross-section of light to a fixed viewpoint (the center of projection), providing the pictorial depth information appropriate for that viewpoint.

The geometric information in pictures is ambiguous: an infinite number of three-dimensional objects could produce the same projection on the retina. For example, an upright trapezoid or a rectangle tilted in depth could both produce the same trapezoidal projection (Ames, 1955/1968; Sedgwick, 1986). However, even without other sources of depth information (such as motion parallax), the visual system readily selects a three-dimensional interpretation of pictures. The visual system appears to interpret pictures by relying on interpretive predispositions, also known as constraints or assumptions (Ames, 1955/1968). Two such interpretive predispositions have been suggested: (a) observers assume that their viewpoint is at eyeheight above the ground (Cutting, 1987); (b) that certain lines on the picture surface represent parallel lines in the pictorial space and that other lines repre-

---

[1] In fact, the etymology of "perspective" in Latin "to see through," as if through a window.

sent perpendicular lines (Kubovy, 1986). The interpretation of the depth information in pictures, in fact, suffices to guide effective action. For instance, in an experiment by Smith and Smith (1961), observers viewed pictures of a room monocularly through an aperture. Thinking that they were looking into a real room while viewing this picture, observers could accurately throw balls at the targets depicted up to 8 m away. In addition, the perception of pictorial depth is not even negated by motion and stereopsis information, which should specify a planar surface (Hochberg & Brooks, 1987).

A picture is a cross-section of the visual rays projecting to one point, O , called the center of projection (Figure 1). If viewed from a different point, O', which is assumed to be the center of projection, the picture would imply a different spatial layout. To illustrate, let us re-create the virtual space for an observer who is viewing the picture from O, by back-projecting light rays from the eye into virtual space (Figure 1a), a procedure first proposed by La Gournerie (1859) and summarized by Cutting (1987, 1988). We will then examine how the square floor ABCD of a box in the scene may be geometrically reconstructed when the observer's viewpoint, O', moves closer to or farther away from the picture plane, P (Figure 1b). The effect of such a displacement is to compress or expand ABCD in depth, forming a virtual object $A'B'C'D'$. Specifically, if $d(OP)$ is the distance from the center of projection to the picture plane, $d(O'P)$ is the distance from the new observer viewpoint to the picture plane, $d(OA)$ is the depth distance from the center of projection to one corner of the square, and $d(O'A')$ is the transformed depth distance from the new observer viewpoint to the transformed point in virtual space, then: $d(O'A) = d(OA)[d(O'P)/d(OP)]$ where $d(O'P)/d(OP)$ is the proportion of magnification/minification. (Moving O' parallel to the picture plane causes the virtual space to undergo affine shear; Figure 1c.)

Various theories have attempted to characterize the perception of pictures from viewpoints other than the center of projection (Cutting, 1987; Rogers, 1995). In this paper we address the compensation theory of perspective robustness.

The evidence regarding the constancy of perceived spatial layout in the face of viewpoint change is inconsistent. Researchers who support the compensation hypothesis (e.g., Pirenne, 1970; Kubovy, 1986; Rosinski & Farber, 1980; Rosinski, Mulholland, Degelman, & Farber, 1980; Goldstein, 1979, 1987) have argued that when the picture surface is not visible, perceived pictorial space changes as observers view a picture from different viewpoints. When the picture surface is visible, however, observers see the spatial layout as if they were viewing the picture from its center of projection. They correctly see the slant of a pictured object (Rosinski et al., 1980), the spatial layout of depicted scenes (Goldstein, 1979, 1987), and the rectangularity of pictured boxes (Perkins, 1968, 1973). Rosinski and Farber (1980) have summarized the compensation viewpoint as follows: "It seems that we perceive a pictorial representation of space veridically, even when the geometric projection to the eye is greatly distorted. Moreover, pictures apparently look the same regardless of the viewing point" (p. 149). Although no
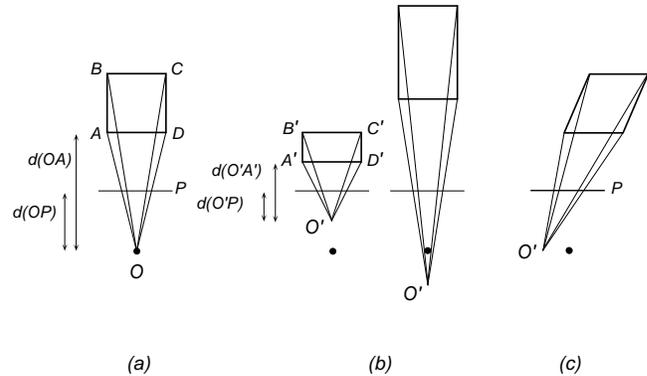

Figure 1. How virtual space changes with observer viewpoint (after Cutting, 1987; Cutting, 1988). The virtual space is seen from top view. $d(OP)$ is the distance from the center of projection to the picture surface. $d(OA)$ is the distance from the center of projection to a point A on the object in virtual space. $d(O'P)$ is the distance from a given observer viewpoint to the picture surface. $d(O'A')$ is the transformed distance from the observer viewpoint to point A' on the object in virtual space. (a) Virtual space as viewed from the center of projection. (b) Viewing from too close or too far respectively causes compression (magnification) and expansion (minification) of virtual space. (c) Viewing from the side causes a shear of virtual space. For all viewing points, $d(O'A')/d(O'P)$ remains constant. $d(O'P)/d(OP)$ is the proportion of magnification/minification.

one has proposed a detailed account of how compensation could occur, Kubovy (1986, p. 137) has suggested that the viewer infers the location of the center of projection, mentally shifts herself to that point (or transforms the picture)—in a process akin to the mental transformations summarized by Shepard and Cooper (1982)—and therefore experiences the picture as if viewing from that point.

Other researchers reject the compensation hypothesis. Camera lenses of different focal lengths create pictures with varying fields of view and centers of projection. These manipulations affect perceived virtual layout despite the visibility of the picture surface (Kraft & Green, 1989). Similarly, varying the observer's viewing distance to a picture changes egocentric and exocentric depth estimates but not estimates of object width and size (Smith, 1958a, 1958b; Smith & Gruber, 1958; Bengston, Stergios, Ward, & Jester, 1980). Furthermore, compression and expansion in perceived space occur even when pictures are viewed binocularly. A related, but not telling observation, is that increasing picture surface visibility seems to increase variability in perceived depth (Lumsden, 1983; Adams, 1972). Finally, Nicholls and Kennedy (1993) asked observers to look through a monocular peephole at pictures of cubes of three perspective convergences at three distances. They found that each picture looked more cube-like when viewed from the center of projection than from other distances.

The preceding review reveals that the evidence in favor of the compensation hypothesis comes from experiments in which the observer's viewpoint was displaced laterally, whereas the evidence against it comes from experiments in which the distance of the observer's viewpoint from the pic-

ture was varied. We therefore chose to test the compensation hypothesis using the latter manipulation. In Experiment 1 we show that when the center of projection is moved away from the observer's viewpoint, observers perceive the transformations of pictorial space. In this experiment we did not know how visible was the picture surface. We therefore conducted a second experiment in which we manipulated the visibility of the pictorial surface. In Experiment 2 we show that—as predicted by the compensation hypothesis—the degree of perceived invariance of pictorial space increases as we increase the visibility of the picture surface. However, the perception of pictorial space is not completely invariant even under the condition in which the picture surface is most visible. We suggest that our data are explainable by a compensation mechanism that is not all-or-none but instead continuously increases in effectiveness as the picture surface becomes more visible.

Experiment 1: Comparing Cubes of Varying Angular Subtenses and Perspectives

Experiment 1 tested whether observers who viewed pictures monocularly from a close distance would perceive the projective transformations of pictorial space. If so, pictures ought to look best from the center of projection and progressively worse as the center of projection moves away from the viewpoint. This experiment replicated and expanded on an experiment by Nicholls and Kennedy (1993). From a fixed distance to the screen, observers viewed pairs of computer-generated line drawings of cubes of varying angular subtenses and perspective convergences.

Geometrically, the larger the angular subtense of a three-dimensional object (the more of one's visual field it takes up) the more perspective convergence the object has regardless of the actual size of the object (Figure 2). In other words, the larger the angular subtense of an object, the greater the ratio of the projected size of the nearer parts to the farther parts. The geometrically appropriate perspective convergence of an object is thus linked to its angular subtense. The same is true with pictured objects viewed from the center of projection.

However, when pictures are magnified (when the viewpoint is closer to the picture than the center of projection) or minified (when the viewpoint is farther from the picture than the center of projection) the angular subtense of pictured objects changes, but the perspective convergence on the picture surface does not. If observers perceive transformations of pictorial space, then pictured cubes should look best viewed from the center of projection when perspective convergence is appropriate for a given angular subtense.

Method

Participants.

Twelve graduate (including the first author) and undergraduate students (nine male and three female) at the University of Virginia viewed the displays. Nine of the participants received payment while three volunteered. All had normal or corrected-to-normal vision.
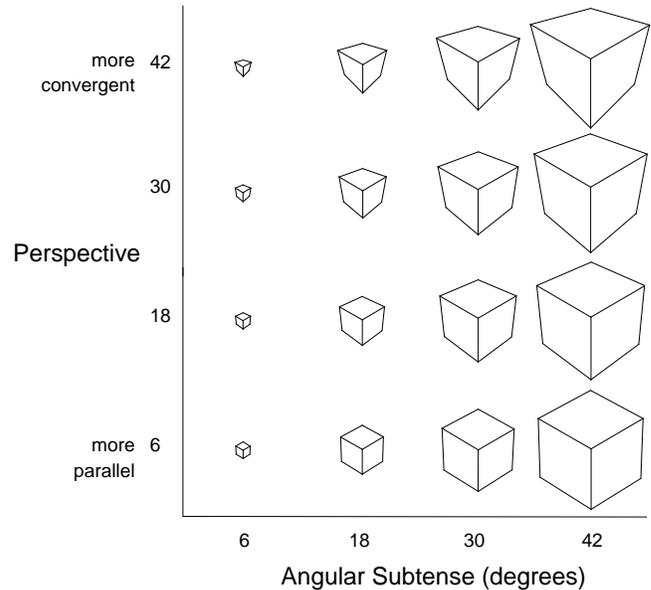

Figure 2. Stimuli for Experiment 1. The sixteen cubes in Experiments 1 and 2 varied in perspective convergence and angular subtense. The units along the Angular Subtense dimension (*x*-axis) refer to the angular subtense of the image at the fixed viewpoint 20 cm from the screen. The units along the perspective dimension (*y*-axis) indicate the amount of perspective convergence appropriate for a cube of that angular subtense in degrees. Thus, the cubes along the Main Diagonal from lower-left to upper-right have centers of projection at 20 cm from the screen and perspective convergences appropriate for their angular subtense. Table 1 lists the locations of the centers of projection of all the pictures.

Apparatus.

The stimuli were created on a Silicon Graphics Iris Indigo 16-inch monitor, whose screen dimensions were 29.3 cm (horizontal) by 23.4 cm (vertical). The observers viewed the drawings monocularly, wearing an eye patch over the eye of their choice. An adjustable headrest positioned the observer's viewing eye 20 cm from the center of the screen. An incandescent light illuminated the room from above. The observer used the two buttons of a computer mouse and one button on the keyboard to indicate responses.

Stimuli.

The stimuli were sixteen black line drawings of cubes on a white background, oriented so that one of the corners of the cube faced the viewer, and a perpendicular from the observer's eye to the picture plane would lie along the main diagonal of the cube (Figure 2). We manipulated two independent variables: *angular subtense* and *perspective convergence*. By angular subtense we mean the visual subtense of the image on the screen when viewed from the fixed viewpoint of the observer. By perspective convergence X we mean the perspective that would be geometrically appropriate for pictures of real cubes subtending an angle X. Our sixteen stimuli were created by crossing four perspective convergences: 6, 18, 30 and 42° with four angular subtenses: 6, 18, 30 and 42°.

Table 1
Distances from the picture plane (in cm) of the centers of projection of the cube pictures used in Experiment 1

angular subtense	Perspective convergence[a]			
	6°	18°	30°	42°
42°	146.3	48.5	28.7	20.0
30°	101.4	33.8	20.0	13.9
18°	60.4	20.0	11.8	8.3
6°	20.0	6.6	3.9	2.7

[a]Convergence appropriate for a cube subtending this visual angle.

Figure 3. Data from Experiment 1. The number in each cell is the number of times each of the sixteen cubes in Experiment 1 was picked as the better-looking cube. To give a graphical overview of the pattern in the data, this amount is also indicated for each cube by a circle whose area is directly proportional to the number of times the cube was picked. Note that the number of times a cube was picked as better-looking is greatest along the Main Diagonal (from lower left cell to upper right cell), where perspective convergence is appropriate for angular subtense. The number of times a cube is picked falls off as angular subtense and perspective convergence become increasingly mismatched (i.e., progressing from the Main Diagonal to either upper left of lower-right corners).

Thus, the cubes along the diagonal of the array in Figure 2 from lower-left to upper-right have centers of projection at 20 cm from the screen and perspective convergences appropriate for their angular subtense. Table 1 lists the locations of the centers of projection of all the pictures. All possible pairings of the cubes yielded $(16 \times 15)/2 = 120$ pairs. Observers saw all pairs, in random order, one pair at a time.

Procedure.

For each trial, one of the two pictures in the pair randomly appeared on the screen first. Observers viewed the two cubes in alternation by pressing a mouse button to toggle back and forth between each picture. The observer's task was to select the picture that looked better as a picture of a cube. We define "better-looking" as follows: Imagine the objects represented by the two pictures and choose the picture that best represents an object with 90° corners and edges of equal length. The observer indicated which of the two pictures looked better by pressing a button in the keyboard while the better-looking picture was on the screen. On a subsequent screen, the observer indicated on a 5-point scale that this cube looked "as good as," "just barely better than," "a bit better than," "a fair amount better than," or "a great deal better than" the other cube in the pair.

Results

A circle plot (Figure 3) depicts the total number of times observers chose each of the sixteen pictures as looking more cube-like, regardless of the what other picture each of the sixteen pictures might have been paired with. In Figures 2 and 3, we refer to the diagonal (from lower left cell to upper right cell), along which perspective convergence matches angular subtense, as the Main Diagonal. Figure 3 shows that the number of times observers picked a picture as better-looking is greatest along the Main Diagonal. The number of times observers picked a picture decreases as angular subtense and perspective convergence become increasingly mismatched (i.e., progressing from the Main Diagonal to either upper left or lower-right corners).

If observers perceive transformations of space in pictures, then cubes closer to the Main Diagonal should look best; cubes sh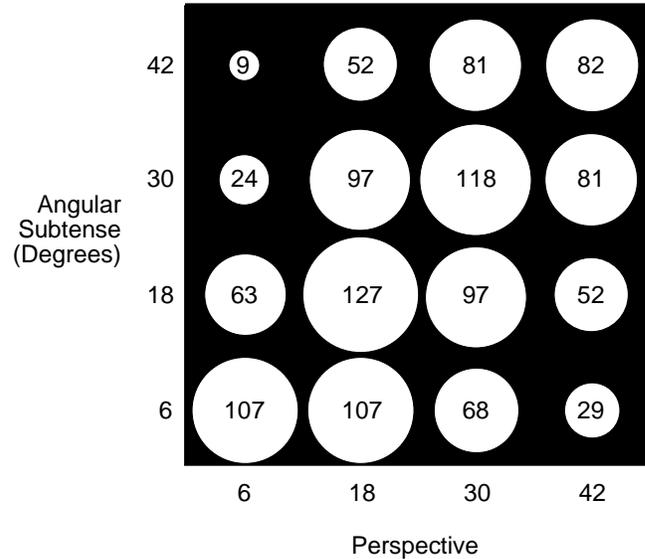ould look progressively worse as the match between perspective convergence and angular subtense worsens. This is indeed what we find. As a measure of this match, we defined a variable Perspective Appropriateness, A ($1 \leq A \leq 4$) where $A = 4$ means that the perspective convergence is most appropriate for the angular subtense of the cube and $A = 1$ means that the perspective convergence is least appropriate. Thus for cubes along the Main Diagonal, $A = 4$ (Table 2). For each trial we computed ΔA, the difference in the Perspective Appropriateness between the cubes presented on that trial ($-3 \leq \Delta A \leq 3$).

We plotted the probability that observers thought that the second cube was better than the first (Figure 4), and found that, as expected, the more ΔA favored the second cube (i.e., the larger ΔA) the more often observers chose the second cube.

We also analyzed the observers' ratings of goodness. Our dependent variable was Difference in Rated Goodness, ΔG, which measured how much better or worse observers thought the first cube of a pair looked as compared to the second cube. The rating scale used by observers ("as good as," "just barely better than," "a bit better than," "a fair amount better than," or "a great deal better than") was coded from 0 to 4 ($-4 \leq \Delta G \leq 4$). $\Delta G = 0$ meant one cube looked "as good as" the other cube, $\Delta G > 0$ meant that the first cube in each pair looked better than the second cube, and $\Delta G < 0$ meant that it looked worse. The boxplots in Figure 5 (see Appendix

Table 2

Values of derived variable of perspective appropriateness of cube pictures used in Experiment 1

angular subtense	Perspective convergence[a]			
	6°	18°	30°	42°
42°	1[b]	2	3	4
30°	2	3	4	3
18°	3	4	3	2
6°	4	3	2	1

[a]Convergence appropriate for a cube subtending this visual angle.
[b]A value of 1 means that the perspective convergence is geometrically least appropriate for the angular subtense, and a value of 4 means that the perspective convergence is appropriate.
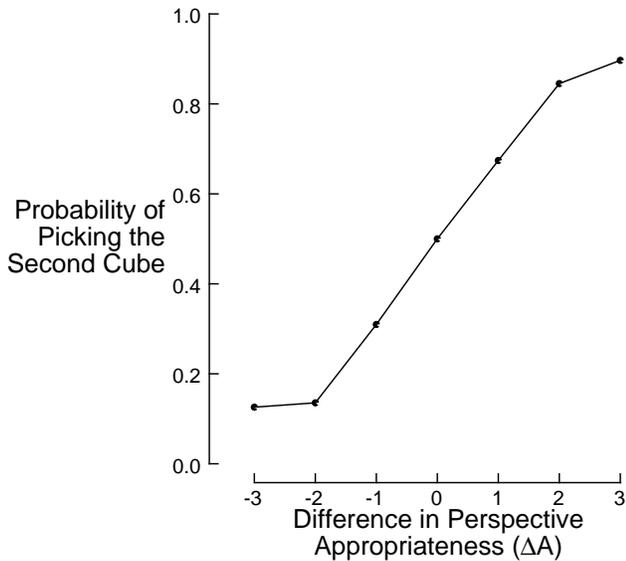

Figure 4. Data from Experiment 1. Proportion of times the second cube in a given pair was picked as a function of Difference in Perspective Appropriateness (ΔA).

Figure 5. Boxplots of Difference in Rated Goodness (ΔG) as a function of the derived variable of Perspective Appropriateness (ΔA) in Experiment 1. The shaded area represents 95% confidence intervals around the median. Appendix A explains the parts of the boxplot in greater detail.

ure 4) and more highly observers rated its relative goodness (Figure 5).

Our results replicate the Nicholls and Kennedy (1993) finding that a cube of a given perspective convergence looks best (most cube-like) at its geometrically appropriate angular subtense. However, we did not replicate their finding that a picture of a cube with moderate perspective convergence looked best regardless of angular subtense. Instead, we found that within each angular subtense, a picture of a cube with the appropriate perspective looked best. We looked to the procedural differences between our Experiment 1 and the Nicholls and Kennedy experiments for an explanation of this difference. For instance, in our experiment we held picture surface information constant by having the observer view all pictures from a fixed distance from the computer monitor; Nicholls and Kennedy used pen-and-paper drawings and varied the viewing distances of pictures to manipulate angular subtense. However, we do not know how these procedural differences could yield different results.

All the cube drawings used in the experiment should look acceptable under Perkins's (1968, 1973) laws.[2] With overhead

A for a specification of the parts of a boxplot) show how ΔA predicts ΔG. When $\Delta A = 0$ (i.e., when comparing two cubes from the Main Diagonal of Figure 2), we predict that $\Delta G = 0$ i.e., that the intercept of the regression of ΔG on ΔA will be 0. Indeed we found that the intercept was not different from 0: -0.03, 95%CI= $(-0.11, 0.06)$. As we move away from the Main Diagonal, ΔA grows or decreases, and we expect ΔG to change monotonically. We found, as expected, that the slope of the linear regression was reliably greater than 0: 0.81, 95% CI= $(0.58, 1.03), R^2 = 28.4\%$.

Discussion

Observers most often chose the pictures whose angular subtense matched their perspective convergence (Figure 3). The more the pictures differed in Perspective Appropriateness, the more frequently the better cube was chosen (Fig-
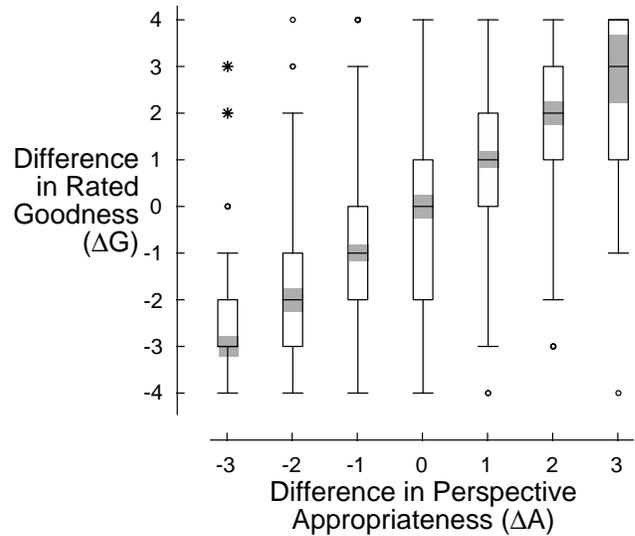
---

[2] Laws that describe the possible parallel projections of the corners of rectangular solids. Corners for which all three faces which comprise the corner are visible are called "fork junctures." Corners for which two of the three faces are visible are called "arrow junctures." Kubovy (1986) states these laws as follows: "Perkins's first law: A fork juncture is perceived as the vertex of a cube if and only if the measure of each of the three angles [as measured on the picture surface] is greater than 90°. Perkins's second law: An arrow juncture is perceived as the vertex of a cube if and only if the measure of each of the two angles [as measured on the picture surface] is less than 90° and the sum of their measures is greater than 90°" (p. 99). Kubovy suggests that for perspective pictures which

illumination and close monocular viewing, observers were undoubtedly aware of the pictorial nature of the stimuli. If these viewing conditions were enough to constitute a visible picture surface, then these stimuli ought to invoke compensatory processes as argued by Kubovy (1986, Chap. 7). However, with the current experimental conditions, observers did not compensate. Since the effect of the visible picture surface is central to the compensation theory, in Experiment 2, we studied the effect of manipulating the visibility of the picture surface.

Experiment 2: The Effect of Picture Plane Visibility on Perceiving Pictorial Space

The compensation hypothesis implies that we either see the surface of a picture or not, and that when we do, we compensate completely for changes in our viewpoint. In Experiment 1, we refuted this hypothesis by showing that when the picture surface was at least moderately visible, observers thought that pictures looked best when viewed from the center of projection. In Experiment 2 we tested a modified version of the compensation hypothesis that acknowledges the possibility of degrees of picture surface awareness, and different degrees of compensation. In this experiment, we expand on Experiment 1 and Nicholls and Kennedy's experiment by examining the perception of line drawings of cubes under differing conditions of picture surface visibility. Observers magnified or minified pictures of varying perspective convergence until they looked best as pictures of cubes.

In our task, an ideal observer who does not compensate would perceive transformed pictorial space and would adjust the picture so that its center of projection was at the eye. Such an ideal non-compensating observer would always choose the geometrically appropriate angular subtense in response to a given perspective convergence. Thus this observer'sfunction relating chosen angular subtense to perspective convergence would be steep and exhibit no variability. On the other hand, an ideal observer who compensated completely would consider all angular subtenses equally appropriate for a given perspective convergence. That is, an ideal compensating observer would not prefer the geometrically appropriate angular subtenses over inappropriate ones. This observer's function relating chosen angular subtense to perspective convergence would be flat and exhibit great variability. Any intermediate result would suggest that the observer was able to compensate partially.

Method

Participants.

Forty undergraduate students at the University of Virginia participated to fulfill a requirement for a psychology class. All participants had normal or corrected-to-normal vision. All participants were naive to the experimental hypothesis.

Design.

Picture surface visibility varied between participants in a 2 eyes (monocular vs. binocular) × 2 lighting (light vs. dark) design to yield 4 conditions of varying picture surface visibility.

Apparatus.

We presented the stimuli on a computer monitor to observers whose viewing position was fixed by a headrest. In some of the conditions (described below) a large flat cardboard sheet with a square viewing hole was interposed between the observer and the monitor. The stimuli were blue line-drawings of cubes on a black background created on a Silicon Graphics Iris Indigo and displayed on a 16-inch monitor. The actual screen dimensions were 29.5 cm wide by 22.1 cm tall. An adjustable headrest positioned the observer's eyes 20 cm from the screen. In the monocular conditions, the headrest positioned the observer's viewing eye 20 cm in front of the center of the screen. In the binocular conditions, the headrest positioned the center point between the observer's eyes 20 cm in front of the center of the screen. The matte black walls of the experiment room absorbed light from the monitor screen.

For the monocular/dark condition, observers viewed the pictures in the dark through a 6 × 6 cm viewing aperture in a black cardboard reduction screen, 77.5 × 77.5 cm, placed 15 cm from the monitor screen. The structure of the set up was never concealed from observers as they sat down for the experiment. We used this reduction screen because the dim light from the monitor was enough to make the edges of the monitor screen faintly visible. With the lights off, the edges of the viewing aperture in the reduction screen cropped the light reflected off the edges of the monitor. In the monocular/dark condition, the picture surface as well as the edges of the monitor screen and the edges of the viewing aperture were invisible during the experiment. In the binocular/dark condition, observers viewed the stimuli in the dark without a reduction screen; the edges of the monitor were faintly visible in this condition. In the two light conditions, there was also no reduction screen. An incandescent light illuminated the room from above and the edges of the monitor were clearly visible. The visible surface texture of this high-resolution monitor display was comparable to a large photograph.

Stimuli.

The computer-generated cubes were similar to those used in the previous experiment, except that the lines were blue and the background of the screen was black. We used cubes of ten different perspective convergences. These convergences were appropriate for cubes that subtended 6, 10, 14, 18, 22, 26, 30, 34, 38 and 42°.

Procedure.

Each observer viewed each of these cubes ten times. The stimuli appeared in random order. For each trial, the initial image of the cube randomly subtended 6, 10, 14, 18, 22, 26, 30, 34, 38, 42° visual angle, as measured from 20 cm from the monitor screen. Observers used two buttons of the mouse

---

follow these laws, observers compensate and therefore do not see these pictures as distorted.

Monocular/Dark Monocular/Light

Angular Subtense (degrees)
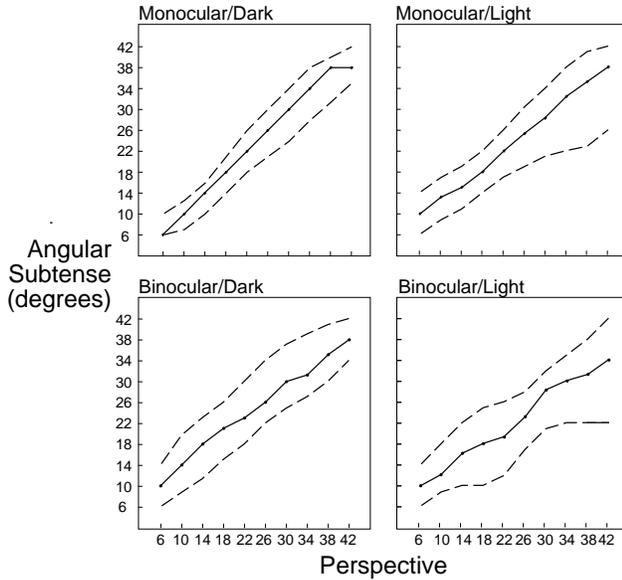
Binocular/Dark Binocular/Light

Perspective

Figure 6. Plots of angular subtense adjustment by perspective for the four viewing conditions. The plots are median traces and hinge traces smoothed using the 3*R* procedure suggested by Tukey (1977). Solid lines connect the smoothed medians and dashed lines connect the smoothed hinges (75th percentiles and 25th percentiles).

to adjust the image size of each cube to be larger or smaller, in discrete steps, to subtend 6, 10, 14, 18, 22, 26, 30, 34, 38, 42° visual angle. Observers hit a key on the keyboard to indicate the image size that looked best, such that if the represented object were real, convex, physical object rather than just a picture, this object would have 90-degree corners and edges of equal length. There were no time restrictions on the task.

Results

In all viewing conditions, observers adjusted cubes of greater perspective convergence to greater angular subtenses. Binocular viewing and illumination of the display increased the variability of the observers' choice of angular subtense and decreased the slope of the function relating Chosen Angular Subtense to perspective convergence (Figure 6 and Table 3).

Binocular Viewing and Display Illumination Increased Response Variability.

Greater response variability indicates that the observers are less sensitive to the appropriate level of angular subtense for each level of perspective, and that their responses conform better to the compensation hypothesis. Binocular viewing and illuminating the display increased the variability of observers' responses. In Figure 6, the distance between the dashed lines, which represent the upper and lower quartiles of the distribution of the data is lowest for the Monocular/Dark condition, higher for the Monocular/Light and the Binocular/Dark conditions, and highest for the Binocular/Light condition. To quantify this increase in variability,

we calculated absolute residuals around the individual linear regression lines for each of the four conditions. A square root transform symmetrized the distribution of the positively skewed residuals (from a skewness of 1.11 to 0.18). We computed the mean root absolute residuals (MRAR) for each observer and entered these into a two-way (Light vs. Dark × Monocular vs. Binocular) ANOVA to obtain standard errors (SE) for the various viewing conditions (Table 3). We found that MRAR(Monocular/Dark) < MRAR(Monocular/Light) ≈ MRAR(Binocular/Dark) < MRAR(Binocular/Light).

Binocular Viewing And Display Illumination Decreased Function Slope.

Recall that the lower the slopes of the regression of the observers responses on Chosen Angular Subtense, the more consistent the data are with the compensation hypothesis. In addition to using the raw dependent variable in our regression, we also performed the regression on the folded-log (flog) transform (Tukey, 1977, Chap. 15) of the data, to overcome problems that may stem from the bounding of the responses at top and bottom. We compared the average slopes of the four viewing conditions. In one analysis, we compared the average slopes of the untransformed data. In another analysis, we compared the estimates of the raw slopes derived from the flog transformed data. The slopes were compared in two-way (Light vs. Dark × Monocular vs. Binocular) ANOVAs to obtain SEs for the various viewing conditions (Table 3). Both for the raw and the transformed dependent variables, slope(Monocular/Dark) > slope(Monocular/Light) ≈ slope(Binocular/Dark) > slope(Binocular/Light).

Discussion

These data join those of Experiment 1 in refuting a strong version of the compensation theory. Even under binocular viewing of a lighted monitor screen, observers adjusted the angular subtenses of the pictures so that the center of projection was approximately at their eye (Figure 6). However, two patterns in the data support a modified compensation theory according to which (a) compensation is never complete, and (b) compensation increases with increasing picture surface visibility: the variability of the data increased and the slope decreased.

Other experiments (e.g., Sedgwick, Nicholls, & Brehaut, 1995; Koenderink, Doorn, & Kappers, 1994; Deregowski & Parker, 1996; Eby & Braunstein, 1995; Hagen & Jones, 1981) have suggested that the visible picture plane flattens the perceived depth of pictorial space (but see Adams, 1972). Since our methods do not ask for a direct estimate of perceived depth, we do not know whether our observers perceived compressed pictorial depth. However, if observers perceived compressed pictorial depth, a more visible picture plane might cause them to prefer drawings with a more parallel perspective. Specifically, if the slant of one of the faces of the cube is perceived as more parallel to the picture plane, then it should look more like a trapezoid than a square face slanted in greater depth, causing observers to prefer a more parallel perspective convergence. If anything, Figure 5

Table 3
Parameter estimates and statistics for Experiment 2

viewing condition	Untransformed data			Flog transformed data	
	slope[a]	MRAR[b]	R^2	Est. raw slope[c]	slope of flog[d]
monoc./dark	0.87	1.97	.71	1.07	0.24
binoc./dark	0.70	2.57	.51	0.81	0.20
monoc./light	0.66	2.75	.46	0.75	0.20
binoc./light	0.41	3.44	.16	0.53	0.14
both monoc.	0.76	2.36	.58	0.91	0.22
both binoc.	0.55	3.00	.30	0.67	0.17
both dark	0.79	2.27	.60	0.94	0.22
both light	0.53	3.10	.28	0.60	0.17

Standard errors (SE) in columns: [a]SE = 0.04. [b]SE = 0.09. [c]SE = 0.02. [d]SE = 0.01.

shows the opposite trend: with the most visible picture surface, the slopes of the regression lines decrease, meaning that observers matched a picture of a given angular subtense is with a greater (less parallel) perspective convergence.

General Discussion

According to the compensation theory a picture surface is either visible or not. We can summarize the theory with two propositions: (a) When the picture surface is invisible, observers do not compensate at all, and perceived space undergoes as much transformation as geometry predicts. (b) When the picture surface is visible, observers compensate fully for shifts in their viewpoint. Whereas proposition (a) is uncontroversial, proposition (b) has not been widely accepted.

The point of departure of the present paper is that picture surfaces do not fall into one of two categories: visible or not. The conclusion of this paper is that in the matter of compensation for transformations of viewpoint, picture perception is not an all-or-none process, as the original compensation theory proposed. Observers neither completely compensate for shifts in their viewpoint nor do changes in viewpoint cause observers to perceive as great a transformation in pictorial space as geometry predicts. How much they compensate depends on how visible we make the picture surface. We call this the *modified compensation theory*.

In Experiment 1, we showed that the original compensation theory is incorrect. Specifically, we created conditions under which we had reason to expect that the observers could see the picture surface, and found that the representation of the cube that they preferred was strongly influenced by the discrepancy between the center of projection of the picture and the observer's viewpoint. The registration of the appropriateness of a particular viewpoint affects the acceptability of a picture as a surrogate for a particular spatial scene.

In Experiment 2, we manipulated the visibility of the picture surface. From this experiment we drew three conclusions that are consistent with a weak compensation hypothesis: (a) The increasing functions shown in Figure 6 imply that observers failed to fully compensate in any of the four conditions of picture surface visibility. (b) On the other hand, the increase in variability in observers' adjustments as surface visibility increased implies that the more visible the picture surface the more observers accept viewpoints that deviate from the center of projection. This variability could reflect the operation of a compensation mechanism. (c) Finally, this pattern of continuous increase of variability with the visibility of the picture surface implies that the operation of such a compensation mechanism is not all-or-none.

The influence of the picture surface on the perception of pictorial space has been discussed in other ways. Notably, Sedgwick and colleagues have proposed that picture perception could be characterized as a process of cross-talk between the perception of pictorial space and the simultaneous perception of the flat projection on the picture surface. The perception of the flat projection could affect the perception of objects in pictorial space in two ways. First, the perceived proportions of the flat projections may bias the perceived proportions of the objects in pictorial space (Sedgwick et al., 1995). Second, the cross-talk account provides an alternative explanation for the robustness of perspective (Sedgwick, 1991). As the observer changes viewing position, the location of points in virtual space changes, but the location of the projection of these points on the picture plane does not. This lack of change on the picture surface would "result in some degree of 'constancy' in the virtual space of the picture in the sense that the virtual layout would not be as distorted as the optic array information would predict" (Sedgwick, 1991, p. 474). It is not clear to us whether this second aspect of the cross-talk hypothesis contributes only to the impression of constancy without changing the shape of perceived pictorial space or whether the combination of the two percepts results in a perceived pictorial space whose shape is closer to the scene that the picture is intended to represent.

If we consider the latter interpretation, then both the cross-talk hypothesis and the modified compensation theory suggest that perceived pictorial space is not completely constant but that the awareness of the picture surface makes perceived

pictorial space less distorted than geometry would predict. The amount of constancy would vary continuously with the visibility of the picture surface. The cross-talk hypothesis attributes the relative constancy of perceived space to the perception of the projected shape on the picture surface. The modified compensation hypothesis attributes this relative constancy to a process of transforming pictorial space into what it would look like if the picture were viewed from the center of projection.

This interpretation of the cross-talk hypothesis would also seem to imply, however, that an awareness of the flat projection would distort perceived pictorial space even when the picture is viewed from the center of projection. The modified compensation theory predicts that when viewed from the center of projection, the shape of perceived pictorial space would not change with the visibility of the picture surface.

It would be important to test whether our conclusions would apply to pictorial scenes with richer information about spatial layout. Specifically, we asked observers to judge the acceptability of pictures based on whether the cubes represented in these pictures had equal sides and right angles. The use of isolated line drawings of cubes may most effectively evoke a compensation mechanism for two reasons. First, such pictures may lack spatial information which would be invariant over change of viewpoint of the picture viewer. Other researchers (e.g., Rogers, 1996; Sedgwick, 1991) have suggested that this invariant information, which would normally be used in perceiving real scenes and realistic pictures, is the basis of the robustness of pictures. Second, distortions might be most easily registered with pictures of rectilinear objects because of their regularity; at the same time, these pictures contain the most explicit geometric information which could allow the visual system to compensate by reconstructing the center of projection (Kubovy, 1986, pp. 89–92).

In scenes of realistic layout, the information needed to reconstruct the center of projection might be less explicit, but realistic pictures might also contain information which is invariant over change in observer viewpoint. The use of such information could work in concert with a compensation mechanism. We speculate that even with information-rich pictures, picture viewers would still register the transformed aspects of pictorial space. Picture viewers would be able to compensate for moderate amounts of transformation. More extreme transformations may not be compensated for, but also may not be as noticeable if their salience depends on the relative amounts of transformed and untransformed information (Goldstein, 1987) as well as on the aspects of the picture observers must attend to for the particular use of the picture. How multiple processes in picture perception might interact is still an open question.

Finally, even though we have shown that the ability to compensate increases (or that the ability to perceive transformed pictorial space decreases) with increasing picture surface visibility, more extreme conditions of picture surface visibility could still be tested (Figure 7). We could push the compensation theory to its limit by testing whether transformed pictorial space is still perceived in situations in which
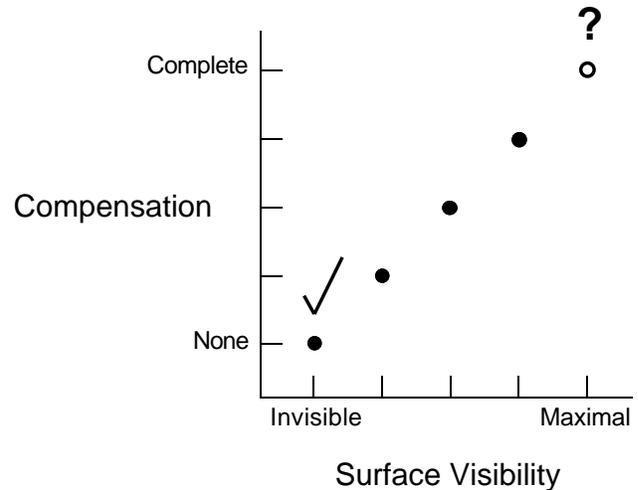

Figure 7. A graphic illustration of our modified compensation theory. The check mark by the bottom-most filled dot represents the consensus that compensation does not occur when the picture surface is invisible. Our experiments have shown that varying levels of picture surface visibility can induce varying levels of compensation (other three filled dots). We have yet to test whether maximal picture surface visibility can result in complete compensation (indicated by the question mark).

surface texture (as with an oil painting on canvas) makes the picture plane even more salient. It may be that compensation in the perception of pictures is never complete, even under these circumstances.

References

Adams, K. R. (1972). Perspective and the viewpoint. *Leonardo*, *5*, 209–217.

Ames, A. (1968). An interpretive manual for the demonstrations in the psychology research center, princeton university: The nature of our perceptions, prehensions and behavior. In W. H. Ittelson (Ed.), *The ames demonstrations in perception* (pp. 1–130). New York, NY: Hafner Publishing Company, Inc. (Original work published 1955)

Bengston, J. K., Stergios, J. C., Ward, J. L., & Jester, R. E. (1980). Optic array determinants of apparent distance and size in pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 751–759.

Cutting, J. E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 323–334.

Cutting, J. E. (1988). Affine distortions of pictorial space: some predictions for Goldstein (1987) that La Gournerie (1859) might have made. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 305–311.

Deregowski, J. B., & Parker, D. M. (1996). The depiction of distance: A Bartellian analysis. *Perception*, *25*, 177–185.

Eby, D. W., & Braunstein, M. L. (1995). The perceptual flattening of three-dimensional scenes enclosed by a frame. *Perception*, *24*, 981–993.

Goldstein, E. B. (1979). Rotation of objects in pictures viewed at an angle: Evidence for different properties of two types of pictorial

space. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 78–87.

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 256–266.

Hagen, M. A., & Jones, R. K. (1981). Picture surface information as a determinant of pictorial perception. In J. Long & A. Baddeley (Eds.), *Attention and performance ix* (pp. 117–134). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hochberg, J., & Brooks, V. (1987). The perception of motion pictures. In E. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 10: Perceptual ecology, pp. 257–304). New York, NY: Academic Press.

Koenderink, J. J., Doorn, A. J. van, & Kappers, A. M. L. (1994). On so-called paradoxical monocular stereoscopy. *Perception*, *23*, 583–594.

Kraft, R. N., & Green, J. S. (1989). Distance perception as a function of photographic area of view. *Perception & Psychophysics*, *45*, 459–466.

Kubovy, M. (1986). *The psychology of perspective and renaissance art.* New York, NY: Cambridge University Press.

La Gournerie, J. d. (1859). *Traité de perspective linéaire contentant les tracés pour les tableaux plans et courbes, les bas-reliefs et les décorations théatrales, avec une théorie des effets de perspective* [Treatise on linear perspective containing drawings for paintings, architectural plans and graphs, bas-reliefs and theatrical set design; with a theory of the effects of perspective]. Paris, France: Dalmont et Dunod.

Lumsden, E. A. (1983). Perception of radial distance as a function of magnification and truncation of depicted spatial layout. *Perception & Psychophysics*, *33*, 177–182.

Nicholls, A. L., & Kennedy, J. M. (1993). Angular subtense effects on perception of polar and parallel projections of cubes. *Perception & Psychophysics*, *54*, 763–772.

Perkins, D. N. (1968). Cubic corners. *Quarterly Progress Report, MIT Research Laboratory of Electronics*(89), 207–214. (Reprinted in Harvard Project Zero Technical Report no. 5, 1971)

Perkins, D. N. (1973). Compensating for distortion in viewing pictures obliquely. *Perception & Psychophysics*, *14*, 13–18.

Pirenne, M. H. (1970). *Optics, painting and photography.* Cambridge, UK: Cambridge University Press.

Rogers, S. (1995). Perceiving pictorial space. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (pp. 119–163). New York, NY: Academic Press.

Rogers, S. (1996). The horizon-ratio relation as information for relative size in pictures. *Perception & Psychophysics*, *58*, 142–152.

Rosinski, R. R., & Farber, J. (1980). Compensation for viewing point in the perception of pictured space. In M. Hagen (Ed.), *The perception of pictures* (Vol. 1: Alberti's window: The projective model of pictorial information, pp. 137–178). New York, NY: Academic Press.

Rosinski, R. R., Mulholland, T., Degelman, D., & Farber, J. (1980). Picture perception: An analysis of visual compensation. *Perception & Psychophysics*, *28*, 521–526.

Sedgwick, H. A. (1986). Space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1: Sensory Processes and Perception, pp. 21-1–21-57). New York, NY: John Wiley and Sons.

Sedgwick, H. A. (1991). The effects of viewpoint on the virtual space of pictures. In S. R. Ellis, M. K. Kaiser, & A. C. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 460–479). London, UK: Taylor and Francis.

Sedgwick, H. A., Nicholls, A. L., & Brehaut, J. (1995, May). *Perceptual interaction of surface and depth in optically minified pictures.* Poster presented at the Annual Meeting of the Association for Research in Vision and Ophthamology, Ft. Lauderdale, FL.

Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformation.* Cambridge, MA: MIT.

Smith, O. W. (1958a). Comparison of apparent depth in a photograph viewed from two distances. *Perceptual and Motor Skills*, *8*, 79–81.

Smith, O. W. (1958b). Judgments of size and distance in photographs. *American Journal of Psychology*, *71*, 529–538.

Smith, O. W., & Gruber, H. (1958). Perception of depth in photographs. *Perceptual and Motor Skills*, *8*, 307–313.

Smith, P. C., & Smith, O. W. (1961). Ball throwing responses to photographically portrayed targets. *Journal of Experimental Psychology*, *62*, 223–233.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Appendix

Extreme Outlier   ✱

1.5 Δh

Outlier   ◦

1.5 Δh

Upper Hinge
≈ 75th percentile

Δh

Median

Lower Hinge
≈ 25th percentile

1.5 Δh

95% Confidence Interval
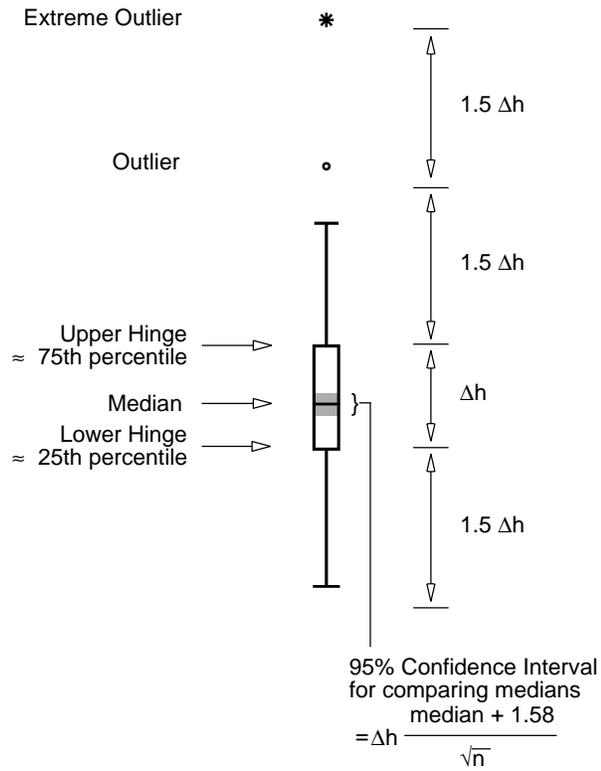for comparing medians
$$=\Delta h\ \frac{\text{median} + 1.58}{\sqrt{n}}$$

Figure 8. Definition of the boxplot.