The perceptual basis of common photographic practice

Emily A. Cooper

Helen Wills Neuroscience Institute,
University of California, Berkeley, Berkeley, CA, USA

Elise A. Piazza

Vision Science Program,
University of California, Berkeley, Berkeley, CA, USA

Martin S. Banks

Helen Wills Neuroscience Institute,
University of California, Berkeley, Berkeley, CA, USA;
Vision Science Program,
University of California, Berkeley, Berkeley, CA, USA;
School of Optometry,
University of California, Berkeley, Berkeley, CA, USA

Photographers, cinematographers, and computer-graphics engineers use certain techniques to create striking pictorial effects. By using lenses of different focal lengths, they can make a scene look compressed or expanded in depth, make a familiar object look natural or distorted, or make a person look smarter, more attractive, or more neurotic. We asked why pictures taken with a certain focal length look natural, while those taken with other focal lengths look distorted. We found that people's preferred viewing distance when looking at pictures leads them to view long-focal-length pictures from too near and short-focal-length pictures from too far. Perceptual distortions occur because people do not take their incorrect viewing distances into account. By following the rule of thumb of using a 50-mm lens, photographers greatly increase the odds of a viewer looking at a photograph from the correct distance, where the percept will be undistorted. Our theory leads to new guidelines for creating pictorial effects that are more effective than conventional guidelines.

Introduction

Every day the Associated Press publishes ~3,000 new photographs, and Facebook users post nearly 250 million (Associated Press, 2009; Shaffer, 2011). Clearly, people rely on pictures to communicate with one another, so it is important to understand how people perceive pictures in typical viewing situations.

Photographers, cinematographers, and computer-graphics engineers create pictorial effects in various ways. For example, photographs of scenes captured with short-focal-length lenses appear expanded in depth, while those captured with long lenses appear compressed. These effects can be seen in still photographs and video. In the latter, the technique creating the effect is called a "dolly-zoom" shot. Figure 1a shows two example photographs from the website Flickr. On the left, the goat looks stretched in depth; on the right, the pitcher and batter appear to be much closer to one another than they actually are. Figure 1b shows two example frames from a dolly-zoom shot in the movie *Goodfellas*. Objects through the window appear much farther away in the frame on the left (from early in the scene) than in the frame on the right (from later). Figure 1c shows how depth compression and expansion can also affect the appearance of a face. Long lenses can make a person look smarter, more attractive, and less approachable; short lenses have the opposite effect (Perona, 2007).

The apparent expansions and compressions in depth are often called perspective distortion, as if these effects are due to a distortion in the physical projection from the scene to the film plane. The effects occur, however, when the projections are geometrically correct. For example, Figures 2 and 3 contain computer-generated (CG) videos of dolly-zoom shots of a kitchen and a face. Both were created using accurate 3D models and correct perspective projection. Thus, the perceptual effects are not caused by physical distortion in the projections. To explain them, one must consider perceptual mechanisms and people's viewing habits, and that is the purpose of our paper.
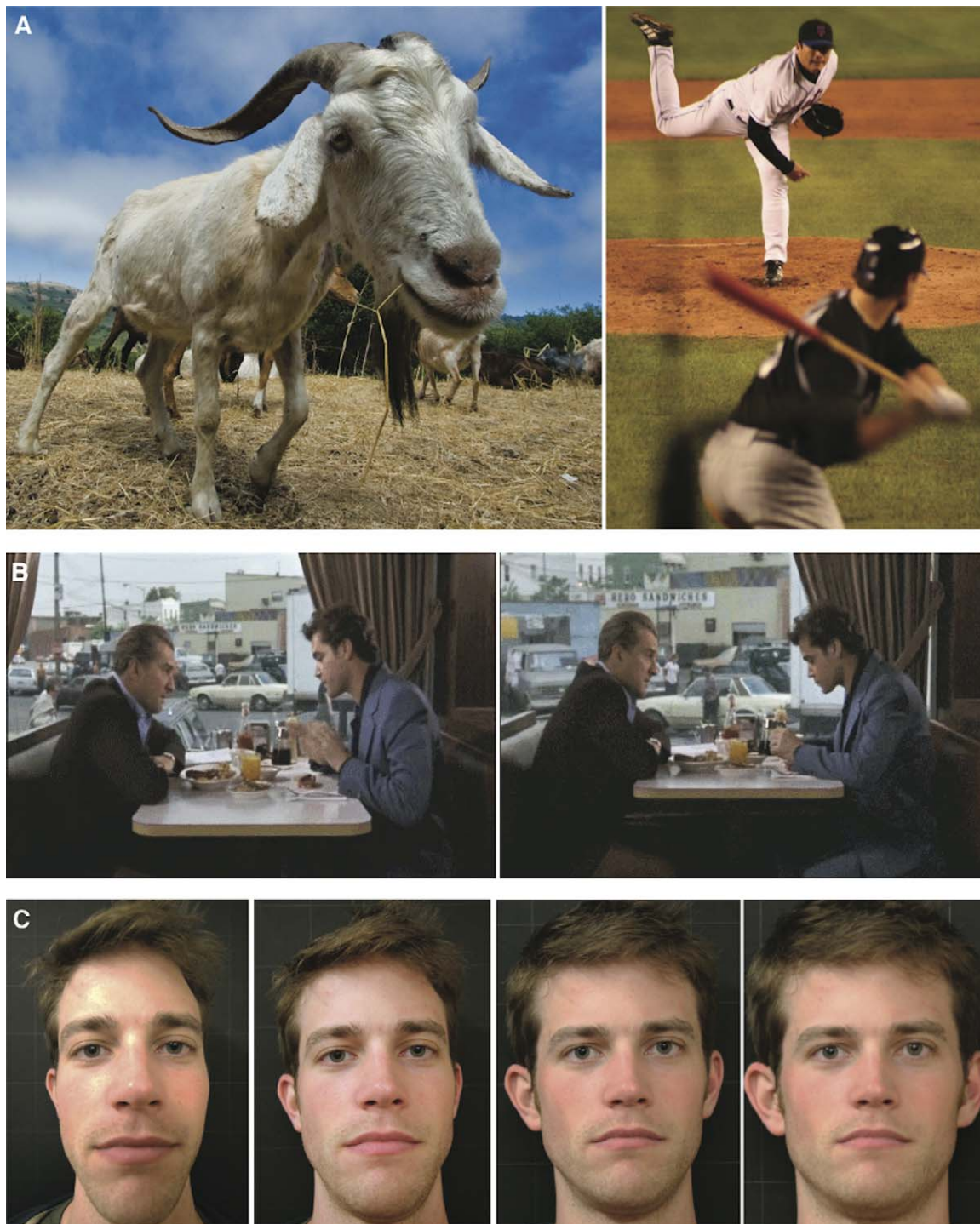
Figure 1. Depth compression and expansion with different focal lengths. (a) Left panel: wide-angle effect (short focal length). This picture was taken with a 16-mm lens (all focal lengths are reported as 35-mm equivalent). The goat looks stretched in depth (©Eliya Selhub; http://www.flickr.com/photos/eliya/2734997796/). Right panel: telephoto effect (long focal length). This picture was taken with a 486-mm focal length. The distance between the pitcher's mound and home plate on an official Major League Baseball field is 18.4 meters. This distance appears compressed (©Mitali Mookerjee; http://www.flickr.com/photos/cool/261259100/). (b) Two stills taken from the movie *Goodfellas* (Warner Brothers). In this scene, the cinematographer slowly moves from being close up and using a short-focal-length lens to being far away and using a long-focal-length lens, while keeping the actors the same size in the frame. The visual effect is a smooth change from depth expansion to compression (a dolly-zoom shot). The effect is particularly noticeable for the car and buildings seen through the window. (c) Photographs of the same person were taken with focal lengths from left to right of 16, 22, 45, and 216 mm. Lens distortion was removed in Adobe PhotoShop, so the pictures are nearly correct perspective projections. Camera distance was

→

Figure 2. (Movie). A dolly-zoom shot of a kitchen. Using Maya (Autodesk), we imaged a 3D model of a kitchen (Birn, 2008) while simultaneously zooming in and dollying out . The angle subtended by the bottle was kept constant. The video begins with a short lens, progresses to long, and returns to short. The bar at the bottom indicates the lens focal length at each moment. The resulting images look very different even though they are all correct perspective projections of a 3D model.

A rule of thumb among professional photographers is to use a focal length of 50 mm for standard 35-mm film (more generally, a focal length equal to the diagonal length of the film or sensor) to create natural-looking images (Kingslake, 1992; Belt, 2008; Modrak & Anthes, 2011; London, Stone, & Upton, 2010). Photography texts offer explanations for this rule's efficacy, but they are either vague or merely restatements of the phenomenon. For example, *Re-framing Photography* claims that using 50-mm lenses "approximates the angle of view and magnification of human vision" (Modrak & Anthes, 2011). *The Elements of Photography* states that "the normal focal length for a given format most closely approximates human sight, and projects an image with the least distortion and compression of space from foreground to background" (Belt, 2008). In an article for the *Photographic Society of America Journal*, Current (1990) suggested that 50 mm is the "normal" lens because most people view pictures from a distance equal to the diagonal length of the picture, and at this distance "the perspective is correct and we are most comfortable" when the picture was captured with a 50-mm lens. We sought a more rigorous explanation of why the 50-mm rule works and why deviations from it yield perceptual distortions.

←

proportional to focal length, so the subject's interocular distance in the picture was constant. The subject's face appears rounder with a short focal length and flatter with a long focal length.
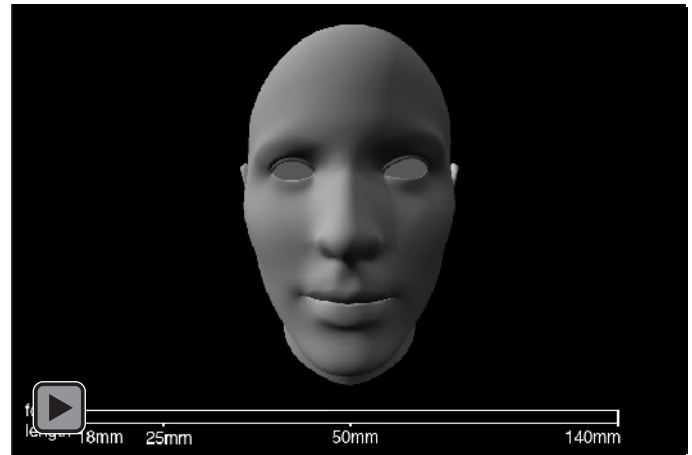

Figure 3. (Movie). A dolly-zoom shot of a face. The same technique from Figure 2 was used. The horizontal angle subtended by the temples of the face was kept constant.

Pictures (i.e., photographs, computer-generated images, and perspective paintings) are created by projecting the light from a 3D scene through a point—the center of projection or COP—onto a flat surface (Figure 4a). This is perspective projection. The field of view of a captured projection is:

$$\theta = 2\tan^{-1}\left(\frac{l_s}{2f}\right) \qquad (1)$$

where l_s is the diagonal length of the film or sensor, f is focal length, and θ is diagonal field of view. If the image on the sensor is magnified by m, the resulting picture has a diagonal length of ml_s. If the viewer's eye is positioned at the picture's COP, the image cast by the picture onto the retina matches the image that would be cast by the original scene. The distance to the COP is:

$$d_{COP} = fm. \qquad (2)$$

Of course, one cannot reconstruct the original scene rigorously from a single retinal image, whether it was generated by a real scene or a picture. But most of the time the brain reconstructs reasonably accurately by using assumptions about perspective (e.g., the chess pieces are the same size, the chessboard is composed of square tiles, the opposite sides of the chessboard are parallel; La Gournerie, 1859; Pirenne, 1970; Sedgwick, 1991; Todorović, 2005). Because viewing a picture from the COP generates the same retinal image as the original scene, it is not surprising that a picture yields a faithful impression of the scene layout or the physical characteristics of a person (Smith & Gruber, 1958; Vishwanath, Girshick, & Banks, 2005; Koenderink, van Doorn, & Kappers, 1994).

However, people do not necessarily position themselves at the COP when viewing pictures; they may be too far or too near. If viewers fail to compensate for an incorrect distance, the interpretation of the pictured
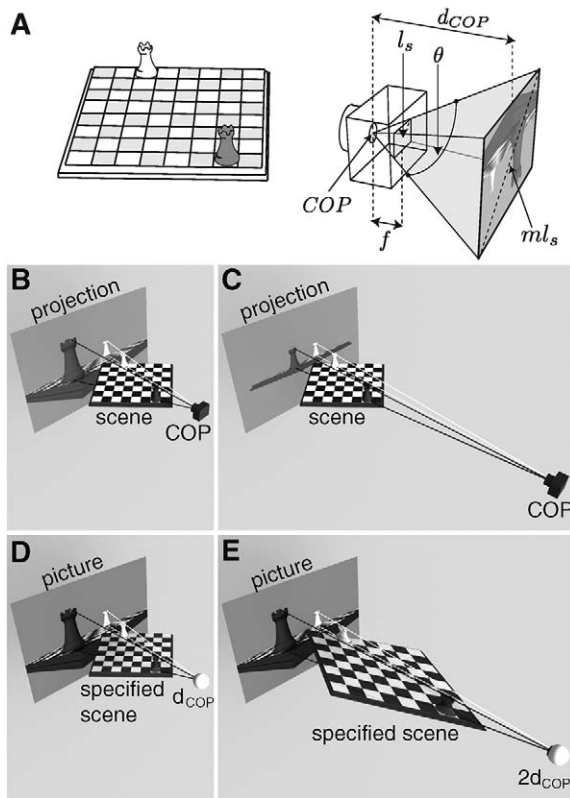
Figure 4. Camera, picture, and viewing parameters. (a) Scene, camera, and picture. A camera with focal length f captures a picture on the sensor. The camera's diagonal field of view is θ. The sensor's diagonal length is l_s, and the print is magnified by m to have a diagonal length of ml_s. The COP is located at the optical center of the camera. The distance to the COP, d_{COP}, is fm, and the diagonal field of view subtended by the picture when viewed from the COP is θ. (b, c) Perspective projection. The original scene—a chessboard—is projected from two different COPs onto a projection plane. (d) If the picture from panel B is viewed from d_{COP}, the specified scene is the same as the original chessboard. (e) If the same picture is viewed from twice the COP distance ($2d_{COP}$), the specified scene is stretched in depth relative to the original chessboard.

scene will be distorted. For example, Figures 4b and c show two pictures of the same scene for two COP distances; the pictures differ. Figures 4d and e show how the apparent 3D scene may differ when one of the pictures (Figure 4b) is viewed from two different distances. When viewed from twice the COP distance, the layout specified by linear perspective is stretched in depth: The near chess piece projects to a larger image than the distant piece and, given the assumption that chess pieces are the same size, they appear farther from each other than they actually are. Similarly, for a viewer positioned too close to a picture, the apparent layout may be compressed in depth.

Previous research found that people do not compensate for incorrect viewing distance (Smith & Gruber, 1958; Bengston, Stergios, Ward, & Jester, 1980; Kraft & Greene, 1989; Todorović, 2009). In fact, Leonardo da Vinci described perceptual distortions when paintings were not viewed from the correct distance and advised painters of realistic scenes to make sure the viewer could view near the COP (da Vinci, 1970). Some research, however, has reported partial compensation for viewing distance, i.e., observers perceived the 3D scene geometry reasonably accurately even when the depicted geometry from linear perspective was distorted due to viewing from distances closer or farther than the COP (Yang & Kubovy, 1999).

We propose a new hypothesis for the effectiveness of the 50-mm rule and for the perceptual distortions from other lenses. The hypothesis incorporates people's viewing habits and the perceptual mechanisms involved in estimating 3D structure from the retinal image. We present two experiments whose results confirm the main tenets of the hypothesis. The first experiment re-examines how people interpret the 3D geometry of a pictured scene in rich, realistic pictures when viewing from the wrong distance. The second one tests how people naturally set their viewing distance when looking at pictures. We then describe appropriate guidelines for constructing pictures when the picture creator's intention is to yield accurate percepts of 3D structure.

Experiment 1: Compensation for viewing distance

Methods

Participants

Five young adults participated. All but one were unaware of the experimental hypotheses. Participants gave informed consent under a protocol approved by the Institutional Review Board of the University of California, Berkeley.

Stimuli

The stimuli were CG images of two rectangular planes joined to form a hinge. The planes were textured with a rectangular grid. The images were rendered using Autodesk Maya and consisted of photographs of wood that were texture-mapped onto the two sides of the hinge, wallpaper in the background, a wood-textured floor, and randomly positioned cubes scattered on the floor (Figure 5). The images were rendered with five different COP distances—11, 16, 28, 55, and 79 cm—and displayed on a CRT (40.6 × 30.4 cm, 2048 × 1536 pixels).

Procedure

Subjects were positioned on a bite bar 28 cm from the CRT. They viewed the screen binocularly with the midpoint of the interocular axis centered in front of the screen. They were told that the two sides of the hinge were rectangular. After each 1.5-second stimulus presentation, subjects indicated whether the hinge angle was greater or less than 90°. A 1-up/1-down adaptive staircase varied the hinge angle symmetrically about the midsagittal axis with four step-size reductions, 10 reversals, and a minimum step size of 2°. Staircases for each of the five COP distances were randomly interspersed and repeated six times. Data were fit with a cumulative Gaussian (psychometric function) using a maximum-likelihood criterion (Wichmann & Hill, 2001). The mean of the best-fitting function was defined as the angle perceived as 90°.

Results

The results of Experiment 1 are shown in Figure 6. If subjects were able to compensate for their viewing distance relative to the COP distance, they would perceive the depicted hinge angle correctly and would set the hinge to 90° in scene coordinates (horizontal dashed line). If subjects failed to compensate for the difference between their viewing distance and the COP distance and instead interpreted the scene directly from the geometry of the retinal image (assuming that the hinge is composed of rectangular planes), they would set the depicted hinge angle to different values in scene coordinates for each COP distance. The predicted settings for this second hypothesis can be calculated from geometric analyses of perspective projections such as those presented in Sedgwick (1991) and Rosinksi and Farber (1980). These analyses show that the angle in scene coordinates and the depicted angle are related by the ratio of the viewing distance to the COP distance. With no compensation, the predicted hinge angle perceived to be 90° is:
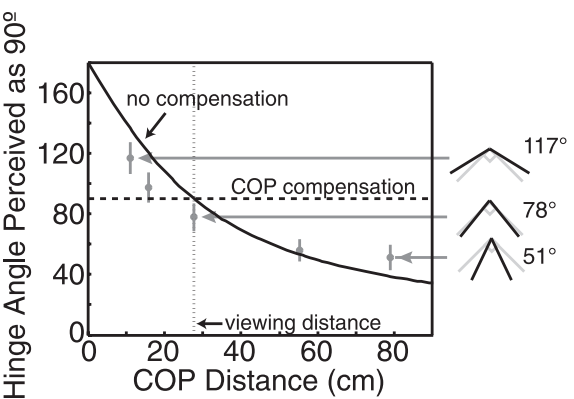

Figure 6. Effect of distance to COP on the angle perceived as 90°. Gray circles represent the mean angle perceived as 90° across subjects; error bars are standard errors. The dotted vertical line indicates the viewing distance. Plan views of the depicted angles that appeared to be 90° are shown in black on the right. The gray hinges indicate 90° for comparison.

$$\omega = 2 \tan^{-1} \left(\frac{d_v}{d_{COP}} \right) \tag{3}$$

where d_{COP} is the COP distance of the picture and d_v is viewing distance (solid curve).

The data are very consistent with the no-compensation prediction. Some subjects had a bias in the angle perceived as 90° when viewing from the COP, but despite this bias, changing the COP distance always had the effect on perceived hinge angle that was predicted by the geometry of the retinal image. When the COP distance was less than the viewing distance, subjects perceived a larger angle as 90°, which means that they experienced depth expansion. When the COP distance was greater than the viewing distance, they perceived a smaller angle as 90°, meaning that they experienced depth compression. When the COP distance and viewing distance were the same, a 90° hinge was perceived as close to 90°, so they experienced neither expansion nor compression.

There were slight, but systematic differences between our data and the no-compensation predictions. Gener-
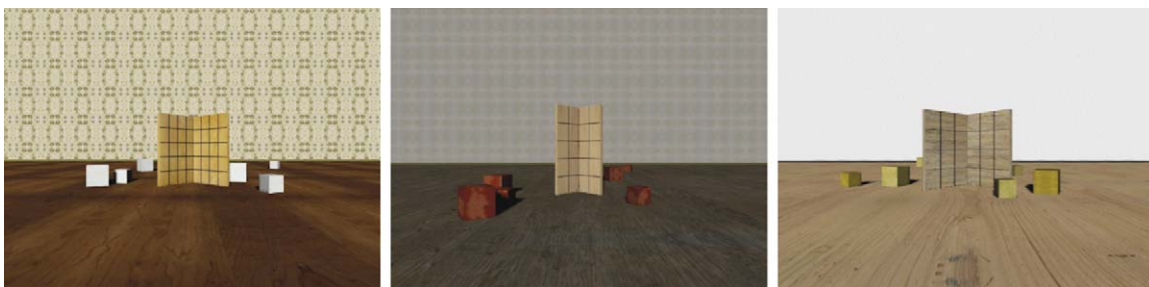

Figure 5. Examples of the hinge stimuli. The environment (background, cubes) and shape of the hinges were randomized to prevent subjects from learning specific pictorial cues to the hinge angle. There were three backgrounds, each with three unique hinge shapes resulting in nine scenes altogether. On each trial, the displayed hinge was selected randomly from these nine scenes.

Figure 7. Example pictures from each scene category. Left to right: indoor, street, outdoor open, outdoor closed, and portrait.

ally, subjects set the hinge angle to slightly less than the predicted value, which means that they perceived the angles as somewhat flatter than dictated by the geometry of the retinal image. (The one exception to this is at the greatest COP distance where they set the angle slightly larger than predicted.) We believe that the cause of this bias is the flatness specified by a number of cues including binocular disparity and focus cues (Watt, Akeley, Ernst, & Banks, 2005). Some previous studies have reported similar, but larger deviations from no-compensation predictions (i.e., the perceptual distortions were significantly less than predicted by perspective geometry) and suggested that there might be partial compensation for incorrect viewing distance (Adams, 1972; Lumsden, 1983; Yang & Kubovy, 1999). Those studies used line drawings or very simple scenes with little perspective information. Flatness cues probably have more effect with line drawings than with realistic stimuli, so we presumably observed less flattening in our experiment (Sedgwick, 1991; Sedgwick & Nicholls, 1994; Todorović, 2009). We conclude that viewers mostly do not compensate for incorrect viewing distance when shown pictures with rich perspective information.

Experiment 2: Preferred viewing distance

In this experiment, we measured people's preferred viewing distance for pictures of different focal lengths, magnifications, and print sizes. The results enabled us to determine whether people use consistent strategies for setting viewing distance and, if so, what those strategies are.

Methods

Participants

Eight young adults participated in the main experiment, and 11 additional young adults partici-

pated in a follow-up experiment. All were unaware of the experimental hypotheses. Participants gave informed consent under a protocol approved by the Institutional Review Board of the University of California, Berkeley.

Stimuli

Scenes for the pictures were selected from five categories described in the scene-recognition literature: indoor, street, outdoor open (e.g., coastline, mountains), outdoor closed (trees), and portrait (Torralba & Oliva, 2003; Torralba, 2009). For each of the first four categories, we used three unique scenes: one photographed scene and two CG scenes. For the fifth category, we used two photographed scenes. Figure 7 provides example pictures of scenes from each category.

The photographs were taken with a Canon EOS 20D SLR camera and saved as low-compression JPEG files. CG images were saved as TIFF files. Both types of image files were saved at 300 dots/inch (dpi) and printed on photographic paper with a resolution of 300 dpi and an aspect ratio of 3:2. All CG images were rendered with infinite depth of field (i.e., no blur) and were illuminated with a combination of a directional and ambient light sources. For the photographs, we used the smallest aperture allowed by the lighting environment to minimize differences in depth of field and exposure between photographs taken with different focal lengths.

There were two primary stimulus manipulations: focal length and magnification. To manipulate focal length, we selected a focal object in each scene and created a series of five images taken with five different focal lengths—22, 32, 56, 112, and 160 mm (35-mm equivalent)—while keeping the camera at one location. All of those pictures were magnified eight-fold and printed at 18 × 12 cm (7 × 5 in). To manipulate magnification, we took photographs with a 56-mm lens and printed them at 18 × 12 cm (same as the previous), and four additional sizes (6 × 4, 9 × 6, 29 × 19, and 39 × 26 cm).

Figure 8. Changing focal length and camera distance to maintain constant size of the focal object (in this case, a pillow). (a) The effect of changing focal length while keeping camera position constant. The focal lengths from left to right are 160, 56, 32, and 22 mm. (b) The effect of changing camera distance while holding focal length constant. From left to right, the camera is moved farther and farther from the focal object. Focal length was always 56 mm. By moving the camera farther from the focal object, the sizes of the focal object are matched to those in the upper row without changing COP distance. Differences between the images in panels a and b are particularly noticeable in the apparent shape of the bed and slant of the wall.

By changing focal length, the focal object became different sizes in the prints (Figure 8a). To determine whether the varying size of that object affected preferred viewing distance, we also created five images in which the focal length was fixed at 56 mm, but the camera was dollied in and out so that the size of the focal object would match those from the five focal lengths (Figure 8b). These were all printed at 18 × 12 cm.

We were curious to see whether these results would generalize to larger picture sizes, so we conducted a follow-up experiment with larger pictures. Eleven new subjects participated. The stimuli were the same with a few exceptions. Only four scenes were used: one indoor, one street, one outdoor open, and one outdoor closed. All pictures were CG. We created pictures with three focal lengths (22, 56, and 160 mm) and printed each at four sizes (18 × 12, 53 × 35, 73 × 49, and 100 × 67 cm). We dollied the camera away from the focal object as we increased the focal length in order to match the size of the object across focal lengths. Subjects were shown each focal length twice and each print size twice with a random selection of two of the four scenes.

Procedure

At the start of each trial, a picture was mounted on a wall at the subject's eye level. Subjects stood initially 5 m from the picture. They were instructed to walk back and forth along a line that was perpendicular to the picture until they were at "the best distance to view the picture from." If they asked for clarification, we told them to select "whatever distance you prefer to view from." Once they indicated that they were at the preferred distance for that picture, the experimenter

recorded the distance with a photograph. The trials were recorded so preferred distances could be measured off-line using the ruler tool in Adobe Photoshop.

Subjects were presented with a picture from each level of each manipulation eight times, with a random selection of 8 of the 14 scenes. Therefore, subjects did not see the same scene/manipulation combination twice. The main experiment took place over four sessions. The order of presentation was randomized. We also wanted to measure test-retest reliability, so we presented eight pictures four times each (once per session). Each subject thus completed a total of 136 trials.

The procedure of the follow-up experiment was identical to the main experiment with a few exceptions. The subjects began each trial standing 6.5 m from the picture and again moved toward and away until they were at their preferred viewing distance. We measured viewing distance with a laser range finder. These measurements occurred in two sessions. Presentation order was randomized. To assess test-retest reliability, we randomly presented three pictures four times (twice per session). Each subject therefore completed a total of 36 trials in this phase of the experiment.

We also investigated whether the manner of picture viewing—standing in front of a wall-mounted picture as opposed to holding a picture while seated—affects preferred viewing distances. Three subjects from the main experiment participated in these measurements. They sat in a chair and held each picture in their hands. They varied distance by adjusting their arms until they achieved the preferred value. We measured that distance using the laser range finder. A subset of the stimuli from the main experiment was used with one focal length (56 mm) and two print sizes (9 × 6 and 18
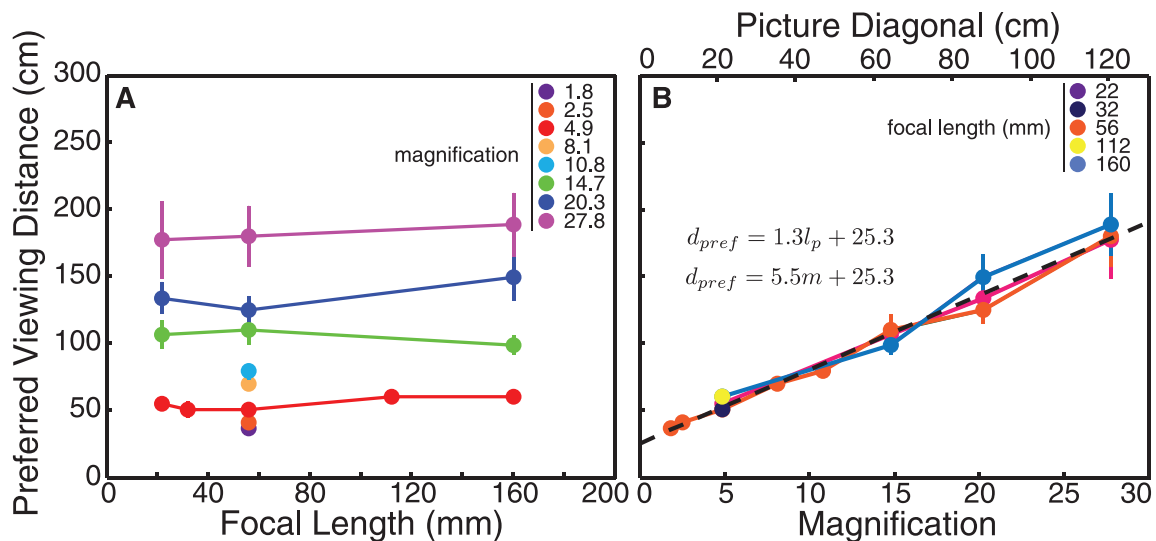
Figure 9. Effects of focal length and magnification on preferred viewing distance. (a) Preferred viewing distance is plotted as a function of focal length for each magnification. Circles represent the data: the mean preferred viewing distance across subjects. Error bars represent standard errors of the mean. Each color represents a different picture magnification (and therefore a different picture size), as indicated by the legend. (b) Data from panel A replotted as a function of magnification for each focal length. The diagonal length of the picture for different magnifications is indicated at the top. A linear regression of the data is represented by the dashed black line and the two equations. All five focal length levels are plotted for magnification = 4.9, but the circles are largely overlapping because there was so little effect of focal length.

× 12 cm). For each print size, 10 of the 14 scenes were randomly selected. Each subject completed a total of 20 trials.

Results

We first asked whether the data from the follow-up experiment differed from the main experiment. A one-way ANOVA performed on the data from overlapping conditions revealed no significant effect ($p = 0.53$), so from here on we combine the data from these two experiments.

The results for the main stimulus manipulations—focal length and magnification—are illustrated in Figure 9. Panel 9a shows mean preferred viewing distance as a function of focal length. The results are plotted separately for each magnification. Some magnifications only have one focal length because the two variables were not completely crossed in the main experiment. There was clearly no effect of focal length on preferred viewing distance for a given magnification. Panel 9b shows the same data, but with mean preferred viewing distance plotted as a function of magnification. There was a strong effect of magnification/picture-size on preferred viewing distance, independent of focal length. The dashed line shows a linear regression of these data ($p < 0.0001$). Equations for the line as a function of picture diagonal (l_p) and magnification (m)

are shown next to the line. Notably, the y-intercept of the line (25 cm) is the same as the nearest comfortable viewing distance for young adults (Ray, 2000).

We replotted a subset of the data in Figure 9 in a way that allows us to examine the picture properties that determine preferred viewing distance. Figure 10a shows two subsets of stimuli for one example scene: five focal lengths for one magnification and eight magnifications for one focal length. Figure 10b shows the average preferred viewing distance for these subsets of all stimuli. If subjects preferred that pictures subtend a particular visual angle, or field of view (FOV), preferred distance would be proportional to print size, and the data would fall along one of the blue lines in Figure 10b, depending on the desired angle. Alternatively, if subjects always moved to the distance of the picture's COP (d_{COP}), the preferred viewing distance would be proportional to focal length and magnification (Equation 2), and the data would lie on the red lines in Figure 10b. The left panel shows that preferred viewing distance was barely affected by COP distance. From the nearest to farthest COP, preferred distance increased by only 20%, significantly less than the 614% change that would have occurred if subjects matched viewing distance to COP distance. The right panel shows that preferred viewing distance was strongly dependent on magnification (or equivalent picture size). But subjects were not establishing a constant field of view; rather, they preferred a small field ($\sim$22°) with small prints and a larger field ($\sim$36°) with large prints.
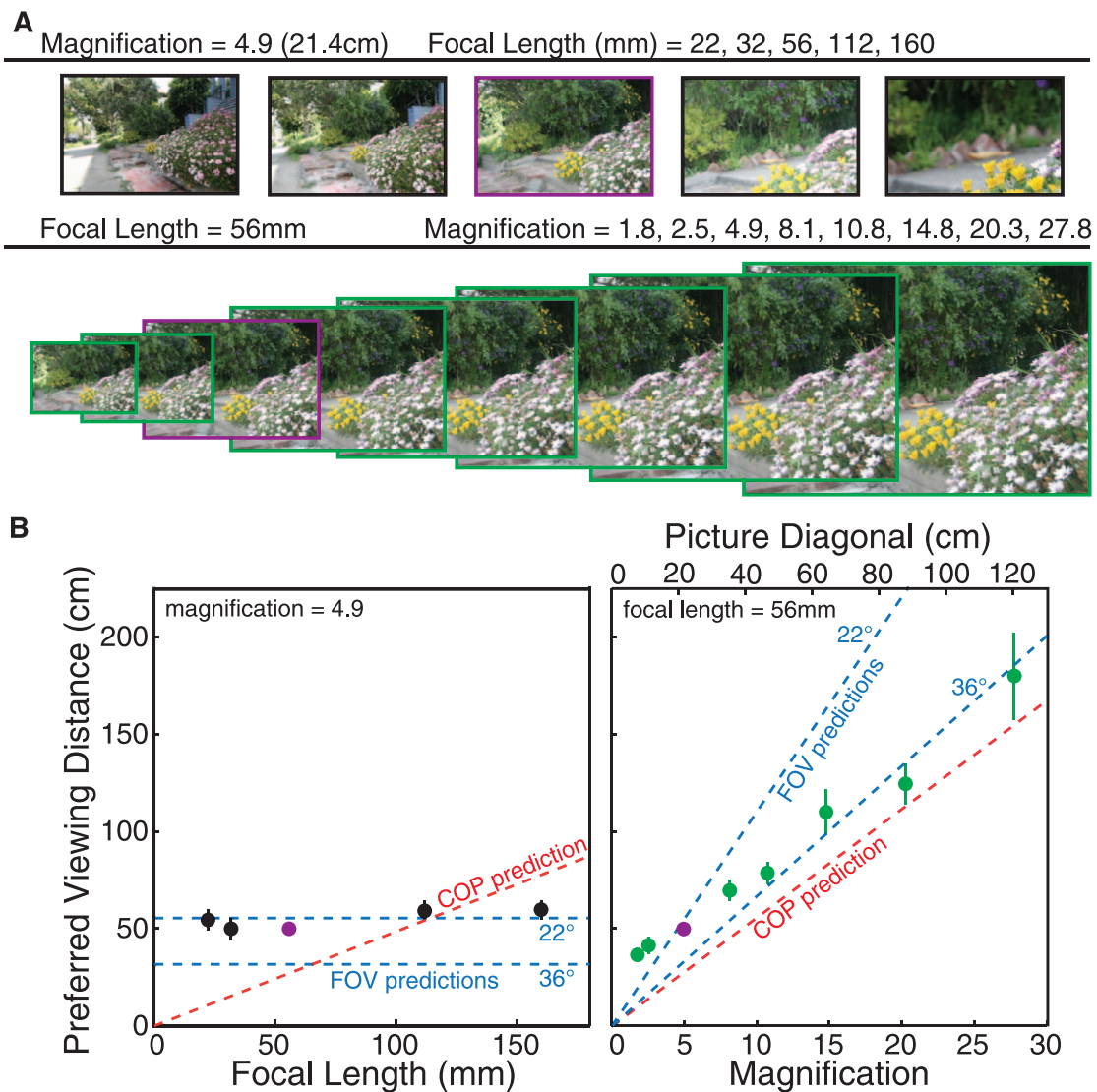
Figure 10. (a) Example stimuli for two subsets of conditions. One subset contains five focal lengths with a magnification of 4.9 (diagonal length of the printed picture was 21.4 cm). The other subset contains eight magnifications with a focal length of 56 mm. The relative sizes of the stimuli actually changed by a factor of 15.4, but we cannot show such a large change in the Figure. Therefore, the change in relative size shown previously is qualitative. The purple boxes around two of the pictures indicate the one that was in both subsets. (b) Two plots of average preferred viewing distance across subjects for each manipulation. Black and green circles represent the focal length and magnification manipulations, respectively, and correspond to the boxes around the pictures in panel a. The purple circles in both plots represent data from one magnification and focal length (4.9 and 56 mm, respectively). Error bars represent standard errors of the mean.

This smaller preferred field of view for small prints likely reflects a trade-off between viewing comfort and angle subtended by the print. We conclude that picture viewers do not naturally set their viewing distance to a picture's COP distance. Instead they adjust distance according to the field of view (albeit smaller fields for small prints and larger fields for large prints). These data are consistent with television-viewing studies, which show that preferred viewing distance is determined by the size of the screen rather than image content or television resolution (Ardito, 1994; Lund, 1993).

We also examined whether there were differences in the effects of focal length and magnification between photographs and CG images. The average difference between the preferred viewing distance for photographs and CG images across the whole data set was only 3.2 cm (standard deviation = 4.6 cm). Because the experimental conditions were not fully crossed, we could not perform an ANOVA on these results. A one-way ANCOVA on the effect of focal length and magnification on preferred distance for photographs versus CG revealed no significant difference for either manipulation ($p = 0.55$ and 0.99).

Although magnification had by far the largest effect on preferred viewing distance, there was a small but significant effect of focal length (for example, the slope of a linear regression of the data plotted in the left panel of Figure 10b was = 0.1, $p = 0.008$). This effect could have been due to the picture's COP distance or to the size of the focal object (i.e., the object centered in the frame). Recall that we included a control condition in which the size of the focal object was manipulated by dollying the camera rather than changing the focal length (Figure 8b). A one-way ANCOVA on preferred distance as a function of normalized focal object size for the two groups (focal length and camera distance) revealed no significant difference between the effects of focal length and camera distance ($p = 0.46$). We conclude that this small effect was due to the size of the focal object and not due to an effect of COP distance on preferred viewing distance.

To assess test-retest reliability, we also calculated the standard deviation of preferred viewing distance for each subject for each of the repeated pictures. The mean standard deviations across all images and subjects were 14 cm for the main experiment and 22 cm for the follow-up experiment. These values are small relative to the means, so the preferred distances were reasonably repeatable.

Finally, we examined the effect of standing (where subjects adjusted their viewing distance by walking to and fro) and sitting (where subjects held the pictures in their hands) on preferred viewing distance. A two-way ANOVA performed on overlapping conditions from the two sets of data revealed no effect ($p = 0.59$), so we conclude that people behave similarly when viewing wall-mounted pictures while standing and when viewing handheld pictures while sitting (provided that picture size is not so large for arm length to limit the ability to set distance to the desired value).

Discussion

We can now explain why focal length affects apparent depth in pictured scenes and facial appearance in portraits. Recall that long- and short-focal-length pictures look respectively compressed and expanded in depth (Figures 1, 2, and 3). We propose that people's preferred field of view when looking at most pictures leads them to view long-focal-length pictures from too near and short-focal-length pictures from too far. Perceptual compression and expansion occur because people do not take their incorrect viewing distances into account. Thus, scenes captured with long lenses look compressed in depth, which makes faces apparently flatter. Likewise, scenes cap-

tured with short lenses appear expanded in depth, which makes faces look rounder.

However, this does not tell us why pictures created with a 50-mm lens look most natural, i.e., neither expanded nor compressed. To investigate this, we calculated for each picture size the focal length for which the subjects' average preferred viewing distance would be equal to the COP distance. We call this the *recommended focal length*:

$$f_{rec} = 43.3 \frac{d_{pref}}{l_p} \qquad (4)$$

where d_{pref} is the average preferred viewing distance, l_p is the diagonal length of the picture, and 43.3 is the diagonal length of standard 35-mm film in millimeters. The recommended values from our data, calculated by averaging the preferred viewing distance across all focal lengths for each picture size from Experiment 2, are plotted in Figure 11. The regression line from Figure 9b is also replotted in terms of recommended focal length. The equation for the line is:

$$f_{rec} = 55 + \frac{1096}{l_p}. \qquad (5)$$

Thus, for prints 35 cm or larger, the recommended focal length is ~50 mm. Most prints, particularly professional ones, are at least that size. We claim therefore that following the 50-mm rule of thumb maximizes the odds of a viewer looking at the photo from the COP distance and thereby makes it most likely that the percept will be undistorted. This rule has presumably evolved over time based on collective experience. Similar recommendations apply for cinematographers, computer-graphics engineers, and painters of realistic images. Some typical image sizes for various formats (Take, 2003) are superimposed as vertical bands in the figure. For most venues, the recommended focal length is ~50 mm (35-mm equivalent). With the small screens of mobile devices, longer focal lengths should be used. If image creators know the size of a typical print or projection of their work, they can use Equation 5 to make a better choice of focal length or to change the distance of the COP in postprocessing (Carroll, Agarwala, & Agrawala, 2010).

Most photography texts advocate the 50-mm rule (Kingslake, 1992; Belt, 2008; Modrak & Anthes, 2011; London et al., 2010), but we wondered whether the rule is actually used in practice. To find out, we collected 3,930 photographs from the website Flickr that were taken with single-lens reflex (SLR) cameras. (These cameras tend to be used by professionals and serious hobbyists.) We obtained the 35-mm-equivalent focal length for those photos from their EXIF data. The median is 68 mm (50% quantile horizontal line in Figure 11). Interestingly, 68 mm is closer than the
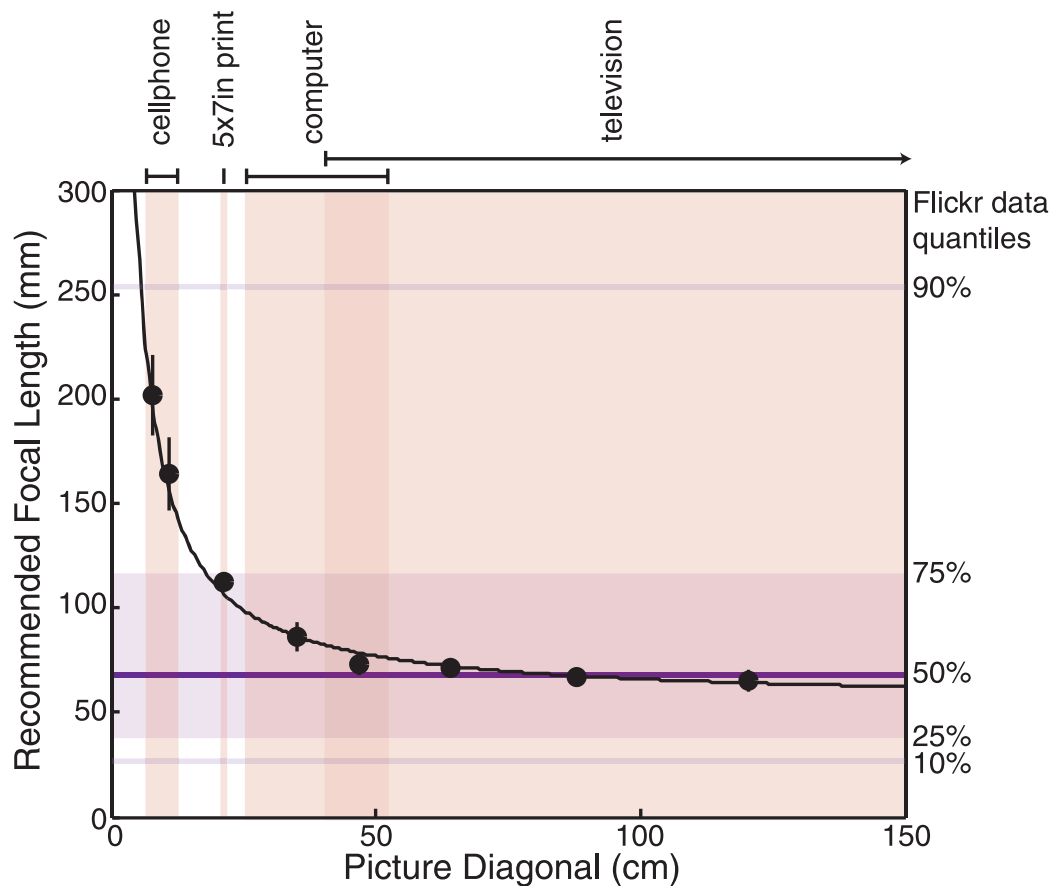
Figure 11. Recommended focal length as a function of picture size. We calculated recommended focal length for each picture size by determining the average preferred viewing distance across all focal lengths from Experiment 2 (Figure 9b) and then calculating the focal length that would produce a COP distance equal to the preferred distance (Equation 4). Circles represent those values, and error bars represent standard errors. The black curve shows the linear regression from Figure 9b replotted in terms of recommended focal length. Vertical bands indicate some typical image sizes for various formats. Horizontal bands indicate quantiles from several cumulative probability values for 3,930 Flickr photographs taken with SLR cameras.

advocated 50 mm to our recommended focal length for a wide range of sizes. Thus, current practice deviates slightly from the 50-mm rule, but is more consistent with our experimental data.

Our recommended focal length is much longer for small picture sizes, such as those on mobile devices. The viewing of images on mobile devices is becoming much more common (Choney, 2009; Carlsson & Walden, 2007). People tend to view smart phones from ~30 cm (Knoche & Sasse, 2008). When standard content is viewed at that distance, the smart-phone user is generally much farther from the display than the COP distance, making the images of objects subtend small angles and producing expansion in apparent depth. Interestingly, smart-phone viewers prefer standard content to be magnified and cropped (Knoche et al., 2007; Song et al., 2010), which increases the COP distance, much like increasing focal length; this practice should make the viewed content appear less expanded than it otherwise would.

Focal length has a strong effect on the perceived personality of subjects in portraits (Perona, 2007). We speculate that such effects derive from correlations between people's actual facial dimensions and personality traits. For example, faces appear narrower when photographed with short lenses and wider when photographed with long lenses (Figures 1c and 3). The actual width-to-height ratio of male faces is positively correlated with aggressive behavior (Carre & McCormick, 2008), so attributions made from apparent ratio changes probably derive from correlations with real ratios. It would be interesting to examine the relationship between other facial dimensions affected by focal length (e.g., nose length, face roundness) and personality traits.

Pictures are useful in part because viewers can gain a faithful impression of the pictured content even when they are not positioned precisely at the COP. However, the ability to compensate for incorrect viewing position differs between being off-axis (i.e.,

off to the side) and at the wrong distance. For off-axis compensation, the visual system estimates the slant of the picture surface and corrects for the expected foreshortening (Pirenne, 1970; Vishwanath, Girshick, & Banks, 2005; Rosinksi et al., 1980). This compensation process is shape constancy (Wallach & Marshall, 1986), which allows one to perceive the dimensions of real objects from various viewpoints. (The observation of nearly complete compensation has occurred when the image content being judged is roughly parallel to the picture surface; when the content is roughly perpendicular to the surface, compensation is much less complete; Goldstein, 1987; Todorović, 2008.) To compensate for incorrect distance, the visual system would have to estimate the correct distance from the picture's contents, but such estimation is prone to error (La Gournerie, 1859; Kubovy, 1986; O'Brien & Farid, 2012). Thus, we argue that compensation for off-axis viewing occurs because the computations involved are useful in everyday vision and generally not prone to error. We argue further that compensation for incorrect viewing distance does not occur because the required computations are not useful in everyday vision and are prone to error.

Conclusion

We claim that the 50-mm rule emerged because of people's tendency to view pictures from a distance that establishes a desirable field of view and their inability to compensate when that tendency yields an incorrect viewing distance. Our data can be used to create better guidelines, based on empirical results, for creating effective pictures for all viewing situations.

Acknowledgments

Commercial relationships: none.
Corresponding author: Emily Cooper.
Email: emilycooper@berkeley.edu.
Address: Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, USA.

References

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

Ardito, M. (1994). Studies of the influence of display size and picture brightness on the preferred viewing distance for HDTV programs. *SMPTE Journal, 103,* 517–522.

Associated Press. (2009). Inauguration Day coverage propels record viewership of AP's online video services [Press release]. Internet site: http://www.ap.org/pages/about/pressreleases/pr_012109a.html (Accessed January 18, 2012).

Belt, A. F. (2008). *The elements of photography: Understanding and creating sophisticated images.* 2nd ed. Burlington, MA: Focal Press.

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 & Performance, 6,* 751–759. [PubMed].

Birn, J. (2008). *Lighting & Rendering in Maya: Lights and Shadows.* Internet site: http://www.3drender.com.

Carlsson, C., & Walden, P. (2007). Mobile TV—To live or die by content. *Proceedings of the 40th Hawaii International Conference on System Sciences, 1,* 1530–1605.

Carre, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behavior in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B, 275,* 2651–2656. [PubMed].

Carroll, R., Agarwala, A., & Agrawala, M. (2010). Image warps for artistic perspective manipulation. *ACM Transactions on Graphics, 29*(4):127, 1–9.

Choney, S. (2009). Online video watching nearly doubles since '06. *Pew Internet,* http://www.pewinternet.org/Media-Mentions/2009/Online-video-watching-nearly-doubles-since-06.aspx (Accessed September 22, 2011).

Current, I. (1990). Best viewing distance for photographers. *PSA Journal,* Sept, 16.

Da Vinci, L. (1970). In J. P. Richter (Ed.), *The literary works of Leonardo da Vinci.* London: Phaidon.

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology, 13*(2), 256–266. [PubMed].

Kingslake, R. (1992). *Optics in photography.* Bellingham, WA: SPIE.

Knoche, H., Papaleo, M., Sasse, M. A., & Vanelli-Coralli, A. (2007). The kindest cut: Enhancing the user experience of mobile TV through adequate zooming. *Proceedings of ACM Multimedia, 1,* 23–29.

Knoche, H., & Sasse, M. A. (2008). The sweet spot: How people trade off size and definition on mobile devices. *Proceedings of ACM Multimedia, 1,* 21–30.

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

Kraft, R. N., & Green, J. S. (1989). Distance perception as a function of photographic area of view. *Perception and Psychophysics, 45*(4), 459–466. [PubMed].

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

La Gournerie, J. D. (1859). *Traité de perspective linéare contenant les tracés pour les tableaux, plans et courbes, les bas-reliefs et les décorations théatrales, avec une théorie des effets de perspective.* Paris: Dalmont et Dunod.

London, B., Stone, J., & Upton, J. (2010). *Photography.* 10th ed. New Jersey: Prentice Hill.

Lumsden, E. A. (1983). Perception of radial distance as a function of magnification and truncation of depicted spatial layout. *Perception & Psychophysics, 33*(2), 177–182. [PubMed].

Lund, A. M. (1993). The influence of video image size and resolution on viewing-distance preferences. *SMPTE Journal, 102,* 406–415.

Modrak, R., & Anthes, B. (2011). *Reframing photography.* New York: Routledge.

O'Brien, J. F., & Farid, H. (2012). Exposing photo manipulation with inconsistent reflections. *ACM Transactions on Graphics, 31*(1), 1–11.

Perona, P. (2007). A new perspective on portraiture. *Journal of Vision, 7*(9):992, http://www.journalofvision.org/content/7/9/992.short, doi:10.1167/7.9.992.

Pirenne, M. (1970). *Optics, painting, & photography.* Cambridge: Cambridge University Press.

Ray, S. F. (2000). The geometry of image formation. In R. E. Jacobson, S. F. Ray, G. G. Atteridge, & N. R. Axford (Eds.), *The manual of photography: Photographic and digital imaging* (pp. 39–57). Oxford: Focal Press.

Rosinski, R. R., Mulholland, T., Degelman, D., & Farber, J. (1980). Picture perception: An analysis of

visual compensation. *Perception & Psychophysics, 28,* 521–526. [PubMed].

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: Taylor & Francis.

Sedgwick, H. A., & Nicholls, A. L. (1994). *Distortions of pictorial space produced by optical minification.* Poster presented at the annual meeting of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, FL.

Shaffer, J. (2011). Bigger, faster photos. *The Facebook Blog.* http://blog.facebook.com/blog.php?post=10150262684247131 (Accessed September 22, 2011).

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

Song, W., Tjondronegoro, D. W., Wang, S., & Docherty, M. J. (2010). Impact of zooming and enhancing region of interests for optimizing user experience on mobile sports video. *Proceedings of the International Conference on Multimedia, 1,* 321–329.

Take, H. (2003). Market and technology trends in LCD TVs. *Sharp Technical Journal, 4,* 1–4.

Todorović, D. (2005). Geometric and perceptual effects of the location of the observer vantage point for linear-perspective images. *Perception, 34,* 521–544. [PubMed].

Todorović, D. (2008). Is pictorial perception robust? The effect of the observer vantage point on the perceived depth structure of linear-perspective images. *Perception, 37,* 106–125. [PubMed].

Todorović, D. (2009). The effect of observer vantage point on perceived distortions in linear perspective images. *Attention, Perception, & Psychophysics, 71*(1), 183–193. [PubMed].

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems, 14,* 391–412. [PubMed].

Torralba, A. (2009). How many pixels make an image? *Visual Neuroscience, 26*(1), 123–131. [PubMed].

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience, 8,* 1401–1410. [PubMed].

Wallach, H., & Marshall, F. J. (1986). Shape constancy in pictorial representation. *Perception & Psychophysics, 39,* 233–235. [PubMed].

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision, 5*(10):834–862, http://www.journalofvision.org/content/5/10/7, doi:10.1167/5.10.7. [PubMed].

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63,* 1293–1313. [PubMed].

Yang, T., & Kubovy, M. (1999). Weakening the robustness of perspective: Evidence for a modified theory of compensation in picture perception. *Perception & Psychophysics, 61,* 456–467. [PubMed].