Temporal integration in the perception and discrimination of solid shape

J. Farley Norman¹ · Jessica M. Dukes¹ · Hannah K. Shapiro¹ · Karli N. Sanders¹ · Sarah N. Elder¹

Published online: 3 April 2020
© The Psychonomic Society, Inc. 2020

Abstract

Two experiments evaluated the importance of temporal integration for the perception and discrimination of solid object shape. In Experiment 1, observers anorthoscopically viewed moving or stationary cast shadows of naturally shaped solid objects (bell peppers, *Capsicum annuum*) through narrow (4-mm wide) slits. At any given moment, observers could only see a very small portion of the overall object shape (generally less than 10%). The results showed that the observers' discrimination performance for the moving cast shadows was much higher than that obtained for the stationary shadows, demonstrating the ability to temporally integrate the piecemeal momentary information about shape that was available through the narrow apertures. In a second experiment, estimates of the strength of the observers' impressions of solid shapes rotating in depth were obtained as well as discrimination accuracies; perceptions of the original moving condition were compared with a new condition where the frames of the apparent motion sequences depicting solid objects in continuous motion (behind the slits) were randomly scrambled. The observers perceived the anorthoscopic displays as depicting solid objects rotating in depth, but only in the continuous motion condition. Interestingly, the discrimination performance in the scrambled condition remained relatively high—observers were still able to integrate information across the multiple scrambled frames in order to produce discrimination performance that was significantly higher than that obtained in the stationary shadow condition. This study was the first to thoroughly evaluate whether and to what extent human observers can effectively discriminate and perceive solid object shape anorthoscopically.

Keywords 3D perception: Depth and shape from X · Motion: In Depth

About 50 years ago, Eriksen and Collins (1968) published an article that evaluated the temporal organization of form. In this pioneering study, participants on any given trial were briefly shown a single pattern of points; viewed by itself this initial pattern appeared to depict nothing (i.e., no spatial structure was visible). However, Eriksen and Collins found that when a second nonoverlapping, pattern of points was presented 25 ms later, participants were able (with about 85% correct accuracy) to perceive and identify a meaningful spatial pattern (a three-letter nonsense syllable, such as “HOV”). What was important about this study was that the nonsense syllables were not visible in either the first or second stimulus half;

the nonsense words only became visible following temporal integration occurring somewhere within the visual system.

A second method for studying temporal integration (for the perception of shape) was developed by Zöllner (1862; also see Stewart, 1900). Zöllner moved (horizontally translated) large-diameter circles and squares behind a narrow 2-mm wide slit; all parts of the object were occluded, except for a narrow slice that was visible through the 2-mm wide aperture. Under circumstances like these, a viewer can see only a small portion of the overall figure at a single time. Zöllner found, however, that when a viewed figure is moved back and forth (i.e., oscillated) behind the slit, an observer can see the entire object (even though the perceived objects' proportions were subject to some distortion). This *anorthoscopic perception* of shape obviously requires the integration of information over time. The basic phenomenon discovered by Zöllner has been studied repeatedly since 1965 (e.g., Aydin, Herzog, & Ögmen, 2009; Parks, 1965; Rieger, Grischow, Heinze, & Fendrich, 2007; Rock, 1981; Shimojo & Richards, 1986). The anorthoscopic viewing of 2-D figures produces neuronal activity in the lateral occipital complex (LOC; see Orlov & Zohary, 2018; Yin, Shimojo, Moore, & Engel, 2002); the temporal integration of spatial information needed to

This article was updated to correct an error in the heading of table 2.

✉ J. Farley Norman
Farley.Norman@wku.edu

¹ Department of Psychological Sciences, Ogden College of Science and Engineering, Western Kentucky University, Bowling Green, KY 42101-2030, USA

perceive anorthoscopic shape is therefore the result of cortical processes beyond V1. It is interesting that LOC neurons are activated during anorthoscopic perception, because it has been shown that the LOC can be reliably activated by both visual and haptic object shape (Amedi, Jacobson, Hendler, Malach, & Zohary, 2002; James et al., 2002; Peuskens, Claeys, Todd, Norman, Van Hecke, & Orban, 2004).

In our laboratory, we previously demonstrated that human observers can effectively integrate information over time to recognize various forms of biological motion (Norman, Payton, Long, & Hawkes, 2004b). In Experiment 2 of this study, a dotted figure moved (i.e., walked, jogged, or skipped) behind occluding surfaces that blocked 50% to 85% of the total stimulus display—narrow vertically oriented slits like those used by Zöllner (1862) permitted brief and partial views of the overall patterns of biological motion. Norman, Payton, et al. (2004b) found that younger adults could perform well at discriminating different forms of biological motion even when 85% of the patterns of biological motion were occluded (see their Fig. 6). In almost all of the previous research involving anorthoscopic perception, only two-dimensional stimulus patterns have been used. The goal of the current study was to extend the study of anorthoscopic perception to the perception and discrimination of 3-D object shape. While a previous demonstration exists to show that 3-D anorthoscopic perception is possible (Fujita, 1990), the current study is the first to thoroughly investigate the anorthoscopic perception and discrimination of solid object shapes defined by the *kinetic depth effect*. The term kinetic depth effect literally refers to the perception of depth (i.e., 3-D structure) from motion (for literature relating to the kinetic depth effect, see Andersen, 1996; Bingham & Lind, 2008; Braunstein, 1966; Lappin, Doner, & Kottas, 1980; Norman et al., 2016; Norman & Raines, 2002; Todd, Akerstrom, Reichel, & Hayes, 1988; Wallach & O'Connell, 1953). In most instances of the kinetic depth effect, the projected views of a solid object rotating in depth contain identifiable features (e.g., texture, sharp corners) that clearly mark individual locations on the object's surface. These identifiable

features are important because their motion in projected images can be tracked, measured, and thus used in the computational recovery of an object's 3-D structure (e.g., Koenderink & van Doorn, 1991; Hoffman & Bennett, 1986; Ullman, 1979). In some important real-world situations, however, such identifiable surface structure is not present in projected images. Cast shadows represent one such real-world situation. When environmental light sources illuminate solid objects, shadows of these objects are cast onto background surfaces; when solid objects move (e.g., rotate in depth), their cast shadows will deform. Traditional structure-from-motion algorithms cannot recover 3-D object structure from such shadow deformations (e.g., see Norman & Todd, 1994; Todd, 1985); nevertheless, human observers do recover much 3-D information from deforming cast shadows and silhouettes (e.g., Cortese & Andersen, 1991; Norman, Dawson, & Raines, 2000; Norman et al., 2009; Norman & Raines, 2002). The current experiments expand upon previous research and investigate whether and to what extent human observers can perceive and discriminate the solid shape of objects from the fragmentary pieces of deforming shadows that are visible through narrow slits (cf. Zöllner, 1862).

Experiment 1

Method

Apparatus The stimulus presentations and the collection of participant responses was performed by an Apple Mac Pro computer (Dual Quad-Core processors, with ATI Radeon HD 5770 hardware-accelerated graphics) using an Apple 27-inch LED Cinema Display. The monitor was located at a 60-cm viewing distance.

Stimulus displays The experimental stimuli were cast shadows of natural solid objects (bell peppers, *Capsicum annuum*). The five particular bell pepper replicas used in the current study (see Fig. 1) were a subset of the 12 objects originally created

Fig. 1 Photograph of the five natural objects (plastic replicas of bell peppers, *Capsicum annuum*) that were used to generate the static and deforming cast shadows that the observers viewed anorthoscopically.

The five stimulus objects (1–5 are arranged left to right) are shown with their rotation axes (metal rods). These five objects are a subset of the 12 objects originally created by Norman, Norman, et al. (2004a)

by Norman, Norman, Clayton, Lianekhammy, and Zielke (2004a). Eighty shadows of each of the five objects (each object was rotated in depth over a range of 360 degrees in angular increments of 4.5 degrees) were cast onto a flat surface (see Norman et al., 2009) and recorded using a Nikon Coolpix 995 digital camera; the resulting images had a resolution of $1,024 \times 768$ pixels. The object shadows were viewed through two stationary narrow slits (always 4-mm wide). On each trial, the slits were randomly offset (left and right) from the center of the stimulus by 20 to 32 mm; therefore, the two slits were randomly separated by 40 to 64 mm (this procedure produced additional stimulus variation that was independent of the depicted object shadows; thus, when a particular object was presented multiple times during the experiment, each stimulus image or apparent motion sequence was unique). The occluding surface into which the slits were cut was colored blue, while the background and cast shadow colors were white and black, respectively. Example stimulus images are shown in Fig. 2. In the representative stimulus images depicted in Fig. 2 (bottom row), the occluding surface hides 92.5% of the object shadows; therefore, only about 7.5% of the stimulus object shadows were visible through the anorthoscopic slits/apertures.

Procedure There were two main conditions. In one condition (stationary), observers were shown a single cast shadow (where the orientation of the object about a Cartesian vertical axis was randomly determined) and were required to identify the object (1–5; see Figs. 1 and 2). In the other condition (moving), observers viewed apparent motion sequences of cast shadows (depicting a solid object rotating in depth about a Cartesian vertical axis; the rotation axis of each object can be seen in Fig. 1) and were required to perform the same identification task (the initial orientation of the objects, at the beginning of the apparent motion sequences, was randomly determined). The order of object presentations

(Objects 1–5) was determined completely at random. The duration of each stimulus presentation (moving or stationary) was 3.0 seconds. In the moving condition, the individual frames of the apparent motion sequences were updated at 60 Hz (thus, 180 total frames were presented), and the depicted objects rotated 2.25 revolutions during the 3.0-second stimulus presentation.

The observers' task was straightforward, simply to view the static or deforming (moving) cast shadows on any particular trial and identify which of the objects (1–5; see Fig. 1) had been presented. However, at the beginning of the experiment, the observers (especially the completely naïve ones) did not know which object was Number 1, which was Number 2, and so forth. At the beginning of each experimental session, repeated blocks of 10 trials were conducted, where moving cast shadows (without the occluders and slits) were presented. During this familiarization, the observers would receive a brief auditory beep for correct judgments (so that they could learn which object was which). Once each observer had reached a performance level of at least 90% correct, the experimental trials (where the depicted object shadows were presented behind occluding surfaces through anorthoscopic slits) began. It is important to note that no feedback was ever provided during the experimental trials (when the occluders and slits were present). These procedures were analogous to those used in our previous relevant research (e.g., Norman et al., 2000; Norman et al., 2009).

Each observer participated in six experimental sessions that included 100 trials each; each observer, therefore, made a total of 600 identification judgments. All of the observers made judgments for both the motion and no motion (i.e., stationary) conditions. Each object (1–5) was presented a total of 60 times for each condition (motion and stationary). Each observer followed a completely random order of experimental sessions (three sessions for the motion condition and three sessions for the stationary condition), except that the first session was

Fig. 2 Representative stimulus images are illustrated in the bottom row; the entire corresponding object shadows are visible in the top row (in both rows, Objects 1–5 are arranged left to right). In these example stimulus

images, the occluding surface hides 92.5% of the object shadows; therefore, only about 7.5% of the stimulus object shadows were visible to the observers during the experimental trials

Table 1 Confusion matrix for moving objects/cast shadows viewed anorthoscopically

	Stimulus Object				
	1	2	3	4	5
Responses					
1:	413	12	124	74	2
2:	6	520	32	12	35
3:	79	19	421	30	10
4:	101	3	22	483	1
5:	1	46	1	1	552

devoted to motion for half of the observers (A.M., H.S., J.F.N., R.L., & S.E.) and was devoted to stationary shadows for the remaining half of the observers (J.D., K.S., K.W., L.M., & P.A.).

Observers There were a total of 10 adult observers, who were students and faculty at Western Kentucky University. Half of the observers were the coauthors, while the remaining five observers were naïve with respect to stimulus generation, purposes of the experiment, and so forth. This sample size (10 adults) was two to 2.5 times as large as some of our previous investigations of solid shape discrimination (e.g., Norman et al., 2000; Norman et al., 2009). While this large of a sample size was technically unnecessary (the usage of three or four observers would produce sufficient power) to answer the basic

questions addressed here, it was helpful in that it allowed us to evaluate whether any differences in discrimination performance exist between experienced and naïve observers. The visual acuity of the observers was excellent; the acuity measured at 1 meter was -0.17 LogMAR (log minimum angle of resolution).

Results and discussion

The overall results are shown in Tables 1 and 2, as well as in Figs. 3 and 4. Overall (and individual observer) confusion matrices for the moving and stationary shadow conditions were constructed from the observers' responses (e.g., see Tables 1 and 2, as well as Fig. 3). From such 5×5 confusion matrices, d' values (the signal-detection measure of perceptual

Table 2 Confusion matrix for stationary objects/cast shadows viewed anorthoscopically

	Stimulus Object				
	1	2	3	4	5
Responses					
1:	199	60	157	103	32
2:	46	299	77	74	107
3:	136	51	217	102	20
4:	181	88	113	293	36
5:	38	102	35	28	405

Fig. 3 Experimental results. The observers' responses (1–5) in Experiment 1 are plotted for each of the five stimulus objects for the moving (left panel) and stationary (right panel) shadow conditions. The exact (numerical) frequencies plotted in this figure can be seen in Tables 1 and 2

sensitivity; see Macmillan & Creelman, 1991) were calculated for each observer, motion condition (moving vs. stationary), and pair of objects (1 & 2, 1 & 3, 1 & 4, 1 & 5, 2 & 3, . . . , 4 & 5). This calculation of pairwise object discriminability (for 10 pairs of objects) from 5×5 confusion matrices was analogous to the procedures used in past research (e.g., Bell & Lappin,

1973; Norman et al., 2017; Norman et al., 2000). Figure 4 plots the observers' shape discrimination performances for the 10 different pairs of objects in terms of d' for both moving and stationary objects/cast shadows. The results for the coauthors and those of the naive observers have been combined, because there was no significant difference in their sensitivity

Fig. 4 Experimental results. The observers' shape discrimination accuracies in Experiment 1 are plotted in terms of d' (Macmillan & Creelman, 1991) for each of the 10 pairs of stimulus objects. The filled bars indicate results obtained for moving objects/cast shadows viewed

anorthoscopically, and the open bars indicate results obtained for stationary objects/cast shadows viewed anorthoscopically. The error bars indicate ± 1 standard error

to shape, according to a three-way split-plot analysis of variance (ANOVA), $F(1, 8) = 1.4$, $p = .27$.

As Fig. 4 clearly shows, there was a large effect of the motion: The observers' d' values in the moving condition (mean $d' = 3.68$) were much higher than those obtained for the stationary condition (mean $d' = 1.87$). This effect of motion was shown to be significant, $F(1, 9) = 316.6$, $p < .000001$; $\eta_p^2 = 0.97$, according to a two-way within-subjects ANOVA. There were also significant effects of the object pair, $F(9, 81) = 36.7$, $p < .000001$; $\eta_p^2 = 0.80$, and the Motion $\times$ Object Pair interaction, $F(9, 81) = 2.4$, $p < .02$; $\eta_p^2 = 0.21$. The easiest pairs of objects to discriminate were Objects 1 and 2, 1 and 5, 3 and 5, and 4 and 5, while the hardest pairs to discriminate were Objects 1 and 3 and 1 and 4. While the effect of motion was large for all pairs of objects, it was nevertheless strongest for Pairs 1 and 2, 2 and 4, and 3 and 4 (and least for Pairs 1 & 3 and 3 & 5).

Experiment 2

The results obtained for the moving condition in Experiment 1 clearly showed that human observers can integrate fragmentary pieces of visible object shadows over time in order to effectively recognize and discriminate solid object shape. The observers' ability to recognize individual objects (and distinguish them from others) from static shadows viewed

through narrow slits was much poorer (compare the black and white bars in Fig. 4). While the observers in Experiment 1 agreed that the moving shadows (viewed through narrow apertures) produced compelling perceptions of solid object shapes rotating in depth, no systematic judgments were recorded. One purpose of the following experiment was to ask observers to numerically evaluate the perceived appearance of the moving shadows viewed anorthoscopically. Another important issue concerns the type of information needed for observers to make accurate discrimination judgments. Do observers need continuous (i.e., smooth) deformations of cast shadows in order to effectively recognize the stimulus objects? Perhaps yes, but possibly no. The good performance obtained in Experiment 1 for the moving condition (black bars in Fig. 4) could have simply been a consequence of the fact that 80 individual cast shadows were visible (each one only very briefly) during a single trial in the moving condition (but only a single, randomly determined cast shadow was visible on each trial in the stationary condition). In the following experiment, performance for the previously investigated moving condition (i.e., same as that used in Experiment 1) was compared with performance obtained for a new “scrambled” condition, where the same apparent motion sequences were presented to the observers, but the individual frames (i.e., individual cast shadows) were presented in a random order. Both of the conditions employed in this experiment presented the same exact images of cast shadows to the observers, but

Fig. 5 Experimental results. The observers' shape discrimination accuracies in Experiment 2 are plotted in terms of d' (Macmillan & Creelman, 1991) for each of the 10 pairs of stimulus objects. The filled bars indicate results obtained for continuously moving objects/cast shadows viewed anorthoscopically, and the open bars indicate results

obtained for the scrambled condition (same stimulus displays as in the continuous motion condition, except that the individual frames in the apparent motion sequences have been randomly rearranged). The error bars indicate ± 1 standard error

continuous motion/cast shadow deformation was only present in one condition (present in the motion condition, but absent in the scrambled condition).

Method

Apparatus The apparatus was the same as that used in Experiment 1.

Stimulus displays The stimulus displays used for the moving shadow condition in the current experiment were identical to those used in Experiment 1. The stimulus displays for the scrambled condition were identical to those used in the moving condition, except that the individual frames of the apparent motion sequences were randomly scrambled; instead of the individual frames being presented in sequential order from 1 to 80, the order of the frames in the scrambled condition was completely random (e.g., an order such as 53, 17, 29, 72, 68, 3, 8, 76, 45, . . . 22).

Procedure The procedures used for the shape discrimination task in Experiment 2 were identical in all respects to those used in Experiment 1. At the end of each of the six experimental sessions (2 motion types [continuous motion vs. scrambled] $\times$ 3 experimental sessions/motion type), the observers were asked to numerically rate the subjective appearance of the stimulus displays using the following scale (5 = *stimulus displays looked consistently like solid objects rotating in depth*; 4 = *stimulus displays looked like solid objects rotating in depth most of the time*; 3 = *stimulus displays looked like solid objects rotating in depth some of the time*; 2 = *stimulus displays looked like solid objects rotating in depth only rarely*; 1 = *stimulus displays never looked like solid objects rotating in depth*).

Observers The four observers (A.M., H.S., J.F.N., & P.A.) who participated in this experiment all had previously participated in Experiment 1. Half of the observers (H.S. & J.F.N.) were coauthors, while the remaining observers (A.M. & P.A.) were naïve with respect to stimulus generation, purposes of the experiment, and so forth. The visual acuity of the observers was excellent; the acuity measured at 1 meter was -0.15 LogMAR.

Results and discussion

The results of Experiment 2 are shown in Fig. 5; they are plotted in a manner that is analogous to that of Fig. 4 (results from Experiment 1). It is readily evident that the observers' discrimination performance was higher for the continuous motion condition than for the scrambled condition (the average d' values for the moving and scrambled conditions were 4.12 and 3.38, respectively). This difference was

statistically significant according to a two-way within-subjects ANOVA, $F(1, 3) = 17.7$, $p = .025$, $\eta_p^2 = 0.86$. Of course, it should be noted that despite this numerical difference, the observers' discrimination performance for the scrambled condition was excellent. There were also significant effects of both object pair, $F(9, 27) = 6.8$, $p < .0001$, $\eta_p^2 = 0.69$, and the interaction of object pair and motion type, $F(9, 27) = 3.5$, $p < .01$, $\eta_p^2 = 0.54$. It is clear from an inspection of Fig. 5 that the effect of motion type (continuous motion vs. scrambled) was much larger for some object pairs (e.g., Objects 1 & 3, 1 & 4, 2 & 3, 3 & 4) than for others (e.g., 1 & 2, 1 & 5, 2 & 5); this is the reason for the obtained significant interaction between object pair and motion type. The observers' performance for the scrambled condition, while lower than that obtained for the continuous motion condition, was nevertheless much higher than the discrimination performance that occurred for the stationary condition in Experiment 1 (average d' values for these four observers in the scrambled and stationary conditions were 3.38 and 2.08, respectively). This difference in performance was significant, as shown by a two-way within subjects ANOVA, $F(1, 3) = 23.4$, $p < .02$, $\eta_p^2 = 0.89$.

At this point, it is important to remember that each of the four observers numerically evaluated the perceptual appearance of the anorthoscopic stimulus displays in both the continuous motion and scrambled conditions. The average ratings (for the scale described in the procedure section) for the continuous motion and scrambled conditions were 5.0 and 1.58, respectively. All observers in all experimental sessions devoted to the continuous motion condition therefore agreed unanimously that those stimulus displays “looked consistently like solid objects rotating in depth.” Given that the continuously deforming cast shadows were viewed through narrow (4-mm wide) slits, this is a remarkable outcome. Observers can successfully integrate small and disparate pieces of cast shadows over time in order to obtain compelling perceptions of a solid object rotating in depth. Such perceptions of a solid object rotating in depth almost never occurred for the scrambled stimulus displays. This difference in rated appearance (across the four observers, 12 ratings were provided for the continuous motion experimental sessions, and 12 ratings were similarly provided for the scrambled experimental sessions) for the stimulus displays used in the continuous motion and scrambled conditions was statistically significant (e.g., according to a sign test, $N = 12$, $x = 0$, $p < .001$).

General discussion

The *Oxford English Dictionary* defines integration (“Integration,” 2020) as “the making up or composition of a whole by adding together or combining the separate parts or elements.” About a half-century ago, the research of Eriksen and Collins (1968) certainly demonstrated integration—they

presented human observers with stimulus images that individually contained no visible shapes. When the dotted images were viewed in quick succession, however, observers were able to perceive and identify simple two-dimensional shapes (nonsense syllables, such as “HOV”) with good accuracy (see Figs. 1 and 2 of Eriksen & Collins). Even though our moving and scrambled stimulus displays were quite different (apparent motion sequences depicting object cast shadows versus the sequences of dot patterns used by Eriksen & Collins), our results also demonstrate the visual system’s remarkable ability to effectively integrate information over time in such a manner that permits the effective identification of objects. The results of our scrambled condition (Experiment 2) show that the individual parts (or elements) of our cast shadows can be integrated sufficiently into whole “objects” despite the fact that the individual views of the moving objects were not presented in a typical sequential order. The fact that effective integration can occur under such circumstances demonstrates the extraordinary resilience of the human visual system in its ability to recognize solid objects. In particular, in this scrambled condition, our observers were able to integrate fragmentary pieces of cast shadows (fragmentation caused by anorthoscopic viewing through narrow slits) over time to produce object discrimination performance that was much higher than that obtained by anorthoscopically viewing stationary shadows (average d' values for the scrambled and stationary conditions were 3.38 and 2.08, respectively). Human observers can thus integrate visual information about objects over time, even under conditions that do not produce compelling perceptions of 3-D shape (remember that the scrambled apparent motion sequences in Experiment 2 rarely produced perceptions of solid shapes rotating in depth).

In contrast to the perceptions obtained in the scrambled condition of Experiment 2, the stimulus displays depicting continuous motion/deformation of cast shadows did produce compelling perceptions of solid objects rotating in depth despite the fact that these deforming cast shadows were anorthoscopically viewed through narrow slits. With regard to perceived 3-D shape from motion (e.g., Andersen, 1996; Braunstein, 1966; Lappin et al., 1980; Norman et al., 2013; Norman & Raines, 2002; Todd et al., 1988), the visual system must integrate the motions of identifiable object surface features/points across at least two successive views (e.g., Koenderink & van Doorn, 1991; Norman & Todd, 1993; Todd & Norman, 1991). It is especially important to note, however, that the stimulus images in the current experiment were cast shadows (see Fig. 2). No identifiable points on object surfaces existed within our experimental stimulus images; only the outer shadow contours (or portions of them) were visible. Since no inner detail (i.e., surface features/texture) was present, one cannot obtain or estimate the feature motions needed for a traditional computational analysis of structure-from-motion (see Cortese & Andersen, 1991; Norman & Todd, 1994; Todd, 1985). Computational algorithms (e.g., Cipolla & Giblin, 2000;

Hernández, Schmitt, & Cipolla, 2007; Wong & Cipolla, 2004) do exist that can recover 3-D object structure from deforming occlusion boundary contours (i.e., deforming silhouettes). This suggests that similar algorithms could potentially be developed that could recover 3-D structure from the deforming shadows that occur when illuminated objects rotate in depth.

The good discrimination performance that we obtained in the current experiments during the continuous movement condition (overall d' was 3.68 in Experiment 1 and 4.12 in Experiment 2) is all the more remarkable because our observers (in the experimental trials) were not even allowed to see the entire object cast shadows. They only saw the stimulus shadows through narrow anorthoscopic (4-mm wide) slits. In typical studies of anorthoscopic perception (e.g., Aydin et al., 2009; Parks, 1965; Rieger et al., 2007; Rock, 1981; Shimojo & Richards, 1986; Zöllner, 1862), two-dimensional figures translate behind slits. Our current study greatly extends this previous research and demonstrates that human observers can effectively perceive and identify three-dimensional object shape by integrating fragmentary pieces of cast shadows over time.

Conclusion: Human observers are exceptionally capable of perceiving and recognizing 3-D objects from cast shadows. Our current results demonstrate that this can be done (1) without surface detail/features, (2) with most of an object’s cast shadow (more than 90%) being hidden from view, and can reliably occur despite (3) the continual appearance and disappearance of shadow regions that accompanies anorthoscopic viewing.

Open practices statement The data and materials for this study are available upon request from the first author (farley.norman@wku.edu).

References

- Amedi, A., Jacobson, G., Hendler, T., Malach, R., & Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cerebral Cortex*, 12, 1202–1212. doi:<https://doi.org/10.1093/cercor/12.11.1202>
- Andersen, G. J. (1996). Detection of smooth three-dimensional surfaces from optic flow. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 945–957. doi:<https://doi.org/10.1037/0096-1523.22.4.945>
- Aydin, M., Herzog, M. H., & Ögmen, H. (2009). Shape distortions and Gestalt grouping in anorthoscopic perception. *Journal of Vision*, 9(3):8, 1–8. doi:<https://doi.org/10.1167/9.3.8>
- Bell, H. H., & Lappin, J. S. (1973). Sufficient conditions for the discrimination of motion. *Perception & Psychophysics*, 14, 45–50. doi:<https://doi.org/10.3758/BF03198616>
- Bingham, G. P., & Lind, M. (2008). Large continuous perspective transformations are necessary and sufficient for accurate perception of metric shape. *Perception & Psychophysics*, 70, 524–540. doi:<https://doi.org/10.3758/PP.70.3.524>
- Braunstein, M. L. (1966). Sensitivity of the observer to transformations of the visual field. *Journal of Experimental Psychology*, 72, 683–689. doi:<https://doi.org/10.1037/h0023735>

- Cipolla, R., & Giblin, P. (2000). *Visual motion of curves and surfaces*. Cambridge, England: Cambridge University Press.
- Cortese, J. M., & Andersen, G. J. (1991). Recovery of 3-D shape from deforming contours. *Perception & Psychophysics*, 49, 315–327. doi:<https://doi.org/10.3758/BF03205987>
- Eriksen, C. W., & Collins, J. F. (1968). Sensory traces versus the psychological moment in the temporal organization of form. *Journal of Experimental Psychology*, 77, 376–382. <https://doi.org/10.1037/h0025931>
- Fujita, N. (1990). Three-dimensional anorthoscopic perception. *Perception*, 19, 767–771. doi:<https://doi.org/10.1068/p190767>
- Hernández, C., Schmitt, F., & Cipolla, R. (2007). Silhouette coherence for camera calibration under circular motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29, 343–349. doi:<https://doi.org/10.1109/TPAMI.2007.42>
- Hoffman, D. D., & Bennett, B. M. (1986). The computation of structure from fixed-axis motion: Rigid structures. *Biological Cybernetics*, 54, 71–83. doi:<https://doi.org/10.1007/BF00320477>
- Integration. (2020). In *OED online*. Retrieved from <http://www.oed.com/view/Entry/97356>
- James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*, 40, 1706–1714. doi:[https://doi.org/10.1016/S0028-3932\(02\)00017-9](https://doi.org/10.1016/S0028-3932(02)00017-9)
- Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, 8, 377–385. doi:<https://doi.org/10.1364/JOSAA.8.000377>
- Lappin, J. S., Doner, J. F., & Kottas, B. L. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, 209(4457), 717–719. doi:<https://doi.org/10.1126/science.7394534>
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.
- Norman, J. F., Adkins, O. C., Dowell, C. J., Hoyng, S. C., Shain, L. M., Pedersen, L. E., ... Gilliam, A. N. (2017). Aging and visual 3-D shape recognition from motion. *Attention, Perception, & Psychophysics*, 79, 2467–2477. doi:<https://doi.org/10.3758/s13414-017-1392-8>
- Norman, J. F., Cheeseman, J. R., Pyles, J., Baxter, M. W., Thomason, K. E., & Calloway, A. B. (2013). The effect of age upon the perception of 3-D shape from motion. *Vision Research*, 93, 54–61. doi:<https://doi.org/10.1016/j.visres.2013.10.012>
- Norman, J. F., Dawson, T. E., & Raines, S. R. (2000). The perception and recognition of natural object shape from deforming and static shadows. *Perception*, 29, 135–148. doi:<https://doi.org/10.1068/p2994>
- Norman, J. F., Lee, Y., Phillips, F., Norman, H. F., Jennings, L. R., & McBride, T. R. (2009). The perception of 3-D shape from shadows cast onto curved surfaces. *Acta Psychologica*, 131, 1–11. doi:<https://doi.org/10.1016/j.actpsy.2009.01.007>
- Norman, J. F., Norman, H. F., Clayton, A. M., Lianekhammy, J., & Zielke, G. (2004a). The visual and haptic perception of natural object shape. *Perception & Psychophysics*, 66, 342–351. doi:<https://doi.org/10.3758/BF03194883>
- Norman, J. F., Payton, S. M., Long, J. R., & Hawkes, L. M. (2004b). Aging and the perception of biological motion. *Psychology and Aging*, 19, 219–225. doi:<https://doi.org/10.1037/0882-7974.19.1.219>
- Norman, J. F., Phillips, F., Cheeseman, J. R., Thomason, K. E., Ronning, C., Behari, K., ... Lamirande, D. (2016). Perceiving object shape from specular highlight deformation, boundary contour deformation, and active haptic manipulation. *PLOS ONE*, 11(2), e0149058. doi:<https://doi.org/10.1371/journal.pone.0149058>
- Norman, J. F., & Raines, S. R. (2002). The perception and discrimination of local 3-D surface structure from deforming and disparate boundary contours. *Perception & Psychophysics*, 64, 1145–1159. doi:<https://doi.org/10.3758/BF03194763>
- Norman, J. F., & Todd, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, 53, 279–291. doi:<https://doi.org/10.3758/BF03205183>
- Norman, J. F., & Todd, J. T. (1994). Perception of rigid motion in depth from the optical deformations of shadows and occlusion boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 343–356. doi:<https://doi.org/10.1037/0096-1523.20.2.343>
- Orlov, T., & Zohary, E. (2018). Object representations in human visual cortex formed through temporal integration of dynamic partial shape views. *Journal of Neuroscience*, 38, 659–678. doi:<https://doi.org/10.1523/JNEUROSCI.1318-17.2017>
- Parks, T. E. (1965). Post-retinal visual storage. *American Journal of Psychology*, 78, 145–147.
- Peuskens, H., Claeys, K. G., Todd, J. T., Norman, J. F., Van Hecke, P., & Orban, G. A. (2004). Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *Journal of Cognitive Neuroscience*, 16, 665–682. doi:<https://doi.org/10.1162/0899892904323057371>
- Rieger, J. W., Gröschow, M., Heinze, H.-J., & Fendrich, R. (2007). The appearance of figures seen through a narrow aperture under free viewing conditions: Effects of spontaneous eye motions. *Journal of Vision*, 7(6):10, 1–13. doi:<https://doi.org/10.1167/7.6.10>
- Rock, I. (1981). Anorthoscopic perception. *Scientific American*, 244(3), 145–153.
- Shimojo, S., & Richards, W. (1986). “Seeing” shapes that are almost totally occluded: A new look at Parks's camel. *Perception & Psychophysics*, 39, 418–426. doi:<https://doi.org/10.3758/BF03207070>
- Stewart, C. C. (1900). Zöllner's anorthoscopic illusion. *American Journal of Psychology*, 11, 240–243.
- Todd, J. T. (1985). Perception of structure from motion: Is projective correspondence of moving elements a necessary condition? *Journal of Experimental Psychology: Human Perception and Performance*, 11, 689–710. doi:<https://doi.org/10.1037/0096-1523.11.6.689>
- Todd, J. T., Akerstrom, R. A., Reichel, F. D., & Hayes, W. (1988). Apparent rotation in three-dimensional space: Effects of temporal, spatial, and structural factors. *Perception & Psychophysics*, 43, 179–188. doi:<https://doi.org/10.3758/BF03214196>
- Todd, J. T., & Norman, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, 50, 509–523. doi:<https://doi.org/10.3758/BF03207535>
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, 45, 205–217. doi:<https://doi.org/10.1037/h0056880>
- Wong, K.-Y. K., & Cipolla, R. (2004). Reconstruction of sculpture from its profiles with unknown camera positions. *IEEE Transactions on Image Processing*, 13, 381–389. doi:<https://doi.org/10.1109/TIP.2003.821113>
- Yin, C., Shimojo, S., Moore, C., & Engel, S. A. (2002). Dynamic shape integration in extrastriate cortex. *Current Biology*, 12, 1379–1385. doi:[https://doi.org/10.1016/S0960-9822\(02\)01071-0](https://doi.org/10.1016/S0960-9822(02)01071-0)
- Zöllner, F. (1862). Ueber eine neue Art anorthoskopischer Zerrbilder [About a new kind of anorthoscopic distorted pictures]. *Annalen der Physik*, 193, 477–484. doi:<https://doi.org/10.1002/andp.18621931108>