tvst

DOI: 10.1167/tvst.4.3.9

Article

# Aniseikonia Tests: The Role of Viewing Mode, Response Bias, and Size–Color Illusions

Miguel A. García-Pérez[1], Eli Peli[2]

[1] Departamento de Metodología, Facultad de Psicología, Universidad Complutense, Campus de Somosaguas, Madrid, Spain
[2] The Schepens Eye Research Institute, Massachusetts Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, MA, USA

**Correspondence:** Miguel A García-Pérez, Departamento de Metodología, Facultad de Psicología, Universidad Complutense, Campus de Somosaguas, 28223 Madrid, Spain; e-mail: miguel@psi.ucm.es

**Received:** 23 January 2015
**Accepted:** 23 April 2015
**Published:** 12 June 2015

**Keywords:** aniseikonia; eye movements; size perception; vernier acuity; size–color illusion

**Citation:** García-Pérez MA, Peli E. Aniseikonia tests: the role of viewing mode, response bias, and size–color illusions. Tran Vis Sci Tech. 2015;4(3):9, doi:10.1167/tvst.4.3.9

**Purpose:** To identify the factors responsible for the poor validity of the most common aniseikonia tests, which involve size comparisons of red–green stimuli presented haploscopically.

**Methods:** Aniseikonia was induced by afocal size lenses placed before one eye. Observers compared the sizes of semicircles presented haploscopically via color filters. The main factor under study was viewing mode (free viewing versus short presentations under central fixation). To eliminate response bias, a three-response format allowed observers to respond if the left, the right, or neither semicircle appeared larger than the other. To control decisional (criterion) bias, measurements were taken with the lens-magnified stimulus placed on the left and on the right. To control for size–color illusions, measurements were made with color filters in both arrangements before the eyes and under binocular vision (without color filters).

**Results:** Free viewing resulted in a systematic underestimation of lens-induced aniseikonia that was absent with short presentations. Significant size–color illusions and decisional biases were found that would be mistaken for aniseikonia unless appropriate action is taken.

**Conclusions:** To improve their validity, aniseikonia tests should use short presentations and include control conditions to prevent contamination from decisional/response biases. If anaglyphs are used, presence of size–color illusions must be checked for.

**Translational relevance:** We identified optimal conditions for administration of aniseikonia tests and appropriate action for differential diagnosis of aniseikonia in the presence of response biases or size–color illusions. Our study has clinical implications for aniseikonia management.

## Introduction

Aniseikonia is a binocular condition by which the perceived size of an object differs between the eyes, often caused by the retinal images of the object differing in size between the eyes. Different retinal image sizes occur naturally in near vision for objects located at meaningfully different distances from each eye[1,2] but differences in perceived size can also be due to anatomical, optical, or neural differences between the eyes.[3] Within a range, optical correction of measured aniseikonia is feasible. Early instruments designed to measure aniseikonia were the Ophthalmo-Eikonometer (OE)[4,5] and the Space Eikonometer (SE),[6] which are complex instruments including a head rest, rigidly fixed cells for trial lenses, and a display device. The OE does not seem to have ever entered production, whereas production of the SE was discontinued decades ago. These instruments are nowadays rarely available in clinical practice.

Simpler aniseikonia tests were proposed soon afterward[7,8] whose use is rarely reported and whose validity was, to our knowledge, never established. An alternative whose use is often reported is the New Aniseikonia Test (NAT).[9,10] The NAT uses a printed set of red and green semicircles (anaglyphs) of different relative sizes that patients view haploscopically with red–green glasses to select the pair in which both semicircles are perceptually equal in size. The

booklet where anaglyphs are printed (at six per page) is held at reading distance and no complex instrument is required to deliver the stimuli. A validity study showed that the NAT underestimates natural aniseikonia as well as aniseikonia induced with size lenses (henceforth, lens-induced aniseikonia) by a factor of approximately 3.[11] The underestimation was confirmed later[12,13] but its cause remained unknown. A tiny part of it can be attributed to an unnoticed design error in the NAT: the Plus Aniseikonia Series was created by reducing the size of the test (green) semicircle by some specified percentages, and these reductions are incorrectly taken directly as the amount of aniseikonia. If the pair is selected in which the test semicircle is $x\%$ smaller, the amount of aniseikonia is not $x$ because an image reduced by $x\%$ must be enlarged instead by $y\%$ to regain its original size, with $y = 100x/(100 - x)$. The resultant nonlinear underestimation, $y - x$, varies with $x$ as $x^2/(100 - x)$, which is less than the reported amount (e.g., aniseikonia measured as 10% with the NAT would actually be 11.11%, which represents underestimation by a factor of only 1.1).

A more recent alternative is the Aniseikonia Inspector (AI),[14] a test based on anaglyphs similar to those making up the NAT but administered via computer. Three versions of the AI were developed over the years. Version 1 reproduced the video test of McCormack et al.,[11] where unlimited time is given to manipulate the size of one of the semicircles until it matches subjectively the size of the other. In Version 2, semicircles were replaced with rectangles, unlimited inspection time was replaced with short presentations, and the method of adjustment was replaced with a task in which observers indicate whether the left, the right, or neither rectangle is subjectively larger than the other. The measure of aniseikonia then arises from fitting a two-parameter psychometric function to the data via treatment of "equal" responses as half "left" and half "right." Finally, Version 3 defaults back to unlimited presentation time although the option for short presentations is still available, with no other major changes.

Studies on the validity of the AI offer conflicting results. In principle, the video test[11] motivating the AI Version 1 proved valid. In the same line, de Wit[14] reported accurate measurements of lens-induced aniseikonia in the horizontal, vertical, and oblique meridia with the AI Version 1. Analogous results were reported for a custom-made computerized test[15] that resembles the video test[11] and the AI Version 1. In contrast, Rutstein et al.[16] reported for the AI Version 1 a meaningful underestimation with respect to the expected amount of lens-induced aniseikonia and also with respect to measurements obtained with the SE in the horizontal and vertical meridia in the light and in the dark. This underestimation was lower in magnitude than that reported by McCormack et al.[11] for the NAT: the slope of the relation of measured to lens-induced aniseikonia with the AI Version 1 was 0.67 in the vertical meridian and 0.58 in the horizontal meridian. Underestimation of lens-induced aniseikonia with the AI Version 1 has been reported in other studies.[17–19] Version 2, on the other hand, seemed to overestimate aniseikonia instead.[18] Finally, the only published study employing Version 3 has reported acceptably accurate measurements of lens-induced aniseikonia in the short presentation mode with central fixation.[20]

Misestimation of aniseikonia has been claimed to be clinically irrelevant[21–23] because further checks can be conducted before or after prescribing iseikonic spectacle lenses, although the process is lengthy and may be expensive. Yet, identifying the cause of the underestimation in the NAT and the AI Version 1 as well as the cause of the conflicting patterns of misestimation across versions of the AI is a problem whose solution could render improved tests that save time and costs to patients in clinical practice, also helping to prevent analogous problems that might arise in the use of psychophysical methods for the assessment of other visual functions. Identifying the causes of underestimation of the NAT and the AI (and, by contrast, the reasons that the OE and the SE provide accurate measurements) can also contribute to a better understanding of the perceptual aspects of aniseikonia and, thus, lead to better treatment options.

The goal of this research was to identify the causes of these patterns of misestimation and to identify the conditions (method, stimuli, viewing mode, etc.) for accurate measurements of aniseikonia. The protocol for data collection used in this study is more thorough than those in the NAT or the AI, and permits the separation of sensory and decisional influences (criterion bias) on performance. Also, data analyses are carried out separately for each observer, as group data typically obscure rather than clarify the outcomes.[24] The next three sections describe the reasons justifying the choices made in the design of the study, namely viewing conditions, stimulus characteristics, and psychophysical procedure.

## Viewing Conditions

Comparative judgments of the size of two objects can be made using direct perceptual estimates of size or using indirect correlates of size. A simple demon-

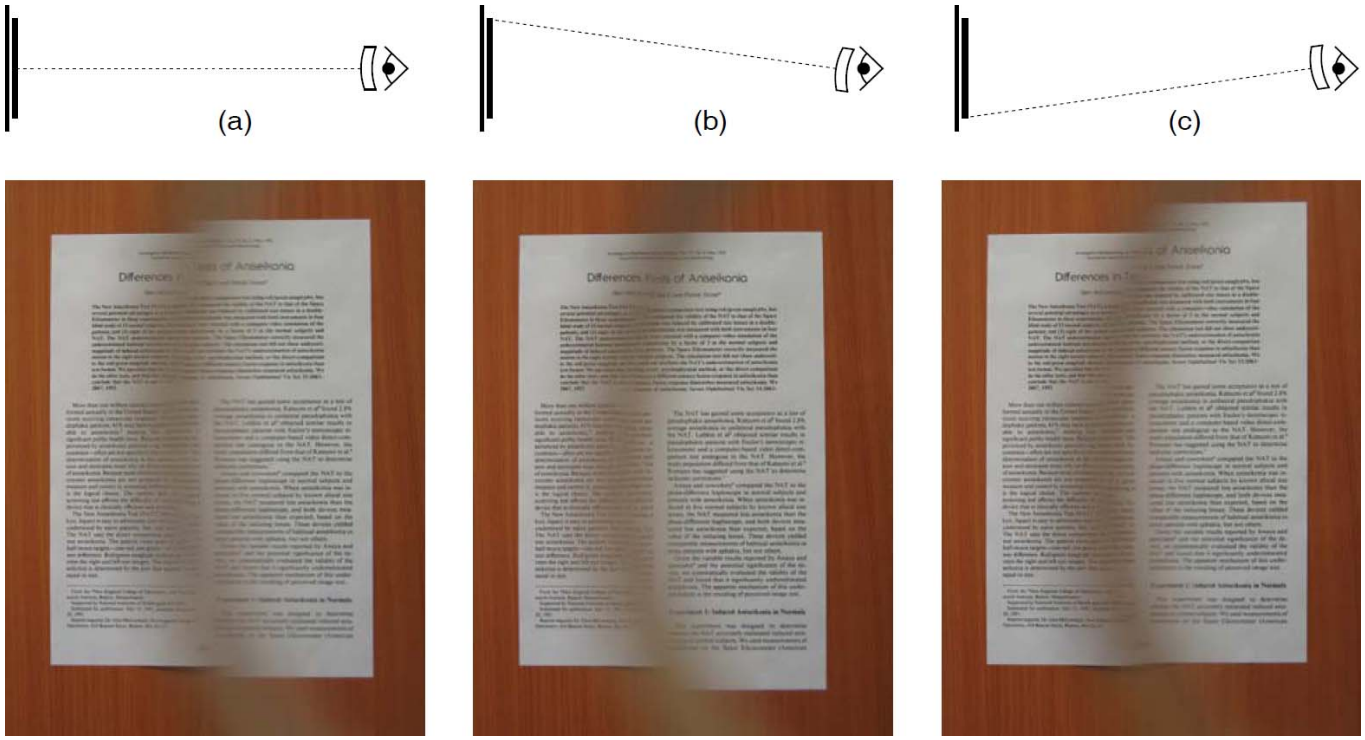


Figure 1. Illustration of the center of expansion of a magnifying lens, determined by the projection of its optical axis onto the object plane. Viewing is assumed to be monocular with the size lens placed stably before the eye in a way that it magnifies only the right side of the scene. (a) With the lens aiming at the vertical center of the sheet on the wall (*upper sketch*, not to scale), expansion of the *right side* of the scene occurs outward from that point so that the *top* and *bottom* edges of the sheet on the *right side* appear to extend vertically beyond the corresponding edges on the *left side*. (b) With the lens tilted so that it aims at the *top edge* (*upper sketch*, not to scale), the *top edge* appears aligned on both sides of the scene and the expansion of the right side makes the *bottom edge* appear to be located farther down in comparison with the same edge on the *left side*. (c) With the lens tilted to aim at the *bottom edge* instead (*upper sketch*, not to scale), the opposite effect occurs and now the *bottom edges* appear aligned on both sides, whereas the *top edge* is misaligned across sides of the scene.

stration will highlight some issues that are relevant to our subsequent discussion. Looking through an afocal size lens, readers can see for themselves the perceptual effects described next. The demonstration involves closing one eye and placing the size lens stably before the other eye in such a way that it magnifies only the right side of the scene. The picture in Figure 1a shows the resultant percept when the optical axis of the size lens projects onto a point near the vertical center of the sheet on the wall. In these conditions, the right side of the sheet seems vertically larger and, importantly, moving the eyes up or down does not alter this percept as long as the optical axis of the size lens projects onto the same point. In other words, changes of visual axis (determined by the position of the eyes) do not alter the appearance of the percept, which is only determined by the optical axis of the (external) size lens. If the optical axis of the lens aims instead at the top or at the bottom edges of the sheet, the percept will change as illustrated in Figures 1b or 1c: the top edge or the bottom edge appears aligned across left and right sides of the scene whereas the other edge appears separated. Eye movements do not alter these percepts either. This illustration shows that (isotropic) optical magnification can be described as occurring around a center of expansion determined by the projection of the optical axis of the magnifier onto the object plane. At the center of expansion, the magnified and unmagnified sides of the scene appear perceptually aligned. The center of expansion readily identified in this illustration has bearings on size comparisons when magnified and unmagnified images are each seen with one eye but sharing some spatial anchor, as in stimuli used in current aniseikonia tests.

Now consider the conventional haploscopic conditions created in the NAT or the AI (in which each eye sees only one of the objects whose sizes are to be compared) in a situation in which aniseikonia is nonmeridional and caused by optical factors residing in or on the eye (so that optical and visual axes as

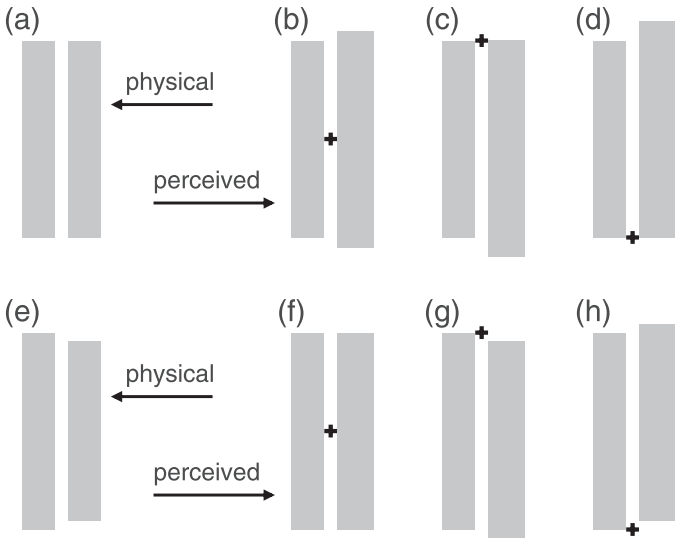


**Figure 2.** Perceived image under aniseikonia varies with fixation location when changes in fixation also relocate the optical axis. This illustration assumes a 10% magnification altering the haploscopically perceived size of the right rectangle. (a) Sample physical stimulus in which the *two rectangles* have the same size. (b–d) Perceived images when fixation location varies as indicated by the *crosses*. (e) Sample physical stimulus in which the *right rectangle* is shorter so as to become equal in length when magnified by 10%. (f–h) Perceived images when fixation location varies as indicated by the *crosses*. Fixation at the center of the configuration conveys size information, which must be judged with peripheral vision; fixation at the *top* or *bottom* edges of the configuration conveys alignment information and discloses what the physical reality of the stimulus is at the foveated points, which can be translated into size judgments that will contradict the judgments arising from central fixations. At the same time, and provided the physical width of *both rectangles* is identical, the magnified rectangle also has a larger perceived width, which further complicates judgments of size.

defined above move together when the eyes move). Equivalently, consider aniseikonia caused by spectacle lenses placed outside the eyes but in a situation in which eye movements are accompanied by adjustments of head position, so that optical and visual axes again move together. Although optical and visual axes could indeed be dissociated when aniseikonia is caused by external lenses (giving rise to the distinction between static and dynamic aniseikonia),[25] coordinated eye–head movements are a natural occurrence presumably aimed at bringing the eyes back to the primary position of gaze.[26]

Figure 2 illustrates the effects that changes in bilateral fixation location that also displace the center of expansion have when the stimuli are rectangles for a hypothetical patient with +10% aniseikonia affecting the haploscopic perception of the right rectangle. When the rectangles are physically identical (Fig. 2a), fixation at the geometrical center of the configuration results in the right rectangle appearing in peripheral vision to extend beyond the top and bottom edges of the left rectangle (Fig. 2b); if the patient then gazes to the top of the configuration to confirm the impression obtained peripherally, both rectangles would appear aligned at their top (Fig. 2c); evidence of alignment at the bottom would also be obtained when the patient gazes to the bottom of the configuration (Fig. 2d). A patient who scans the stimulus to gather the evidence needed to make a size judgment will have a hard time reconciling the conflicting perceptual information obtained peripherally and foveally across fixation locations: the right rectangle seems larger peripherally (Fig. 2b) but the top and bottom edges of both rectangles are foveally perceived to be aligned (Figs. 2c, d). Conflicting perceptual information arises also in a configuration that physically compensates for aniseikonia (Fig. 2e). In this case, rectangles would appear equal in length when fixation is on the center of the configuration (Fig. 2f), but fixations at the extremes would reveal foveally that the top and bottom edges of the right rectangle are located inwards relative to the edges of the left rectangle (Figs. 2g, h).

The preceding discussion assumed that the perceived size of the right rectangle is 10% larger than that of an identical rectangle on the left. Similar considerations apply when the right rectangle has instead a smaller perceived size (i.e., if aniseikonia is −10%), except that the strength of the conflicting evidence may differ. In general, such strength varies with the vertical angular subtense of the rectangles. If angular subtense is small, peripherally obtained information would me more reliable perceptually and, thus, fixation at the center (Fig. 2b) would provide dependable peripheral evidence of a larger size for the right rectangle whereas fixation at the extremes (Figs. 2c, d) would reveal alignment at the fixated extreme along with differences in size at the other extreme, seen farther peripherally. If the angular subtense is instead large, central fixations (Fig. 2b) may simply give a weaker peripheral impression of differences in size whereas fixations at the extremes (Figs. 2c, d) will only provide evidence of alignment at the foveated extreme: The other extreme falls too far out into the periphery to provide usable or reliable evidence of size.

It is unclear how observers will handle this conflicting information to make their judgments, whether they will use a consistent criterion across presentations, or whether all observers will handle it

identically. If observers relied more on global size judgments gathered via peripheral information upon central fixations, aniseikonia could be measured as accurately as the testing conditions and psychophysical procedure allow; yet, if they inferred size using also alignment information gathered via foveal evidence upon fixations at the edges, aniseikonia would be underestimated because judgments of size would be biased toward the physical reality of the stimulus. It is interesting to note that underestimation of lens-induced aniseikonia has been reported for the two tests (the NAT and the AI Version 1) in which presentation time is unlimited and observers are allowed (or requested) to scan the configurations without restraining head movements or otherwise ensuring a fixed center of expansion (which would certainly vary with head movements, as size lenses in those validity studies were placed on a trial frame worn by observers). In these conditions, test results will be greatly affected by whether or not observers' scanning strategies keep a fixed center of expansion. The left-eye and right-eye images are perceptually misaligned when the center of expansion is not at the fixation point, something that has been incorrectly referred to as fixation disparity.[14] The AI includes functionality to shift the left- and right-eye stimuli relative to one another to restore perceived alignment, but this strategy is effective only until the center of expansion is relocated again by head movements.

This research tested the conjecture that viewing conditions are a major determinant of the accuracy of size comparisons used to measure aniseikonia caused by optical factors. Data were thus collected in conditions in which presentation time was very short (so as to prevent observers from scanning the stimulus and changing the center of expansion) or unlimited and with instructions to scan the stimulus and check for alignment (so as to bring in the effects of changes in the center of expansion). It should be noted that measurements taken with the OE or the SE are free of these effects, for two reasons. First, the center of expansion is fixed in these instruments: lenses are placed in cells that are rigidly attached to the instrument, which in turn keeps a fixed position relative to the stimulus plane. Second, measurements are not based on judgments of size but on the consequences of different image sizes on perceived alignment (in the OE)[4,5] or on distortion of binocular space perception (in the SE).[6,27] This may explain why measurements of lens-induced aniseikonia with the SE are more accurate than those obtained with the NAT or with the AI.

## Stimulus Shape, Color, and Binocular Fusion

The semicircular stimuli used in the NAT and in the AI Version 1 have two desirable characteristics. One is that semicircles of different diameters are scaled in both spatial dimensions. Thus, overall and isotropic magnification of a small semicircle brings its size and shape to match exactly those of a larger semicircle. The second characteristic is that two semicircles facing each other do not superimpose by dilation and translation, making binocular fusion of the haploscopic percept impossible.

Rectangular stimuli in Version 2 and Version 3 of the AI lack these characteristics. Magnification of the physically shorter rectangle will make its length match that of the other rectangle, but their widths will then differ; some observers may thus be tempted to make their judgments weighing also the perceived widths of the rectangles perifixation, rather than using only perceived lengths judged with peripheral vision. Use of this joint criterion would result in overestimation, as there would be a bias toward the configuration in which rectangle widths are perceptually identical (i.e., foveal evidence of larger width would not be countered until there is compelling peripheral evidence of shorter length). Interestingly, a meaningful and often significant overestimation of aniseikonia has been reported for the AI Version 2.[18]

Two further complications arise from the use of anaglyphs. One of them is the size–color illusion: the perceived size of two physically identical areas varies with their hue.[28–32] The other is the related phenomenon that hue and brightness affect perceived depth and, indirectly, perceived size.[33–35] These two perceptual phenomena can masquerade as aniseikonia when red–green anaglyphs are used: differences in the perceptual (post retinal) processing of the color with which each eye is stimulated will be mistaken for differences between the eyes themselves.

The research reported here used semicircles as in the NAT and the AI Version 1. Also, the size–color illusion was assessed in an experimental condition in which semicircle sizes were compared in full binocular vision, that is, without red–green glasses so that both semicircles were seen with both eyes. With the same purpose, data were also collected in a haploscopic condition without lens-induced aniseikonia that swapped the red and green filters between the eyes, a strategy that could reveal other color effects affecting measurements (e.g., lateral chromatic aberrations of the eye, color-related differences in focus, effects of brightness differences between the two colors, etc.). For simplicity, we will

use the term "size–color illusion" to refer to any aspect of color vision that affects perceived size.

## Psychophysical Procedure

The method of adjustment used in Version 1 of the AI and implicitly required by the NAT has been criticized for being prone to criterion bias.[36,37] Versions 2 and 3 of the AI implement alternative methods that have long been presumed to be criterion-free, but they are identically susceptible to contamination from response and decisional biases.[38–41] Alternative psychophysical methods used across aniseikonia tests are, thus, equally prone to bias. Individual biases of such type would not explain the systematic and differential patterns of misestimation reported in the literature, but the problem of bias needs to be resolved.

A comparative study of methods used to measure spatial bisection ability for the evaluation of unilateral visual-field defects[42] determined the cause of persistent discrepancies across methods and identified an optimal method via which response biases are eliminated and sensory aspects are separated out from decisional biases affecting observed performance. Because the psychometric properties of that method are independent of the sensory modality or attribute (brightness, blur, size, etc.) subject to judgment, this research applied only this optimal method, namely, a comparative task with three response options. Each trial under this method involves a stimulus presentation analogous to those in the AI although the red semicircle is presented on the left or on the right at random in each trial; also, observers report on each trial whether the two stimuli appear identical in size or, rather, the one on the left or the one on the right appears subjectively larger. Finally, "equal" responses are treated in the analysis as a distinct category instead of being evenly split as half "left larger" and half "right larger" as is done in the AI,[43] nor are the data from the two presentation arrangements (i.e., red on the left and red on the right) aggregated into a single data set. How the resultant data are analyzed will be fully described below.

# Methods

This research followed the tenets of the Declaration of Helsinki and its protocol was approved by the institutional review board. Informed consent was obtained from each subject prior to participation.

## Observers

Seven male (including one of the authors) and three female observers participated in the study. Their ages ranged from 24 to 77 years. They were refracted prior to the study and had normal or corrected-to-normal acuity of 20/20 or better in each eye. All observers were nonstrabismic, had normal stereo acuity (<70 sec of arc) by the Randot Stereo Test (Precision Vision, Lasalle, IL), had normal binocular vision (fusion) under dichoptic anaglyphic viewing with the Worth 4-dot Test, and (except for the author) were naïve as to the goals of the study. One of the observers was expected to have meaningful aniseikonia due to uniocular cataract extraction and implantation of an intraocular lens with an intentional change in refractive error to reduce myopia in the operated eye, in anticipation of an analogous adjustment to be made when the fellow eye is operated. Two other observers showed a mild aniseikonia in our measurements that could not be attributed to evident optical causes, as they had an only minimally anisometropic prescription. A further ophthalmic measurement of these observers did not identify meaningful differences in axial length, keratometry, anterior chamber depth, or lens thickness between the eyes, so their mild natural aniseikonia did not seem to have an optical origin.

## Apparatus and Stimuli

All experimental events were controlled by MATLAB scripts under the Psychophysics Toolbox Version 3 (http://psychtoolbox.org). Responses were collected via the computer keyboard.

Stimuli were displayed at 60 Hz on a SAMSUNG SyncMaster 192N liquid crystal display monitor 37.5-cm horizontally and 30-cm vertically (Samsung Electronics, Co., Suwon, South Korea). The 1280 × 1024–pixel image area subtended 28.07° × 22.62° of visual angle at the viewing distance of 75 cm, which was secured with a chin-and-forehead rest that could not prevent observers from making small changes in head position that would displace the center of expansion. The stimuli were semicircles displayed side by side with a separation of 25 pixels (0.55°) and with their centers symmetrically placed with respect to the center of the image area (the sketch in Fig. 3). One of the semicircles was green and the other was red so that each was visible only to one eye when color filters were worn. Red and green semicircles differed in luminance but no attempt was made to match them because anaglyphic viewing implies that they will both

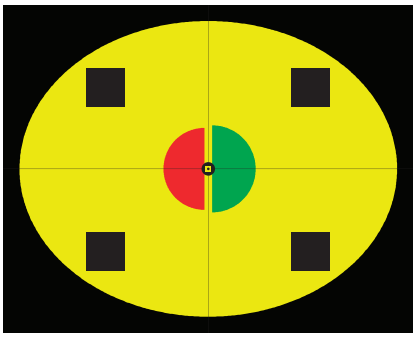

Figure 3. Sample stimulus, peripheral fusion locks, and fixation aids drawn to scale.

be seen (monocularly) as equally dark patches on a bright background, with color becoming apparent only under binocular vision as a result of cortical processing. Leakage through the filters was prevented by adjusting the yellowish color of the immediate elliptical background using a matching procedure.[44] The elliptical area had a major axis of 1179 pixels (26.04°) and a minor axis of 923 pixels (20.39°). The outer red background used in the AI was replaced here with an analogous black background because a pilot study showed that a red background disrupted perceptual stability substantially. It also became obvious in the pilot study that extra binocular locks were needed and these were provided by black squares 121 pixels in side and placed at the centers of the four quadrants of the image area.

The red semicircle was arbitrarily designated as the reference (standard) and had the same size on all trials. Its diameter was 257 pixels (5.68°) to match the recommended angular subtense in the NAT and, approximately, the angular subtense used by Fullard et al.[18] in the AI Version 1.

Black (binocularly visible) fixation aids were designed to be analogous to those in the AI and unlike the simple fixation crosses used in the NAT. They consisted of a 1-pixel wide cross spanning the entire image area and a patterned central fixation mark with an outer diameter of 43 pixels (0.95°). Fixation aids were present throughout the session, even in the intertrial intervals.

With the exceptions described below, green and red filters were, respectively, placed in front of the left and right eyes. Aniseikonia was induced by placing afocal size lenses (Chadwick Optical Inc., Souderton, PA) with nominal magnifications of 2%, 3%, and 4% before the right eye; for negative magnifications, the lens was flipped in front of the same eye. Thus, the standard (red semicircle) had the same retinal size in all conditions, as it was viewed with the left eye through the green filter without lens-induced aniseikonia. The only exception was the condition to control for the size–color illusion, in which color filters were swapped between the eyes. Size lenses, color filters, and observers' prescriptions (if needed) were mounted on a trial frame that observers wore during the sessions. Identical position and adjustment of the trial frame across sessions with the same observer could not be guaranteed. Observers wearing presbyopic correction in their habitual prescription were corrected for display distance using a +1.25 diopter (D) add, which should minimize changes of magnification arising from the use of afocal size lenses in near vision.[45]

Based on the tools and variables used in the grinding of our nominal 2%, 3%, and 4% size lenses, their magnifications were calculated to be 2.05%, 3.05%, and 4.07%, respectively, although the manufacturing process can only achieve these calculated values to some precision. Magnification was assessed directly by measuring the change in the deviation of laser beams passing at different angles through the center of the lenses, which yielded measured magnifications of 2.02%, 2.93%, and 3.97%. For all practical purposes, we used nominal magnification values in all calculations.

## Procedure

For each viewing mode (short presentation or free viewing), measurements were taken in nine conditions: a full binocular condition (i.e., without color filters), two haploscopic conditions without lens-induced aniseikonia (swapping color filters between the eyes), and six haploscopic conditions with right-eye magnifications of ±2%, ±3%, and ±4%. Observers indicated in each trial whether the left, the right, or neither semicircle was perceptually larger than the other. In the short presentation mode, they were instructed to keep their gaze steady on the fixation mark and judge overall size; in the free viewing mode, they were asked to scan the configuration and check for alignment at the top and at the bottom to make their judgment (recall that the goal of this study is not to find out whether observers would have done this naturally, but to determine the consequences of doing it). Assurance that free viewing had the intended effects was obtained during the breaks needed to change size lenses, in which observers were informally debriefed about their experiences and difficulties with the task. With the exception of observer 10, who admitted to scanning the configuration only rarely, all observers reported spontaneously that the two semi-

circles appeared to slide against each other in opposite directions as they gazed to the top and to the bottom, a clear indication that eye movements were accompanied by head movements and relocation of the center of expansion, thus providing conflicting foveal and peripheral size information.

Data in the free viewing mode were collected first, as instructions given in the short presentation mode (and the observers' ensuing experiences) might have prompted them to maintain central fixation also in the free viewing mode. Data collection under each viewing mode consisted of several sessions, starting always with the full binocular condition, followed by the two haploscopic conditions without size lenses in a random order, and continuing with the six remaining haploscopic conditions also in a random order. In short presentation conditions, trials consisted of a get-ready period of 500 ms followed by the 600-ms stimulus presentation; in free viewing conditions, each trial also included a get-ready period of 500 ms but the stimulus was not extinguished until observers had given their response.

Data in short presentation conditions were collected in three (occasionally two or four) blocks of 128 trials, preceded by a practice session of at least 30 trials. In each trial, the size of the test semicircle was determined by adaptive methods optimized for efficient and accurate estimation of nonmonotonic psychometric functions.[46] The 128 trials in each session arose from sixteen 8-trial, randomly interwoven adaptive staircases, with four staircases starting at each of four test diameters. In conditions without size lenses, starting diameters were −8, −4, 4, and 8 pixels longer than the standard. When size lenses were used, these values were shifted according to the induced magnification, with additional shifts introduced to collect data from observers suspected of natural aniseikonia. The 16 staircases represented two otherwise identical sets of eight, differing in that the test target was displayed on the left in one of the sets and on the right in the other. When the observer's response reflected that the test was subjectively larger (shorter), test size was reduced (increased) by one step for the next trial along that staircase, which was not necessarily the next trial in the session; responses reflecting that test and standard sizes were perceptually equal changed the size of the test in two steps in a direction decided at random with equiprobability. Step size was 2 pixels: the least possible change in diameter without shifting the center of the semicircle.

With free viewing, two consecutive 96-trial blocks were used consisting of twelve 8-trial staircases, two at each of three starting test diameters (−6, 0, and 6 pixels longer than the standard diameter, also shifted according to lens-induced or natural aniseikonia) for each of the two possible test positions. Step size was also 2 pixels and staircase rules were identical to those described above.

Data collection proceeded at the observer's pace. Only a response to the current trial triggered the next trial. In short presentation conditions, a further key was enabled for observers to decline responding if they had missed the stimulus (e.g., due to blinks, inadequate fixation, etc.), but observers were instructed to refrain from using this key as a means to give themselves a second chance with a stimulus they had not missed. Trials thus discarded were repeated at a later time.

## Expected Location of the Point of Subjective Equality

The point of subjective equality (PSE) is the physical magnitude that a test stimulus must have for its perceived magnitude to match the perceived magnitude of a reference (standard) stimulus. Formally, if $x_s$ is the physical magnitude of the standard, $\mu_s$ is the function describing how perceived magnitude varies with physical magnitude for the standard, and $\mu_t$ is the corresponding function for the test, the PSE is the physical magnitude $x_{PSE}$ at which $\mu_t(x_{PSE}) = \mu_s(x_s)$ so that $x_{PSE} = \mu_t^{-1}(\mu_s[x_s])$.

Under full binocular viewing in the absence of a size–color illusion, $\mu_t = \mu_s$ and the expectation is that $x_{PSE} = x_s$. A size–color illusion makes $\mu_t \neq \mu_s$ and, thus, significant shifts of the measured PSE with respect to $x_s$ reflect the magnitude of the size–color illusion. In haploscopic conditions without lens-induced (or natural) aniseikonia and without size–color illusions, $\mu_t = \mu_s$ also and the expectation is again that $x_{PSE} = x_s$. Significant shifts of the measured PSE with respect to $x_s$ then reflect either natural aniseikonia or a size–color illusion. These two potential causes can be distinguished because our full binocular condition provides a measure of the magnitude of a hypothetical size–color illusion.

Finally, in haploscopic conditions with only lens-induced aniseikonia affecting perceived size (i.e., no natural aniseikonia and no size–color illusions), $x_{PSE}$ should be shifted away from $x_s$ in quantities that can be anticipated from the aniseikonia induced by the lens, which is what makes $\mu_t$ and $\mu_s$ differ in these cases. With $x_s = 257$ pixels, the expected $x_{PSE}$ for magnifications of 2%, 3%, and 4% are 251.96, 249.51, and 247.12 pixels, respectively; for magnifications of −2%, −3%, and −4%, the expected $x_{PSE}$ are 262.14, 264.71, and 267.28 pixels, respectively. Note that

$|x_{PSE} - x_s|$ differs for positive and negative magnifications of the same size. Without natural aniseikonia or size–color illusions, shifts of the measured PSEs away from these expectations will reflect a failure to measure lens-induced aniseikonia adequately.

## Model-Based Analysis

Psychometric functions were fitted to the data from each subject in each condition and viewing mode to measure the PSE. The theoretical form of these functions was derived from Thurstonian assumptions about the sensory and decisional processes involved in the ternary task. Although the application of this type of model to analogous tasks has been described in detail elsewhere,[39,40,42,47–49] the model is sketched in Figure 4 for two sample situations and a brief description follows.

The psychophysical function μ describing the relation between perceived and physical length has been shown to be approximately linear via magnitude production tasks[50] and magnitude estimation tasks,[51–53] whereas the standard deviation σ of perceived length increases with physical length.[50] The same holds for perception of circle size.[50–52] Consider haploscopic conditions with perceived size varying between the eyes due to natural or lens-induced aniseikonia. The perceived size $S_i$ of a semicircle of diameter $x$ seen with eye $i \in \{OS, OD\}$, for left eye (OS) or right eye (OD), is thus a random variable whose mean $\mu_i$ and variance $\sigma_i^2$ are functions of $x$ given by

$$\mu_i(x) = \alpha_i + \beta_i x \tag{1a}$$

$$\sigma_i^2(x) = \kappa_i + \gamma_i x \tag{1b}$$

Without loss of generality, assume that observers compare the sizes of the semicircles on the left and right sides of the screen and base their decision on the perceived difference $D = S_R - S_L$, where subscripts denote position on the screen (right or left) and will later be replaced with references to the eye with which each semicircle is seen. Observers give a "left larger" response when $D < \delta_1$, they give a "right larger" response when $D > \delta_2$, and they give a "can't tell" response when $\delta_1 \le D \le \delta_2$. With normally distributed $S_i$, the probability of each type of response is

$$\text{Prob}(\text{"left larger"}) = \Phi[(\delta_1 - \mu_d)/\sigma_d] \tag{2a}$$

$$\text{Prob}(\text{"can't tell"}) = \Phi[(\delta_2 - \mu_d)/\sigma_d] - \Phi[(\delta_1 - \mu_d)/\sigma_d] \tag{2b}$$

$$\text{Prob}(\text{"right larger"}) = 1 - \Phi[(\delta_2 - \mu_d)/\sigma_d] \tag{2c}$$

where Φ is the unit-normal cumulative distribution function and $\mu_d$ and $\sigma_d$ are the mean and standard deviation of $D$, which vary with test size $x$ according to which semicircle was on each side and which eye was each semicircle seen with (Appendix A in the Supplementary Material).

By substitution in Equation 2, the psychometric functions $\Psi_j$ (for "left larger" responses) and $\Upsilon_j$ (for "can't tell" responses) when the test (green) semicircle is displayed at position $j$ (with $j \in \{L, R\}$, for left and right position) and the green filter is before the left eye are (Figs. 4b, c)

$$\Psi_L(x) = \Phi\left(\frac{\delta_1 - (\mu_{OS}(x_s) - \mu_{OD}(x))}{\sqrt{\sigma_{OD}^2(x) + \sigma_{OS}^2(x_s)}}\right) \tag{3a}$$

$$\Psi_R(x) = \Phi\left(\frac{\delta_1 - (\mu_{OD}(x) - \mu_{OS}(x_s))}{\sqrt{\sigma_{OD}^2(x) + \sigma_{OS}^2(x_s)}}\right) \tag{3b}$$

$$\Upsilon_L(x) = \Phi\left(\frac{\delta_2 - (\mu_{OS}(x_s) - \mu_{OD}(x))}{\sqrt{\sigma_{OD}^2(x) + \sigma_{OS}^2(x_s)}}\right) - \Phi\left(\frac{\delta_1 - (\mu_{OS}(x_s) - \mu_{OD}(x))}{\sqrt{\sigma_{OD}^2(x) + \sigma_{OS}^2(x_s)}}\right) \tag{3c}$$

$$\Upsilon_R(x) = \Phi\left(\frac{\delta_2 - (\mu_{OD}(x) - \mu_{OS}(x_s))}{\sqrt{\sigma_{OD}^2(x) + \sigma_{OS}^2(x_s)}}\right) - \Phi\left(\frac{\delta_1 - (\mu_{OD}(x) - \mu_{OS}(x_s))}{\sqrt{\sigma_{OD}^2(x) + \sigma_{OS}^2(x_s)}}\right) \tag{3d}$$

Equations for the reverse arrangement of filters (red before the left eye) are analogous except that subscripts OS and OD are swapped; in the full binocular condition, equations are also analogous except that subscripts OS and OD are respectively replaced with subscripts "red" and "green" (Appendix A in the Supplementary Material). Note that $\delta_1 = \delta_2$ renders $\Upsilon_L(x) = \Upsilon_R(x) = 0$ and, hence, yields observers who never give "can't tell" responses. On the other hand, $\delta_1 = -\delta_2$ renders $\Upsilon_L(x) = \Upsilon_R(x)$ and, thus, yields observers without decisional bias. Then, the lateral separation between psychometric functions for each arrangement (test presented on the left or on the right) is a pure indicator of decisional bias.

Substituting Equation 1 into Equation 3, the arguments of Φ become

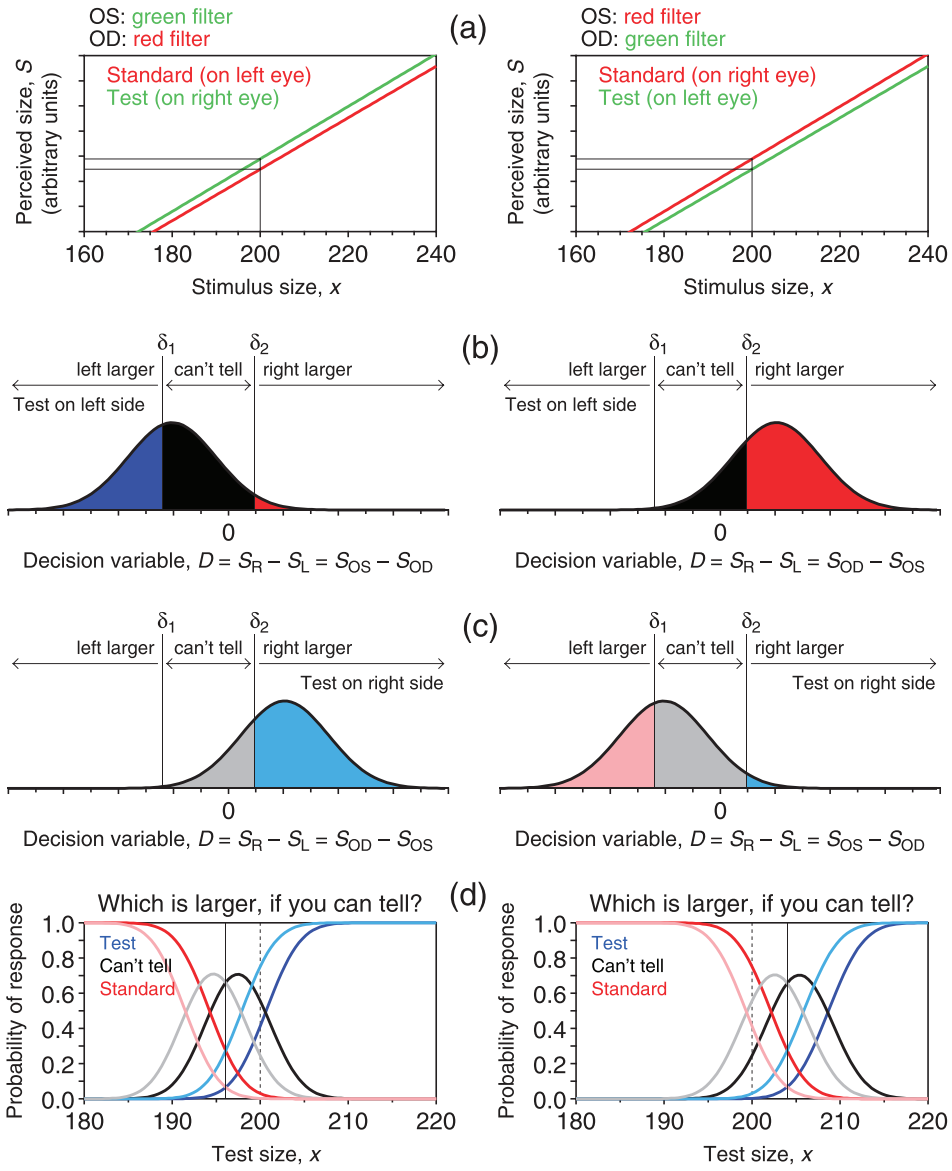


**Figure 4.** Model psychometric functions for an observer whose right retinal image is 2% larger than the left retinal image. Color filters for haploscopic viewing without additional lens-induced aniseikonia are assumed to be placed before the eyes in opposite arrangements in each column, as indicated at the *top*. (a) Psychophysical functions describing perceived size with each eye, $\mu_{OS}(x) = 6.4x$ and $\mu_{OD}(x) = 6.528x$ (*red and green lines*). The perceived size of the standard (*red*) semicircle of size $x = 200$ (arbitrary units) varies according to the eye with which it is seen. The scale of the vertical axis is irrelevant and has been omitted. The variance of perceived size is assumed to be equal with both eyes and given by equation 1b with $\kappa_{OS} = \kappa_{OD} = 0$ and $\gamma_{OS} = \gamma_{OD} = 0.16$. (b) Distribution of the decision variable $D$ when the test stimulus has the same size as the standard and the test is presented on the left side of the screen. *Vertical lines* indicate how the decision space is partitioned and mapped onto judgments. The horizontal scale is irrelevant and has been omitted. (c) Identical to the previous except that the test stimulus is presented on the *right side*. (d) Resultant psychometric functions in the ternary task. Each curve refers to the response (left larger, can't tell, or right larger) and condition (test on the left or test on the right) indicated by the corresponding color in (b) and (c). The *vertical dashed line* indicates the point of objective equality (POE); the *solid vertical line* indicates the PSE, which differs from the POE due to the natural aniseikonia of the observer and whose position varies with the arrangement of color filters before the eyes. The sets of curves for test-on-left-side and test-on-right-side presentations are pushed apart by decisional bias (i.e., when $\delta_1 \neq -\delta_2$).

$$\frac{\delta_k - (\mu_{OS}(x_s) - \mu_{OD}(x))}{\sqrt{\sigma^2_{OD}(x) + \sigma^2_{OS}(x_s)}} = \frac{\delta_k - (\alpha_{OS} - \alpha_{OD} + \beta_{OS}x_s - \beta_{OD}x)}{\sqrt{\kappa_{OS} + \kappa_{OD} + \gamma_{OS}x_s + \gamma_{OD}x}} \quad (4a)$$

$$\frac{\delta_k - (\mu_{OD}(x) - \mu_{OS}(x_s))}{\sqrt{\sigma^2_{OD}(x) + \sigma^2_{OS}(x_s)}} = \frac{\delta_k + (\alpha_{OS} - \alpha_{OD} + \beta_{OS}x_s - \beta_{OD}x)}{\sqrt{\kappa_{OS} + \kappa_{OD} + \gamma_{OS}x_s + \gamma_{OD}x}} \quad (4b)$$

for $k \in \{1, 2\}$. Because our experiments involve a single standard level $x_s$ and a narrow range of test levels $x$ around it, fitting this full model revealed that the term $\gamma_{OD}\,x$ in the argument of the denominator of Equation 4 makes a negligible contribution and does not improve the fit compared with a simplified version of the model in which $\sigma^2_{OD}(x) + \sigma^2_{OS}(x_s)$ is regarded as a constant. We thus fitted a simplified model in which $\gamma_{OD} = \gamma_{OS} = 0$ so that the arguments of $\Phi$ in Equation 3 finally become

$$\frac{\delta_k - (\mu_{OS}(x_s) - \mu_{OD}(x))}{\sqrt{\sigma^2_{OD}(x) + \sigma^2_{OS}(x_s)}} = \delta^*_k - \alpha^* + \beta^* x \quad (5a)$$

$$\frac{\delta_k - (\mu_{OD}(x) - \mu_{OS}(x_s))}{\sqrt{\sigma^2_{OD}(x) + \sigma^2_{OS}(x_s)}} = \delta^*_k + \alpha^* - \beta^* x \quad (5b)$$

where $\alpha^* = (\alpha_{OS} - \alpha_{OD} + \beta_{OS}x_s)/\sqrt{\kappa_{OS} + \kappa_{OD}}$, $\beta^* = \beta_{OD}/\sqrt{\kappa_{OS} + \kappa_{OD}}$ and $\delta^*_k = \delta_k/\sqrt{\kappa_{OS} + \kappa_{OD}}$ for $k \in \{1, 2\}$ are the four parameters to be estimated. Analogous equations with changes in the subscripts are obtained for the remaining conditions (Appendix A in the Supplementary Material).

## Parameter Estimation and Approximate Nonrejection Regions

Maximum-likelihood estimates for $\alpha^*$, $\beta^*$, $\delta^*_1$, and $\delta^*_2$ were obtained from data for each observer in each condition using the NAG[54] subroutine E04JYF under the constraints $\alpha^* > 0$, $\beta^* > 0$, and $\delta^*_1 \le \delta^*_2$. It has been shown[48] that the PSE is the point at which the ensemble of psychometric functions has an axis of bilateral symmetry, which can be obtained from parameter estimates as $x_{PSE} = \alpha^*/\beta^*$. The PSE, measured in units of $x$, can be readily transformed into any required measure of aniseikonia, either as a percentage or as a factor.

Expected and measured PSEs are unlikely to match exactly due to sampling error, but meaningful discrepancies between expected and measured PSEs should be identifiable. In subject-by-subject analyses in which a single PSE is measured per condition, sampling models for confidence intervals (CIs) around parameter estimates are not applicable. Nevertheless, rejection regions around the expected PSE can be determined via simulation methods. Previous studies have shown that the standard error of estimates of the location of a psychometric function for discrimination (the PSE) vary with the adaptive procedure implemented, the number of trials and, often, other parameters such as the slope of the psychometric function.[55–61] Appendix B in the Supplementary Material describes the bootstrap procedure that was used to determine approximate nonrejection regions so as to check for the statistical significance of the deviation of a measured PSE with respect to the expected PSE.

# Results

## Size–Color Illusion and Natural Aniseikonia

Figure 5 shows results for each observer in the full binocular condition under each viewing mode. With the exception to be discussed below, the size–color illusion is understandably absent for all observers under the free viewing mode (i.e., the measured PSE is virtually at the standard level) because size was determined indirectly by checking for alignment at the extremes in what becomes a foveal vernier task. The exception was observer 10, who reported compliance with instructions to scan the configuration "only when needed to make a judgment," adding that such need did not arise very often. In contrast, in the short presentation mode where observers must perform a size discrimination task at a glimpse of the stimulus, traces of a size–color illusion (i.e., the PSE is 1–3 pixels above $x_s$) are apparent for all observers except 3, 4, 8, and 9, with the red semicircle appearing perceptually larger than the green semicircle.

All observers gave "can't tell" responses relatively frequently under both viewing modes, with the exception of observer 3. In general, the psychometric functions for "can't tell" responses (black and gray curves) were narrower and taller in the free viewing mode than in the short presentation mode, suggesting more ability to discriminate small differences in size under free viewing. Finally, under the free viewing mode, all observers except 5 and 10 showed little or no decisional bias (lateral displacement of psychometric functions across test positions; darker versus paler data points and curves) compared with the

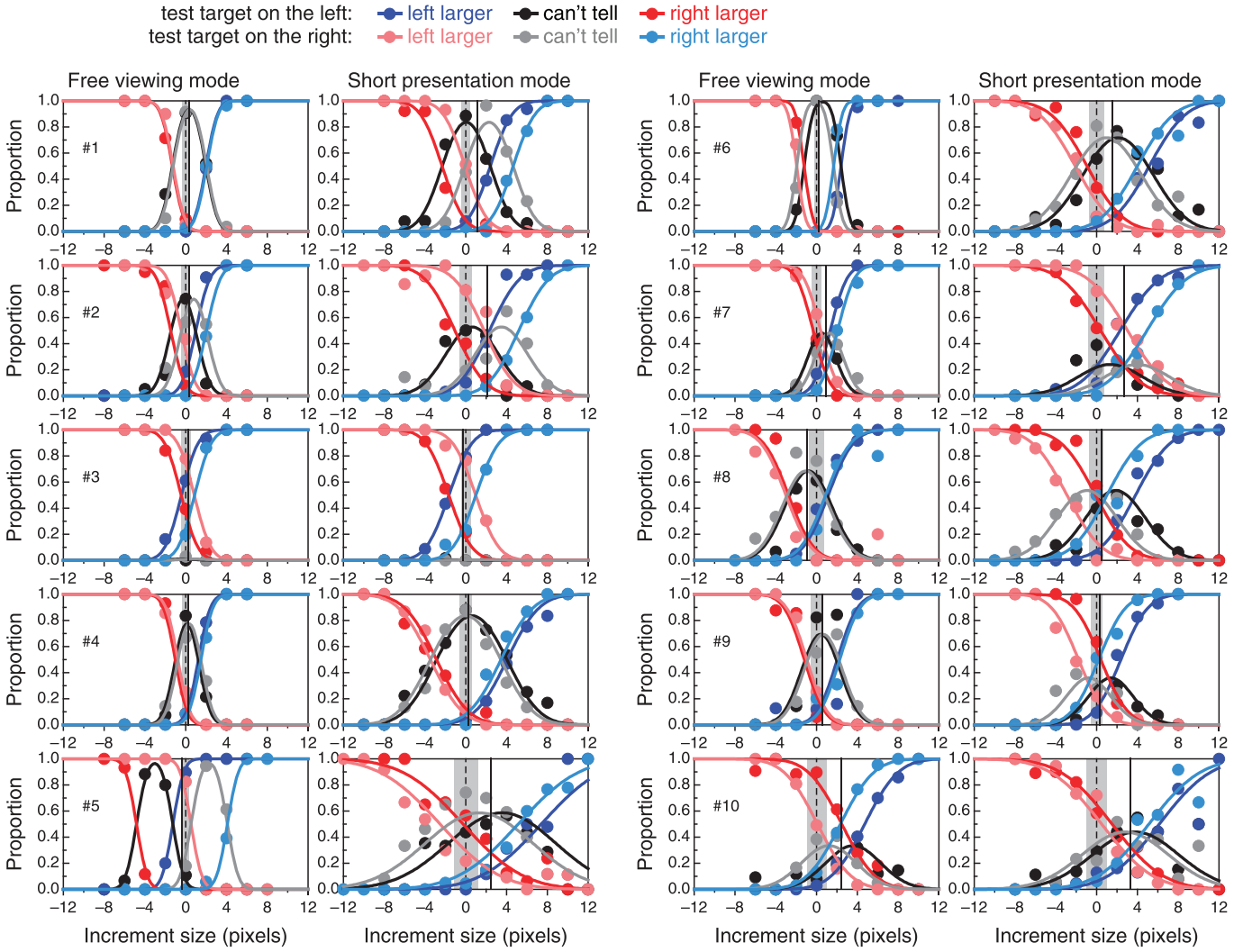


**Figure 5.** Psychometric functions in the full binocular condition under each viewing mode. In each panel, the *dashed vertical line* gives the expected location of the PSE if perceived size does not vary with color; the *solid vertical line* indicates the actual location of the PSE. Because both semicircles are seen with both eyes in these conditions, any meaningful lateral shift of the PSE away from its expected location reflects differences in perceived size associated to color and independent of any natural aniseikonia that observers might have; the *gray area* around the expected PSE denotes the nonrejection region.

larger decisional bias that they generally exhibited under the short presentation mode (recall that observer 10 reportedly used fixation under both viewing modes so the only difference for him was the unlimited versus short presentation duration).

Figure 6 shows results for each observer and viewing mode in the haploscopic conditions without size lenses and color filters in two arrangements before the eyes. Observers display decisional biases and patterns of usage of the "can't tell" option analogous to those that they displayed in the full binocular condition. As illustrated in Figure 4, natural aniseikonia without size–color illusions shows as ensembles of psychometric functions that are shifted away from $x_s$ by the same amount but in opposite directions when color filters are swapped between the eyes. In line with this argument, in the free viewing mode (left half of Fig. 6), observer 9 shows weak traces of natural aniseikonia (about 1%) in the form of a larger retinal image in the left eye and observer 5 shows evidence of a larger aniseikonia (about 4%) in the opposite form. These measures may have been underestimated because free viewing allows performing a foveal vernier task only disrupted by conflicting peripheral information about overall size. In fact, data from the short presentation mode in which observers perform a size discrimination task (right half of Fig. 6) suggest larger natural aniseikonia

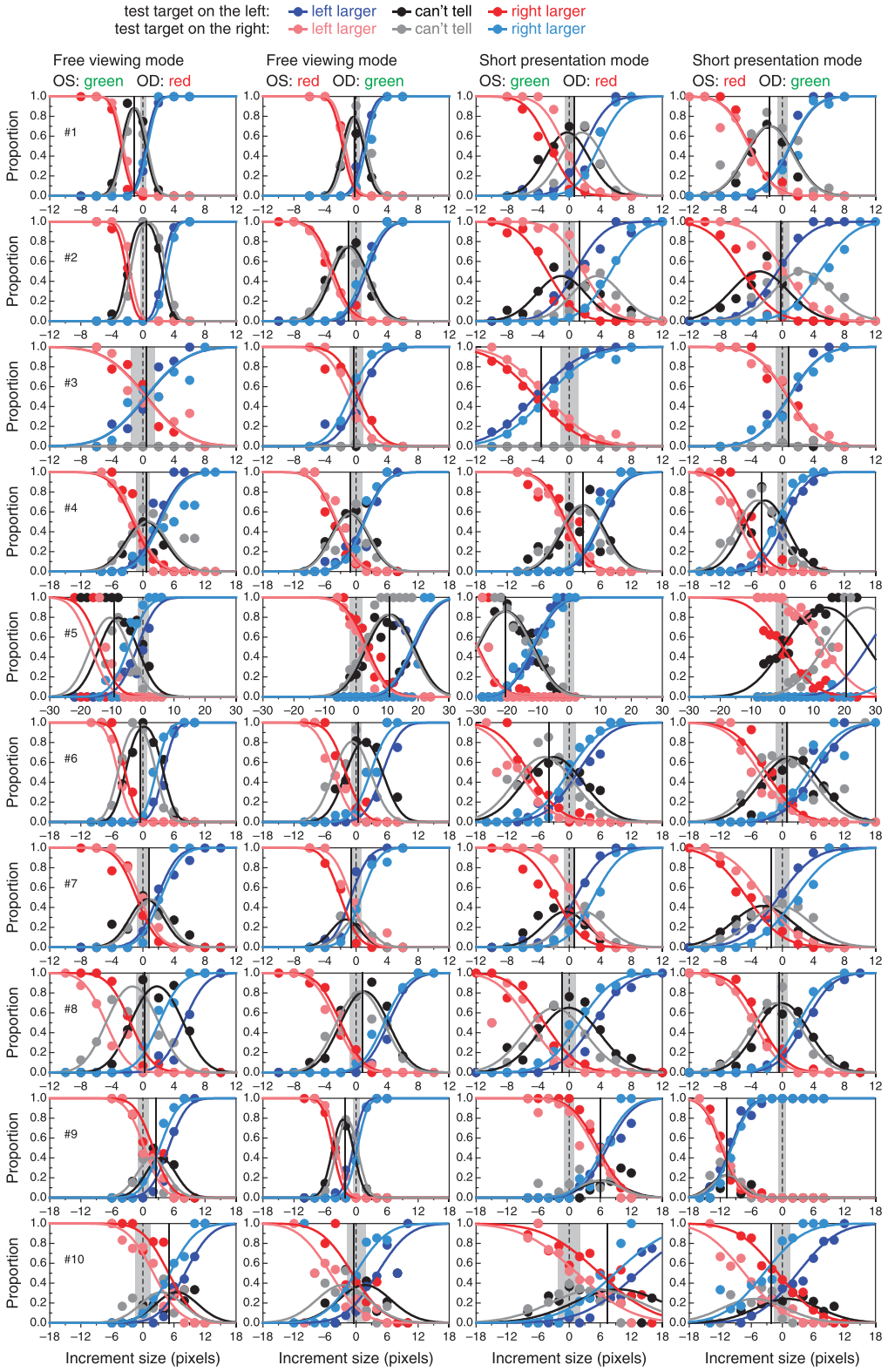


Figure 6. Psychometric functions for each observer and viewing mode in the two haploscopic conditions without lens-induced aniseikonia, providing a further test of size–color illusion and natural aniseikonia. *Dashed and solid vertical lines and gray areas* represent expected and actual PSE and nonrejection regions as in Figure 4, where the expectation reflects no size–color illusion and no aniseikonia. Note that the horizontal range is broader for observers 4, 6, 9, and 10, and even broader for observer 5.

for these observers: 2% or 4% for observer 9 (depending on which arrangement of color filters is considered) and 8% for observer 5. The short presentation mode also revealed a weak natural aniseikonia (about 1.5%) for observer 4, which went undetected in the free viewing mode.

A size–color illusion without natural aniseikonia has a different manifestation: the measured PSE would be at the same point for both arrangements of color filters before the eyes, the reason being that the red and green lines in Figure 4a (describing perceived size with each eye) would still follow different paths within each panel due to differences in perceived size according to color but their paths would be identical in the two panels because the eye with which each semicircle is seen does not make any difference in the absence of natural aniseikonia. As a result, the ensembles of psychometric functions will be shifted away from $x_s$ by the same amount and in the same direction when color filters are swapped between the eyes. No observer's data showed evidence to this effect.

Finally, and by the same argument, joint effects of natural aniseikonia and size–color illusion show as ensembles of psychometric functions displaced away from $x_s$ by different amounts and possibly in different directions when color filters are swapped between the eyes (recall that many color-related issues may produce these asymmetric displacements and that we refer to them as size–color illusions for simplicity). Data from some observers (e.g., 1, 2, and 7) show traces of this asymmetry with short presentations that are consistent with the traces of a size–color illusion that were identified for them in the full binocular condition also under the short presentation mode (Fig. 5). On the other hand, observers 3, 6, 9, and 10, show stronger signs of a mixture of natural aniseikonia and size–color illusion in the short presentation mode, although the lack of evidence of a size–color illusion for observers 3 and 9 in the full binocular condition suggests other causes.

## Lens-Induced Aniseikonia

Results for haploscopic conditions with size lenses are shown in Figure 7 for two sample observers under each viewing mode, revealing features that can only be extracted with a ternary task. Consider observer 2 first (left half of Fig. 7). This observer gave "can't tell" responses more often in the free viewing mode (left column) than in the short presentation mode (right column): At all levels of lens-induced aniseikonia, the psychometric functions for "can't tell" responses (black and gray curves) are taller under free viewing than with short presentations, in agreement with an analogous behavior displayed by this observer without size lenses (Fig. 6). Secondly, and also in agreement with results obtained without size lenses, psychometric functions are steeper under free viewing than with short presentations. Thirdly, decisional bias is larger in the short presentation mode than under free viewing and its form is consistent across levels of lens-induced aniseikonia (darker data and curves are always shifted to the left of lighter data and curves, as was the case also in Fig. 6). Finally, and more importantly, underestimation of lens-induced aniseikonia as indicated by a measured PSE (solid vertical line in each panel) that is displaced toward 0 relative to the expected PSE (dashed vertical line in each panel) is generally larger under free viewing than with short presentations.

In comparison, observer 3 (right half of Fig. 7) gave very few "can't tell" responses (and only under free viewing), has psychometric functions that only differ slightly in slope across viewing modes, rarely displays any decisional bias, seems more immune to lens-induced aniseikonia for negative magnifications under free viewing, and seems instead about equally prone to lens-induced aniseikonia under both viewing modes for positive magnifications.

This range of individual differences in usage of the "can't tell" response option, slope of the psychometric functions, decisional bias, and susceptibility to lens-induced aniseikonia is representative of analogous features consistently displayed by the data from other observers across conditions (including the condition without size lenses). Other observers also showed systematic differences in these respects between the free viewing and short presentation modes. Some of these features (specifically, differences in usage of the "can't tell" response option, decisional bias, and slope of psychometric functions across viewing modes) were clearly apparent in Figures 5 and 6 also.

For a compact display of results about susceptibility to lens-induced aniseikonia, Figure 8 plots measured aniseikonia against expected lens-induced aniseikonia for each observer under each viewing mode, including also data from the null condition (i.e., haploscopic presentation without size lenses). Observer 5 shows systematic signs of a relatively strong natural aniseikonia measured at about 4% in the free viewing mode but at about 8% in the short presentation mode. Observer 10 also shows traces of a weaker aniseikonia of the opposite sign, measured at about 3% under both viewing modes (recall that this

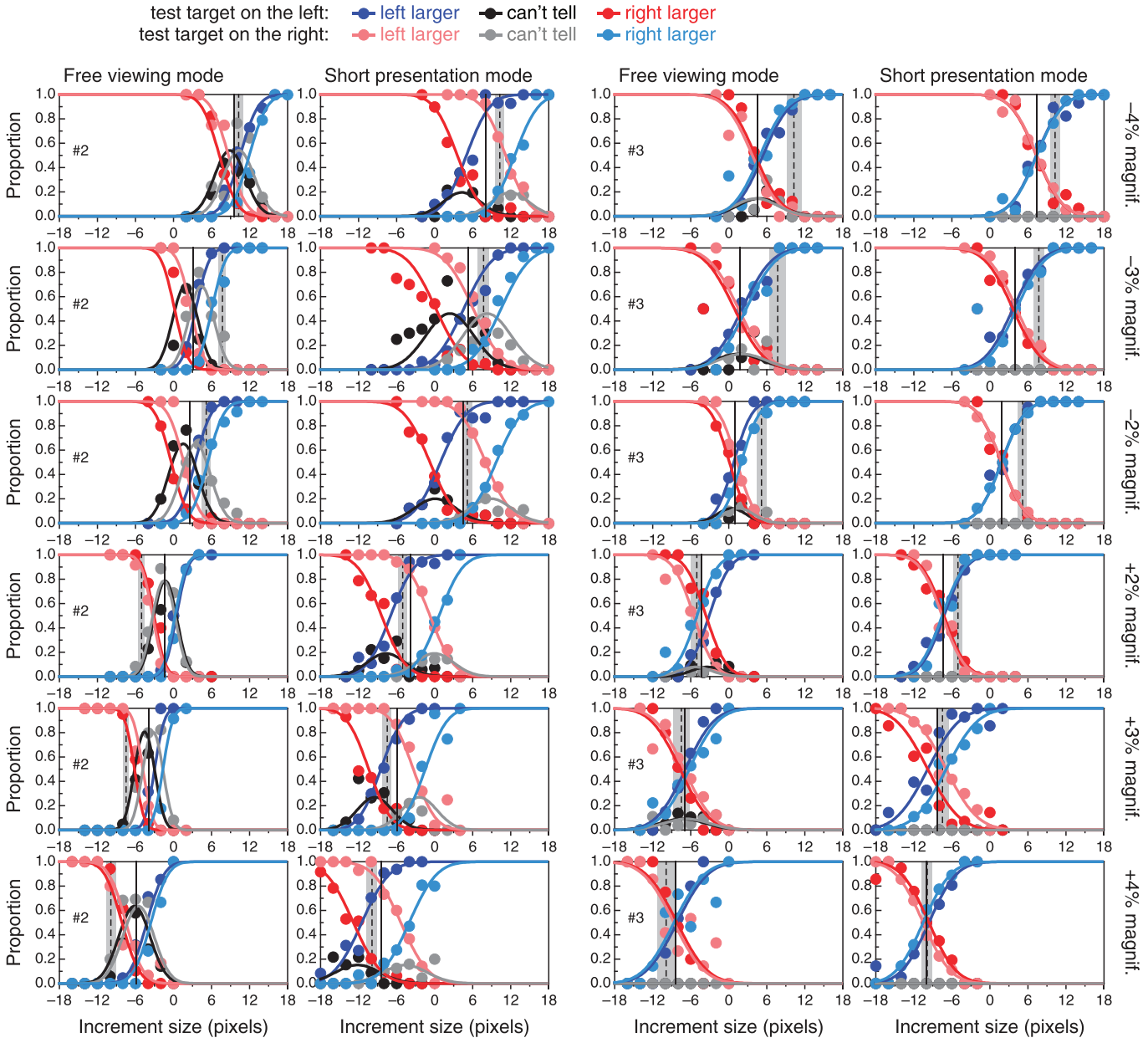


Figure 7. Full set of data and fitted psychometric functions for two observers (*left and right halves* of the figure) across viewing modes (*left and right column* within each half) and levels of lens-induced aniseikonia (*rows*).

observer used fixation also under the free viewing mode), and observer 9 shows also traces of an even weaker aniseikonia of the same sign and measured at about 1% in the free viewing mode but at about 2% in the short presentation mode. This natural aniseikonia did not seem to affect the accuracy with which these observers make size comparisons in the full binocular condition under free viewing (Fig. 5): Data from observers 5 and 10 in the short presentation mode with full binocular vision were only noisier and revealed broader psychometric functions, suggesting a size–color illusion in observer 10 and less decisional bias than in the free viewing mode for observer 5; data from the also aniseikonic observer 9 showed none of these characteristics under full binocular viewing, where data and psychometric functions were indistinguishable from those of iseikonic observers. Leaving the aniseikonic observers 5, 9, and 10 aside, Figure 8 reveals some additional interesting features under each viewing mode and across them.

Consider first the free viewing mode (left panel in Fig. 8). Data points fall everywhere under the

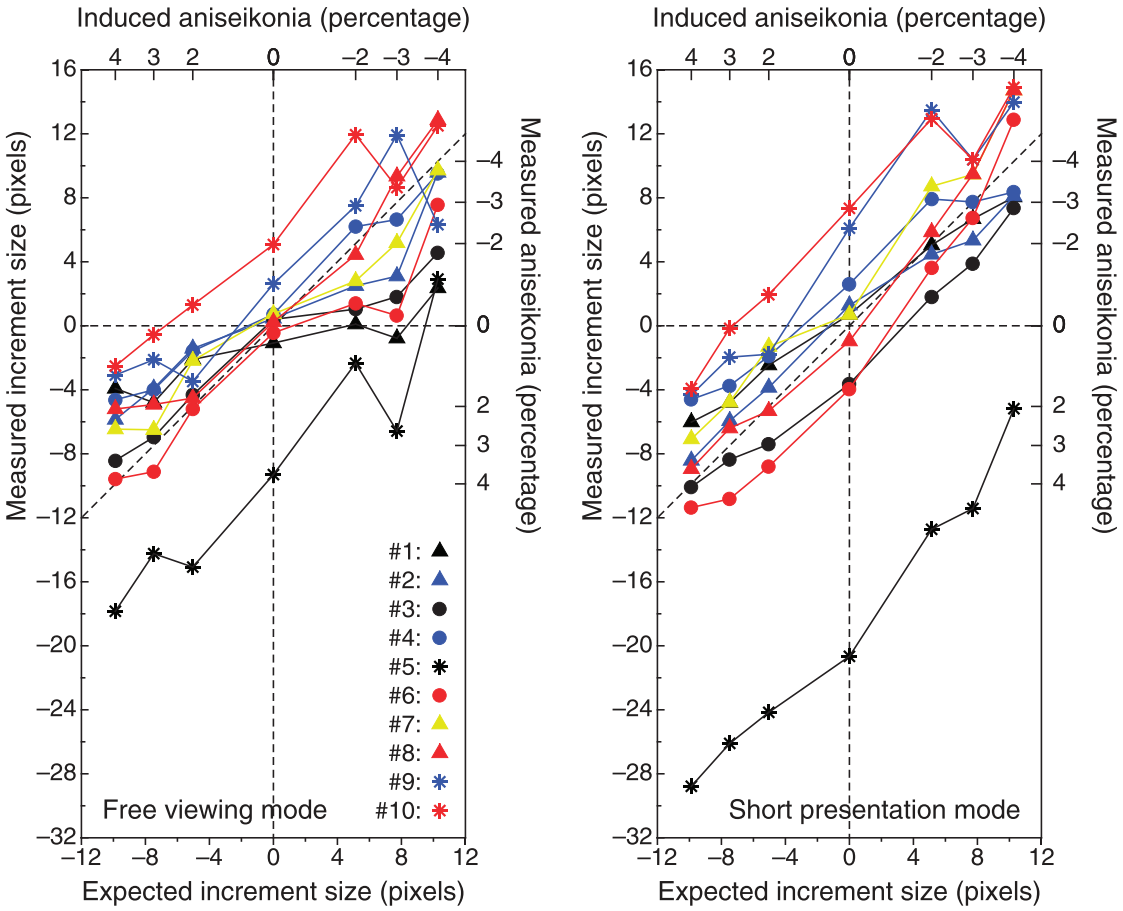


Figure 8. Relation between expected and measured aniseikonia. Expected and measured aniseikonia are given as the increment size of the diameter of the test semicircle (in pixels) for its perceived size to match that of the standard semicircle, and also as a percentage. The *left panel* shows results under the free viewing mode; the *right panel* shows results under the short presentation mode. Data from each individual observer are identified by the color described in the *legend* inside the *left panel*.

diagonal in the first quadrant, indicating that measured aniseikonia can be anywhere between null and the expected value when size lenses are flipped to produce negative magnifications. This space is covered in such form because data from different observers follow distinctly different paths, suggesting that they carry out the task in different ways that produce different outcomes under the free viewing mode. A similar scatter occurs above the diagonal in the third quadrant (when size lenses are placed to produce positive magnifications), except that the area near the horizontal axis is less covered.

Results under the short presentation mode differ remarkably (right panel of Fig. 8). Data from observers 3, 4, and 6 describe paths parallel to the diagonal but meaningfully above or below it. These patterns were not observed for them under the free viewing mode, and they are consistent with a natural aniseikonia in which either the right-eye image (observers 3 and 6) or the left-eye image (observer 4) is larger than the other eye's image. It is also remarkable that the apparent aniseikonia of observers 3 and 6 might have gone unnoticed if color filters had been placed in reverse before the eyes, due to the asymmetry described earlier on discussing results shown on the right half of Figure 6. The natural aniseikonia of observer 4, on the other hand, did show upon swapping color filters between the eyes, as this observer's measured PSEs in such cases were symmetrically placed with respect to the expected PSE (the right half of Fig. 6).

With or without natural aniseikonia, data from all observers under the short presentation mode describe a path that is very approximately parallel to the diagonal and, thus, indicate absence of any systematic underestimation. Although these data paths speak for themselves, we conducted the conventional regression analysis to check for (overall) unit slope in each

viewing condition. For free viewing, the regression slope was 0.697, which was significantly different from unity ($t_{68} = -4.43$; $P < 0.001$), and thus indicates underestimation of lens-induced aniseikonia (95% CI for the slope: [0.560, 0.833]); in the short presentation mode, the regression slope was 0.937, which was not significantly different from unity ($t_{68} = -0.56$), and thus indicates adequate estimation of lens-induced aniseikonia (95% CI for the slope: [0.712, 1.162]). Note also that under free viewing the slope at positive magnifications seems closer to unity whereas the slope at negative magnifications seems farther from unity. This reflects the role of contradictory perceptual information discussed earlier, which has different strengths for positive and negative magnifications under the free viewing mode. In sum, compared with significant underestimation of lens-induced aniseikonia in the free viewing mode, the short presentation mode does not result in underestimation and allows detecting weak natural aniseikonia that is not apparent under free viewing.

### Additional Conditions

The results discussed thus far prompted us to collect data from some observers under additional conditions to look for the origin of some of the features discussed earlier. Because those data address side issues, their description and discussion are placed in Appendix C in the Supplementary Material.

## Discussion

This study investigated factors causing the underestimation reported in the literature when aniseikonia is measured with the NAT or the AI. We included control conditions to check for contaminating factors such as size–color illusions or decisional biases, which might masquerade as aniseikonia (or mask an existing aniseikonia) under the conventional protocol. Our results are summarized next, followed by commentary on their implications for the clinical measurement of aniseikonia, for the empirical validation of aniseikonia tests, and for the clinical management of aniseikonia.

### Annotated Summary of Results

Viewing mode seems to be the main factor contributing to the underestimation of lens-induced aniseikonia (Fig. 8). With size lenses mounted on a trial frame worn by observers, free viewing with unlimited time to give a response allows observers to relocate the center of expansion and gather contradictory foveal and peripheral perceptual information about size. The resolution of the conflict appears to involve weighing the two sources of contradictory information, which can only favor underestimation. The strength of conflicting information decreases when size lenses are placed to produce negative instead of positive magnifications. In contrast, short presentations eliminate changes in the center of expansion while the stimulus is available for inspection, although such changes may occur across trials. But, even with short presentations, use of anaglyphs to measure aniseikonia brings in other contaminants whose role we also investigated.

We collected data in a full binocular condition in which both semicircles are seen with both eyes (Fig. 5) so as to check for a size–color illusion that might be mistaken for aniseikonia when each color semicircle is seen with one eye. We found no traces of a size–color illusion under free viewing, surely due to the fact that observers carried out a vernier task instead of a size discrimination task. Yet, we found significant traces of a size–color illusion in 6 of 10 observers under the short presentation mode: the green semicircle (which was brighter than the red semicircle when seen without filters) needed to be up to about 1% larger than the red semicircle for both semicircles to be perceived as having the same size. Although the size–color illusion seems small in magnitude (when present) in our sample, it will only contribute an additive error to measurements of aniseikonia. The error has bearings on clinical practice although such errors would cause a systematic misestimation of lens-induced aniseikonia, not the multiplicative underestimation reported in the literature.[11–13,16–19]

In all experimental conditions, data were collected with the red and green semicircles displayed in both presentation arrangements (red on the left and green on the right, or vice versa) to check for decisional bias. We found strong evidence of decisional bias in most observers, which manifests as psychometric functions that are laterally displaced for each presentation arrangement. The magnitude and form of the bias varied across observers and across viewing modes for some observers, but it was otherwise constant for each observer across levels of lens-induced aniseikonia (Fig. 7). The presence of decisional bias also has bearings on clinical practice, as it represents a strong contaminant when data are collected using a single arrangement (i.e., green always on the right or on the left). Consider the data for observer 2 under the short presentation mode in the haploscopic condition without size lenses and with

red and green filters placed before the left and right eyes, respectively (rightmost panel in the second row of Fig. 6). If the clinical test arbitrarily places the green (test) semicircle only on the left of the screen, only the darker symbols and curves would be obtained so that observer 2 would be taken to perceive a larger image with the left eye; if the test semicircle were placed instead on the right of the screen, only the lighter symbols and curves would be obtained so that observer 2 would instead be taken to perceive a smaller image with the left eye. These contradictory outcomes are only the result of decisional bias, whose presence can only be unveiled by using both presentation arrangements and analyzing the data separately. In any case, decisional bias introduces additive errors of clinical relevance that would cause a systematic misestimation of lens-induced aniseikonia, but again not the multiplicative underestimation reported in the literature.[11–13,16–19]

Our haploscopic conditions also included measurements of aniseikonia without size lenses but with color filters placed in both possible arrangements (i.e., green filter before the left eye and red filter before the right eye, and vice versa), to distinguish true aniseikonia from size–color illusions. In most cases we found the expected symmetric effects of natural aniseikonia (or lack thereof), but data from three of ten observers under the short presentation mode revealed what appeared to be a mild but significant aniseikonia in one of the filter arrangements and no aniseikonia whatsoever in the other (see the results for observers 3 and 6 in the right half of Fig. 6). Data from two other observers in the same situation reflected a mild aniseikonia in one of the arrangements of color filters and a stronger aniseikonia in the other (results for observers 9 and 10 in the right half of Fig. 6). These asymmetries were also apparent when we measured lens-induced aniseikonia under both filter arrangements for three of these observers (Fig. C2 in Appendix C of the Supplementary Material), suggesting a combination of color effects and aniseikonia. Although these asymmetries cannot cause the underestimation reported in the literature,[11–13,16–19] they indicate that the test protocol should include both arrangements of color filters.

## Development and Validation of Clinical Tests for Aniseikonia

Our results have implications for the development, administration, and validation of aniseikonia tests of the NAT type (i.e., those based on direct comparisons of the perceived size of haploscopically presented stimuli), and also on the use of color in such tests. These are summarized and commented on next.

Firstly, short presentations and viewing under central fixation are advisable to ensure that observers truly perform a size discrimination task. Short presentations preclude major changes in the center of expansion during the short time in which the stimulus is perceptually available, although the center of expansion may nevertheless change across trials if the observer does not maintain a steady head position. These changes across trials introduce noise in the measurements by hampering size comparisons on trials in which the two stimuli are perceived to be vertically displaced in opposite directions. It is important to stress that these problems were entirely absent in the OE and the SE, not only because those instruments did not assess size perception but because the instruments themselves ensured a fixed center of expansion. Preferably, the AI and analogous tests should be administered with a phoropter or a similar device that can be rigidly placed and aligned with the stimulus plane to ensure a fixed and appropriately located center of expansion just as the OE and SE did. In these conditions, the free viewing mode could be safely used, as observers' scanning strategies would be inconsequential.

Secondly, stimulus shape should be such that binocular fusion of the two haploscopically presented stimuli is prevented. The semicircles originally used in the NAT or in Version 1 of the AI are thus more appropriate than the rectangles used in versions 2 and 3 of the AI.

Thirdly, a test using color stimuli for haploscopic presentations should clearly differentiate aniseikonia from size–color illusions. Although such illusion was not very strong in our sample and it was not apparent for all observers, it will be mistaken for aniseikonia when each color is seen with one eye. Adding a full binocular condition to assess the magnitude of such illusion should be part of the protocol.

Fourthly, size–color illusions or other aspects of color perception may produce different measures of aniseikonia according to which color is presented to each eye. Therefore, the test protocol should include separate measurements with color filters in the two possible arrangements before the eyes.

Fifthly, decisional biases produce lateral shifts of the psychometric functions according to which positions test and standard stimuli occupy on the screen. For each arrangement, these shifts have opposite directions away from the PSE (the measure

of aniseikonia) and they can only be differentiated from true aniseikonia if the protocol includes separate measurements and analyses under both position arrangements.

All of these recommendations are supported by our results and render a test protocol that includes the conditions and analyses underlying the results plotted in Figures 5 and 6 under the short presentation mode (or with free viewing if a fixed center of expansion is ensured). Note that some of these aspects of the test protocol are motivated by the use of color stimuli for haploscopic presentations and the need to differentiate aniseikonia from other color-related issues. A feasible alternative would involve haploscopic presentation using a polarized three-dimensional display system (or liquid crystal shutters in a temporal haploscopic system) so that each stimulus is again seen with only one eye. Both stimuli could thus be presented in black on a gray background, rendering a test that is free of color vision. Research on the feasibility and use of aniseikonia tests based on those principles is still needed, but they would definitely eliminate from the test protocol the need for a full binocular condition and, perhaps, the need for a swap of polarizers between the eyes. Decisional biases should still need to be checked for by measuring aniseikonia under each position arrangement for standard and test stimuli, as decisional bias affects all psychophysical measurements.

All tests must also be validated and the validation of aniseikonia tests requires a proof that lens-induced aniseikonia is adequately estimated. Because placing size lenses on a trial frame brings about contamination from changes in the center of expansion across trials with the same size lens and across conditions involving different size lenses (even under the short presentation mode), using a method that ensures a fixed center of expansion is highly advisable. This might be easily achieved by using a phoropter instead of a trial frame. Phoropters already incorporate color filters, polarized lenses, and all that is needed to fix each observer's prescription, although placing size lenses on them may require a custom-made attachment.

## Management of Aniseikonia

We have shown that the location of the center of expansion has important consequences on the accuracy with which aniseikonia can be measured. This observation, in turn, suggests some perceptual consequences when natural viewing conditions are involved, that is, when both eyes see the same object.

When the cause of aniseikonia is in or on the eye, changes of fixation in natural viewing conditions also relocate the center of expansion so that the different perceived sizes manifest mostly in the peripheral visual field while the foveated area on the stimulus plane always stays aligned in both eyes across eye movements. In contrast, when aniseikonia is caused by external lenses (e.g., on glasses), changes of fixation by simply moving the eyes behind the lenses have distinct consequences: because the center of expansion remains at a location on the stimulus plane that is only determined by the position of the external lens, moving the eyes brings to the foveae images of different sizes and in disparate positions, requiring fusional efforts. Our limited ability to accomplish this especially in the vertical direction may be responsible for the visual discomfort and symptoms that are often reported (mainly upon reading) by patients whose aniseikonia is caused by spectacle lenses. Remole[25] seemed to refer to these two situations when he pointed out the difference between static and dynamic aniseikonia, showing also that dynamic aniseikonia produces more visual discomfort than static aniseikonia, particularly in the vertical direction.

Such distinction does not apply under the conditions of our measurements, where the images seen with each eye do not need to (and cannot) be fused into a single percept. Nevertheless, our observations suggest a simple way to treat dynamic aniseikonia: prompt patients to tilt their head (and not just to move their eyes) as needed when they change fixation. The fact that some aniseikonic patients do not report symptoms or visual discomfort may reflect that they learned to do so by themselves. Moving the head to accompany changes in fixation is a strategy that users of progressive addition spectacle lenses implement to look through the part of the lens that is appropriate for the distance; using a similar strategy should not be difficult for prepresbiopic aniseikonic patients, removing entirely the extra discomfort produced by dynamic aniseikonia. Of course, this strategy may be unfeasible for presbyopic users of bifocal or multifocal lenses. With bifocals, reading with anisometropia results in dynamic aniseikonia or near-point prismatic effect and requires prism slab off to assist fusion. Other options that may make head tilt useful for managing aniseikonia include the recourse to separate reading glasses, clip-on frames with a full lens near add in the magnetic or swing-away clip, or deformable adjustable spectacle lenses.[62] Various designs of adjustable lenses have been recently commercialized.[63] An empirical study on the efficacy of these

management options is worth carrying out, but such study is beyond the scope of this paper.

## Acknowledgments

The authors thank Azma Rehman for assistance with data collection, and Jae-Hyun Jung for measurements of the magnification of size lenses. Parts of the computations were carried out on EOLO, the MECD- and MICINN-funded HPC of Climate Change at Moncloa Campus of International Excellence, Universidad Complutense.

Supported by Grant PSI2012-32903 from Ministerio de Economía y Competitividad (MAGP) and by National Institutes of Health Grants R01EY05957 and R01EY023385 (EP).

Disclosure: **M.A. García-Pérez,** None; **E. Peli,** None

## References

1. Ogle KN. Relative sizes of ocular images of the two eyes in asymmetric convergence. *Arch Ophthalmol*. 1939;22:1046–1067.
2. Peli E. Optometric and perceptual issues with head-mounted displays. In: Mouroulis P, ed. *Visual Instrumentation: Optical Design and Engineering Principles*. New York, NY: McGraw-Hill; 1999:205–276.
3. Bradley A, Rabin J, Freeman RD. Nonoptical determinants of aniseikonia. *Invest Ophthalmol Vis Sci*. 1983;24:507–512.
4. Berens C. A study of 288 patients examined with the ophthalmo-eikonometer. *Br J Ophthalmol*. 1937;21:132–142.
5. Berens C, Loutfallah M. Aniseikonia: a study of 836 patients examined with the ophthalmo-eikonometer. *Trans Am Ophthalmol Soc*. 1938; 36:234–267.
6. Ames A. The space eikonometer test for aniseikonia. *Am J Ophthalmol*. 1945;28:248–262.
7. Brecher GA. A new method for measuring aniseikonia. *Am J Ophthalmol*. 1951;34:1016–1021.
8. Brecher GA, Winters DM, Townsend CA. Image alternation for aniseikonia determination. *Am J Ophthalmol*. 1958;45:253–258.
9. Awaya S. *New Aniseikonia Tests*. Tokio: Handaya Co., Ltd.; 1957.
10. Awaya S, Sugawara M, Horibe F, Torii F. The "New Aniseikonia Tests" and its clinical application [In Japanese]. *Nippon Ganka Gakkai Zasshi*. 1982;86:217–222.
11. McCormack G, Peli E, Stone P. Differences in tests of aniseikonia. *Invest Ophthalmol Vis Sci*. 1992;33:2063–2067.
12. Yoshida M, Sato M, Awaya S. Evaluation of the clinical usefulness of the New Aniseikonia Tests [In Japanese]. *Nippon Ganka Gakkai Zasshi*. 1997; 101:718–722.
13. Antona B, Barra F, Barrio A, Gonzalez E, Sánchez I. The validity and repeatability of the New Aniseikonia Test. *Optom Vis Sci*. 2006;83: 903–909.
14. de Wit GC. Evaluation of a new direct-comparison aniseikonia test. *Binocul Vis Strabismus Q*. 2003;18:87–94.
15. Ugarte M, Williamson TH. Aniseikonia associated with epiretinal membranes. *Br J Ophthalmol*. 2005;89:1576–1580.
16. Rutstein RP, Corliss DA, Fullard RJ. Comparison of aniseikonia as measured by the aniseikonia inspector and the space eikonometer. *Optom Vis Sci*. 2006;83:836–842.
17. Antona B, Barra F, Barrio A, Gonzalez E, Sanchez I. Validity and repeatability of a new test for aniseikonia. *Invest Ophthalmol Vis Sci*. 2007;48:58–62.
18. Fullard RJ, Rutstein RP, Corliss DA. The evaluation of two new computer-based tests for measurement of aniseikonia. *Optom Vis Sci*. 2007;84:1093–1100.
19. Weise KK, Marsh-Tootle W, Corliss D. Evaluation of computer-based testing for aniseikonia in children. *Optom Vis Sci*. 2010;87:883–889.
20. Kehler LAF, Fraine L, Lu P. Evaluation of the Aniseikonia Inspector Version 3 in school-aged children. *Optom Vis Sci*. 2014;91:528–532.
21. de Wit GC. Comparison of aniseikonia as measured by the Aniseikonia Inspector and the Space Eikonometer. *Optom Vis Sci*. 2007;84: E535–E536.
22. de Wit GC. Validity and repeatability of a new test for aniseikonia: discussion. *Invest Ophthalmol Vis Sci*. 2007;48:eLetter 3620.
23. Currie D. Partial correction of irregular aniseikonia secondary to retinal traction. *Optom Vis Sci*. 2012;89:1081–1086.

24. Estes WK. The problem of inference from curves based on group data. *Psychol Bull*. 1956;53:134–140.
25. Remole A. Dynamic versus static aniseikonia. *Aust J Optom*. 1984;67:108–113.
26. Stahl JS. Amplitude of human head movements associated with horizontal saccades. *Exp Brain Res*. 1999;126:41–54.
27. Ogle KN. Theory of the space-eikonometer. *J Opt Soc Am*. 1946;36:20–32.
28. Oyama T, Nanri R. The effects of hue and brightness on the size perception. *Jpn Psychol Res*. 1960;2:13–20.
29. Tedford WH, Bergquist SL, Flynn WE. The size-color illusion. *J Gen Psychol*. 1977;97:145–149.
30. Cleveland WS, McGill R. A color-caused optical illusion on a statistical graph. *Am Stat*. 1983;37:101–105.
31. Nakano M, Tanabe S, Mori Y, Ikegami B, Fujita I. Expansive and contractive size perception with color patches. *J Vis*. 2005;5(8):1027.
32. Yoo HS, Smith-Jackson TL. Colour size illusion on liquid crystal displays and design guidelines for bioinformatics tools. *Behav Inform Tech*. 2011;30:775–785.
33. Edwards AS. Effect of color on visual depth perception. *J Gen Psychol*. 1955;52:331–333.
34. Egusa H. Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual field. *Perception*. 1983;12:167–175.
35. Vos JJ. Some new aspects of color stereoscopy. *J Opt Soc Am*. 1960;50:785–790.
36. Burbeck CA. Criterion-free pattern and flicker thresholds. *J Opt Soc Am*. 1981;71:1343–1350.
37. Heckmann T, Schor CM. Panum's fusional area estimated with a criterion-free technique. *Percept Psychophys*. 1989;45:297–306.
38. Alcalá-Quintana R, García-Pérez MA. A model for the time-order error in contrast discrimination. *Q J Exp Psychol*. 2011;64:1221–1248.
39. García-Pérez MA, Alcalá-Quintana R. The difference model with guessing explains interval bias in two-alternative forced-choice detection procedures. *J Sens Stud*. 2010;25:876–898.
40. García-Pérez MA, Alcalá-Quintana R. Interval bias in 2AFC detection tasks: sorting out the artifacts. *Atten Percept Psycho*. 2011;73:2332–2352.
41. García-Pérez MA, Alcalá-Quintana R. Shifts of the psychometric function: distinguishing bias from perceptual effects. *Q J Exp Psychol*. 2013;66:319–337.
42. García-Pérez MA, Peli E. The bisection point across variants of the task. *Atten Percept Psycho*. 2014;76:1671–1697.
43. Optical Diagnostics. *Aniseikonia Inspector. Version 3. User's Manual*. Available at: http://www.opticaldiagnostics.com/products/ai/ai_manual.pdf. Accessed 19 September 2014.
44. Mulligan JB. Optimizing stereo separation in color television anaglyphs. *Perception*. 1986;15:27–36.
45. Wang G-J. Principles and designs of the iseikonic lenses for near vision. *Optom Vis Sci*. 2006;83:E930–E937.
46. García-Pérez MA. Adaptive psychophysical methods for nonmonotonic psychometric functions. *Atten Percept Psycho*. 2014;76:621–641.
47. García-Pérez MA, Alcalá-Quintana R. On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model. *Psychon B Rev*. 2012;19:820–846.
48. García-Pérez MA. Does time ever fly or slow down? The difficult interpretation of psychophysical data on time perception. *Front Human Neurosci*. 2014;8:415.
49. Sridharan D, Steinmetz NA, Moore T, Knudsen EI. Distinguishing bias from sensitivity effects in multialternative detection tasks. *J Vis*. 2014;14(9);16.
50. Seizova-Caji T. Size perception by vision and kinesthesia. *Percept Psychophys*. 1998;60:705–718.
51. Rule SJ. Subject differences in exponents of psychophysical power functions. *Percept Motor Skill*. 1966;23:1125–1126.
52. Rule SJ. Subject difference in exponents from circle size, numerousness, and line length. *Psychon Sci*. 1969;15:284–285.
53. Armstrong L, Marks LE. Differential effects of stimulus context on perceived length: implications for the horizontal–vertical illusion. *Percept Psychophys*. 1997;59:1200–1213.
54. Numerical Algorithms Group. *NAG Fortran Library Manual, Mark 19*. Oxford; 1999.
55. García-Pérez MA, Alcalá-Quintana R. Sampling plans for fitting the psychometric function. *Span J Psychol*. 2005;8:256–289.
56. Alcalá-Quintana R, García-Pérez MA. A comparison of fixed-step-size and Bayesian staircases for sensory threshold estimation. *Spatial Vis*. 2007;20:197–218.
57. García-Pérez MA, Alcalá-Quintana R. Bayesian adaptive estimation of arbitrary points on a

psychometric function. *Brit J Math Stat Psy*. 2007;60:147–174.
58. García-Pérez MA, Alcalá-Quintana R. Empirical performance of optimal Bayesian adaptive estimation. *Span J Psychol*. 2009;12:3–11.
59. Kershaw CD. Statistical properties of staircase estimates from two interval forced choice experiments. *Brit J Math Stat Psy*. 1985;38:35–43.
60. Leek MR, Hanna TE, Marshall L. Estimation of psychometric functions from adaptive tracking procedures. *Percept Psychophys*. 1992;51:247–256.
61. Maloney LT. Confidence intervals for the parameters of psychometric functions. *Percept Psychophys*. 1990;47:127–134.
62. Peli E, Rabinovich J, Barnea D. A new deformable adjustable lens for presbyopia. *Optom Vis Sci*. 1992;69(Suppl.);115.
63. *Variable Focus Eyeglasses*. Available at: http://www.allaboutvision.com/lenses/variable-focus.htm. Accessed March 25, 2015.

# Aniseikonia tests: The role of viewing mode, response bias, and size–color illusions
# Electronic Supplementary Material

Miguel A. García-Pérez
Departamento de Metodología, Facultad de Psicología, Universidad Complutense,
Campus de Somosaguas, 28223 Madrid, Spain

Eli Peli
The Schepens Eye Research Institute, Massachusetts Eye and Ear,
Department of Ophthalmology, Harvard Medical School, Boston, MA 02114

## Appendix A: Model psychometric functions

This appendix derives the mathematical expression of the psychometric functions in Eqs. 3 and 5 and lists the variants that hold under alternative conditions. Table A1 lists how the mean $\mu_d$ and the standard deviation $\sigma_d$ of $D$ vary with test size $x$ according to which semicircle was on each side of the screen and which eye was each semicircle seen with. As seen in the table, for any given arrangement of color filters before the eyes, interchanging the positions of the semicircles on the screen leaves $\sigma_d^2$ unchanged but reverses the sign of $\mu_d$.

Equations 3 in the main paper apply under the conditions in the top half of Table A1. For the conditions in the bottom half, where the red filter is placed before the left eye and the green filter is placed before the right eye, Eqs. 3 become

Table A1. Mean and variance of $D$ as a function of the color of the filter placed before each eye and the position of the color semicircles on the screen. We assume that the red semicircle is the standard and the green semicircle is the test. The fixed size of the standard is denoted $x_s$ and the variable size of the test is denoted $x$.

| Filter | Position of each color semicircle (color on left ; color on right) | mean, $\mu_d$ | variance, $\sigma_d^2$ |
|---|---|---|---|
| OS: green<br>OD: red | green ; red | $\mu_{OS}(x_s) - \mu_{OD}(x)$ | $\sigma^2_{OD}(x) + \sigma^2_{OS}(x_s)$ |
| | red ; green | $\mu_{OD}(x) - \mu_{OS}(x_s)$ | |
| OS: red<br>OD: green | green ; red | $\mu_{OD}(x_s) - \mu_{OS}(x)$ | $\sigma^2_{OS}(x) + \sigma^2_{OD}(x_s)$ |
| | red ; green | $\mu_{OS}(x) - \mu_{OD}(x_s)$ | |

$$\Psi_L(x) = \Phi\left( \frac{\delta_1 - \left(\mu_{OD}(x_s) - \mu_{OS}(x)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} \right), \quad \text{(A1a)}$$

$$\Psi_R(x) = \Phi\left( \frac{\delta_1 - \left(\mu_{OS}(x) - \mu_{OD}(x_s)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} \right), \quad \text{(A1b)}$$

$$\Upsilon_L(x) = \Phi\left( \frac{\delta_2 - \left(\mu_{OD}(x_s) - \mu_{OS}(x)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} \right) - \Phi\left( \frac{\delta_1 - \left(\mu_{OD}(x_s) - \mu_{OS}(x)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} \right), \quad \text{(A1c)}$$

$$\Upsilon_R(x) = \Phi\left( \frac{\delta_2 - \left(\mu_{OS}(x) - \mu_{OD}(x_s)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} \right) - \Phi\left( \frac{\delta_1 - \left(\mu_{OS}(x) - \mu_{OD}(x_s)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} \right), \quad \text{(A1d)}$$

and Eqs. 5 become

$$\frac{\delta_k - \left(\mu_{OS}(x) - \mu_{OD}(x_s)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} = \delta_k^* + \alpha^* - \beta^* x, \quad \text{(A2a)}$$

$$\frac{\delta_k - \left(\mu_{OD}(x_s) - \mu_{OS}(x)\right)}{\sqrt{\sigma_{OS}^2(x) + \sigma_{OD}^2(x_s)}} = \delta_k^* - \alpha^* + \beta^* x, \quad \text{(A2b)}$$

with $\alpha^* = (\alpha_{OD} - \alpha_{OS} + \beta_{OD} x_s)/\sqrt{\kappa_{OD} + \kappa_{OS}}$, $\beta^* = \beta_{OS}/\sqrt{\kappa_{OD} + \kappa_{OS}}$, and $\delta_k^* = \delta_k/\sqrt{\kappa_{OD} + \kappa_{OS}}$.

In the full binocular condition, the psychophysical function embodies joint influences from both eyes but it may vary with stimulus color due to the size–color illusion. When the standard stimulus is red, Eqs. 3 become

$$\Psi_L(x) = \Phi\left( \frac{\delta_1 - \left(\mu_{red}(x_s) - \mu_{green}(x)\right)}{\sqrt{\sigma_{green}^2(x) + \sigma_{red}^2(x_s)}} \right), \quad \text{(A3a)}$$

$$\Psi_R(x) = \Phi\left( \frac{\delta_1 - \left(\mu_{green}(x) - \mu_{red}(x_s)\right)}{\sqrt{\sigma_{green}^2(x) + \sigma_{red}^2(x_s)}} \right), \quad \text{(A3b)}$$

$$\Upsilon_L(x) = \Phi\left( \frac{\delta_2 - \left(\mu_{red}(x_s) - \mu_{green}(x)\right)}{\sqrt{\sigma_{green}^2(x) + \sigma_{red}^2(x_s)}} \right) - \Phi\left( \frac{\delta_1 - \left(\mu_{red}(x_s) - \mu_{green}(x)\right)}{\sqrt{\sigma_{green}^2(x) + \sigma_{red}^2(x_s)}} \right), \quad \text{(A3c)}$$

$$\Upsilon_R(x) = \Phi\left( \frac{\delta_2 - \left(\mu_{green}(x) - \mu_{red}(x_s)\right)}{\sqrt{\sigma_{green}^2(x) + \sigma_{red}^2(x_s)}} \right) - \Phi\left( \frac{\delta_1 - \left(\mu_{green}(x) - \mu_{red}(x_s)\right)}{\sqrt{\sigma_{green}^2(x) + \sigma_{red}^2(x_s)}} \right), \quad \text{(A3d)}$$

and Eqs. 5 become

$$\frac{\delta_k - \left(\mu_{green}(x) - \mu_{red}(x_s)\right)}{\sqrt{\sigma^2_{green}(x) + \sigma^2_{red}(x_s)}} = \delta_k^* + \alpha^* - \beta^* x, \quad \text{(A4a)}$$

$$\frac{\delta_k - \left(\mu_{red}(x_s) - \mu_{green}(x)\right)}{\sqrt{\sigma^2_{green}(x) + \sigma^2_{red}(x_s)}} = \delta_k^* - \alpha^* + \beta^* x, \quad \text{(A4b)}$$

with $\alpha^* = (\alpha_{red} - \alpha_{green} + \beta_{red} x_s) / \sqrt{\kappa_{red} + \kappa_{green}}$, $\beta^* = \beta_{green} / \sqrt{\kappa_{red} + \kappa_{green}}$, and $\delta_k^* = \delta_k / \sqrt{\kappa_{red} + \kappa_{green}}$. Subscripts "red" and "green" are swapped if the standard stimulus is green instead.

## Appendix B: Approximate non-rejection regions for measured PSEs

A *P*% non-rejection region is the range of values for measured PSEs that can be obtained via some psychophysical procedure on *P*% of the occasions when the true PSE is at some specified location. Bootstrap simulation methods were used to obtain central 95% non-rejection regions around the theoretical PSEs in each of the nine experimental conditions under each viewing mode. Because the precise location of the true PSE only contributes an anchor point, without loss of generality simulations assumed $\mu_{OD} = \mu_{OS}$ and $\sigma_{OD} = \sigma_{OS}$ so that the true PSE is at the standard level. Simulations also assumed $\alpha_i = \gamma_i = 0$ in Eqs. 1, as $\alpha_i$ only sets an arbitrary and inconsequential anchor in the subjective continuum (see García-Pérez, 2014) whereas $\gamma_i$ produces negligible effects when the range of stimulus levels is relatively narrow. This leaves $\beta_{OS} = \beta_{OD} = \beta$ and $\kappa_{OS} = \kappa_{OD} = \kappa$ as the only relevant sensory parameters. Parameter values needed to generate the data ($\beta$, $\kappa$, $\delta_1$, and $\delta_2$) varied within ranges that were consistent with the spread of the psychometric functions observed in our empirical data and with the patterns of usage of the "can't tell" response option also observed in our empirical data. Regarding the spread of psychometric functions, three different scenarios were considered: broad, narrow, and assorted. In the "broad-spread" scenario, $\beta$ was uniformly distributed on [0.6, 0.8] and $\kappa$ was uniformly distributed on [1.2, 1.4]; in the "narrow-spread" scenario, $\beta$ was uniformly distributed on [1.2, 1.4] and $\kappa$ was uniformly distributed on [0.6, 0.8]; in the "assorted" scenario, $\beta$ was uniformly distributed on [0.6, 1.4] and $\kappa$ was uniformly distributed on [0.6, 1.4]. Regarding patterns of usage of the "can't tell" response option, decision boundaries across replicates were drawn so that their midpoint $(\delta_1 + \delta_2)/2$ was

uniformly distributed on $[-2.5\sqrt{2\kappa},\ 2.5\sqrt{2\kappa}]$ and their span $\delta_2 - \delta_1$ was uniformly distributed on $[0,\ 5\sqrt{2\kappa}]$, using the value for $\kappa$ drawn for the corresponding replicate.

In each scenario, data were generated for 10,000 replicates undergoing the psychophysical procedure used in the free viewing mode (amounting to 192 trials per replicate, as described in Method) and, separately, the psychophysical procedure used in the short presentation mode (amounting to 384 trials per replicate, as described also in Method). Data from each replicate in each case were used to obtain parameter estimates as was done with empirical data from real observers and the measured PSE was analogously obtained. The outcome measure for these analyses is the relative error of estimation, defined as the difference between true PSE and measured PSE scaled according to the estimated spread of the psychometric functions. For each replicate, parameter estimates were used to compute the relative error of estimation, which is thus $(\alpha^*/\beta^* - x_s)\beta^* = \alpha^* - \beta^* x_s$. The 2.5-th and the 97.5-th percentiles of the distribution of relative errors of estimation across replicates were then determined, which define the 95% non-rejection region around the true PSE.

Figure B1 shows the distribution of relative errors of estimation in each scenario (rows) under each psychophysical procedure (columns), with the boundaries of the non-rejection region indicated by the

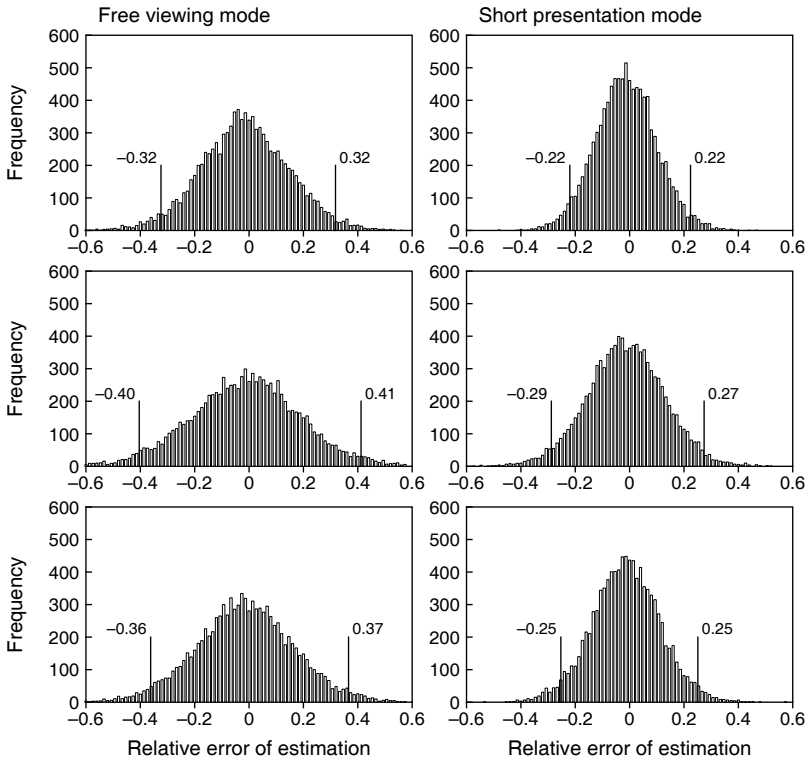


**Fig. B1**. Distributions of relative errors of estimation of the PSE with the staircase design used to collect data in the free viewing mode (left column) or in the short presentation mode (right column) when the psychometric functions are broad (top row), narrow (center row) or with assorted spreads (bottom row). The vertical lines delimit the central 95% non-rejection region around the expected PSE, with boundaries given by the numerals next to them.

short vertical lines and the numerals next to them. The width of these regions varies across viewing conditions (compare left and right panels in each row), a natural consequence of the fact that the short presentation mode gathered data from a larger number of trials and, thus, allowed more accurate parameter estimates. Because of the normalization with respect to the spread of the estimated psychometric functions, the width of the distributions is relatively unaffected by large differences in the range of true spreads of the underlying psychometric functions in each scenario. We will use 95% non-rejection regions from the "assorted" scenario (bottom row in Fig. B1) to judge the significance of deviations between measured and expected PSEs in our empirical study. Thus, in the free viewing mode, measured PSEs (estimated as $x_{PSE} = \alpha^*/\beta^*$, as described in Method) are regarded as significantly different from the expected PSE at $x_{exp}$ if they lie outside the interval $[x_{exp} - 0.36/\beta^*, x_{exp} + 0.37/\beta^*]$; similarly, in the short presentation mode, measured PSEs are regarded as significantly different from expected PSEs if they lie outside the interval $[x_{exp} - 0.25/\beta^*, x_{exp} + 0.25/\beta^*]$. Individual non-rejection regions are computed using the estimate of $\beta^*$ for each observer in each condition and they are shown as gray areas around the theoretical PSE in plots of data and fitted psychometric functions (Figs. 5–7).

## Appendix C: Additional conditions

This appendix presents and discusses the results obtained in two additional conditions that test speculations about the origin of some side issues that arose from our main experiments.

The first issue is that, in the free viewing mode, patterns of strong underestimation of lens-induced aniseikonia at negative magnifications broke at nominal −4% for some observers (see the results for observers #1, #2, and #6 in the left panel of Fig. 8 in the paper). This is presumably caused by the relatively small retinal size involved in this study: Even when alignment at the foveated extreme might be apparent, the shorter-than-standard length of the test at the other extreme is also very conspicuous in near peripheral vision and observers may have relied more on peripheral information on size than on foveal information on alignment. An analogous situation occurs in reverse at large positive magnifications, but the effect would be weaker because now the test has a larger-than-standard size and its endpoint at the other extreme falls farther into the periphery. The tenability of this surmise was checked out using a larger configuration (721 pixels for the standard diameter) so as to match the

retinal size of the (rectangular) standard stimulus in Fullard et al.'s (2007) study with the AI Version 2. With an angular subtense of 15.93 deg, the poorer (farther) peripheral information may not counter the evidence of alignment gathered foveally. The inconclusive results are shown in Fig. C1. Underestimation at large negative magnifications was indeed stronger for two observers (#1 and #2). This suggests that judgments of size based on alignment information gathered foveally and undisturbed by contradictory but distant peripheral evidence are relatively immune to the effects of size lenses, resulting in a stronger underestimation of lens-induced aniseikonia. For observer #6, on the other hand, the use of larger stimuli did not produce any discernible effect. This could result from a failure of the larger stimuli to reach sufficiently peripheral regions for this observer, but also from a failure to comply with the instructions to judge size by checking for alignment at the top and at the bottom, as these data were collected after the observer had completed all of the short presentation conditions with instructions to maintain central fixation.

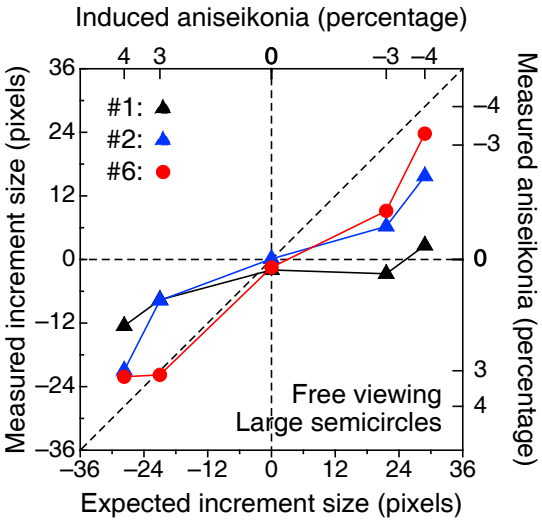


**Fig. C1**. Relation between expected and measured aniseikonia under the free viewing mode with large stimuli.

The second issue of concern is the relatively strong asymmetry displayed by observers #3, #6, and #9 in the short presentation mode under the two haploscopic conditions without size lenses but color filters swapped between the eyes (right half of Fig. 6 in the paper). With the green and red filters respectively placed before the left and right eyes, data showed what appears to be a mild natural aniseikonia for observers #3 and #6 and a relatively stronger natural aniseikonia in the opposite direction for observer #9 (see their data in the right panel of Fig. 8 in the paper). Recall from Fig. 5 in the paper that, in the short presentation mode, observers #3 and #9 did not show any traces of a size–color illusion and that observer #6 only showed weak traces of it. Then, if this were actually natural aniseikonia, the shift of the data paths for these observers with respect to the diagonal should turn into a shift by the same amount but in the opposite direction when color filters are swapped

between the eyes. This symmetric shift was not observed in the only condition (without size lenses) in which color filters had been swapped between the eyes (see the results in this condition for observers #3, #6, and #9 in the right half of Fig. 6 in the paper). Additional data were thus collected for these observers with color filters swapped between the eyes and size lenses placed before the left eye instead. The results are shown in Fig. C2. Data now lie virtually on the diagonal for observers #3 and #6 and far below the diagonal for observer #10. The asymmetry persists, attesting to unidentified aspects of color perception contaminating aniseikonia measurements.

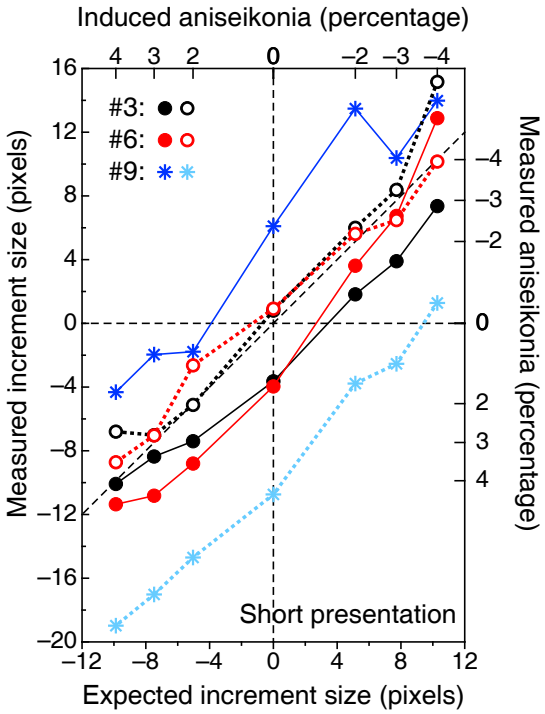


**Fig. C2**. Relation between expected and measured aniseikonia with each possible placement of color filters before the eyes under the short presentation mode. For ease of comparison, solid symbols and continuous lines replot data from the right panel of Fig. 8 in the paper for the observers of concern. Open symbols and dotted lines plot data for the same observers when color filters were swapped between eyes (red before the left eye and green before the right eye).

## References

Fullard, R. J., Rutstein, R. P., & Corliss, D. A. (2007). The evaluation of two new computer-based tests for measurement of aniseikonia. *Optometry and Vision Science*, *84*, 1093–1100. doi:10.1097/OPX.0b013e31815b9e4c

García-Pérez, M. A. (2014). Does time ever fly or slow down? The difficult interpretation of psychophysical data on time perception. *Frontiers in Human Neuroscience*, 8:415. doi:10.3389/fnhum.2014.00415