Better discrimination for illusory than for occluded perceptual completions

Jiawei Zhou School of Life Sciences, University of Science and Technology of China, China

Bosco S. Tjan Department of Psychology and Neuroscience Graduate Program, University of Southern California, USA

Yifeng Zhou School of Life Sciences, University of Science and Technology of China, China

Zili Liu Department of Psychology, University of California Los Angeles, USA

We applied the thin–fat Kanizsa shape discrimination task invented by D. L. Ringach and R. Shapley (1996) to study perceptual completion by measuring whether the discrimination was more accurate for illusory than for occluded shapes. Differently from Ringach and Shapley, we tested naive observers with stereoscopic displays. Discrimination was consistently more accurate for illusory than for occluded shapes under a variety of stimulus conditions. However, the absolute performance was worse than Ringach and Shapley's experienced observers, who discriminated illusory and occluded shapes equally well. When our naive observers were trained, their performance approached that in Ringach and Shapley, and their performance difference diminished between the illusory and occluded. The more precise discrimination of the illusory shapes by untrained observers is consistent with the subjective impression that illusory contours appear clearer and positionally better defined. This makes sense from the perspective of Bayesian decision theory: the location of an illusory contour that is closer to an observer might be more important than an occluded contour, and hence obligatorily represented more precisely. We conclude the paper by discussing implications of our results on the current debate on mechanisms of perceptual completion (M. K. Albert, 2007; B. L. Anderson, 2007; P. J. Kellman, P. Garrigan, T. F. Shipley, & B. P. Keane, 2007).

Keywords: modal, amodal, illusory, perceptual completion, Kanizsa, discrimination

Citation: Zhou, J., Tjan, B. S., Zhou, Y., & Liu, Z. (2008). Better discrimination for illusory than for occluded perceptual completions. *Journal of Vision, 8*(7):26, 1–17, http://journalofvision.org/8/7/26/, doi:10.1167/8.7.26.

Introduction

Perceptual grouping and segmentation is a fundamental problem in visual perception. The problem is well illustrated by the Kanizsa (1979) square when part of the square's boundary is optically absent (Figure 1A). The Kanizsa square, when perceived as sitting on top of four disks, is called illusory, subjective, or modal. When perceived as seen through four holes, it is called occluded or amodal. One common way to generate the amodal stimulus is to add a ring to close the "mouth" of each inducer, as shown in Figure 1B. The amodal percept can also be achieved without changing the stimulus (Figure 1A). An observer can perceive a white square against black background, when only the corners of the square are visible through four holes in a white occluding plane. Because this amodal percept is harder to achieve than its illusory counterpart, binocular disparity is often added to set the contours of Kanizsa square behind the circular contours of the inducers (Nakayama, Shimojo, & Silverman, 1989). Its illusory counterpart can be obtained by simply swapping the left and right images. The two percepts are hence generated with the same images, and an ideal observer, whose performance is solely determined by stimulus information (Green & Swets, 1974; Knill & Kersten, 1991), will yield identical performance for the illusory and amodal stimuli alike. Hence, any difference in human performance between the illusory and amodal discrimination has to be due to the brain, not to the stimulus.

Why should one expect any performance difference between perceiving illusory vs. amodal shapes? Indeed, the *identity hypothesis* (Kellman, Garrigan, & Shipley, 2005; Kellman & Shipley, 1991) postulates that the same contour interpolation process is at work for illusory and amodal shapes. In the current study, we measured human performance in discriminating illusory versus amodal Kanizsa shapes. The central result of our study is that performance was more accurate for illusory than for amodal discrimination in a variety of stimulus conditions.

doi: 10.1167/8.7.26 Received September 10, 2007; published September 24, 2008 ISSN 1534-7362 © ARVO

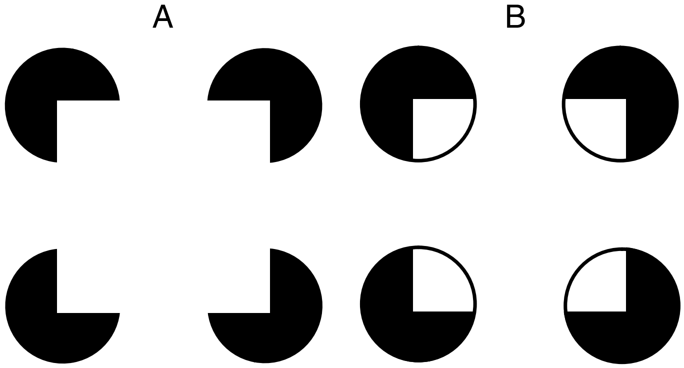

Figure 1. (A) Illustration of the Kanizsa square, when a white square is perceived to sit on top of four black disks (an illusory percept). An alternative, though less dominant, percept is that a white square against black background is seen through four holes (an amodal percept). (B) A variation of the Kanizsa square when the amodal percept is unambiguous.

Perceptual precision in illusory and amodal shape discrimination

How does one objectively and parametrically study the difference between the illusory and amodal percepts that arise when the stimulus images are the same? Ringach and Shapley's (1996) solution was to rotate the inducers of a standard Kanizsa square to create a "fat" shape (i.e., vertical contours bowed outward and horizontal contours inward) and to rotate the entire "fat" shape by 90° to produce a "thin" shape (Figure 2). By using this thin–fat discrimination paradigm, they found evidence both supporting and contradicting the identity hypothesis. Supporting it, they found that discrimination thresholds were similar for illusory and amodal shapes. They also demonstrated that perceptual completion helped extracting inducer orientation information. Namely, when all inducers were identically oriented such that there was no perceptual completion, and when those inducers either all rotated clockwise or counterclockwise, discrimination of the orientation of the inducers, which for an ideal observer is equivalent to the thin–fat discrimination, became more difficult for human observers. Similarly, when all inducers were flipped from a thin–fat configuration to face outward, such that bilateral symmetry was preserved without any perceptual completion, discrimination also worsened. They further demonstrated the importance of boundary completion (Figure 2C): line segments coinciding with the boundaries of the Kanizsa square impeded thin–fat discrimination for illusory and amodal shapes alike. All these were consistent with the identity hypothesis. It is important to note, however, that most observers in Ringach and Shapley were well practiced, and all observers had low discrimination thresholds.

Ringach and Shapley (1996) also found evidence disconfirming the identity hypothesis. When a briefly presented (117 ms) stimulus was immediately followed by a mask, the threshold for amodal discrimination nearly doubled that for illusory discrimination. This difference diminished only when the presentation time was increased to 167 ms. They suggested that there is a fundamental difference between illusory and amodal contour completions but cautioned that this difference may only reflect differential processing time rather than differential completion processes.

Kellman, Yin, and Shipley (1998) also applied the thin–fat discrimination task to study illusory, amodal, and quasi-modal completions. The quasi-modal display was created either by closing the "mouths" of two of the four inducers or by stereoscopically putting two inducers

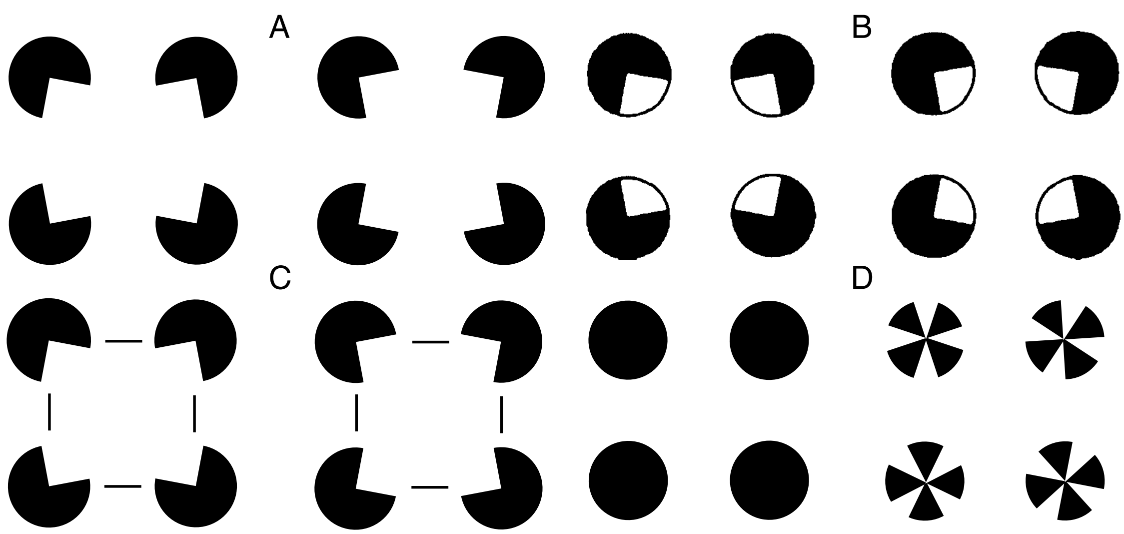

Figure 2. (A) An illusory "fat" shape and its 90° rotational "thin" counterpart. (B) Amodal "fat" and "thin" shapes. (C) The line interference condition for the illusory shapes. (D) Two masks (disks and pinwheels). The contrast polarity of the experimental stimuli was reversed from the illustrations here. Adapted from Ringach and Shapley (1996).

behind, and two in front of, the Kanizsa shape. As a result, the quasi-modal percept is that two disks are partially occluded by a Kanizsa shape, whose two other corners are visible through two holes, while the boundaries are unspecified between the occluding plane in front of the Kanizsa shape and the background plane behind it. The stimulus presentation time was unlimited, and naive participants responded as quickly and accurately as they could. The response times across the three stimuli were comparable, with 98% accuracy. From this equal performance and the argument that contour completion appears continuous between an illusory and an amodal contour in quasi-modal displays, Kellman et al. (1998) concluded that "amodal and modal completion depend on a common underlying mechanism that connects edges across gaps" (p. 859). This conclusion implies that if the performance is different, then it does not support the common mechanism hypothesis.

Using the same paradigm of thin–fat discrimination, Gold, Murray, Bennett, and Sekuler (2000) showed results that, at first glance, supported the identity hypothesis. They used classification images (Ahumada, 1967) to reveal the strategies in thin–fat discrimination. Specifically, illusory or amodal Kanizsa shapes (Figures 2A and 2B) were imbedded in independent additive luminance noise, and well-practiced observers discriminated thin vs. fat shapes. The resultant classification images were virtually identical for the illusory and amodal conditions, when noise pixels along the illusory or amodal contours affected equally an observer's response.

However, using four L corners to replace the four inducers in the Kanizsa configuration, Murray (2002) found that the classification image was similar to those in Gold et al. (2000), even though the L-corners perceptually generate no contour completions. Murray suggested that the classification images of the contours may have little to do with perceptual completion, but with observers' internal templates instead, regardless whether the stimulus was illusory, amodal, or L-corners. Extrapolating from Murray, we speculate that observers in Gold et al. might not be able to tell whether the shape was illusory or amodal because the noise was relatively high in a classification image study, which might render the circular outline intended for inducing the amodal percept hardly visible. Furthermore, the participants were not required to make an illusory-amodal judgment.

In summary, most of the experiments of thin–fat discrimination of Kanizsa shapes found no difference between the illusory and amodal conditions. Except in Kellman et al. (1998), most observers were well practiced, and two-dimensional stimuli were used, where an amodal shape was created from its illusory counterpart by closing the "mouth" of each inducer. There is evidence that illusory and amodal contour completions develop with different time scales. There is also evidence that similar classification images of illusory and amodal discriminations may have little to do with perceptual completion. It should be noted that although Kellman et al. found no performance difference from the naive observers, the accuracy performance at ceiling (98%) might have masked any differences. Moreover, their unlimited stimulus presentation time might have allowed eye movement, which might be different for different stimuli.

In the remainder of this paper, we present experimental results using naive observers and stereoscopic displays. We also used as controls non-stereoscopic displays. As much as we could, we used experimental parameters similar to those in Ringach and Shapley (1996); conditions beyond those were also tested. To anticipate, we found that thin–fat discrimination was consistently worse for amodal than for illusory shapes (discrimination threshold increased by 82%), in apparent contradiction to the findings of Ringach and Shapley. We were able to replicate the findings in Ringach and Shapley by training our naive participants to reach discrimination thresholds similar to those in Ringach and Shapley, while the threshold difference between the illusory and amodal discriminations diminished.

Our results speak to the mechanism or mechanisms of contour completion, particularly in the context of the current debate about the identity hypothesis. They also inform us about the goal of the computation, which can be assessed from the perspective of Bayesian decision theory. We will elaborate on both in the discussion.

Methods

We conducted nine experiments. The stimuli were presented either in stereo using red-green filters or non-stereo. We tested four types of stimuli: illusory, amodal, illusory with interfering lines, and amodal with interfering lines (Figure 2C).

Apparatus

The experiments in China used a Sony CPD-G220 display (medium luminance, stereo), a Philips 107P20 display (high luminance, stereo and non-stereo), and a FlexScan T561 display (high luminance, non-stereo). All displays had a resolution of 1280 × 1024 pixels and a refresh rate of 60 Hz. At a viewing distance of 51 cm, each pixel subtended 1.54 min of arc. A chinrest helped stabilize the viewing distance. In the stereo condition, a dark viewing box abutted the display. In the non-stereo condition, no viewing box was used in accordance with Ringach and Shapley (1996). The experiments were conducted in dark rooms.

In the stereo condition, both a high and a medium luminance condition were used (background: 59 or 2.5 cd/m^2, inducers: 76.7 or 4.7 cd/m^2, Weber contrast: +30% or

+88%). All luminance values in the stereo condition were measured after the red-green filters. The CIE chromaticity of the red and green images were (0.62, 0.34) and (0.29, 0.61), respectively, on the Sony CPD-G220 display. The transmittance of the red filter was 0.42 for red and 0.014 for green, and that of the green filter was 0.048 for red and 0.27 for green. After luminance calibration and gamma correction, the luminance ratio of the red to green images viewed through the respective filters was 0.99.

In the non-stereo condition, the stimuli were in grayscale. The luminance was 58.6 cd/m^2 for the background, and 76.3 cd/m^2 for the inducers, matching closely the stereo condition, and with a Weber contrast of +30%.

The non-stereo condition was also tested at UCLA using a ViewSonic Graphics Series G225f display with a same resolution and refresh rate as in China. The viewing distance was 63 cm, giving rise to the same visual angle of 1.54 min of arc per pixel. The luminance of the background was 59 cd/m^2, and the Weber contrast of the inducers was +30%.

Stimuli

The diameter of each inducer was 4.37° in visual angle (170 pixels). The center of each inducer was 12.35° eccentric from the fixation, giving rise to a support ratio of 0.25 (the diameter of the inducer over the width of the Kanizsa square). In a different condition, the eccentricity of an inducer center was 8.98° and the diameter was 3.18° (124 pixels) in order to maintain the same support ratio of 0.25. Anti-aliasing was done by projecting an inducer to the screen from a digital matrix four times as large.

In all but one stereo condition, the stimuli were arranged in two depth planes with a disparity of 92 min of arc. Excluding the fixation mark and the lines in the "line interference" conditions, the illusory and amodal stimuli were generated by swapping the left- and right-eye images. In one particular stereo condition, three depth planes were used. The monitor frame was in the middle; the other two planes were ±92 min of arc in disparity.

Since it is ecologically invalid to see through an opaque plane, the fixation mark was always in the front plane. It was a disk with a diameter of 61.6 min of arc (40 pixels). Inside the disk, there was a "×" sign, whose arms were 4 by 40 pixels. The interfering lines, when present, were also in the front plane. Each line was four pixels thick and half as long as the Kanizsa square. In the non-stereo condition, the amodal stimuli were created by closing each inducer from the outside with a ring of six pixels thick, in order to maintain the support ratio. The disk masks were used in most of the experiments. They shared the same 3D position and size as the inducers. Pinwheel masks will be specified when used. They were used rarely because our naive participants found them exceedingly difficult. The luminance of the inducers, fixation disk, interfering lines, and masks was the same. The luminance of the fixation sign and background was also the same.

Procedure

A psychometric function was measured in each experimental condition using the method of constant stimuli. Ten angles of inducer rotation were used: 0.1, 0.3, 0.5, 0.8, 1, 2, 3, 4, 5, and 6°. There were four, blocked experimental conditions: (illusory, amodal) × (interfering lines, no lines). For each block of 500 trials, the inducer rotational angles were randomly interleaved. The order of the experimental conditions (blocks) was counterbalanced between participants by assigning a new, random block sequence to an odd-numbered participant and the reversed sequence to the next, even-numbered participant. Each block took about 10 to 15 minutes to complete. A participant typically rested for a few minutes between blocks.

The presentation sequence in each trial was as follows: fixation from the preceding trial, 117 ms of the Kanizsa stimulus, 50-ms fixation, 300 ms of the mask, and fixation until response. The participant responded with a key press to indicate if the stimulus was "thin" or "fat." The next trial started immediately after the response. Auditory feedback was provided per trial only in Experiments 2 and 6. The fixation mark and, in the relevant conditions, the interfering lines were present during the entire block to sustain binocular fusion. Participants were reminded to maintain fixation at all times. They could take an optional rest after every 25 trials. They were told that accuracy was important, but were not told to respond as quickly as possible.

Before an experiment, the participants were presented with an illustration of the illusory and amodal stimuli and were explained what "thin" and "fat" meant. They then practiced the task with the stimuli of 6° inducer rotation, with and without lines, until they were correct on 9 out of the last 10 trials. The practice typically took 75–100 trials to complete.

Data analysis

Data were analyzed in two ways. First, to facilitate direct comparison with Ringach and Shapley (1996), we plotted the probability of correct identification as a function of the absolute value of inducer rotation. Whereas Ringach and Shapley used Quick function (Watson, 1979) to fit the data to estimate the threshold at 81.6% correct, this approach was problematic for some of our inexperienced participants, whose performance was often below the 81.6% criterion (particularly in the interfering lines conditions). Therefore, instead of fitting a psychometric function to estimate threshold, we applied

ANOVA to the unfitted data to test for any differences between conditions, treating inducer rotation, occlusion, and interfering lines as within-subject factors. Thresholds were estimated in the overall analysis in Meta-analysis of Experiments 1–9: Practice trials and response time section in addition to the learning experiment in Experiment 9.

Second, we fitted a cumulative Gaussian function to the frequency of responding "thin" as a function of the signed inducer rotational angle. The mean of the Gaussian distribution represents the point of subjective equality (PSE) for "thin" vs. "fat" and thus quantifies the response bias of an observer. The reciprocal of the standard deviation represents the observer's sensitivity in discriminating "thin" versus "fat." The sensitivity and bias from the cumulative Gaussian fittings were found to be consistent with the results from the ANOVA. In the interest of space, however, we will show the cumulative Gaussian analysis only in Experiment 2. For the same reason, although we will report all ANOVA effects that are statistically significant, we will focus on the main effect of illusory versus amodal discriminations.

Participants

Eighty-seven students from the Chinese University of Science and Technology, naive to the purpose of the experiments, participated. Except for the learning study, each student participated in only one experiment. Author JWZ, author ZL, and participant DLR from the Ringach and Shapley (1996) study also participated. All participants had normal or corrected-to-normal monocular vision, and normal binocular vision.

Experiments

Experiment 1: Discrimination in stereo at medium contrast (+30%)

Figure 3 shows the results from 10 participants. All ANOVA effects were at least marginally significant. The main effect of occlusion was marginally significant: $F(1, 9) = 4.14, p = .07$. The main effect of line interference was significant: $F(1, 9) = 23.47, p = .001$. The main effect of inducer rotation was highly significant, as expected, $F(9, 81) = 139.94, p \ll .001$. The interaction between occlusion and angular rotation was significant: $F(9, 81) = 2.40, p = .018$, indicating that when the angular rotation was sufficiently large, illusory shapes were easier to discriminate than amodal shapes. The interaction between occlusion and line interference was marginally significant: $F(1, 9) = 4.72, p = .058$. The interaction between inducer rotation and interference was significant: $F(9, 81) = 3.78, p = .001$. The three-way interaction between occlusion,

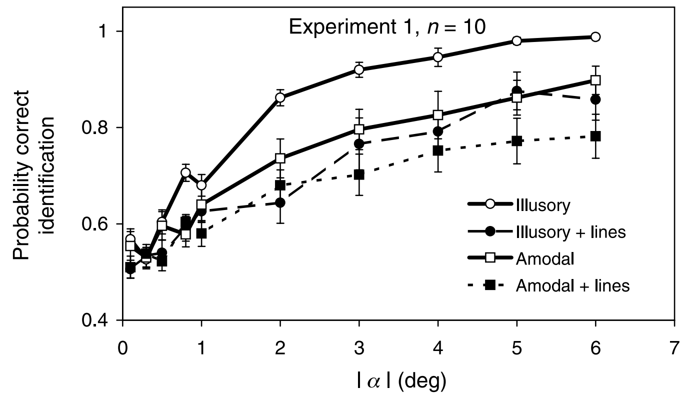

Figure 3. Thin–fat discrimination accuracy as a function of the angular rotation $|\alpha|$ of the inducers, in illusory and amodal conditions, from 10 naive participants. In the range of rotational angles where the performance was above chance, illusory discrimination was better than amodal discrimination. The interfering lines also impeded discrimination. Error bars represent *SEM*.

inducer rotation, and line interference was also significant: $F(9, 81) = 2.91, p = .005$. Since the comparison between illusory and amodal discriminations without interfering lines was of high interest, we compared these two, the result was significant: $F(1, 9) = 6.89, p = .028$.

The two major results are therefore (1) discrimination was better for illusory than for amodal shapes, and (2) the presence of interfering lines impeded discrimination.

Experiment 2: In stereo at high contrast (+88%) with feedback

In order to test whether the results in Experiment 1 were specific to the stimulus contrast, in this experiment, we lowered the background luminance from 59 to 2.5 cd/m^2, and the inducer luminance from 76.7 to 4.7 cd/m^2, giving rise to a Weber contrast +88% for the inducers. Eleven naive observers and author JWZ participated, with auditory feedback. The experiment was otherwise identical to Experiment 1.

As shown in Figure 4, the results from the 11 naive observers were similar to those in Experiment 1. Namely, the main effect of inducer rotation $|\alpha|$ was significant, $F(9, 90) = 143.24, p \ll .001$. The main effect of occlusion was significant: $F(1, 10) = 9.80, p = .011$. The main effect of line interference was significant: $F(1, 10) = 7.93, p = .018$. Finally, the interaction between inducer rotation and line interference was also significant: $F(9, 90) = 3.30, p = .0016$. The remaining three interactions were not significant.

We fitted a cumulative Gaussian function to each of the 12 participant's frequency-of-thin-response data per condition (Figure 5) and estimated the participant's bias and sensitivity.

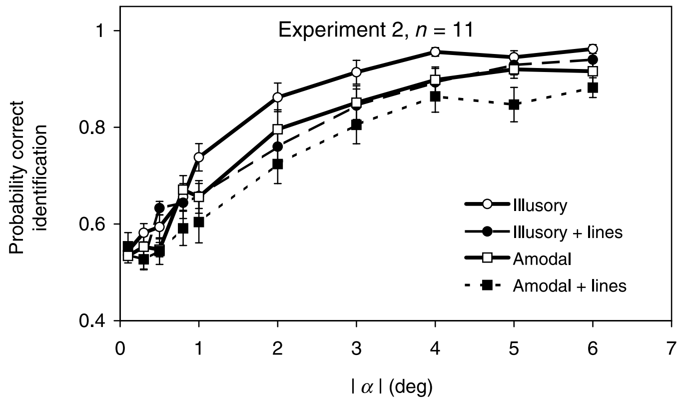

Figure 4. Thin–fat discrimination psychometric functions, when the stimulus Weber contrast was +88% as compared to +30% in Experiment 1. Feedback was provided per trial.

The ranges of R^2 as a result of the fitting were: illusory [.91, .99], amodal [.63, .97], illusory + lines [.73, .96], and amodal + lines [.55, .95]. ANOVA was applied to each condition's 12 standard deviations (1/sensitivity) and means, respectively. Illusory discrimination was more sensitive than amodal discrimination: $F(1, 11) = 10.86$, $p = .0071$. Discrimination with interfering lines was less sensitive than without: $F(1, 11) = 7.42$, $p = .020$. The interaction was not significant. There was a statistically significant bias responding "fat" with interfering lines than without: $F(1, 11) = 5.22$, $p = .043$. This bias was consistent with the same finding in Ringach and Shapley (1996), although its cause is not understood. No other effect was significant.

When author JWZ's data were excluded, exactly the same pattern of results and statistical significance remained. Since author JWZ coded and pre-tested all experimental programs, he was relatively more experienced with the stimuli and task. Indeed, his overall accuracy was 82%, as compared to 74% from the remaining 11 participants. Nevertheless, his thresholds at 81.6% correct showed the similar ordering as the rest of the participants: illusory, 0.93°; amodal, 1.68°; illusory + lines, 1.69°; and amodal + lines, 3.23°. Thresholds are reported here to illustrate that although JWZ's illusory threshold was comparable to those in Ringach and Shapley (1996), his amodal threshold was 80% greater.

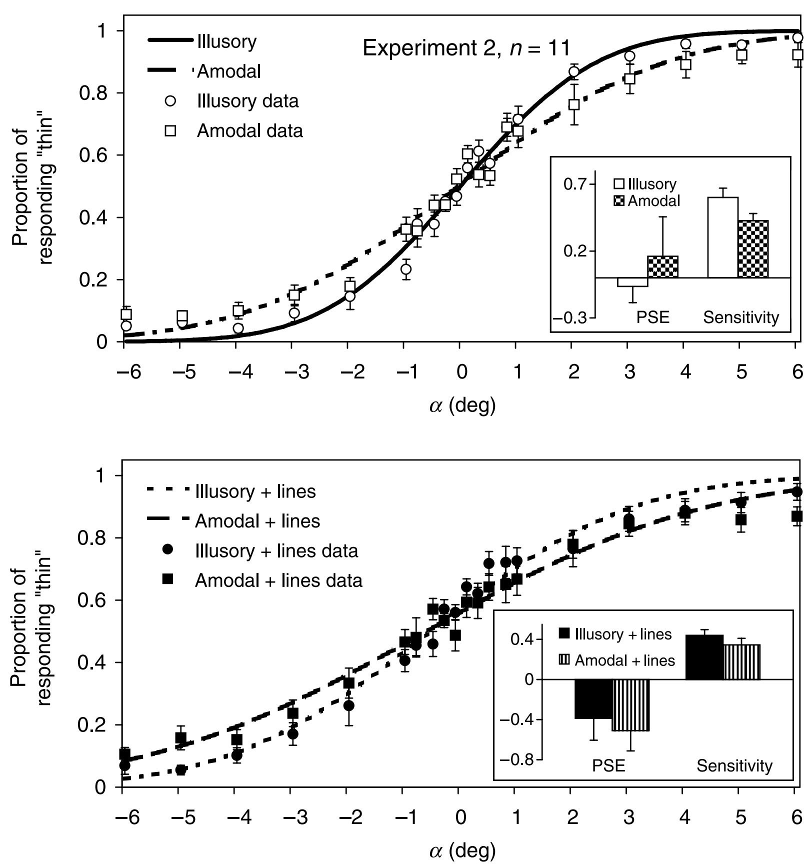

Figure 5. Frequency-of-thin-response psychometric functions, without and with interfering lines, fitted with cumulative Gaussians. Each cumulative Gaussian used the average mean and standard deviation from the 11 naive observers in the corresponding condition, excluding author JWZ, who was tested with a different set of α angles. The analyses gave rise to the same results with or without JWZ's data. In the insets, average PSE (mean) and sensitivity (reciprocal of standard deviation) were plotted.

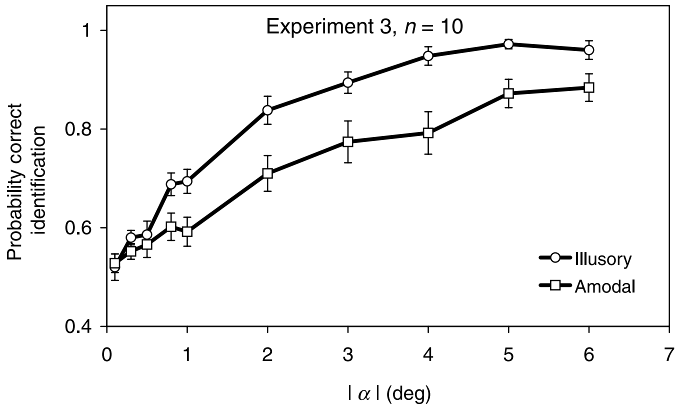

Figure 6. Repeating Experiment 2 without feedback and without interfering lines.

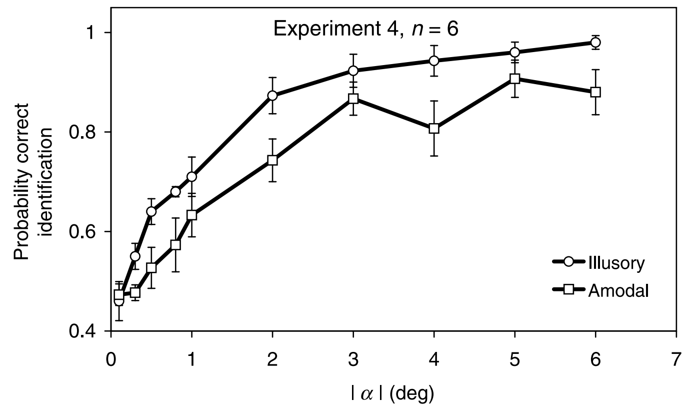

Figure 7. Psychometric functions of illusory and amodal thin–fat discriminations when the eccentricity of each inducer was 8.98° instead of 12.35°. The support ratio of 0.25 was retained. The experiment was otherwise identical to Experiment 3.

Experiment 3: In stereo at high contrast (+88%) without feedback

In Experiment 2, not only contrast was changed from Experiment 1, but feedback was added also. In order to ensure that contrast change alone could retain the effect, we repeated Experiment 2 without feedback, without the interfering lines, and with 10 fresh participants. Otherwise this experiment was identical to Experiment 2.

As shown in Figure 6, all effects were significant. The main effect of occlusion was significant: $F(1, 9) = 9.62$, $p = .013$. The main effect of inducer rotation was significant: $F(9, 81) = 114.29$, $p \ll .001$. The interaction was also significant: $F(9, 81) = 2.79$, $p = .0067$.

Experiment 4: With a smaller stimulus

Ringach and Shapley (1996) used both 12.35° and 8.98° as inducer eccentricity and found no threshold difference between the illusory and amodal conditions. Here, we repeated Experiment 3 with six fresh participants. The eccentricity of each inducer was reduced from 12.35° to 8.98°, and the diameter from 4.37° to 3.18° to maintain the support ratio of 0.25.

ANOVA yielded very similar results as in Experiment 3 (Figure 7). The main effect of inducer rotation was significant, $F(9, 45) = 78.75$, $p \ll .001$. The main effect of occlusion was also significant, $F(1, 5) = 17.24$, $p = .0089$.

Experiment 5: Discrimination at a constant depth

So far in our stereo depth manipulation, there were two depth planes. In the illusory condition, the Kanizsa shape was in the front plane, and the disks were in the back. In the amodal condition, the depth ordering was reversed. In order to test whether the absolute depth of the Kanizsa shape was responsible for the better discrimination in the illusory condition, we employed in this experiment three depth planes.

Specifically, the perceived Kanizsa shape was always in the middle, monitor plane. In the illusory condition, the four inducers were presented in the back plane. In the amodal condition, the four inducers were presented in the front plane. Therefore, the absolute depth of the Kanizsa shape remained unchanged between the illusory and amodal conditions. The absolute stereo disparity between the Kanizsa shape and the inducers remained at 92 min of arc. Six fresh students participated in the two conditions without interfering lines. Figure 8 shows the results.

ANOVA again yielded results consistent with previous experiments. The main effect of inducer rotation was significant, $F(9, 45) = 56.82$, $p \ll .001$. The main effect of occlusion was also significant, $F(1, 5) = 27.53$, $p = .0033$.

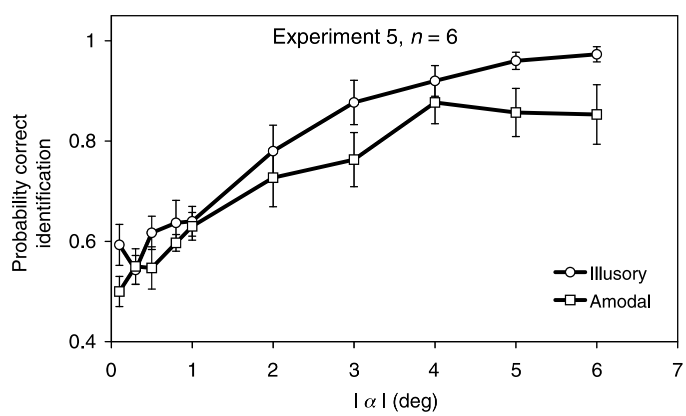

Figure 8. Thin–fat discrimination of illusory and amodal Kanizsa shapes when the shapes' stereoscopic depth was fixed across conditions. This was done by placing the inducers behind the fixed plane of a Kanizsa shape to create the illusory condition, and placing the inducers in front of the Kanizsa shape to create the amodal condition. The eccentricity of the inducers was 12.35°.

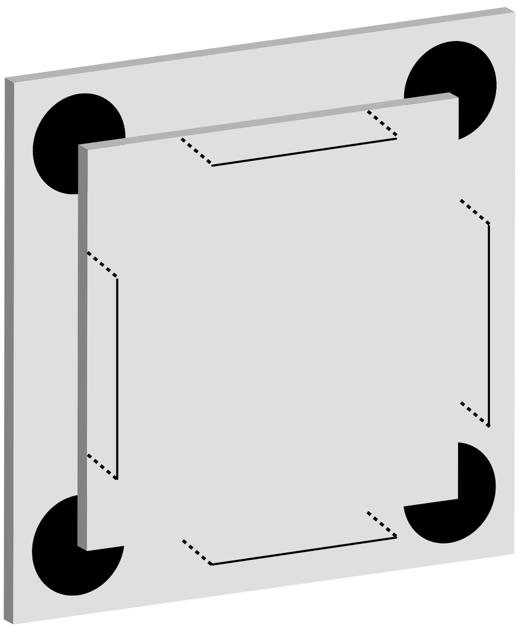

Figure 9. Schematic illustration of a percept in the condition "illusory + lines2." In order to test the effect of interfering lines when they were not in the same depth plane as Kanizsa shapes, the four interfering lines were above the illusory shape. The inducers and illusory shape had zero stereoscopic disparity.

Experiment 6: Zero binocular disparity between the Kanizsa shape and inducers

The purpose of this experiment was twofold. First, we wanted to know whether the better illusory discrimination was limited to stereo displays. To this end, binocular disparity in Experiment 5 was reduced to zero in this experiment (while still using the red-green filters), and the amodal condition was maintained by closing the "mouth" of each inducer.

Second, we had observed that in the interfering conditions, although the illusory and amodal discriminations were both comparably worsened by the lines, the amodal shape was behind the four lines whereas the illusory shape and the four lines were co-planar. Hence, the interference in the amodal condition might be underestimated. To test this, we created a fifth condition, illusory + lines2 (Figure 9), by putting the four interfering lines stereoscopically in front of the illusory shape. The disparity between these lines and the illusory shape was the same as that between the lines and the occluded shape before. In this way, we could test whether the interference was reduced when the four lines and the illusory shape were not co-planar.

Sixteen fresh observers participated with counterbalancing in the five conditions: illusory, amodal, illusory with lines, amodal with lines, and illusory with lines2. Feedback was provided per trial.

ANOVA of the first four conditions yielded results similar to the experiments before. The main effect of inducer rotation was significant, $F(9, 135) = 355.58$, $p \ll .001$. The main effect of occlusion was significant, $F(1, 15) = 23.19$, $p = .00022$. The main effect of line interference was significant, $F(1, 15) = 23.97$, $p = .00019$. The interaction between line interference and inducer rotation was significant, $F(9, 135) = 3.20$, $p = .0015$. Finally, the interaction between occlusion and line interference was also significant, $F(1, 15) = 4.94$, $p = .042$. With the interference, discrimination worsened more for the illusory (78% to 72% correct) than for the amodal conditions (73% to 71% correct).

The last result indicated that the interfering lines impacted the illusory more than amodal shapes. This is probably because, although an amodal shape and the interfering lines had the same disparity of zero, perceptually the amodal shape by definition was behind the plane of the interfering lines. To test this conjecture more directly, we analyzed the three conditions when an illusory shape had no interfering lines, had the lines in the same plane, or had the lines in front. ANOVA revealed a significant main effect of the line conditions, $F(2, 30) = 11.07$, $p = .00025$. The main effect of inducer rotation was significant, $F(9, 135) = 344.05$, $p \ll .001$. The interaction was also significant, $F(18, 270) = 1.84$, $p = .021$. A closer look revealed that the interfering lines being in a different plane had a smaller impact (76% correct) than being co-planar (72% correct), $F(1, 15) = 7.68$, $p = .014$. Yet, lines of different depth still interfered as compared with no lines (76% vs. 78% correct), $F(1, 15) = 4.42$, $p = .053$. Figure 10 shows the results.

Finally, we combined data from the first four conditions in this experiment ($n = 16$) and Experiment 2 ($n = 11$) to check whether stereo manipulation was any different from closing inducer "mouths" when stereo disparity was zero. The main difference between the two experiments was non-zero vs. zero binocular disparity. To our surprise, no

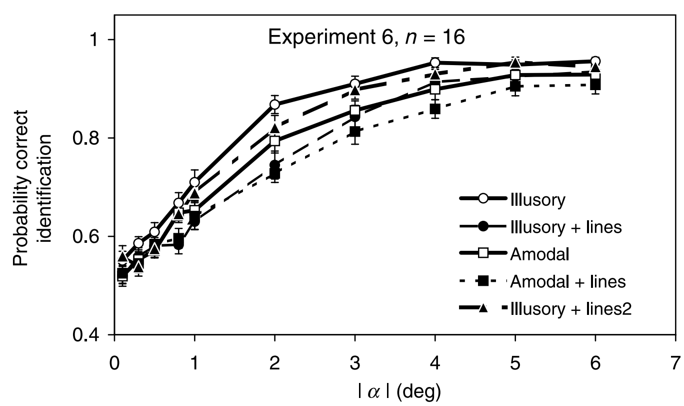

Figure 10. This experiment was similar to Experiment 2, except that the stereoscopic depth difference between the Kanizsa shape and inducers was zero. To create the amodal condition, a ring was added to close each inducer "mouth." A fifth condition was created by placing in depth the four interfering lines in front of the illusory shape. This new condition was termed "illusory + lines 2."

difference whatsoever was found with the stereo disparity manipulation, $F(1, 25) = .003 \ll 1$. Apparently, the perceived occlusion relationship was primarily responsible for the difference between illusory and amodal discrimination, no matter if the occlusion relationship was created stereoscopically or pictorially.

The other effects remained: the main effect of occlusion was significant, $F(1, 25) = 27.19$, $p \ll .001$. The main effect of line interference was significant, $F(1, 25) = 27.99$, $p \ll .001$. The main effect of inducer rotation was highly significant, $F(9, 225) = 457.39$, $p \ll .001$. The interaction between line interference and inducer rotation was also significant, $F(9, 225) = 5.79$, $p \ll .001$.

Experiment 7: Replicating Ringach and Shapley (1996) with naive observers

Our experiments so far yielded better illusory than amodal discriminations, whereas Ringach and Shapley found no difference. One possible explanation of the discrepancy is that our participants were psychophysically inexperienced, whereas Ringach and Shapley's were well practiced (except one, participant VRB, who was inexperienced but aware of the experimental purpose). Numerically, the angular threshold at 81.6% correct was approximately 1° in Ringach and Shapley, whereas that in our Experiment 2 was 2.5°. We also had to use the easier disk masks in our experiments because our inexperienced participants had thresholds exceeding 7° in a pilot study, when the highly effective pinwheel masks were used. In comparison, Ringach and Shapley found that both the disks and pinwheels yielded similar thresholds.

In the current experiment, we tested whether illusory discrimination remained better than amodal discrimination for inexperienced participants, using stimuli identical to those in Ringach and Shapley (1996), but with the easier disk masks. Specifically, we used the same Weber contrast (+30%), luminance, and stimulus size (12.35°). Ten inexperienced observers participated in the two no-interfering conditions, without the red-green filters and with grayscale stimuli.

ANOVA again yielded very similar results as before (Figure 11). The main effect of inducer rotation was significant, $F(9, 81) = 167.90$, $p \ll .001$. The main effect of occlusion was significant, $F(1, 9) = 5.72$, $p = .040$. Their interaction was also significant, $F(9, 81) = 2.27$, $p = .025$.

We noticed that the difference between illusory and amodal discriminations, though statistically significant, appeared smaller than in Experiments 1–6. To investigate the possible causes, we analyzed jointly the data from this experiment and those from Experiment 1. Both experiments shared the same stimulus contrast (Weber +30%). The differences are that occlusion in Experiment 1 was manipulated with binocular disparity, whereas here it was manipulated pictorially in two dimensions (2D). The main effect of stereo (70% correct) versus 2D viewing (74% correct) was marginally significant, $F(1, 18) = 3.33$, $p = .085$. The main effect of illusory versus amodal conditions remained significant, $F(1, 18) = 7.83$, $p = .012$. The main effect of inducer rotation and its two-way interactions with the other two factors were all significant ($p < .001$).

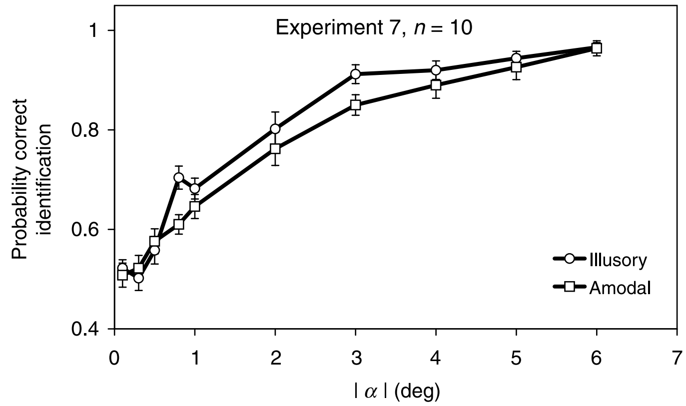

Figure 11. Replicating Ringach and Shapley (1996) when the stimuli were grayscale and no red-green filters were used. The amodal condition was created by closing up each inducer's "mouth" with a ring. No interfering lines were used. The experiment was otherwise identical to Experiment 1.

We extended this investigation by also considering the high-contrast experiments (Weber contrast of +88%): Experiment 6 ($n = 16$, zero disparity) and Experiment 2 ($n = 11$, non-zero disparity). In order to balance the degrees of freedom in the ANOVA analysis (so that $n = 11$ in Experiments 2 and 6), we averaged randomly selected six participants' data in Experiment 6 and treated them as one participant's. ANOVA was conducted on data from Experiments 1, 2, 6, and 7, with the between-subject factors of contrast (+30% vs. +88%), stimulus disparity (zero vs. non-zero), and the within-subject factor of occlusion (illusory vs. amodal). Only the main effect of illusory vs. amodal was highly significant, $F(1, 38) = 21.96$, $p \ll .001$. We concluded therefore that the apparent smaller effect in the current experiment was due to random variations of data, not to any systematic experimental manipulations such as stimulus contrast or stereopsis.

Experiment 8: 200-ms (as opposed to 117) stimulus presentation time, in stereo

Thus far, all experiments in this study used 117-ms stimulus presentation time, which was followed by 50-ms fixation and then the disk mask. The same 117 ms was used in Ringach and Shapley (1996) that yielded comparable thresholds for illusory and amodal conditions. Experiment 8 tested whether any difference between illusory and amodal discriminations would diminish with a longer stimulus presentation time. Two hundred milliseconds was chosen, but not longer, to minimize eye movements. The smaller stimulus of 8.98° inducer eccentricity was also used because there was evidence that the smaller stimulus was easier to discriminate (see Experiment 9). Except the 200-ms stimulus presentation time, this experiment was identical to Experiment 4 (8.98° inducer eccentricity, stereo). Twelve fresh students participated who went through the same practice procedure.

As shown in Figure 12, illusory discrimination was better than amodal discrimination, $F(1, 11) = 7.28$, $p = .021$. The main effect of inducer rotation was, as expected, statistically significant also, $F(9, 99) = 117.16$, $p < .001$. The interaction was not significant.

Data from this experiment were then compared with those in Experiment 4. The only difference between the two experiments was the stimulus presentation time (200 vs. 117 ms). ANOVA revealed that the main effect of stimulus presentation time was not statistically significant ($F < 1$), even though the overall performance with 200-ms presentation was slightly better than with 117-ms presentation (74% vs. 73% correct). The main effect of occlusion (illusory vs. amodal) was significant, $F(1, 16) = 18.04$, $p = .001$. The main effect of inducer rotation was highly significant, as expected, $F(9, 144) = 191.25$, $p \ll .001$. No other effect was significant.

The main result of this experiment was that when the stimulus presentation time was lengthened from 117 to 200 ms, the discrimination advantage of illusory over amodal was hardly reduced. There is never doubt that as the stimulus is presented increasingly longer in time (ignoring eye movement for now), the performance of illusory and amodal discriminations will converge (both are also bounded by the ceiling of 100% correct). The longer is the presentation time, the less significant the results in the current study will be because the illusory versus amodal difference would have existed only in a narrow range of stimulus presentation time. Results from

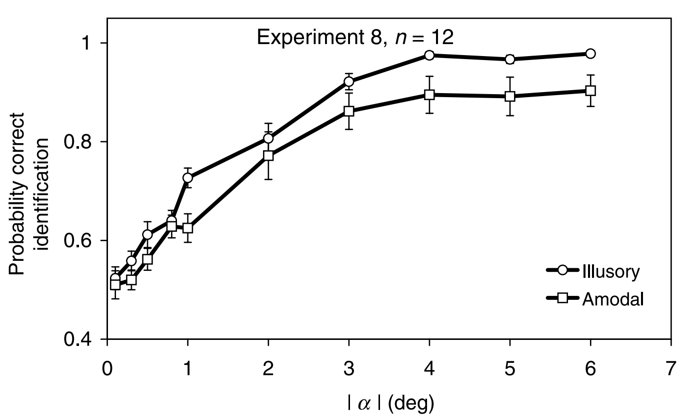

Figure 12. Instead of 117 ms, 200-ms stimulus presentation time was used that was followed by 50-ms fixation and then the disk mask. The inducer eccentricity was 8.98°, and the stimuli were presented in stereo. Except the 200-ms stimulus presentation time, this experiment was identical to Experiment 4.

Experiment 8 indicate that the discrimination advantage of illusory over amodal was not reduced from 117- to 200-ms stimulus presentation.

Experiment 9: Interaction between learning and occlusion

The purpose of this experiment was to test the hypothesis that the discrepant results between our experiments and Ringach and Shapley's (1996) were due to learning. Namely, while inexperienced participants better discriminated illusory than amodal shapes, this difference would be reduced through extensive practice.

Two students were trained who had piloted in replicating Ringach and Shapley's (1996) pinwheel mask experiment. Author JWZ and another student from Experiment 2 were also trained. Prior to training, all participants practiced the task with the inducer rotation $|\alpha| = 6°$, with and without the interfering lines, until they were correct 9 out of the last 10 trials. Prior to training, they also completed the four 500-trial blocks of the experiment to measure psychometric functions, (illusory, amodal) × (interfering lines, no lines). Author JWZ had also the additional experience of coding and testing all experiments.

Similarly as in Ringach and Shapley (1996), non-stereo grayscale stimuli were used, with disk masks, and with inducer eccentricity of 12.35°. In each daily training session, psychometric functions of illusory and amodal discrimination were measured in two blocks (500 trials each), without interfering lines. The order of the two blocks was counterbalanced from one session to the next.

As shown in Figure 13, all four participants' discrimination thresholds decreased through training. The final thresholds for illusory and amodal discriminations were similar to each other and were comparable to those in Ringach and Shapley (1996).

Interestingly, we noticed that the illusory and amodal thresholds for the four trainees in Figure 13 were comparable even in their first training session, after their initial, full-length psychometric function measurement in illusory and amodal discrimination. The comparable illusory and amodal thresholds throughout training raised two possibilities. (1) The threshold difference found in Experiments 1–8, albeit with much practice, was only transient, making the entire results less interesting. (2) There was substantial transfer between the illusory and amodal conditions since all four participants trained with both conditions daily. In order to tease apart these possibilities, two additional naive participants trained with exactly the same conditions as the four trainees above, except that only one of the two conditions would be trained. We had pre-determined that the condition with a lower initial threshold would be trained because transfer to the higher initial threshold condition would be easier to demonstrate than otherwise.

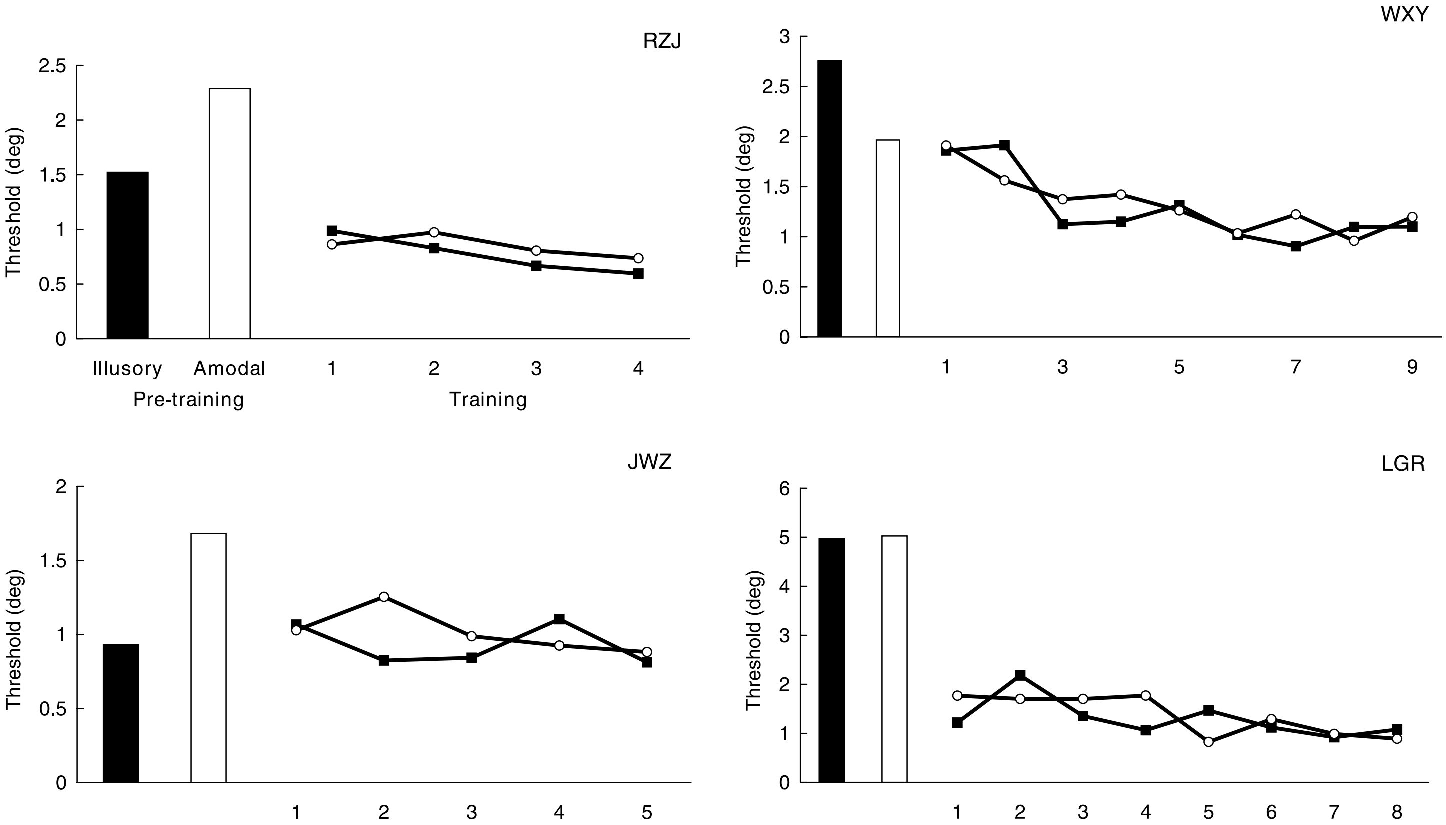

Figure 13. Learning of four practiced participants, including author JWZ. Trainee RZJ and JWZ had completed Experiment 2. WXY and LGR had participated in the pilot experiment with the grayscale stimuli and pinwheel masks, which partially explained the large thresholds. The *y*-axis is the discrimination threshold defined as the inducer rotation that gave rise to 81.6% correct discrimination. All four participants subsequently trained with the 2D grayscale stimuli and with the 12.35° inducer eccentricity. All participants' thresholds ended around 1°, comparable to those in Ringach and Shapley (1996).

For both trainees, the initial, pre-training psychometric functions for both illusory and amodal discriminations were measured with counterbalance (ABBA for one trainee and BAAB for another). For both trainees, the illusory threshold was lower, consistent with previous results. As shown in Figure 14, there was much individual difference between the two trainees. One trained only three sessions for the illusory threshold to reach 1°, whereas another trained 16 sessions. Importantly, there was substantial transfer from the illusory to the amodal discrimination post-training. Furthermore, although the amodal thresholds were reduced from pre- to post-training due to the transfer, the post-training thresholds remained higher than the illusory thresholds. This further indicates that although illusory and amodal completion may share much in common to allow for the transfer, their difference remains.

To further illustrate the difficulty of the task for an unpracticed observer, author ZL tested himself at UCLA with the non-stereo grayscale stimuli, identical to those in Experiment 1 of Ringach and Shapley (1996). In particular, the eccentricity of an inducer was 12.35°. ZL was psychophysically experienced but not with the thin–fat task. His initial thresholds were worse than the six student trainees. ZL trained six sessions in each of the two conditions, with counterbalance. At the end of the training, while his illusory discrimination no longer improved, his amodal discrimination remained below 80% correct. Only after he switched to the smaller amodal stimuli with the 8.98° inducer eccentricity, could he reach the 81.6% correct (Figure 15C). ZL initially trained with the pinwheel masks, with 12.35° inducer eccentricity, with illusory and amodal discriminations, and with three sessions each and counterbalancing. His discrimination was at chance (data not shown).

One author in Ringach and Shapley (1996), DLR, also measured, 12 years later, his psychometric functions at UCLA in illusory (first) and amodal (second) discrimination, with 8.98° inducer eccentricity. His thresholds were illusory, 1.52°, and amodal, 1.95°, consistent with the rest of the participants.

To summarize, we found that with extensive practice with both the illusory and amodal stimuli, naive trainees could reduce their thresholds to about 1° and the threshold difference between illusory and amodal discrimination diminished. This result is consistent with Ringach and

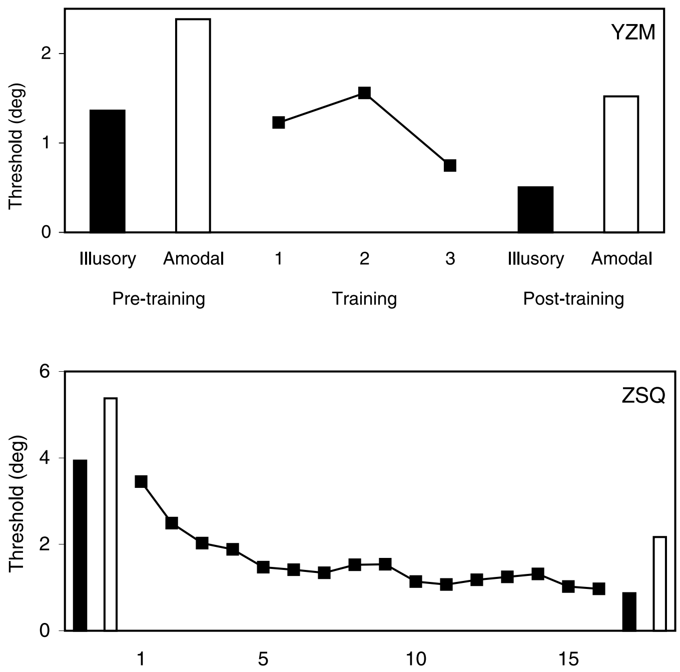

Figure 14. Discrimination thresholds in pre-training, training, and post-training from two naive trainees. Only illusory discrimination was trained, which was chosen because its pre-training threshold was lower than the amodal's. In both pre- and post-training, illusory and amodal psychometric functions were measured with counterbalance. Although there was substantial transfer from illusory to amodal discrimination, the amodal thresholds remained larger than their illusory counterparts.

Shapley (1996). Interestingly, there was substantial transfer between illusory and amodal discriminations, indicating that they indeed share much in common. However, this transfer was incomplete. Participants ZL and DLR, both experienced psychophysics observers, also showed the threshold difference post-training, a result consistent with the rest of the participants and of experiments in the current study.

Meta-analysis of Experiments 1–9: Practice trials and response times

Recall that in each experiment, a participant had to go through practice trials with the inducer rotation at 6° and be correct 9 out of the last 10 trials before starting the main experiment. The number of trials needed to complete the practice before each experiment can further demonstrate the difference between illusory and amodal discriminations. We pooled the practice data, when interfering lines were absent, from all experiments except Experiments 2 and 6 because the data were not saved. Data were excluded when the pinwheel masks were used because no participants could pass the criterion of 9 correct trials out of the last 10, even though 7° and 8° inducer rotations were used instead of 6°. ANOVA yielded a significant difference between the average numbers of trials, illusory: 36, amodal: 83, $F(1, 55) = 9.74$, $p = .003$.

To illustrate the size of the key effect, we also computed the thresholds at 81.6% correct for the illusory and amodal

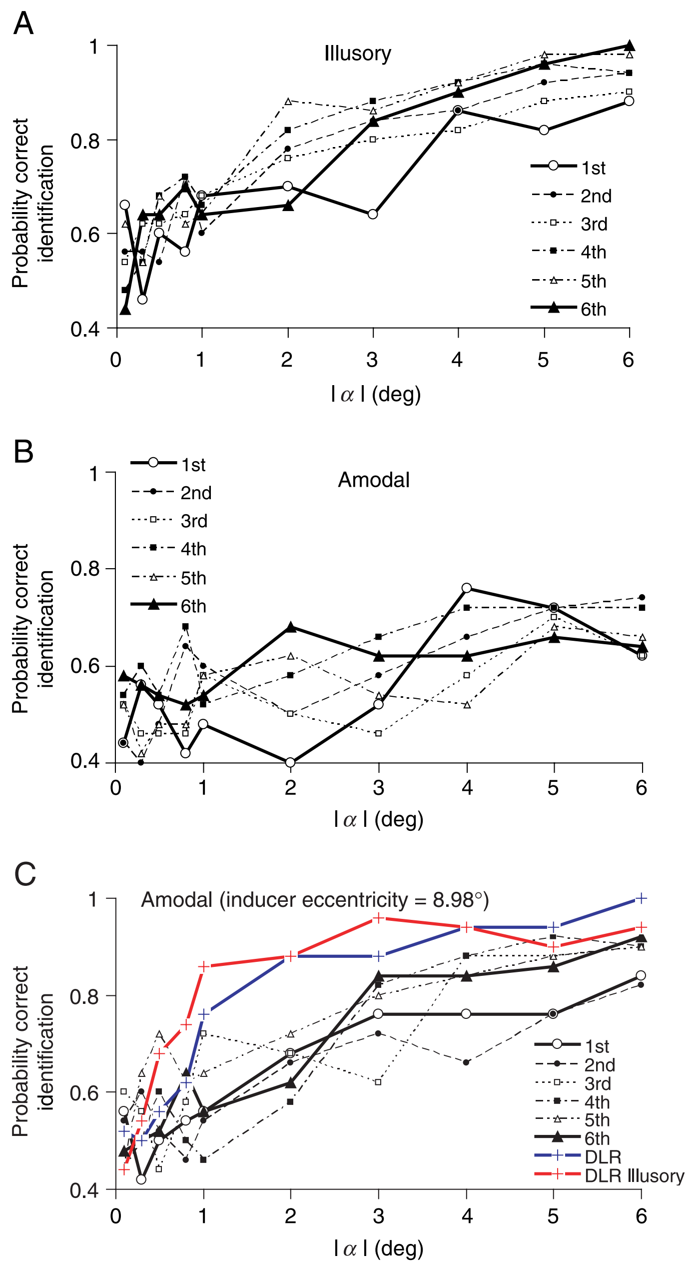

Figure 15. (A, B) Author ZL's psychometric functions when trained with the illusory and amodal discriminations with counterbalance, six sessions each, with the inducers at eccentricity 12.35°. Although his illusory discrimination was acceptable, his amodal discrimination remained below 80% correct. (C) A smaller amodal stimulus, with inducers at eccentricity 8.98° but a constant support ratio of 0.25, was tested afterwards for six additional sessions. ZL finally reached 80% correct. Participant DLR's psychometric function in blue is also shown. For comparison, DLR's psychometric function in illusory discrimination is shown in red.

discriminations from Experiments 1–8, from a total 81 naive participants. The average illusory and amodal thresholds were 1.95° and 3.56°, respectively, a threshold increase by 82%.

Finally, we analyzed response time data and found that an illusory trial was faster than an amodal one, and that correct responses were faster than incorrect responses. No speed-accuracy trade-off was apparent. Specifically, we analyzed response time data from Experiments 1, 4, 5, 7, 8, and 9 that saved them. Response times were faster for illusory trials than for amodal trials (0.74 vs. 0.90 sec), $F(1, 45) = 14.58$, $p \ll .001$. Correct responses were also faster than incorrect responses (0.68 vs. 0.95 sec), $F(1, 45) = 39.90$, $p \ll .001$. We also analyzed response times of correct trials only. Correct illusory trials were faster than correct amodal trials (0.60 vs. 0.76 sec), $F(1, 45) = 17.31$, $p \ll .001$.

Discussion

We set out to answer an empirical question: is the precision in shape discrimination affected by occlusion? In nine experiments, we showed that with identical amount of stimulus information, thin–fat shape discrimination was more accurate and faster for an unoccluded than for an occluded shape. The discrimination difference diminished only after practice.

Our empirical question was motivated by the current debate regarding the identity hypothesis. After elaborating the relationship between our empirical findings and the identity hypothesis, we will propose a different theoretical framework, from a Bayesian perspective, on the computational goals and functional mechanisms for perceptual completion and shape perception.

The identity hypothesis revisited

The identity hypothesis by Kellman and colleagues (Kellman et al., 2005; Kellman & Shipley, 1991) is a prominent theory on perceptual completion. They proposed that modal and amodal contour completions share an identical mechanism, and they supported their hypothesis by demonstrating similar performance with modal and amodal stimuli alike. The theoretical appeal of this hypothesis can be illustrated by the following example (Figure 16A). When stereoscopically fused, the white rectangle is perceived to be slanted in depth that occludes the left column and is visible behind the right column. The argument that illusory and amodal completion share the same mechanism can be summarized as follows. The left or right column, when in isolation, cannot be determined either as illusory or amodal. Such determination is only possible when the white rectangle's location in depth is considered. Likewise, the rectangle's contour on either the left or right side cannot be known to be illusory or amodal until the rectangle's depth relative to the columns are considered. In this sense, according to Kellman and colleagues, it is logically difficult to consider modal and amodal processes as distinct.

Recently, the identity hypothesis has been challenged (Anderson, Singh, & Fleming, 2002; Singh, 2004), leading to an engaging debate (Albert, 2007; Anderson, 2007; Kellman, Garrigan, Shipley, & Keane, 2007). For instance, in a disparity-induced illusory-amodal display, when the left and right images are swapped, the perceived shapes can change qualitatively. In Figure 16B, one percept is that each black disk is occluded by a white corner. The other percept is a black cross visible only through four holes. Hence, the figure and ground (or border ownership) switch roles when binocular disparity is reversed. Using this type of stimuli, Anderson et al. (2002) argued that such qualitative change is inconsistent

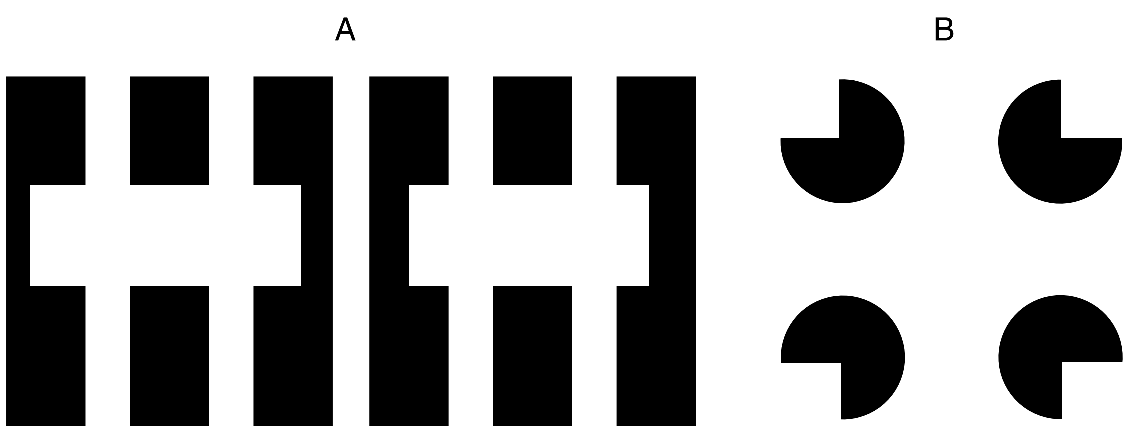

Figure 16. (A) When fused in stereo, a horizontal rectangle is perceived to be slanted in depth, occluding the left column and visible behind the right column. (adapted from Figure 14 of Kellman et al., 2005). (B) An illustration that illusory and amodal percepts can be qualitatively different (Anderson et al., 2002). When the straight contours are closer in depth than the curved, the percept is a white corner occluding each disk. Otherwise, the percept is a black cross visible only at the four inner corners seen through four holes.

with the identity hypothesis since the corresponding illusory and amodal shapes are qualitatively different and it is unclear how the identity hypothesis breaks the symmetry.

The second and related component of the debate is the subjective visibility of illusory contours that appear qualitatively different from amodal contours. For example, observers often report that the illusory contours of a Kanizsa square appear better defined than the amodal contours ("clearer," "sharper") (Petry & Meyer, 1987). Our study quantifies this component: Is the perceived location of an illusory contour better defined such that it gives rise to more accurate shape discrimination? The answer from our data is affirmative, with untrained observers. If one takes the identity hypothesis literally as the perceptual process being absolutely identical for modal and amodal percepts alike, then our results contradicted the hypothesis since the performance difference was most likely due to the brain, not to the stimulus. However, even with minimal stimulus difference, behavioral differences between modal and amodal percepts do not necessarily imply two distinct processes. Ringach and Shapley (1996) suggested that it could be the same process with different parameters.

While keeping these perspectives in mind, we present an alternative on why different performance between illusory and amodal perception makes sense. We believe that this perspective can fruitfully extend the current line of investigation.

The goal of shape computation and its implementation—a Bayesian perspective

Any neural computation can be considered from at least two perspectives (Marr, 1982): the goal of the computation and its implementation. A Bayesian decision theory makes explicit the goal of a computation. With respect to this goal, we can infer how the computation may be implemented. Our conjecture is that misperceiving the occluded region of a shape is ecologically less costly than misperceiving the unoccluded region because occlusion prevents any direct interaction with the occluded region. If the computational goal of the visual system is to keep the expected ecological cost of a perceptual error below an acceptable threshold, then it can allocate less computational resource (neurons and processing time) where errors are less costly. This line of speculation can be formalized in terms of the Bayesian decision theory. Although the speculation contains components that are not yet constrained by the data, we believe that the value of this speculation lies in its ability guiding future research.

The "best" Bayesian percept is the one that *minimizes* the expected cost. Assuming the depth ordering has already been determined, the expected cost $E[C_D(x \mid I)]$ for a particular percept x given an input I can be calculated by summing the cost of all the possible ways of making an perceptual error, weighed by the probability of such an error:

$$E[C_D(x|I)] = \int C_D(x,u)p(u|I)du = \int C_D(x,u)\frac{p(I|u)p(u)}{p(I)}du. \tag{1}$$

For the thin–fat discrimination, x and u represent the shape of a contour; $p(u \mid I)$ is the posterior probability density; $p(I \mid u)$ is the likelihood function describing the precision of the sensory measurements (i.e., given the contour u, the probability of observing the inducers at their orientations); $p(u)$ is the prior probability distribution over all possible contours; $p(I)$ is the probability density for seeing the image; and $C_D(x, u)$ is the cost when the perceptual judgment is x while the reality is u. D denotes occlusion relationship, making explicit that the cost depends on depth ordering. The key premise of our conjecture assume that $C_{\text{occluded}}(\cdot) < C_{\text{unoccluded}}(\cdot)$.

Determining the best percept by Equation 1 can be computationally expensive because the function $E[C_D(x \mid I)]$ can have many local minima. Fortunately, it is often unnecessary to *minimize* the expected cost—it suffices to keep the expected cost below a certain threshold c. This relaxation turns an optimization problem to a constraint-satisfaction problem, which often greatly reduces computational demands (Lasdon, 2002). This is because many of the local minima become irrelevant. Rather than computing the true expected cost function $E[C(\cdot)]$, the system now computes a simpler expected cost function $E[C'(\cdot)]$ with fewer local minima, as long as

$$E[C'(x)] < c \Rightarrow E[C(x)] < c. \tag{2}$$

An immediate behavioral consequence of this approximation is that performance accuracy will decrease because the range of allowable responses is now larger (Figure 17A). More importantly, with the lower cost function assumed for an amodal percept, more computational simplification is possible (Figure 17B). That the visual system may devote less computational hardware in amodal perception is consistent with our finding that amodal discrimination was less accurate and took longer.

For thin–fat discrimination, we assume that the cost of misjudging an occluded shape is less. When the cost of making an error is less, the system can cut corners by approximating contour shapes and measuring sensory information less precisely. We assume that this is the natural state of an unpracticed observer.

With practice, the cost functions ($C_D(x, u)$ of Equation 1) are modified. As the cost functions for the illusory and

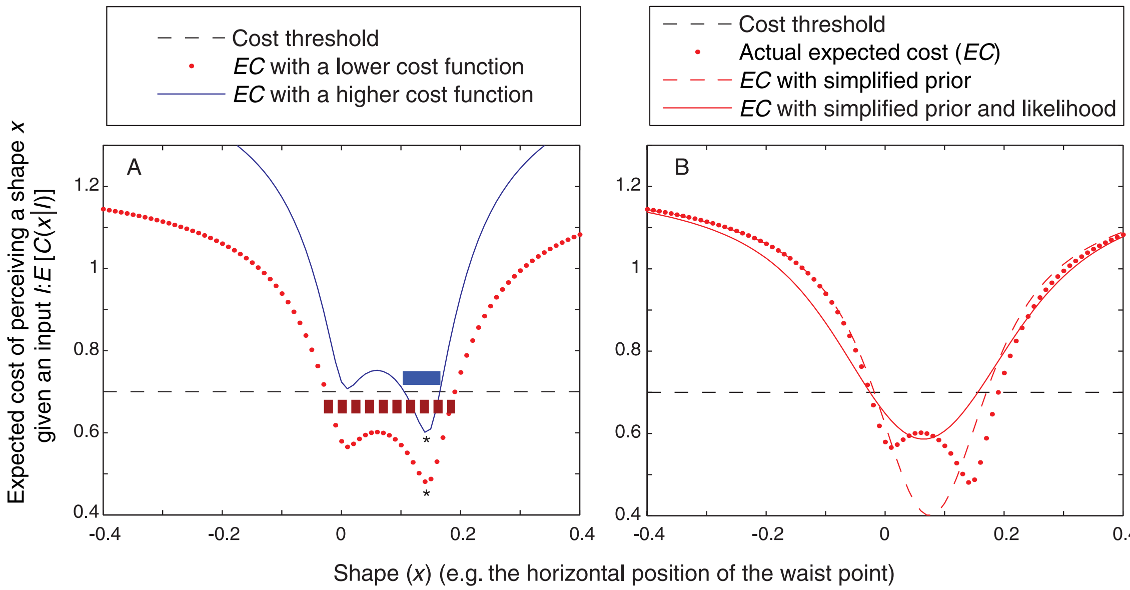

Figure 17. (A) Examples of two hypothetical expected-cost functions $E[C(x \mid I)]$ for two experimental conditions. These functions share the same posterior probability distribution, but one (in • symbol) has a scaled-down cost function ($C_D(x, u)$ of Equation 1). For a system that minimizes the expected cost, the response (in * symbol) is identical for both conditions. In comparison, for a system that keeps the expected cost below a certain threshold (thin dashed line), the ranges of allowable responses (thick solid and dash lines) are different for the two conditions. A larger range of responses (thick dash line) results from the less stringent cost function. This larger range of responses leads to a higher error rate. (B) The response range of a constraint-satisfying system can be approximated by a simpler function to replace the true expected-cost function (in • symbol). Attractive approximations include a simpler prior distribution (dashed curve) and a combination of a simpler prior distribution and a simpler likelihood function (solid curve).

amodal conditions become identical, the simplification in the amodal condition becomes inappropriate, forcing the visual system to adjust its computation in order to bring performance inline with the new cost functions. Behaviorally, equal cost functions lead to equal accuracy, which is consistent with findings by us and by Ringach and Shapley (1996) with trained observers.

When the cost functions become identical, the internal computations could become identical. However, it is more likely that the different computational components have different degrees of flexibility such that the computations remain distinct even after practice. Ringach and Shapley's (1996) finding that shortened stimulus presentation more impeded the amodal condition even for the trained observers is consistent with this scenario, particularly if the putative improvement in sensory measurement is implemented by increasing the sampling duration. That is, the visual system improves its likelihood computation $p(I \mid u)$ in the amodal condition with training, while at the same time retaining an over-simplified prior probability distribution $p(u)$, which was established when the cost function was low.

The idea that training leads to an improved sensory measurement in the amodal condition is testable. For example, one may perturb the inducer orientations with external noise and measure in thin–fat discrimination the rotational threshold as a function of the external noise. This threshold vs. noise function can be used to estimate the equivalent input noise of the system and its sampling efficiency (Gold, Bennett, & Sekuler, 1999; Pelli, 1990; Tjan, Braje, Legge, & Kersten, 1995). A lower equivalent input noise in the amodal condition is indicative of a more precise sensory measurement in inducer orientations.

To reiterate, we have argued that internal computations are different if the cost functions are different. This statement relies on two assumptions: (1) the visual system aims to keep the expected cost of misperception under a certain criterion, and (2) computations are simplified whenever possible due to limited computational resource. With these assumptions, we can explain results of ours and of Ringach and Shapley (1996) by assuming that it costs less to misjudge an occluded than an unoccluded shape and that training alters this cost function.

Our analysis provides a new perspective in the current debate about the identity hypothesis. If different cost functions lead to different computations, then our cost conjecture is inconsistent with a literal interpretation of the identity hypothesis. On the other hand, since the different computations are simplified versions of the true cost function computation, the nature of the computations remains the same. Hence, one could argue that our conjecture is consistent with the identity hypothesis in that the computational differences are quantitative rather than qualitative.

Conclusions

Our results from stereoscopic manipulations strongly implicated a difference between illusory and amodal shape processes in the visual system since there is little difference in the stimulus information between the two conditions. Given that equal performance has been used as supporting evidence for the identity hypothesis, our results contradicted this hypothesis since the performance difference was due to the brain, not to the stimulus. On the other hand, if the performance difference is zero, does equal performance imply an identical mechanism? Not necessarily, because it remains theoretically possible that different mechanisms give rise to the same behavior. We provided a concrete possibility in the Bayesian decision framework: that for trained observers, a more precise sensory measurement compensates for a less precise prior assumption about contour shapes.

Our analysis compares the similarities and differences between perceiving occluded and unoccluded shapes. We assumed that the computational goal is the same in both conditions, but that it costs less to misjudge an occluded shape. Under the constant pressure of reducing computational demands, computing an occluded shape is more simplified. The simplifications include making less precise sensory measurements (i.e., simplifying the likelihood function) and assuming fewer details in representing shapes and their probability distributions (i.e., simplifying the priors). Furthermore, the simplifications may be of varying degrees of permanency such that even when the cost of making a perceptual error is equated through training, the underlying computations can still be different. We believe that these theoretical conjectures are empirically testable. The implied empirical framework provides a systematic and rigorous approach towards testing and interpreting the identity hypothesis.

Acknowledgments

We thank, first and foremost, Dario Ringach for detailed and helpful discussions throughout this study. We thank Phil Kellman for helpful discussions and theoretical debates. We thank the reviewers (including Bart Anderson), Brian Keane, and Nestor Matthews, for helpful comments on the science and English. This research was supported in part by the National Natural Science Foundation of China (No. 30630027 to YZ), the Council on Research of the Academic Senate of the Los Angeles Division of the University of California (ZL), and NSF of USA (BCS 0617628 to ZL).

Commercial relationships: none.
Corresponding author: Yifeng Zhou.
Email: zhouy@ustc.edu.cn.
Address: School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, PR China.

References

Ahumada, A. J. (1967). Detection of tones masked by noise: A comparison of human observers with digital-computer-simulated energy detectors of varying bandwidths. Doctoral dissertation, University of California, Los Angeles.

Albert, M. K. (2007). Mechanisms of modal and amodal interpolation. *Psychological Review, 114,* 455–469. [PubMed]

Anderson, B. L. (2007). The demise of the identity hypothesis and the insufficiency and nonnecessity of contour relatability in predicting object interpolation: Comment on Kellman, Garrigan, and Shipley (2005). *Psychological Review, 114,* 470–487. [PubMed]

Anderson, B. L., Singh, M., & Fleming, R. W. (2002). The interpolation of object and surface structure. *Cognitive Psychology, 44,* 148–190. [PubMed]

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature, 402,* 176–178. [PubMed]

Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology, 10,* 663–666. [PubMed] [Article]

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Huntington, New York: Robert E. Krieger Publishing Company.

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger Publishers.

Kellman, P. J., Garrigan, P., & Shipley, T. F. (2005). Object interpolation in three dimensions. *Psychological Review, 112,* 586–609. [PubMed]

Kellman, P. J., Garrigan, P., Shipley, T. F., & Keane, B. P. (2007). Interpolation processes in object perception: Reply to Anderson (2007). *Psychological Review, 114,* 488–508. [PubMed]

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23,* 141–221. [PubMed]

Kellman, P. J., Yin, C., & Shipley, T. F. (1998). A common mechanism for illusory and occluded object completion. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 859–869. [PubMed]

Knill, D. C., & Kersten, D. (1991). Ideal perceptual observers for computation, psychophysics and neural networks. In R. Watt (Ed.), *Pattern recognition by man and machine*. London: MacMillan Press.

Lasdon, L. S. (2002). *Optimization theory for large systems*. Mineola, NY: Dover.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Murray, R. F. (2002). Perceptual organization and the efficiency shape discrimination. Doctoral dissertation (p. 164), University of Toronto.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception, 18,* 55–68. [PubMed]

Pelli, D. G. (1990). The quantum efficiency of vision. In *Vision: Coding and efficiency* (pp. 3–24). Cambridge: Cambridge University Press.

Petry, S., & Meyer, G. E. (1987). *The perception of illusory contours*. New York: Springer-Verlag.

Ringach, D. L., & Shapley, R. (1996). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision Research, 36,* 3037–3050. [PubMed]

Singh, M. (2004). Modal and amodal completion generate different shapes. *Psychological Science, 15,* 454–459. [PubMed]

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research, 35,* 3053–3069. [PubMed]

Watson, A. B. (1979). Probability summation over time. *Vision Research, 19,* 515–522. [PubMed]