Acta Psychologica 134 (2010) 40–47

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Does monocular visual space contain planes? ☆

Jan J. Koenderink [a,d,*], Liliana Albertazzi [b], Andrea J. van Doorn [c], Raymond van Ee [d], Wim A. van de Grind [e], Astrid M.L. Kappers [d], Joe S. Lappin [f,1], J. Farley Norman [g], A.H.J. (Stijn) Oomes [a], Susan P. te Pas [k], Flip Phillips [h], Sylvia C. Pont [c], Whitman A. Richards [i], James T. Todd [j], Frans A.J. Verstraten [k], Sjoerd de Vries [l]

[a] *Delft University of Technology, Faculty of EEMCS, Mekelweg 4, 2628 CD Delft, The Netherlands*
[b] *Department of Cognition and Education Sciences, Trento University, Rovereto Branch, Corso Bettini 31, 38068 Rovereto, Italy*
[c] *Delft University of Technology, Faculty of Industrial Design, Landbergstraat 15, 2628 CE Delft, The Netherlands*
[d] *Department Physics of Man, Helmholtz Institute, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*
[e] *Behavioural Biology Group, Utrecht University, Padualaan 8, De Uithof, 3584 CH Utrecht, The Netherlands*
[f] *Department of Psychology and Vanderbilt Vision Research Center, Vanderbilt University, TN, USA*
[g] *Department of Psychology, Western Kentucky University, Bowling Green, KY 42101-1030, USA*
[h] *Vision Laboratories, Skidmore College, 815 N Broadway, MS 2106, Tisch Learning Center 155, Saratoga Springs, NY 12866-1632, USA*
[i] *CSAIL, Massachusetts Institute of Technology, 32-364, Cambridge, MA 02139, USA*
[j] *Psychology Department, The Ohio State University, 216 Lazenby Hall, 1827 Neil Avenue, Columbus, OH 43210, USA*
[k] *Experimental Psychology, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands*
[l] *TNO Defensie en Veiligheid, Kampweg 5, 3769 DE Soesterberg, The Netherlands*

ARTICLE INFO

Article history:
Received 4 September 2009
Received in revised form 28 November 2009
Accepted 1 December 2009
Available online 6 January 2010

PsycINFO classication:
2323

Keywords:
Visual space
Monocular stereopsis

ABSTRACT

The issue of the *existence of planes*—understood as the carriers of a nexus of straight lines—in the *monocular* visual space of a stationary human observer has never been addressed. The most recent empirical data apply to *binocular* visual space and date from the 1960s (Foley, 1964). This appears to be both the first and the last time this basic issue was addressed empirically. Yet the question is of considerable conceptual interest. Here we report on a direct empirical test of the existence of planes in *monocular* visual space for a group of sixteen experienced observers. For the majority of these observers monocular visual space lacks a projective structure, albeit in qualitatively different ways. This greatly reduces the set of viable geometrical models. For example, it rules out all the classical homogeneous spaces (the Cayley–Klein geometries) such as the familiar Luneburg model. The qualitatively different behavior of experienced observers implies that the generic population might well be inhomogeneous with respect to the structure of visual space.

1. Introduction

The geometrical structure of "monocular visual space" has rarely been the subject of formal, geometrical research, whereas the empirical studies are also comparatively scarce. Indeed, "stereopsis" is usually taken to be synonymous with "*binocular* stereopsis" by many dictionaries and informative web-sites ("WIKI's"). "*Monocular* stereopsis" is frequently treated as an oxymoron as titles of scientific papers referring to "paradoxical monocular stereopsis" illustrate (Enright, 1991). Yet experientially "visual space" does in no way collapse into the (flat!) "visual field" if one closes one eye. Moreover, many animals, such as cows and rabbits, have no, or only very minor, binocular overlap but nevertheless show signs of possessing a well-developed visual space. Thus the topic certainly deserves the attention of vision science.

The few current models of monocular visual space are frequently based on the classical homogeneous spaces that are the Cayley–Klein spaces of constant curvature (Yaglom, 1979). Such spaces have been proposed by Luneburg (1947) as describing the structure of binocular visual space, which Luneburg identifies as Lobachevky's hyperbolic space. But even generic Riemann spaces lack a projective structure (Berger, 2007), showing that the Cayley–Klein spaces are quite special. The existence of a projective structure should not be *assumed* as Luneburg (1947) did, but should be treated as an *empirical issue*.

One of the few empirical studies is Foley's (1964), dating from the nineteen sixties. At the time it was assumed that readers would be familiar with the required geometrical background. Since this cannot be assumed for our modern readers we start the paper with a summary review of this material.

☆ This work resulted from an ad lib collaboration at the occasion of the (forced) retirement of the first author.

* Corresponding author. Address: Delft University of Technology, Faculty of EEMCS, Mekelweg 4, 2628 CD Delft, The Netherlands. Tel.: +31 152 84145.
E-mail address: jan.koenderink@telfort.nl (J.J. Koenderink).
[1] Address: 5083 Hanging Moss Ln., Sarasota, FL 34238, USA.

0001-6918/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2009.12.002

The primitive elements of a three-dimensional projective space are "points", "lines" and "planes". Any plane is by itself a two-dimensional projective space. Axiomatically the role of "points" and "lines" in planar or "points" and "planes" in spatial (3D) projective geometry are mutually interchangeable (the so-called property of *duality*), thus these formal entities are quite distinct from the visual qualities usually associated with them (e.g., the "extendedness" of lines as compared to points). Basic axioms of incidence have that—in the projective plane—any two distinct points define a unique line and any two distinct lines define a unique point, whereas—in a three-dimensional projective space—any three distinct points define a unique plane and any three distinct planes define a unique point. Moreover, any two distinct planes define a unique line, a plane and a line define a distinct point, and so forth. In cases of interest to vision lines carry a continuous sequence of points and planes contain a continuous nexus of criss-crossing lines (thus we disregard finite geometries here). In particular, if one considers the nexus of all lines connecting any vertex to any point on its opposite side in a triangle, any two lines from different vertices meet in a point of the interior, thus the nexus is *planar*. This is the property addressed in this study.

In a generic Riemannian space such planar webs of geodesics do not exist (Berger, 2007; Buseman, 1955). The required projective structure occurs only in the classical homogeneous spaces, the Cayley–Klein spaces, in which the curvature is constant. Examples of such spaces are not only the familiar Euclidean space, but also all the spaces considered by Luneburg (elliptic and hyperbolic geometries), including the one he singled out as descriptive of "visual space". Although Luneburg considered mainly *binocular* stereopsis, a close analysis of his arguments reveals that these arguments apply equally to *monocular* visual space. Luneburg considers "free movements of objects" (Helmholtz, 1868) obvious, this indeed limits the possibilities to spaces of constant curvature. It is his key assumption from which most of his formalism follows. But the existence of such a group of congruences is quite independent of the binocularity issue. It is easy enough to come up with very reasonable models of the geometry of visual space that fail to be Cayley–Klein geometries and for which no planar webs of geodesics exist. In order to illustrate this we discuss an instance of such a model in the Appendix. The upshot of this discussion is that *the issue of the existence of a projective structure, as exemplified by the existence of "planes" (as coherent webs of geodesics), is up to empirical verification.*

One particularly convenient set of axioms (at least in the context of this paper) that define a "projective space" as a synthetic geometry is the following (Bennett, 1995):

Definition A *projective space* is an ordered pair $(\mathscr{P}, \mathscr{L})$, where $\mathscr{P}$ is a nonempty set whose elements are called *points* and $\mathscr{L}$ is a nonempty collection of subsets of $\mathscr{P}$ called *lines* subject to the following axioms:

[**PS1**] Given any two distinct points P and Q, there is one and only one line (called $\ell(P,Q)$) containing them.
[**PS2**] (The Pasch Axiom) If A, B, C, and D are distinct points such that there is a point E in $\ell(A,B) \cap \ell(C,D)$, then there is a point F in $\ell(A,C) \cap \ell(B,D)$ (see Fig. 1).
[**PS3**] Each line contains at least three points; not all points are collinear.

Axiom **PS2** (see Fig. 1) is due to Moritz Pasch (1843–1930) (but in a different context) and was first used by Ostwald Veblen (1880–1960) as a clever way to avoid postulating the existence of planes (Bennett, 1995). Instead, planes are defined as *flats*:

Definition Let M be a point of $(\mathscr{P}, \mathscr{L})$ not in line ℓ. The *plane* α determined by M and ℓ is given by

$$\alpha = \bigcup_{A \in \ell} \ell(M,A) = \{P : \ell(P,M) \cap \ell \neq \emptyset\} \cup \{M\}. \quad (1)$$

We will denote this plane by (M,ℓ). Notice that a plane is defined as the closure of a fan of lines that connect the points of a line with a point not on that line, a very intuitive notion (Fig. 2).

One proves that the line defined by two points in a plane lies in that plane and that two lines intersect if and only if they are coplanar. One also shows that the point M that appears in the definition is nothing special: Any point T (say) in the plane α not on the line ℓ may be substituted for M and one still obtains the same plane. One also proves that any three non-collinear points in a projective space lie on a unique plane. Thus the "planes" defined in this manner indeed possess all the familiar properties.

The definition of a plane as a "flat" can be generalized to higher dimensional flats. This allows the definition of *dimension*. For in-

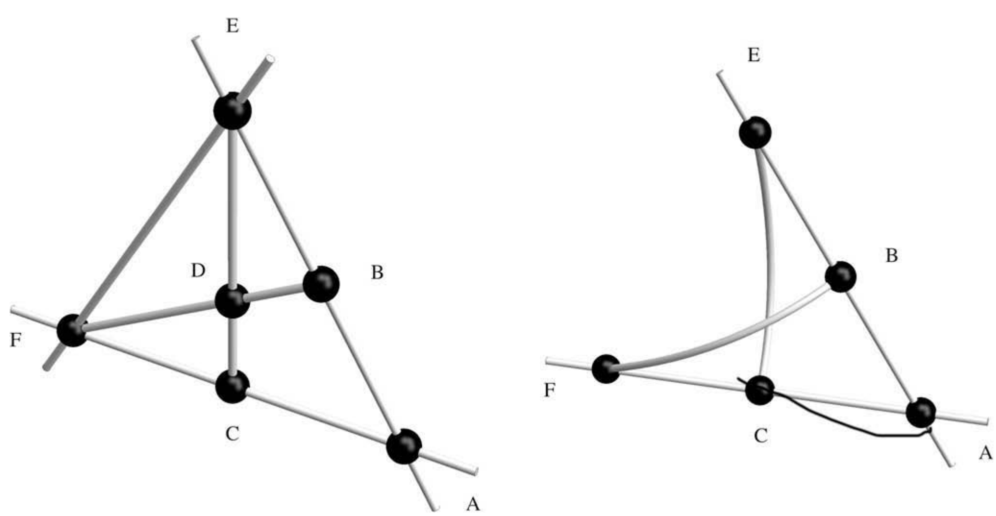

Fig. 1. The "Pasch Axiom": *Given four points* $\{A,B,C,D\}$, *let the lines determined by the points* $\{A,B\}$ *and* $\{C,D\}$ *possess a common point E, then the lines determined by the points* $\{A,C\}$ *and* $\{B,D\}$ *possess a common point F (say)*. In this figure we also indicated the line determined by the points $\{E,F\}$ in order to show that the points $\{A,E,F\}$ are meant to determine a *flat* triangular area in the sense that lines that connect a vertex to the opposite side (like the lines determined by the points $\{E,C\}$ and $\{F,B\}$) meet in a common point "inside the triangle" (here D), in other words, the lines connecting the vertices to the opposite sides "mesh" to form a planar nexus. This is the condition tested in the experiment, shown at left. Intuitively, it is easy to conceive of EC and FB as slightly curved ("out of the plane of the triangle") such as *not* to meet in a common point D (figure at right). Of course this would imply the non-existence of the triangular "plane"! This shows that the Pasch axiom really has non-trivial content (Euclid should have caught this!).

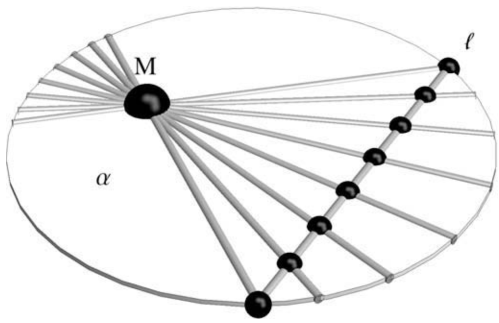

Fig. 2. Given a point M and a line ℓ (M not on ℓ), one constructs the fan of lines that connect M with ℓ. The fan describes a plane α.

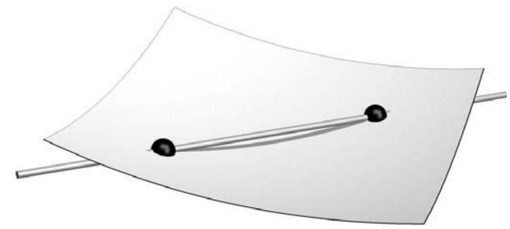

Fig. 4. A really bad case as "planes" go. The line has two distinct points in common with the "plane", yet it fails to lie fully in it. In such a case one concludes that in this space "planes do not exist". In Foly's terms the space "fails the Desarguasian property". The Pasch Axiom (Axiom **PS2**; Fig. 1) is obviously violated in such a geometry.

stance, three-dimensional space is a flat obtained from a plane and a point not on it. One proves the familiar properties, e.g., two distinct planes intersect in a line, and so forth. The crucial fact for our experiment is that Desargues' Theorem for projective space (two triangles that are "point perspective" are also "line perspective" and *vice versa*; see Fig. 3) can be *proven* for dimensions three and higher, but not in planar geometry. For planar geometry it needs to be included in the axioms, albeit perhaps in some roundabout way.

Thus the only non-Desarguesian projective spaces are necessarily two-dimensional, that is to say *planes*, the Moulton plane perhaps being the most familiar example (Moulton, 1902). Any projective space which points are not all coplanar is Desarguesian. Thus any three-dimensional space in which the theorem of Desargues fails (that does not have the "Desarguasian property") cannot be a projective space. Conversely, if all planes are Desarguesian, then they can be embedded into a three-dimensional projective geometry. This illustrates the conceptual load of the "Desarguesian property".

There are various ways to check the existence of planes. In practice one looks for a method that is easily implemented psychophysically. In this paper we check Pasch's Axiom (Axiom **PS2**). This is an apt choice, because in case it fails visual space lacks a projective structure due to the fact that lines fail to fully coincide with planes that have at least two distinct points in common with them. Because of the geometrical facts mentioned above axioms **PS1**...**3** imply Desargues Theorem, this is why the (sparse) literature on the topic uses the "Desarguesian property" (meaning that Desargues' theorem holds) as a key word. It does in no way imply the actual empirical check of the Desargues Theorem though, instead one checks the Pasch Axiom instead, which is a much easier task. This is what Foley (1964) did and it is also our paradigm.

As said above, the only attempt to check for the existence of planes in a visual space known to us is the most remarkable paper by Foley (1964). What Foley means by the "Desarguesian property" is that any line that meets a plane in two distinct points lies fully in that plane (see Fig. 4). This addresses the very existence of planes, in Foley's case in a binocular visual space. Here we study the existence of planes in a monocular visual space.

Foley worked with luminous points in a dark room, at finite distance and observed with both eyes. Foley's aim was to empirically check the axiom of Blank's axiomatization of visual space (Blank, 1953) that boils down to the existence of planes. The outcome of this attempt was unfortunately indecisive. In his abstract Foley concludes:

> "...*that the visual spaces of a significant proportion of the observers [2 out of 6 observers] are Desarguesian; those of others may be non-Desarguesian*".

In this paper we report on an attempt to settle the issue for the case of monocular visual space. It is a priori likely that one might find very significant interindividual differences, as we have encountered them in previous investigations (Norman, Crabtree, Clayton, & Norman, 2005).

2. Essential rationale for the experiment

We implement a purely monocular visual space based upon the single cue of *apparent size*. This is done by considering a space that is empty except for a number of white, Lambertian spheres of

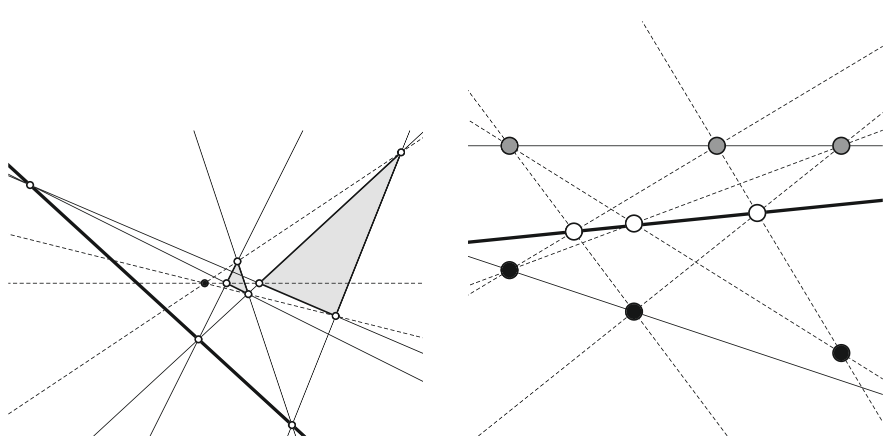

Fig. 3. Two major theorems of "standard" (or perhaps "intended") projective plane geometry: left, an example of the "Desargues configuration";right, an example of the "Pappus configuration". In the Desargues configuration one starts with two triangles that are 'point perspective" (the sides concur in the single black point) and notices that they are also "line perspective" (the sides meet in collinear points on the single black line). In the Pappus configuration one starts with two triples of collinear points (black and gray) and notices that the cross-wise intersections of points taken from each triple happen to be collinear (the white points on the black line). In 1905 Hessenberg proved that in projective plane geometry the Pappus property implies the Desarguesian property, thus the Pappus is the stronger condition. Notice that these configurations can be drawn in many ways that would at first sight appear almost unrecognizable.

identical diameter. The size cue is enabled by instructing the observer that the spheres are indeed of equal sizes. This appears to be natural enough to our observers (equal size appears to be the default assumption of the visual system in cases of the complete absence of prior information), though we evidently have to rely on their applying this prior knowledge to the stimulus configurations used in the experiment.

In this setting the size cue tends to function immediately and transparently, that is to say, if one varies the relative size of objects many observers experience a relative *distance variation*, rather than *size changes*. There is a definite spectrum here though, several observers noticed combined distance and size variations, whereas one observer experienced all objects close to the frontoparallel plane with mostly size variations. The observers thus varied between the extremes of (intuitively and automatically) ascribing most of the variation to depth and to ascribing most of the variation to size. This is no doubt an important factor in the interpretation of the results.

We assume that the visual field as well as visual space is invariant with respect to rotations about the vantage point. This should apply at least approximately, although a variety of objections could easily be raised. We use this assumption so as to be able to combine observations obtained in a canonical configuration. In the experiment we consider various configurations of three spheres that are coplanar with the vantage point. In the canonical configuration we use the horizontal plane, with the center sphere in the primary (frontal) direction. Moreover, we repeat any configuration in the left–right reversed configuration and average over the results. These conventions serve to *force isotropy* of the results. It thus increases the confidence one might have in the results by factoring out possible complications due to horizontal–vertical–oblique differences and so forth. Although all these restrictions could be removed, this would involve a major effort and would render the results less decisive.

Theoretically, any violations of Pasch's Axiom (Axiom **PS2**) are expected to vanish for small configurations and monotonically grow with their size. Thus we design the experiment for a triangular configuration in the visual field of the maximum feasible size. The limit is set by the extent of the visual field and the difficulties many observers experience with extremely eccentric targets. We decided on a spherical triangle in the visual field (see Fig. 5 left) *ABC* with three equal sides of an arclength equal to a quarter great circle (a right angle, that is $\frac{\pi}{2}$ radians or 90°). As a consequence the interior angles of the triangle are also three right angles. We place the spheres in the directions of the vertices of this spherical triangle, but at a (very) different distance from the origin. Of the resulting triangle ABC in space (see Fig. 5 right) the distance from the origin (that is the vantage point) to *B* is twice the distance to *A* and the distance to *C* is even four times the distance to *A*.

At this point a short digression is in order because of a likely confusion: Notice that we aim to address the projective structure (if it exists) in a visual space, yet in the above we use metrical descriptions in terms of Euclidean distances and angles. It might be thought that this introduces methodological inconsistencies. However, one has to distinguish sharply between *physical space* in which we describe the structure of the stimulus and *visual space* which is the arena of descriptions of the responses. These spaces have different ontologies, physical space may simply be taken to be "the space we move in", whereas visual space is a mental entity. Keeping these distinctions in mind at all times avoids possible confusion considering consistency. In some discussions we use a third type of space, namely a formal, hypothetical model of visual space. This model is not Euclidean, but a space with a simple Riemannian metric. In this model we have a nexus of well-defined geodesics which allows us to predict entities in visual space. Of course such predictions have only hypothetical value, they are up for empirical

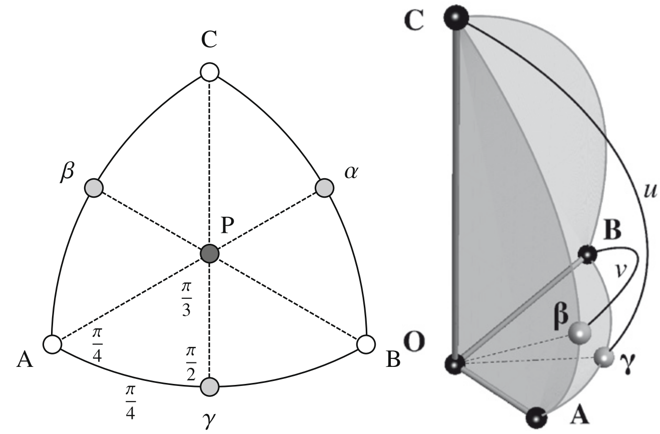

Fig. 5. left: A spherical triangle *ABC* with lines connecting the vertices to the midpoints of the opposite sides. The common intersection at *P* is to be thought of as potentially *three different points*, one on αA, one on βB and one on γC. The triangle is in the visual field, thus the three points might be at different depths in visual space. In the configuration used in the experiment *ABC* is an equilateral spherical triangle with sides $\frac{\pi}{2}$ (such a triangle is also equiangular with interior angles $\frac{\pi}{2}$). In the triangle $AP\gamma$ the side $A\gamma$ is $\pi/4$ (because half the side *AB*), the angle $A\gamma P = \pi/2$ (for half of π, the straight line $A\gamma B$), the angle $\gamma AP = \pi/4$ (for half of $CAB = \pi/2$) and the angle $AP\gamma = \pi/3$ (because the six triangles meeting at *P* are congruent and the interior angles at *P* add up to 2π). This suffices to solve for the remaining sides using the conventional rules for the right angled spherical triangle. right: The configuration in space. Notice that OB is twice OA and OC is four times OA. All the curved lines are geodesics (shortest connecting "lines") according to the model discussed in the appendix. The direction Oβ bisects the angle AOC and the direction Oγ bisects the angle AOB. The geodesics *u* and *v* evidently lack a common point, thus (visually!) violating the Pasch axiom.

check. The key conceptual fact is that *the psychophysical task is of a purely projective nature*. The observer is presented with two lines, one defined by a pair of points that are separated in the visual field and the other a visual ray, specified by a third, single point. The three points are collinear in the visual field. The "lines" in a visual space are mental entities produced by the observer. The task is to indicate the intersection of these lines by moving the third point along the visual ray (purely in depth) so as to make it appear collinear with the point pair in visual space.

To continue the construction: let the midpoints of the arcs *AB*, *BC* and *CA* (in the visual field) be, respectively, γ, α and β (see Fig. 5). The arcs γC, αA and βB intersect *in the visual field* in the common point *P*. In *visual space* the point *P* is actually represented by three *different* ("parallel") points at distinct distances from the eye, one on the arc γC, one on the arc αA and one on the arc βB. In the experiment a single session is composed of the following subtasks (notice that arcs are *oriented*, for instance, the arc *AC* is the same geometrical locus as the arc *CA*, but in the opposite orientation):

1. find the point β on the arc *CA*;
2. find the point β on the arc *AC*;
3. find the point γ on the arc *AB*;
4. find the point γ on the arc *BA*;
5. find the point *P* on the arc βB;
6. find the point *P* on the arc $B\beta$;
7. find the point P' on the arc γC;
8. find the point P' on the arc $C\gamma$.

(See Fig. 6.) Notice that subtasks #1 and #2 are identical except for the orientation. In the experimental paradigm the points *A*, *C* and β are presented on a horizontal line, the arcs $A\beta$ and βC at the correct angular sizes (see Fig. 7). The orientation difference then implies a left–right mirror image. This also applies to subtasks #3, #4, subtasks #5, #6 and subtasks #7, #8.

Notice that subtasks #1 and #2 have to be completed before subtasks #5 and #6, and that subtasks #3 and #4 have to be completed before subtasks #7 and #8. The depth of β is found as the mean of the depths found in subtasks #1 and #2 and the depth of γ as the mean of the depths found in subtasks #3 and #4. Thus there exist constraints on the order in which the subtasks can be performed. Subject to these constraints the order was randomized for each full task.

The result of the experiment is the ratio of the mean of the depths of P (obtained from subtasks #5 and #6) to that of the depth of P' (obtained from subtasks #7 and #8).

As mentioned above, all subtasks are presented in the canonical configuration, a horizontal linear presentation. Fig. 7 gives an impression of a typical situation. The observer is permitted to shift the center sphere in depth (by changing its size) and is required to make it collinear (in visual space!) with the two outer spheres. Since this task has to be performed *in visual space*, it is far from being trivial.

A full session yields a direct test of the validity of Pasch's Axiom (Axiom **PS2**) in a monocular visual space. Of course the structure of the session implies that various settings are dependent upon each other (for instance the result of subtask #5 evidently depends upon the result of subtask #1). However, *repeated sessions are independent of each other*. Therefore we use the mean of a number of sessions in order to increase the accuracy of the estimate of any possible violation. Such a repetition also allows us to find a measure of the significance of the result.

Our prior experience with the structure of monocular visual space for both restricted (Koenderink & van Doorn, 2008) and very wide (Koenderink, van Doorn, & Todd, 2009) visual fields suggests that large interindividual differences are likely to occur. This expectation, coupled with the fact that observers find spatial tasks involving very wide visual fields difficult, almost forced us to find an occasion where a relatively large number (16) of very experienced visual observers were available. All except three of the observers were initially kept unaware of the structure and aim of the experiment. This seemed at least *prudent*, although we do not at all believe this to be *relevant*. After the conclusion of the experiment the observers (all well-known visual scientists) switched their role and assumed their responsibilities as that of a coauthor.

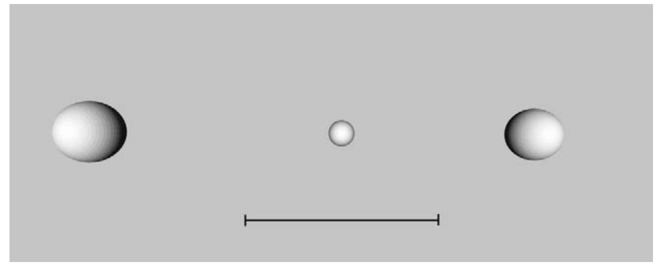

Fig. 7. A typical stimulus configuration. This is a view of three equal sized spheres in the horizontal plane through the eye (the "horizon") at different distances from the observer. The spheres have white Lambertian surfaces and are illuminated by a uniform, parallel beam from behind the observer. As is evident from the shading, the center sphere is illuminated frontally. It is perhaps less obvious that the outer spheres are illuminated in exactly the same way, to many observers it appears like they were illuminated from different directions. The reason is that the field of view is very large as can be judged from the black line segment which indicates the correct viewing distance. From the correct vantage point the perspective deformations vanish, but (to most observers) the shading patterns still look somewhat unexpected. The observer has control over the distance of the center sphere and has the task to place it collinear with the outer spheres *in visual space*.

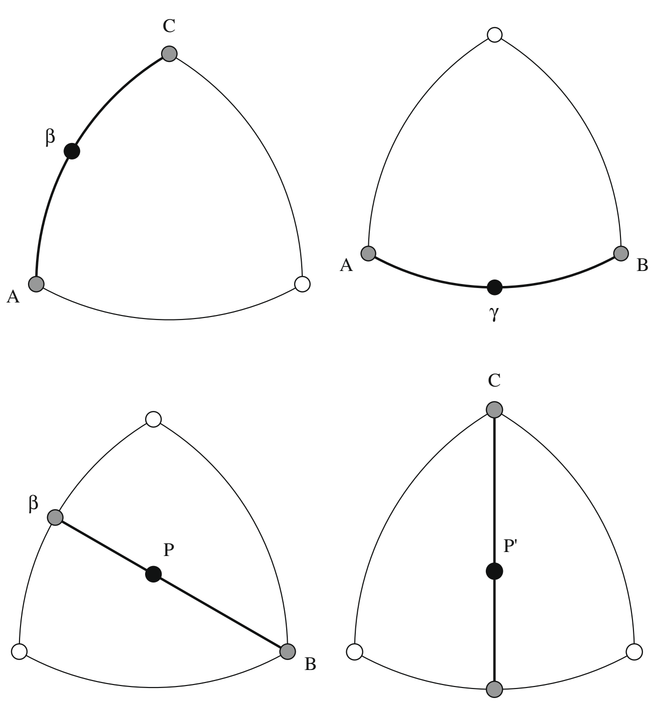

Fig. 6. From left to right, top to bottom, these are subtasks #1, #2 (symmetrical pair), subtasks #3, #4 (symmetrical pair), subtasks #5, #6 (symmetrical pair), subtasks #7, #8 (symmetrical pair). The actual appearance as a stimulus is shown in Fig. 7.

3. Methods

Stimuli were presented on a large plasma display (conventional 96 cm diagonal TV monitor) driven by a macintosh powerbook computer. The display was fitted with a frame designed to aid the observer in using the preferred eye and keeping it at approximately the same location. In order to be able to do the task at all the observer frequently had to perform head movements. The eye position was kept within a 2-cm diameter area by way of a "key hole" aperture. The frame provided an additional aperture so as to be able to leave room for the nose at the screen-side of the key hole. The available field of view was 90° in the vertical and appreciably larger (about 120°) in the horizontal direction.

The stimuli were programmed in OpenGL via the Cocoa object-C environment under the macintosh OS X Leopard operating

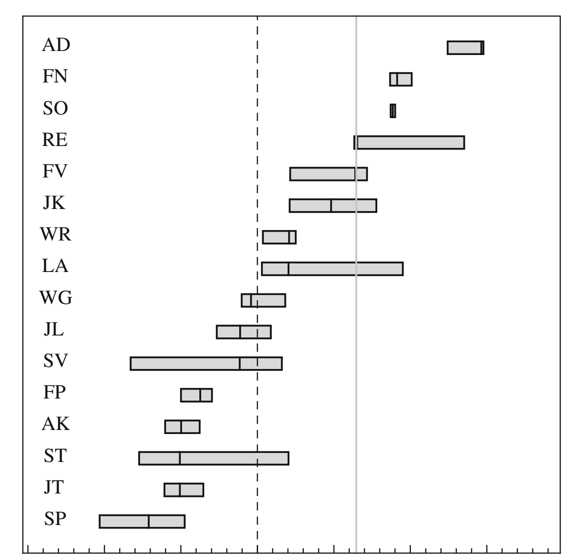

Fig. 8. Results for all observers. The boxes show the 25% and 75% quartiles as wel as the medians for the ratio of depths. In the planar case the value would be unity, as indicated by the black vertical line. The model prediction is indicated by the gray vertical line at right.

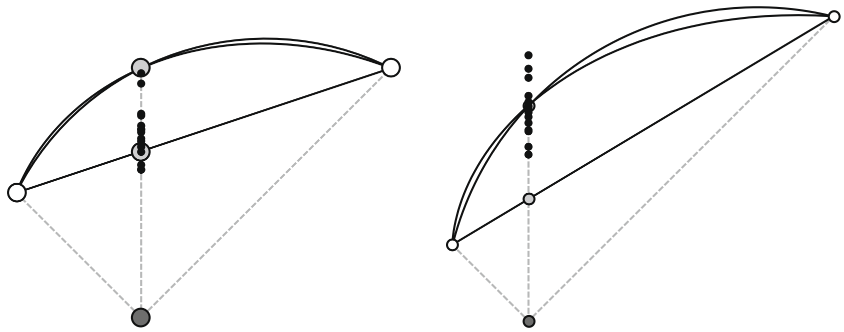

Fig. 9. The geometry for the bisection of arcs *AB* (left, distance ratio 2) and *CA* (right, distance ratio 4). The dark dot is the position of the eye, the white dots are the vertices. The median settings of all observers have been indicated by the black points. They may be compared to the theoretical predictions for a veridical (Euclidian, the straight line), the Luneburg hyperbolic and the dilation–rotation invariant models. For the bisection task the two latter models yield identical predictions. The results are rather different for the distance ratios two and four: In the former case the data cluster on the Euclidian, in the latter case on the non-Euclidian predictions. In the large majority of cases the "geodesics" used by the observers turn their concave side towards the vantage point.

system. The spheres were illuminated with a parallel, uniform beam coming from behind the observer and (visibly) hitting the center-sphere head on. This sphere thus appears as very bright with a darkish edge due to Lambert's cosine law. The other spheres were illuminated in exactly the same way, but appeared rather different, because seen almost in profile due to their appreciable eccentricity. Several observers spontaneously remarked on this. The shading looks somewhat unfamiliar because human observers tend to refer the shading to the local "visual ray" instead of the global direction of illumination (a very striking effect that nevertheless appears to be undocumented in the literature).

All observers performed one "dry run" session in order to familiarize themselves with the task and the controls. They then completed five formal sessions. All observers remarked on the fact that it might be possible to perform the task purely in the visual field by merely interpolating angular size, but added that they did not do so, but did their settings on the basis of a "gut feeling" of depth in three-dimensional visual space. This was indeed the aim (whether we were indeed successful in this is another matter, see below) and one important reason for the choice of a group of experienced observers.

Most observers managed to finish a session in 5–15 min, but one observer took 30 min, another an hour and a half. There was little reason for surprise here since these individual traits, which are perhaps remarkably constant over time spans of an academic career, were well-known from previous collaborations.

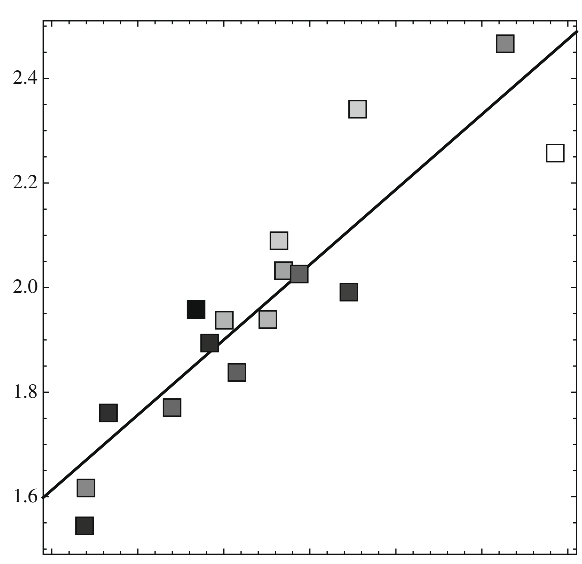

Fig. 10. Linear regression (with constant term) for the distances of the two bisection tasks. The graylevel indicates the deviation from the Pasch Axiom (Axiom **PS2**). The correlation between the depths is rather strong, whereas the deviation from the Pasch Axiom is roughly independent of the depth settings.

4. Results

The main result is shown in Fig. 8. The boxes show the 25% and 75% quantiles with the centerline indicating the median. The results have been sorted (from bottom to top) by increasing value of the median. The dashed black line indicates the unit ratio, that is characteristic for the validity of Pasch's Axiom (Axiom **PS2**).[2] The gray line indicates the prediction from the dilation–rotation invariant Riemannian metric discussed in the Appendix A.

A two-sided test of the mean indicates that eight of the sixteen observers violate Pasch's Axiom (Axiom **PS2**) at the 5% confidence level. Of these violations four involve values larger than one, whereas four involve values less than one.

We also collected data on the depths that were set for the bisection of the arcs *AB* and *CA* (see Fig. 9). For the distance ratio of two the settings cluster on the Euclidian prediction, perhaps with a trend to interpolate between the Euclidian and non-Euclidian predictions. For the distance ratio of four the data clearly cluster about the non-Euclidian prediction. In almost all cases the implied "geodesic" is curved with its concave side turned towards the observer. Notice that it is not possible to distinguish between the Luneburg prediction and that for the dilation–rotation invariant metric since the predictions are identical.

The settings for the two different depth ratios are highly correlated (see Fig. 10), the R^2 value being 0.81. Thus observers appear to be different, but quite constant in their behavior.

The depth settings for the two bisection tasks hardly correlate with the magnitude of the violation of Pasch's Axiom (Axiom **PS2**; R^2 values 0.30 and 0.29). Thus we obtain no handle on the resolution of the diverse results.

5. Conclusions

Our conclusion is in some respects similar to that of Foley (1964) who concludes that the visual spaces of a significant

[2] The geodesics in the Luneburg model are planar curves whose osculating planes generically fail to pass through the vantage point. This (depending on one's mind set) appears to rule out this metric as a viable model. However, if one accepts this, one should use the *induced metric* in planes that contain the vantage point. This applies to cases like the setting of the mid points of the sides, where we force the observer to attend to the plane containing the vertices and the vantage point. This only slightly complicates the analysis, it leads to a predicted ratio of 1.09... rather than 1, which does not affect our conclusions.

proportion of the observers are Desarguesian; whereas those of others may be non-Desarguesian. Of course Foley's conclusion applies to *binocular* visual space, whereas ours applies to *monocular* visual space.

There can be no doubt that the monocular visual space of many observers does not allow the definition of "visual planes", since Pasch's Axiom (Axiom **PS2**) is (frequently *very*) significantly violated. This suggests that it might be prudent to refrain from reference to "visual planes" in any case. This is not necessarily problematic since infinitesimal planar elements can always be defined locally which is enough to be able to make sense of the local curvatures of smooth surfaces. It is typical for general Riemannian spaces to have violations of Pasch's Axiom (Axiom **PS2**), the classical homogeneous Cayley–Klein spaces being very special, or *singular*, in this respect. The lack of *planes* by no means excludes the existence of *surfaces*, which is what really matters in vision. The main impact of our finding is as a constraint on conceptual geometrical models of the structure of visual space.

The violations of the Desarguesian property occur in two qualitatively different ways in that the ratios either exceed or fall short of unity. This might well indicate true qualitative differences within the normal population. In earlier experiments (Koenderink et al., 2009) we have found that such variations exist and are surprisingly large. It will no doubt be rewarding, although very cumbersome, to study this in greater detail, for many more observers and with a considerable battery of different tests.

There can be no doubt that all observers—quite independent of their psychophysical results—have no problems with their optically guided behavior in the physical world. This might seem surprising given the fact that their visual experiences may differ appreciably. Here one has to remember that even very different "user interfaces" (e.g., a UNIX command line interface and the Windows GUI on PC's) might very well prove to serve equally well in various tasks (e.g., deleting or duplicating files in the example) as also argued by Hofmann (in press).

Is it possible that the observers used a purely visual field (two-dimensional) based strategy? Suppose the observer would perform a linear interpolation of angular sizes, using angular distance in the visual field to do the interpolation. For instance the objects at distances one and two would yield angular sizes proportional to one and one-half. At the midpoint the interpolated size would be three-quarters, yielding (using the size cue) a predicted distance of four-thirds (that is 1.33 ...). Continuing this type of calculation one arrives at a ratio of 1.119 ..., very close to the prediction (1.12934 ...) of the model discussed in the Appendix (the gray line shown in Fig. 8). We cannot exclude that some (at least four) of the observers used this strategy. This is in accord with the fact that variations were sometimes experienced as depth sometimes as size variations. However, one should remember that in the final analysis *any* strategy is based upon (two-dimensional) visual field structure. Whether an observer "performs the task in visual space" instead of "performing it in the visual field" is a fundamentally undecidable issue. Pictorial "depth" is a *quale* that has no psychophysical equivalent.

Thus far we have found no likely model that would predict violations *less* than unity. This means little though, since it is very difficult to come up with such models in a principled manner. Given any Riemann metric it is a mere matter of calculation to find the predicted violation, but given a violation there exists no principled method to construct a Riemann metric that would predict it. Given the data such violations no doubt are a *bona fide* category.

Acknowledgement

Hans Kolijn helped us to adapt a wide screen plasma display to requirements of the experiment.

Appendix A. Prediction based on a simple dilation–rotation invariant metric

The predictions of any model of monocular visual space should be invariant with respect to arbitrary rotations–dilations about the ego center. The reason is that neither absolute distance nor absolute direction is optically specified. In order to be able to arrive at a prediction one needs a metric. An obvious, simple model is thus a Riemannian space with the required symmetry. Let the ego-center be taken at the origin of the conventional Cartesian space $\mathbb{R}^3$ with coordinates $\{x, y, z\}$. The required "line element" (or metric) is

$$ds^2 = \frac{2}{C}\left(\frac{dx^2 + dy^2 + dz^2}{x^2 + y^2 + z^2}\right), \qquad (2)$$

where C denotes a constant. The metric is (by construction) invariant against arbitrary rotation–dilations. One easily checks that it defines a curved space with constant positive scalar curvature equal to C.

A transformation into the conventional polar coordinates $\{\varrho, \vartheta, \varphi\}$ (ϱ the distance from the egocenter, ϑ the elevation with respect to the horizon and φ the azimuth) yields

$$ds^2 = \frac{2}{C}\left(d\xi^2 + d\vartheta^2 + \cos^2\vartheta\, d\varphi^2\right), \qquad (3)$$

where $\xi = \log(\varrho/\overline{\varrho})$ (with $\overline{\varrho}$ an arbitrary unit length). Apparently the space is the product of the 2-sphere and a line. In the polar coordinate system the Riemann curvature tensor has two distinct, independent components (others follow from the antisymmetry under exchange of the last two indices), namely $R^2_{332} = -\cos^2\vartheta$ and $R^3_{232} = 1$.

Since the metric is centrally symmetric the geodesics have to be planar curves, confined to planes through the egocenter. Taking a plane $\varphi =$ constant one has

$$ds^2 = \frac{2}{C}(d\xi^2 + d\vartheta^2), \qquad (4)$$

which has the structure of the Euclidean metric. Thus one sees immediately that the geodesics are logarithmic spirals of the form

$$\varrho = \varrho_0\, e^{\lambda(\vartheta - \vartheta_0)}, \qquad (5)$$

thus there is no need to integrate the geodesic equations explicitly[3].

This geometry is sufficient to find a prediction for the experimental results. We set $\overline{\varrho} = 1$ for convenience, the actual value being irrelevant to the problem. Consider a spherical triangle with

[3] Although the symmetry argument appears compelling to us, a referee of the manuscript expressed doubts. In case the reader has similar doubts, here is a formal derivation: The geodesic equations (using the summation convention)

$$\frac{d^2x^\lambda}{dt^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt} = 0, \qquad (6)$$

in polar coordinates are (the mutually independent, non-zero Christoffel symbols (others follow from the symmetry under exchange of the last two indices) are $\Gamma^1_{11} = -1/\varrho$, $\Gamma^2_{33} = \sin\vartheta\cos\vartheta$ and $\Gamma^3_{32} = -\tan\vartheta$):

$$\ddot{\varrho} = \frac{\dot{\varrho}^2}{\varrho}, \qquad (7)$$

$$\ddot{\vartheta} = -\sin\vartheta\cos\vartheta\ \dot{\varphi}^2, \qquad (8)$$

$$\ddot{\varphi} = 2\tan\vartheta\ \dot{\vartheta}\,\dot{\varphi}, \qquad (9)$$

where the dots indicate derivatives with respect to arc length. Given an initial point, and an initial direction, select the coordinates system such that the coordinates of the point are $\{\varrho_0, 0, 0\}$ and the direction $\{\sin\mu, \cos\mu, 0\}$(μ slope of the geodesic). This is always possible because of the spherical symmetry. Then Eq. (9) becomes simply $\ddot{\varphi} = 0$, thus one has $\varphi = 0$ throughout, that is to say, the geodesic is a planar curve in the plane $\varphi = 0$. Eq. (8)becomes $\ddot{\vartheta} = 0$, with the immediate solution $\vartheta = \cos\mu\sigma$, where σ denotes the geodesic distance from the initial point. The remaining Eq. (7), that is $\varrho\ddot{\varrho} - \dot{\varrho}^2 = 0$ is solved by $\varrho(\sigma) = C_1\exp C_2\sigma$. Solving for the constants of integration $C_{1,2}$ one obtains $\varrho(\sigma) = \varrho_0\exp\sin\mu\sigma$. Elimation of the geodesic distance σ from this parametric representation of the geodesic one finds $\varrho(\vartheta) = \varrho_0\exp\tan\mu\vartheta$, which is the form (5) again. Thus the geodesics are indeed planar logarithmic spirals. Notice that selecting an appropriate orientation of the polar coordinate frame effectively formalizes the symmetry argument used above.

vertices in the directions $\mathbf{e_1} = \{1, 0, 0\}$, $\mathbf{e_2} = \{0, 1, 0\}$ and $\mathbf{e_3} = \{0, 0, 1\}$. This triangle is equilateral with sides $\pi/2$ and also equiangular with interior angles $\pi/2$. The three bisectors of the angles (and thus also of the opposite sides) meet in a single point and divide each other in arcs with sines in the ratio $1:\sqrt{2}$, as application of the sine rule in one of the six interior triangles shows. Next we introduce unequal depths at the vertices of a spatial triangle *ABC* with $A = \mathbf{e_1}$, $B = 2\mathbf{e_2}$ and $C = 4\mathbf{e_3}$. The depths at the midpoints of the sides are simply the geometrical means of the depths at the vertices, thus the depth at the midpoint of *CA* is 2 and that at the midpoint of *AB* is $\sqrt{2}$. The depth in the direction of the center of the triangle on the geodesic that connects the vertex *C* to the midpoint of the side *AB* is (by a linear interpolation in the $\{\xi, \vartheta\}$-plane)

$$2^{2-\frac{3\arccos\frac{1}{\sqrt{3}}}{\pi}} = 2.12541\ldots, \tag{10}$$

whereas the depth in the direction of the center of the triangle on the geodesic that connects the vertex *A* to the midpoint of the side *BC* is (again by linear interpolation in the $\{\xi, \vartheta\}$-plane)

$$8^{\frac{\arccos\frac{1}{\sqrt{3}}}{\pi}} = 1.88199\ldots. \tag{11}$$

Thus the two points are seen in the same direction but have different depths, the depths being in the ratio (this is where the arbitrary unit ϱ_0 cancels out)

$$2^{2-\frac{6\arccos\frac{1}{\sqrt{3}}}{\pi}} = 1.12934\ldots. \tag{12}$$

This is the prediction of the model.

Notice that even such a very simple conceptual model, based on nothing more than invariance against arbitrary rotations (eye movements) and dilations (mental scale adjustments, since no scale is optically specified) yields a metric that violates Pasch's Axiom (Axiom **PS2**).

References

Bennett, M. K. (1995). *Affine and projective geometry*. New York: Wiley.
Berger, M. (2007). *A panoramic view of Riemannian geometry*. New York: Springer (Note 1.1.3.1. and Section 6.1.4, pp. 231–233).
Blank, A. A. (1953). The Luneburg theory of binocular visual space. *Journal of the Optical Society of America, 43*, 717–721.
Buseman, H. (1955). *Geometry of geodesics*. New York: Academic Press.
Enright, J. T. (1991). Paradoxical monocular stereopsis and perspective vergence. In S. R. Ellis (Ed.), *Pictorial communication in virtual and real environments* (pp. 567–576). Bristol, PA, USA: Taylor & Francis, Inc.
Foley, J. (1964). Desarguesian property in visual space. *Journal of the Optical Society of America, 54*, 684–692.
Helmholtz, H. von (1868). Über die Thatsachen, die der Geometrie zu Grunde liegen. *Göttinger Nachrichten, 15*, 193–221.
Hoffman, D. (in press). The interface theory of perception: Natural selection drives true perception to swift extinction. In S. Dickinson, M. Tarr, A. Leonardis, B. Schiele (Eds.), *Object categorization: Computer and human vision perspectives*. Cambridge, UK: Cambridge University Press.
Koenderink, J. J., & van Doorn, A. J. (2008). The structure of visual spaces. *Journal of Mathematical Imaging and Vision, 31*, 171–187.
Koenderink, J. J., van Doorn, A. J., & Todd, J. T. (2009). Wide distribution of local sign in the normal population. *Psychological Research, 73*, 14–22.
Luneburg, R. K. (1947). *Mathematical analysis of binocular vision*. Princeton, NJ: Princeton University Press.
Moulton, F. R. (1902). A simple non-desarguesian plane geometry. *Transactions of the American Mathematical Society, 3*, 192–195.
Norman, J. F., Crabtree, C. E., Clayton, A. M., & Norman, H. F. (2005). The perception of distances and spatial relationships in natural outdoor environments. *Perception, 34*, 1315–1324.
Yaglom, I. M. (1979). *A simple non-Euclidean geometry and its physical basis: An elementary account of Galilean geometry and the Galilean principle of relativity*. New York: Springer (translated from the Russian by A. Shenitzer).