



Published in final edited form as:

*Ophthalmology*. 2009 January ; 116(1): 145–153. doi:10.1016/j.ophtha.2008.08.031.

## Fixation Preference and Visual Acuity Testing in a Population-based Cohort of Preschool Children with Amblyopia Risk Factors

Susan A. Cotter, OD, MS<sup>1</sup>, Kristina Tarczy-Hornoch, MD, DPhil<sup>1</sup>, Erin Song, OD<sup>1</sup>, Jesse Lin, MS<sup>2</sup>, Mark Borchert, MD<sup>1,2</sup>, Stanley P. Azen, PhD<sup>1</sup>, and Rohit Varma, MD, MPH<sup>1,2</sup> on behalf of the Multi-Ethnic Pediatric Eye Disease Study Group

<sup>1</sup>Doheny Eye Institute and the Department of Ophthalmology, Keck School of Medicine, University of Southern California, Los Angeles, California

<sup>2</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California

### Abstract

**Purpose**—To compare the clinical assessment of fixation preference (FP) to visual acuity (VA) in a population-based sample of preschool children with amblyopia risk factors.

**Design**—Evaluation of diagnostic test in a population-based study.

**Participants**—Two hundred forty-three children with amblyopia and/or strabismus, aged 30–72 months, living in Los Angeles County, California.

**Methods**—Before measuring VA, FP testing was performed at near and usually without correction, using the binocular fixation pattern in children with strabismus >10 diopters (D), or the induced tropia test for children with strabismus ≤10D, or without strabismus. We determined the sensitivity and specificity of FP testing for predicting unilateral amblyopia, defined by optotype VA, among children with amblyopia risk factors.

**Main Outcome Measure**—Grade of FP.

**Results**—Sensitivity of FP testing for amblyopia among children with anisometropia was 20% (9/44) and specificity was 94% (102/109). Among strabismic children, sensitivity was 69% (9/13; worse in children 30–47 than 48–72 months old) and specificity was 79% (70/89), with similar findings for esotropia and exotropia.

**Conclusion**—The ability of FP testing to correctly identify amblyopia in preschool children with amblyopia risk factors is poor. Clinicians should be wary of using FP as a surrogate measure of intraocular difference in VA in young children.

---

Amblyopia is the most common cause of unilateral visual impairment in children with a prevalence of approximately 2%.<sup>1</sup> Pediatric eye care providers strive for early detection and treatment in the hopes of improved outcomes. Valid monocular visual acuity (VA) measures

---

Correspondence: Rohit Varma, MD, MPH, Doheny Eye Institute, Department of Ophthalmology; 1450 San Pablo St. Rm. 4900, Los Angeles, CA 90033-9224. E-mail: E-mail: rvarma@usc.edu.

Presented in part at: Association for Research in Vision and Ophthalmology, Fort Lauderdale, Florida, May 2007.

Financial Disclosures: The authors have no proprietary or commercial interest in any materials discussed in this article.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

are necessary for the definitive diagnosis and treatment of amblyopia. In clinical practice, however, there are few VA tests that can be rapidly and easily administered in a routine clinical setting to infants and children too young to undergo subjective VA testing. Thus, many eye care providers rely on the clinical assessment of fixation behavior to assess monocular visual function in children too young to cooperate with subjective VA testing.

A rapid clinical procedure that requires minimal equipment, fixation preference (FP) testing is a qualitative measure of monocular visual function during binocular viewing, and the most widely used technique for measuring relative VA in preverbal children. Described as a “mainstay” in the detection of unilateral amblyopia in the preverbal child,<sup>2</sup> fixation patterns are routinely used as the basis for initiating and monitoring amblyopia treatment in preverbal children.

The objective of FP testing, a technique originally described by Knapp and Moore,<sup>3</sup> is to compare the fixation of 1 eye relative to the other to determine whether the child demonstrates a preference for fixating with 1 eye or the other, or whether the child will look equally with either eye. The assumption is that a child with amblyopia, when given the choice, preferentially fixates with his or her sound eye rather than amblyopic eye. Conversely, it is assumed that a child without amblyopia is likely to show no preference in fixation, or spontaneous alternation of fixation. When a FP exists, the preference is graded in a semiquantitative way, thereby providing a presumptive measure of the depth of amblyopia.

Test administration differs depending on whether or not the child has a manifest strabismus. The binocular fixation pattern (BFP) measures the degree with which the habitually fixating eye is preferred for fixation under binocular conditions in those with manifest deviations by determining the length of time the nonpreferred eye can hold fixation. This is done simply by covering the child’s habitually fixating eye to force the deviating eye to fixate. After determining that fixation with the formerly deviating eye is stable, the cover is removed and the length of time required for the habitually fixating eye to resume fixation is observed and graded.

Because it has been shown that children with small-angle strabismus often demonstrate a strong FP on BFP testing in the absence of amblyopia,<sup>4</sup> FP is assessed in these patients using the 10-diopter fixation or induced-tropia test (ITT).<sup>5-7</sup> For this procedure, a 10- to 12-D BD loose prism is introduced before 1 eye to induce vertical diplopia. Once image misalignment has been induced, the patient’s fixation pattern is evaluated similarly to BFP testing in that the ability of each eye to hold fixation under binocular conditions is assessed. This procedure is also used to evaluate FP in patients without strabismus.

Although there are variations in the ways to perform FP testing and numerous permutations of grading scales having been proposed,<sup>2,4-11</sup> all have similar interpretations in that the failure to hold fixation for a certain length of time (variously but similarly described as holding several seconds, or through a smooth pursuit movement or a blink) is suggestive of amblyopia. Regardless of the exact technique and grading scale used, however, FP testing is the most widely used method for assessing monocular visual function in preverbal children despite evidence that both false-positive errors of diagnosis (beyond those from using BFP for microstrabismus) and false-negative errors can result.<sup>2,6-10,12</sup>

The objective of the study reported herein was to compare the clinical assessment of FP to amblyopia status as defined by optotype acuity in a population-based cohort of preschool children with strabismus and/or anisometropia, to determine whether FP can be used as a surrogate measure of interocular difference (IOD) in VA for the diagnosis of unilateral amblyopia in children with amblyopia risk factors.

## Methods

This study was supported through a cooperative agreement with the National Eye Institute of the National Institutes of Health and was conducted in Los Angeles County, California. The protocol and informed consent forms were reviewed and approved by the Institutional Review Board/Ethics Committee of the Los Angeles County University of Southern California Medical Center and a parent or guardian (referred to subsequently as “parent”) of each study participant gave written informed consent. Study oversight was provided by an independent data monitoring and oversight committee.

### Study Cohort

The study population consisted of children, aged 30–72 months, with strabismus and/or anisometropia, living within 44 census tracts in the city of Inglewood and adjacent communities in Los Angeles County, California, who were participants in the Multi-ethnic Pediatric Eye Disease Study (MEPEDS). Details of the study design and sampling plan have been described previously.<sup>13</sup> In brief, after a door-to-door census of all dwelling units within the targeted census tracts, parents of all eligible children 5–70 months of age at the time of the interview were provided with an explanation of the study and invited to participate. Written informed consent was obtained from those who agreed to participate and an appointment was then scheduled for a comprehensive eye examination by a MEPEDS pediatric eye care professional at the local MEPEDS examination center.

### Clinic Ocular Examination

The clinic visit included a structured interview and a comprehensive eye examination performed by MEPEDS optometrists or ophthalmologists who specialize in pediatric eye care and who were trained and certified using standardized protocols.<sup>13</sup> Details of the interview can be found elsewhere.<sup>13</sup> All age-appropriate components of clinical testing were attempted on all children, without exclusions based on developmental delay or disability.

Ocular alignment was tested by unilateral cover (cover–uncover) testing at 6 m (cartoon video) and 40 cm (small colored sticker or picture), without and with correction (if worn). Strabismus was measured by the simultaneous prism cover test and an alternate cover test of standard duration with prism neutralization. Hirschberg and Krimsky testing at near were used, respectively, when the unilateral or simultaneous cover and prism tests could not be performed (Krimsky not used if <10 prism diopters [D]).

Fixation preference testing was performed on all participants (protocol described below) with the child’s existing refractive correction (if any) before measuring presenting distance VA to reduce examiner bias and after alignment testing to determine the appropriate test. Monocular distance VA was measured using single-surround HOTV optotypes on the electronic VA tester<sup>14</sup> according to the Amblyopia Treatment Study protocol,<sup>15</sup> using naming or matching of letters. The specific MEPEDS VA protocol has been described in detail.<sup>16</sup> Initial VA was tested with correction, if worn. If decreased in either eye (<20/50 for children aged 30–47 months or <20/40 for children aged ≥48 months), or if there was a 2-line IOD with VA of ≤20/32 in the worse-seeing eye, acuity was retested through full correction after cycloplegic refraction. Taking the best of all test results for each eye, children who still had subnormal VA in either eye, or a 2-line IOD with VA of ≤20/32 in the worse eye and an amblyopia risk factor (see below), were scheduled for return visit retesting within 2 weeks. Participants who missed their VA retest visit were rescheduled no more than 3 months after the initial visit. Children requiring a new spectacle correction were referred to their eye care provider or family physician to obtain a prescription, and rarely acquired new spectacles before the return visit VA retest. At the return visit, VA was retested without cycloplegia and with spherocylindrical correction

placed in a trial frame to confirm that poor corrected VA was not an artifact of testing after dilation (FP testing was not retested). The correction for the noncycloplegic retest acuity was full cycloplegic spectacle correction of myopia and astigmatism; hyperopia was not corrected if  $<1.50$  D, and if  $\geq 1.50$  D the hyperopia was undercorrected by 1.50 D to allow for incomplete relaxation of accommodation. Children unable to perform VA testing at the original visit were scheduled for a return visit and sent home with HOTV letters for practice. The final, “best-measured” VA for each eye was the best of all test results recorded for that eye.

Cycloplegic refraction was performed with the Retinomax Autorefractor (Right Manufacturing, Virginia Beach, VA)  $\geq 30$  minutes after cycloplegia with 2 drops of 1% cyclopentolate administered 5 minutes apart. Cycloplegic retinoscopy was performed if Retinomax readings with confidence ratings of  $\geq 8$  were not obtained in both eyes after 3 attempts. If parents refused cyclopentolate eye drops, noncycloplegic retinoscopy was performed. Anterior segment and fundus examinations followed refraction.

### Fixation Preference Testing Protocol

The MEPEDS FP testing protocol was performed by pediatric eye care providers who were trained and certified using the standardized protocol shown in Figure 1 and described below. Examiners were recertified every 6 months. Younger children were seated comfortably on the parent’s lap and older children in an examination chair. If a child wore spectacles, they were worn for testing. After ocular alignment testing, examiners performed FP testing by presenting a small captivating near target approximately 40 cm from the child in primary gaze. The protocol for FP testing (Fig 1) differed depending on whether or not the child had strabismus  $>10$ D in magnitude by simultaneous prism cover testing.

The BFP was evaluated for a child with strabismus  $>10$  D. First, the child’s attention was directed to the near target and he or she was encouraged to attend to it. The examiner observed which eye was strabismic and which eye was fixating, noting if the child spontaneously alternated between the 2 eyes or if 1 eye seemed to be preferred for fixation. If alternation between eyes was seen, the examiner estimated the proportion of time each eye was used for fixation; if it seemed that alternation between the 2 eyes was equal, testing was stopped. If 1 eye seemed to be preferred for fixation, the examiner occluded that eye with a cover paddle or the palm of his or her hand, thereby forcing the nonpreferred eye to fixate on the target. Once satisfied that fixation was stable, the examiner then assessed the quality of fixation of the nonpreferred eye by removing the occluder (thus allowing the child to view binocularly) and observing how well and how long the nonpreferred eye was able to maintain fixation before the preferred eye refixated. Fixation preference was graded as A, B, C, or D according to how well the eye could hold fixation according to the criteria in Table 1.

For a child without strabismus or with a small-angle ( $\leq 10$ D) strabismus, the ITT was used. Children with intermittent strabismus measuring  $>10$ D but who were fusing at the time of FP testing were also measured with the ITT. A 12D base-down (BD) loose prism was placed in front of the child’s right eye (creating vertical diplopia) and the child’s eyes were observed. If the child spontaneously alternated fixation between the 2 eyes (which required observing at least 3 re-fixation movements after introduction of the prism), the procedure was repeated by placing the prism over the right eye again to confirm that the initial alternation response was repeatable; if it was consistent, grade A was recorded and testing was complete. If spontaneous alternation was not a consistent response and the child preferred 1 eye, the procedure was continued.

If in response to the BD prism placed over the right eye, both eyes moved upward toward the apex of the prism and stayed upward (apparent right eye preference), or moved up and then down (return to left eye fixation) without further alternation (apparent left eye preference), the

prism was then placed in front of the left eye. If this switch of the prism caused the apparent FP to reverse, grade A was recorded and testing was complete. If FP did not reverse with the switch of prism, then the examiner determined the child's ability to hold fixation with the non-preferred eye. After placing the prism in front of the preferred eye, the examiner occluded the preferred eye to force fixation to the non-preferred eye. Once satisfied that fixation with the non-preferred eye was stable, the examiner removed the cover paddle from the preferred eye and graded the fixation behavior of the non-preferred eye using the B, C, D grading scale. A FP grade was not assigned if it could not be determined.

On the other hand, if the child's eyes did not move upward in response to the initial placement of BD prism over the child's right eye, the target was moved laterally to confirm, by observation of tracking, that the child was showing an apparent left eye preference and was not simply inattentive. If this was the case, then the examiner placed the BD prism in front of the left eye and observed refixational movements. If this switch of the prism caused the apparent FP to reverse (to right eye preference), a grade A was recorded and testing was complete. If apparent left eye FP was maintained, then the examiner tested the child's ability to hold fixation with the non-preferred eye, as described above.

In cases of ambiguous responses, repeat observations were attempted until a clear response or predominant pattern was apparent. FP was classified into 4 grades (A, B, C, D) or unable as indicated in Table 1. If grade B, C, or D were indicated, right or left eye preference was noted. Grades A and B were considered to be cases "without strong FP" and grades C and D were considered to be cases of "strong FP."

### Definitions of Amblyopia and Strabismus

Unilateral strabismic and/or anisometropic amblyopia was defined as a 2-line IOD in best-measured VA with  $\leq 20/32$  in the worse eye, and  $\geq 1$  of the following unilateral amblyopia risk factors: strabismus on examination, a history of strabismus surgery, or amblyogenic anisometropia consistent with the eye with worse VA. Strabismus was defined as constant or intermittent tropia of any magnitude at distance and/or near fixation. Amblyogenic anisometropia was defined as  $\geq 1.00$  D SE anisohyperopia,  $\geq 3.00$  D SE anisomyopia, or  $\geq 1.50$  D anisoastigmatism.

### Data Analysis

Only data from the MEPEDS children who successfully completed the comprehensive eye examination with both FP testing and monocular VA testing and who had amblyopia risk factors of strabismus and/or anisometropia are included in this report. For different groups of children with specified amblyopia risk factors, we calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of FP testing, using the optotype VA-based diagnosis of amblyopia as the gold standard. For a defined subgroup, sensitivity is the ratio of true-positive FP test results to diagnoses of amblyopia, and specificity is the ratio of true-negative FP test results to cases not diagnosed with amblyopia. The PPV is the ratio of true-positive FP test results to all positive FP test results, and NPV is the ratio of true-negative FP test results to all negative FP test results. We also explored whether changing the definition of "strong FP" to either (1) grades B, C, and D or (2) only grade D to denote a strong FP would improve the sensitivity or specificity of FP testing.

*P*-values for comparisons of proportional measures such as sensitivity and specificity between groups were calculated using the Fisher exact test.

## Results

The overall study participation rate in MEPEDS for the targeted census tracts was 77%.<sup>1</sup> Of the 3430 children 30–72 months of age who underwent a comprehensive eye examination including successful completion of FP testing and monocular VA determination, 243 had either amblyogenic anisometropia (henceforth referred to as simply anisometropia;  $n = 153$ ) or strabismus ( $n = 102$ ; 12 children had both conditions) and were used for the analyses reported herein. Of these children, 46% were African American, 51% were Hispanic, and 3% were of another ethnicity; 57% were female. The age distribution was 5%, 24%, 37%, and 34% for children 30–35, 36–47, 48–59, and 60–72 months, respectively. Fourteen percent (34/243) of the children wore spectacles.

The FP grade versus magnitude of IOD in VA for the 54 strabismic and/or anisometropic children with unilateral amblyopia based on a  $\geq 2$ -line IOD in VA are shown in Table 2. Seventeen (31%) of these amblyopic children demonstrated grade C or D FP. This represents the overall sensitivity of FP testing for amblyopia among children with strabismus and/or anisometropia. Overall sensitivity did not differ according to race/ethnicity (African-American vs Hispanic;  $P = 0.53$ ) or gender ( $P = 0.14$ ; data not shown). Fewer amblyopic children aged 30–47 months (13%, 2/16) than aged 48–72 months (39%, 15/38) demonstrated a grade C or D preference; however, the difference was not statistically significant ( $P = 0.06$ ). There was no significant difference in the proportion of children with a grade C or D preference between amblyopic children with an IOD in VA of 2 lines, and those with  $\geq 3$  lines ( $P = 0.24$ ), or between children with 2 lines versus  $\geq 4$  lines IOD in VA ( $P = 0.16$ ). The lack of statistical difference between those with a 2-line difference in IOD in VA and those with  $\geq 3$  lines or  $\geq 4$  lines remained for the 40 children who had isolated anisometropic amblyopia ( $P = 0.43$  and  $P = 0.32$ , respectively; data not shown).

### Sensitivity and Specificity of Fixation Preference Testing in Detecting Amblyopia as a Function of Type of Risk Factor

Table 3 compares FP testing results to the diagnosis of amblyopia in 153 children considered at risk because of anisometropia (in isolation or in combination with strabismus). The overall sensitivity was low at 20% (9/44); sensitivity was 8% for the 30- to 47-month-old children ( $n = 47$ ) and 26% for the 48 to 72-month-old children ( $n = 106$ ;  $P = 0.24$ ). Conversely, specificity was high overall at 94% (102/109); it was 97% in children 30–47 and 92% in those 48–72 months of age ( $P = 0.43$ ).

Also found in Table 3 are the results of FP testing as compared with the diagnosis of amblyopia in 102 children who were considered at risk for amblyopia because they had strabismus (in isolation or in combination with anisometropia). For this group, the sensitivity of FP testing for detecting strabismic amblyopia was 69% (9/13), which was significantly higher than for anisometropic amblyopia ( $P = .002$ ). The specificity of detecting strabismic amblyopia was 79% (70/89). It was significantly higher (100%) in the 48- to 72-month-old children ( $n = 75$ ) than in those 30–47 months of age ( $n = 27$ ; 20%;  $P = 0.007$ ). Specificity was 75% and 91% in the older and younger age groups, respectively, and the difference was not significant ( $P = 0.14$ ). Overall, the specificity of FP testing was significantly lower in children with strabismus than in those with anisometropia ( $P = 0.003$ ).

Lowering the FP threshold to include grade B along with C and D as an indicator of strong FP increased the sensitivity for amblyopia from 20% to 30% in children with anisometropia, but did not alter sensitivity (69%) in children with strabismus. Specificity, meanwhile, decreased slightly (from 94% to 93% in anisometropic children, and from 79% to 72% in strabismic children). Alternatively, using grade D exclusively to indicate strong FP, decreased sensitivity from 20% to 14% and from 69% to 54% for children with anisometropia and strabismus,

respectively. Conversely, specificity increased from 94% to 97% for anisometropia and from 79% to 89% for strabismus.

### **Sensitivity and Specificity of Fixation Preference Testing in Detecting Strabismic Amblyopia as a Function of Strabismus Type**

To evaluate whether the direction of horizontal strabismus influenced FP testing results, we compared the ability of FP testing to correctly identify amblyopia in children with esotropia and exotropia (Table 3), and found no difference according to type of strabismus. Sensitivity was 70% for esotropia and 67% for exotropia ( $P = 1.0$ ), and specificity was 72% for esotropia and 84% for exotropia ( $P = 0.19$ ). A strong FP in the absence of amblyopia was observed in 15% of exotropic children and 22% of esotropic children ( $P = 0.37$ ). Among the esotropic children, sensitivity was higher (100%) in the 47- to 72-month-old children than in children 30–46 months old (25%;  $P = 0.03$ ); specificity was 69% and 86% in the older and younger age groups, respectively ( $P = 0.64$ ). Among the children with exotropia, only 3 of whom had amblyopia, both of the older children had a strong FP and the younger child did not, but the difference was not statistically significant ( $P = 0.33$ ); specificity was 81% and 93% in older and younger children, respectively ( $P = 0.41$ ).

To evaluate whether false-positive errors occurred predominantly among children with constant strabismus or intermittent strabismus, we evaluated the results of FP testing separately in these 2 groups (data not shown). False-positive errors were found in both groups, with specificity values of 62% and 85% for constant and intermittent strabismus, respectively, although specificity was better in intermittent strabismus ( $P = 0.02$ ). To evaluate whether sensitivity and specificity of FP testing in strabismic patients were related to the type of FP testing performed, we compared FP grade with amblyopia diagnosis in those with either a small-angle strabismus or no strabismus at near, who were always tested with the ITT ( $n = 68$ ), and those having a constant strabismus of  $\geq 10$  D at near, who were always tested with the BFP method ( $n = 21$ ). (Children with intermittent large-angle strabismus at near might be tested with either method depending on the frequency of manifestation of strabismus, and were therefore excluded from this analysis.) Sensitivity and specificity did not differ between the 2 groups: Sensitivity was 71% for strabismic cases assessed with the ITT method and 80% for cases assessed by BFP ( $P = 1.0$ ), and specificity was 82% for the ITT method and 75% for the BFP method ( $P = 0.49$ ).

### **Predictive Value of Fixation Preference Testing**

For FP testing among the children with strabismus, the PPV of a strong FP for amblyopia was 32%; this was similar for younger and older children. Even among children with esotropia, the PPV was only 41%. The overall NPV was 95%; this dropped to 83% among children <48 months of age. In the children with anisometropia, the PPV and NPV of FP testing were 56% and 74%, respectively, for a diagnosis of amblyopia, with no notable differences between younger and older children.

### **Discussion**

We prospectively compared the clinical assessment of FP to amblyopia status in a population-based cohort of preschool children to determine if FP could be used as a surrogate measure of IOD in VA for the assessment of monocular visual function in young children with amblyogenic risk factors. Fixation preference testing using a standardized protocol was performed before VA testing and the diagnosis of amblyopia was made based on the results of a comprehensive eye examination that included monocular optotype VA testing. Overall, the ability of FP testing to reliably predict the presence of amblyopia was disappointing.

More than two thirds of the children with amblyopia did not show a strong FP. Cases of amblyopia missed by FP testing were especially common among children with anisometropia, and they were as likely to occur in the setting of more severe amblyopia as in milder amblyopia. Thus, in preschool children with amblyopia risk factors, management decisions based on FP testing alone could deprive many amblyopic children of needed treatment. On the other hand, children who showed a strong FP, especially those who were strabismic, usually did not have amblyopia as defined by IOD in VA. Thus, relying on FP testing alone to assess monocular visual function could lead to the erroneous assumption of amblyopia and result in unnecessary patching or pharmacologic therapy.

The sensitivity and specificity of FP testing for diagnosing amblyopia differed depending on the amblyopia risk factor in question. Among children with anisometropia, FP testing had very poor sensitivity but high specificity, regardless of age group. Among strabismic children, sensitivity was higher than for those with anisometropia, although it was still poor in younger children. Specificity, however, was lower in children with strabismus than among those with anisometropia. Because sensitivity and specificity rates were not substantially altered when we used the other possible cutoff points (grade D only or grades B and C and D) to denote a strong FP, our findings seem robust to variations in FP test interpretation.

For strabismus, specificity and sensitivity did not vary according to either the direction (eso versus exo) of the strabismus, or whether the ITT was used for straight-eyed and microtropic children, as opposed to the BPF technique for those with manifest strabismus >10 D. Specificity was slightly higher for intermittent strabismus than constant strabismus, but false-positive errors occurred in both.

In comparing our results with those reported previously, it is essential to recall that our data are population based, whereas all other reports are from clinical populations. Clinic populations are biased samples, likely to underrepresent children without strong FP because these children may not be followed as closely as children suspected of having amblyopia; this could result in an overestimation of sensitivity. Additionally, there are innumerable differences among previous studies with regard to FP technique, FP grading criteria, measurement of VA, definition of amblyopia, characterization of strabismus, inclusion of persons with anisometropia, inclusion of children without amblyopia risk factors, and age of patients.

Zipf<sup>4</sup> was the first to report that that BPF testing in patients with strabismus could be problematic because of false-positive errors; however, he indicated that this was an issue primarily for patients with microtropia. Wright et al<sup>5</sup> then reported that false-positive errors found in patients with microtropia could be eliminated by using the ITT. By contrast, our data show that false-positive errors in the setting of strabismus occur despite taking the precaution of using ITT for small-angle strabismus, and occur in large-angle as well as small-angle strabismus. Our findings substantiate similar observations of false-positive errors in work by Frank<sup>7</sup> and others.<sup>8,9,12</sup>

Laws et al<sup>2</sup> reported the sensitivity and specificity of BFP testing in strabismic children using 3 different cutoffs for FP testing for 53 children 5–13 years of age. However, none of the 3 cutoffs had both good sensitivity and specificity. The grading cutoff that was most similar to ours showed a sensitivity of 78% and specificity of 86% as compared with the 100% sensitivity and 75% specificity found in our older (48–72 months of age) strabismic children. These authors reported the PPV of a strong FP for amblyopia as 89% in their sample. However, unlike the much lower PPV (even for the esotropic subgroup) derived from our population-based data, their value does not reflect the predictive value for a typical child with strabismus because a clinic population with strabismus is biased toward a higher prevalence of amblyopia than the



general population of strabismic children, thus inflating the apparent PPV of a test with suboptimal specificity.

In 111 patients with manifest strabismus and a mean age of 10.4 years, Sener et al<sup>9</sup> reported that the BFP test had a sensitivity and specificity of 85% and 79%, compared with the 100% and 75%, respectively, that we found in our older children. Hakim<sup>8</sup> reported a sensitivity of 67% and a specificity of 23% for a strong preference on BPF testing in 80 children 3–8 years of age with manifest strabismus of unreported magnitudes.

Our findings in the setting of strabismus describe a phenomenon that has been described as “pseudoamblyopia” (a strong FP in the absence of amblyopia).<sup>8</sup> Because FP, unlike monocular optotype acuity, is assessed when both eyes are viewing, a strong preference has been theorized to reflect an exaggerated form of eye dominance in which the brain chooses to fixate with one eye over the other.<sup>8</sup> Another way to view this finding is that the nonpreferred eye has deep suppression during binocular viewing circumstances, even in the absence of a monocular VA deficit. A strong FP has also been reported to occur in patients with formerly treated amblyopic eyes even when there is little residual difference in VA between the eyes.<sup>12,17</sup> Our findings probably cannot be explained in terms of persistence of FP after amblyopia treatment, however, because none of our participants reported having been previously diagnosed with or treated for amblyopia (although some may have effectively received treatment in the form of refractive correction).

In evaluating our false-positive errors among children with strabismus, we must consider the possibility that some children were in fact genuinely amblyopic, and it was the optotype acuity that was “incorrect.” Indeed, young children sometimes do not perform maximally when their nonamblyopic eye is tested, which can mask the actual IOD in VA. Because we observed a trend toward fewer false positives and higher specificity in younger strabismic children than in older ones, which is the opposite of what one would predict if unreliable VA testing in younger children were masking amblyopia, we doubt that this is the explanation for the false-positive test results in strabismic children. Another factor to consider is that if children acquired spectacles between the time of FP testing and the final return visit VA retest visit, amblyopia that was present initially may have resolved and thus not been diagnosed based on best-measured VA. However, none of the children in this cohort reported receiving spectacles between the initial and return visits. Finally, we used the best-recorded VA measure from each eye at the conclusion of the retesting protocol to assess for an IOD in VA; thus, it is conceivable that some children who were truly amblyopic were not diagnosed as such by optotype acuity because of spuriously good performance on 1 of the acuity tests.

Far less has been written about the insensitivity of FP testing for amblyopia in anisometric children than about FP testing in strabismic children. Wright et al<sup>5</sup> reported poor sensitivity of the ITT for amblyopia in a presumably mixed group of children with anisometropia or microtropia (sensitivity 41% for amblyopia with IOD in VA of  $\geq 2$  lines), but claimed that missed cases, at least of more severe amblyopia ( $\geq 3$  lines IOD in VA), could be prevented by modifying the ITT to require that fixation be held with either eye through a smooth pursuit to meet the criterion for equal vision.<sup>6</sup> However, the authors did not address the question of how such a modification would in turn alter the specificity of the test. It is possible that if our protocol had specifically tested the ability to hold fixation with each eye separately in all children who showed spontaneous alternation, we could have increased test sensitivity in our participants with anisometric amblyopia, although probably not without decreasing the specificity of the test.

Another factor that could have contributed to the low sensitivity for anisometric amblyopia is that, at the time of FP testing, examiners were unaware of the presence of anisometropia;

thus, they may have been more likely to expect an alternation response. In contrast, by the nature of BFP testing, examiners had to have already identified a manifest strabismus before using the BFP test to evaluate FP. This knowledge could heighten their suspicion that a preference in fixation might exist, contributing to a higher detection rate and also to a higher false-positive error rate in strabismic participants. Although our protocol was highly standardized, leaving little room for subjectivity, it is difficult to rule out the possibility that the sensitivity for anisometropia would have been higher if examiners had known that anisometropia was present. However, although a higher prior level of suspicion might occur in the clinic setting when a child is already known to have anisometropia, such a bias in favor of seeing a FP would result in more strong FP results for nonamblyopic as well as anisometropic amblyopic patients. In other words, prior suspicion in and of itself can only increase test sensitivity at the cost of lowering specificity.

From a clinician's viewpoint, specificity and sensitivity alone do not answer the question of the practical implications of a test result; that is, what are the chances that a particular test result on FP testing in a particular child reflects true amblyopia or the absence thereof? This is where it is helpful to consider the PPV and NPV of the test. In our population-based sample of preschool children with strabismus, the PPV of 32% and NPV of 95% indicate that whereas a strabismic child without a strong FP probably does not have amblyopia, only one third of strabismic children with a strong FP actually have amblyopia (even among children with esotropia, the PPV is only 41%). This could lead to unnecessary amblyopia therapy for the remaining two thirds of these strabismic children. Conversely, our findings of a PPV of 56% and NPV of 74% for children with anisometropia indicate that if such a child shows a strong FP in the office, he or she is about as likely as not to have amblyopia, and of those not demonstrating a strong FP on testing, approximately 1 in 4 will have undetected amblyopia. Therefore, children with anisometropia cannot be assumed to be amblyopia free until they are old enough to perform monocular optotype VA testing.

It is important to recall that PPV and NPV depend on the prevalence of amblyopia in the strabismic or anisometropic population in question. However, it is unlikely that the probability of amblyopia in the setting of a given risk factor varies widely among population-based samples of children, so that our observations should be applicable to strabismic and anisometropic children in general. Unlike PPV and NPV derived from biased clinic samples, our predictive values for FP testing are thus applicable to the child with newly identified strabismus or anisometropia presenting to a pediatric eyecare provider for evaluation. As population-wide screenings become more prevalent, leading to earlier and wider identification of children with strabismus and refractive error, it is all the more important to understand the validity of FP testing in a population-based sample of children at risk for amblyopia.

One caveat to consider is that our observations of discrepancies between IOD in VA and FP testing in the setting of known amblyopia risk factors were all in children >30 months old. To extrapolate these findings to children <30 months of age, one would have to assume that FP behavior does not change markedly with age. In our study, the sensitivity of FP testing for anisometropic amblyopia tended to be even lower in younger children within the age range we studied; thus, our finding that FP testing is insensitive for anisometropic amblyopia is probably robust with regard to age. In the case of strabismic children, we did observe a slightly (although not significantly) higher specificity of FP testing in younger children, so it is conceivable that false-positive results are less of an issue in the young populations in whom we usually employ FP testing. On the other hand, we expect that any hypothetical gain in specificity in strabismic children <30 months old would likely be offset by a decrease in sensitivity, based on the significant age dependence of FP sensitivity for strabismic amblyopia seen in our study.

A second methodologic factor that must be considered is that FP testing was performed without spectacle correction in most children, because spectacle wear was uncommon in the study population. This means that most children with anisometropia were tested without optical correction of the underlying anisometropia. However, we do not believe that performing FP testing with spectacle correction would have altered our overall conclusions, because testing with correction would be expected to result, if anything, in a lesser number of children demonstrating a strong FP, leading to even worse sensitivity for amblyopia.

We could identify no sources of bias to explain our findings. Data were collected prospectively in a large, population-based cohort of children by pediatric eye care professionals who were specifically trained and certified for the FP testing protocol described herein. Examiners were masked to VA results at the time of FP testing, with monocular VA subsequently measured using a reliable optotype testing protocol that was developed as a primary outcome measure for controlled trials of amblyopia treatment.<sup>14,15</sup> Amblyogenic risk factors were identified based on a comprehensive eye examination that included an objective cycloplegic refraction and eye health assessment. In addition, because of previously reported differences in performance based on eye alignment, the ITT was used for participants with no strabismus and strabismus  $\leq 10$  D and the BFP test for those with manifest deviations  $>10$  D. These data were analyzed separately as were data for participants with strabismus and those with anisometropia. Finally, because our study is population based, the PPV and NPV of FP testing are applicable to children with these risk factors who present to the eye care provider for an evaluation.

Our findings highlight the need for a reliable method of assessing monocular VA in preverbal children that would be practical and appropriate for routine clinical use, particularly as technological advances in automated screening instruments facilitate the early identification of children with strabismus and/or significant degrees of refractive error. In the meantime, there is currently no way to evaluate the true risk of amblyopia associated with various putative amblyopia risk factors present in early childhood, short of prospective, longitudinal evaluations. Such prospective studies may help us to stratify the real risk associated with a given type of strabismus or level of anisometropia, making it possible for eye care providers to use epidemiologic information in conjunction with clinical data to guide the management of common childhood disorders of visual development.

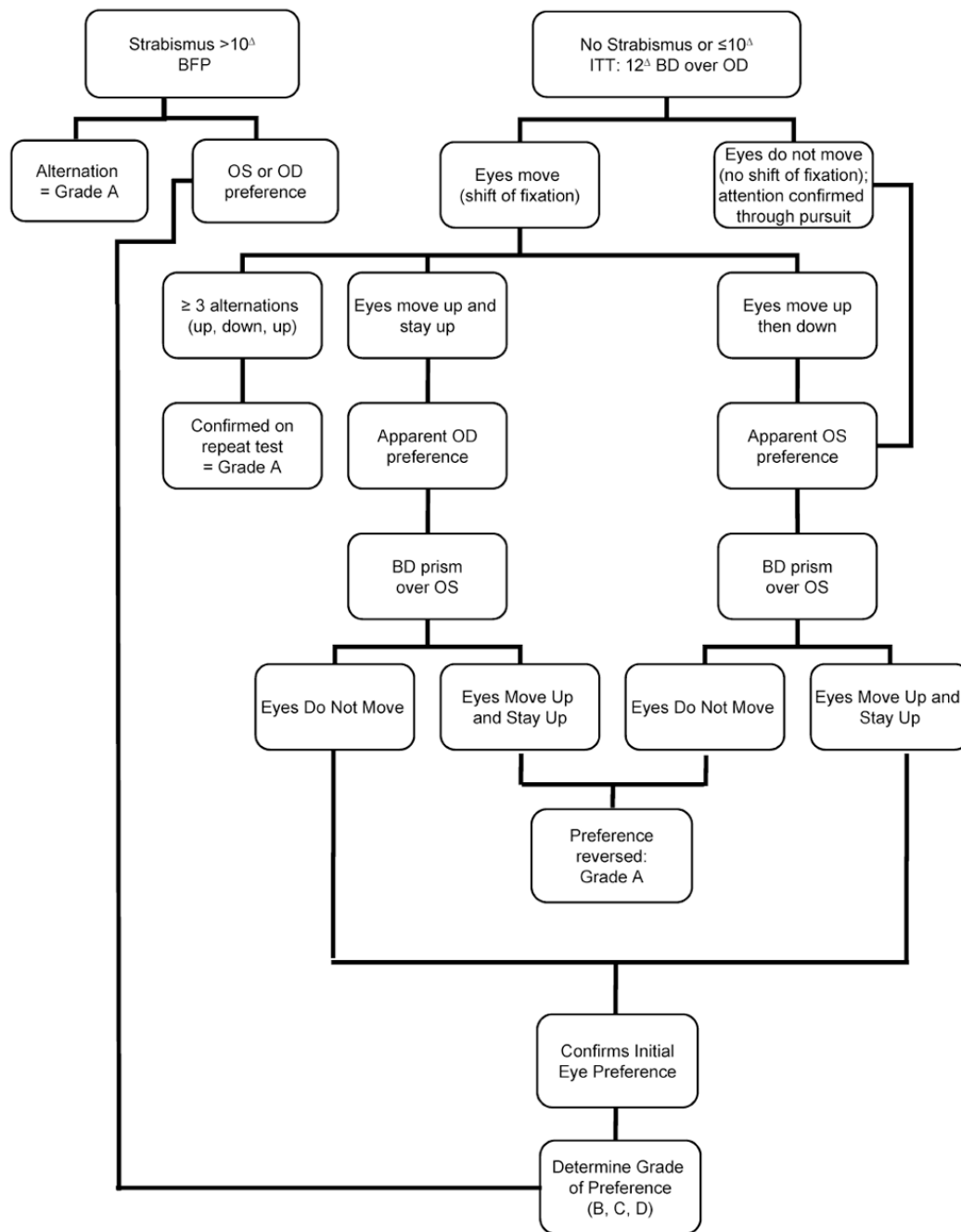
In conclusion, fewer than half of children with strabismus and/or anisometropia who show a strong FP really have amblyopia, whereas a child with anisometropia who does not show a strong FP still has a 1 in 4 chance of having amblyopia. Thus, FP test results alone should not be viewed as a solid basis for deciding whether or not to initiate amblyopia treatment in a preverbal child with strabismus or anisometropia. Providers who initiate or withhold amblyopia treatment on the basis of FP testing alone should be cognizant of the likelihood that they may be unnecessarily prescribing treatments with which compliance is often difficult and frustrating, as well as of the potential for delaying treatment that is genuinely needed. Fixation preference testing should be interpreted in the context of all other available information, such as the characteristics of the amblyopia risk factor in question (e.g., degree and type of anisometropia, frequency, and type of strabismus), and every effort should be made to obtain monocular VA measures as soon as possible.

## Acknowledgments

Supported by the National Eye Institute of National Institutes of Health, Department of Health and Human Services (EY14472 and EY03040) and an unrestricted grant from the Research to Prevent Blindness, New York, New York.

## References

1. Multi-ethnic Pediatric Eye Disease Study Group. Prevalence of amblyopia and strabismus in African-American and Hispanic children ages 6 to 72 months: The Multi-ethnic Pediatric Eye Disease Study. *Ophthalmology* 2008;115:1229–36. [PubMed: 17953989]
2. Laws D, Noonan CP, Ward A, Chandna A. Binocular fixation pattern and visual acuity in children with strabismic amblyopia. *J Pediatr Ophthalmol Strabismus* 2000;37:24–8. [PubMed: 10714692]
3. Knapp P, Moore S. Diagnostic procedures in an orthoptic evaluation. *Am Orthopt J* 1962;12:63–9. [PubMed: 14457192]
4. Zipf RF. Binocular fixation pattern. *Arch Ophthalmol* 1976;94:401–5. [PubMed: 1267923]
5. Wright KW, Walonker F, Edelman P. 10-Diopter fixation test for amblyopia. *Arch Ophthalmol* 1981;99:1242–6. [PubMed: 7259597]
6. Wright KW, Edelman PM, Walonker F, Yiu S. Reliability of fixation preference testing in diagnosing amblyopia. *Arch Ophthalmol* 1986;104:549–53. [PubMed: 3954659]
7. Frank JW. The clinical usefulness of the induced tropia test for amblyopia. *Am Orthopt J* 1983;33:60–9.
8. Hakim OM. Association between fixation preference testing and strabismic pseudoamblyopia. *J Pediatr Ophthalmol Strabismus* 2007;44:174–7. [PubMed: 17542440]
9. Sener EC, Mocan MC, Gedik S, et al. The reliability of grading the fixation preference test for the assessment of interocular visual acuity differences in patients with strabismus. *J AAPOS* 2002;6:191–4. [PubMed: 12075297]
10. O'Reilly C, Smith DR. The 10-diopter base-down fixation test for amblyopia: comparison of techniques. *Can J Ophthalmol* 1984;19:303–5. [PubMed: 6525576]
11. Wallace DK. Tests of fixation preference for amblyopia. *Am Orthopt J* 2005;55:76–81.
12. Fischer N, Brooks SE. The effect of fixation target on fixation preference testing. *Am Orthopt J* 1999;49:105–10.
13. Varma R, Deneen J, Cotter S, et al. Multi-ethnic Pediatric Eye Disease Study Group. The Multi-ethnic Pediatric Eye Disease Study: design and methods. *Ophthalmic Epidemiol* 2006;13:253–62. [PubMed: 16877284]
14. Moke PS, Turpin AH, Beck RW, et al. Computerized method of visual acuity testing: adaptation of the Amblyopia Treatment Study visual acuity testing protocol. *Am J Ophthalmol* 2001;132:903–9. [PubMed: 11730656]
15. Holmes JM, Beck RW, Repka MX, et al. Pediatric Eye Disease Investigator Group. The Amblyopia Treatment Study visual acuity testing protocol. *Arch Ophthalmol* 2001;119:1345–53. [PubMed: 11545641]
16. Cotter SA, Tarczy-Hornoch K, Wang Y, et al. Multi-ethnic Pediatric Eye Disease Study Group. Visual acuity testability in African-American and Hispanic children: the Multi-ethnic Pediatric Eye Disease Study. *Am J Ophthalmol* 2007;144:663–7. [PubMed: 17868633]
17. Cassin B. Alternate fixation in the non-strabismic child. *Am Orthopt J* 1982;32:111–6.



**Figure 1.** Fixation preference testing protocol for the Multi-Ethnic Pediatric Eye Disease Study. BFP = binocular fixation pattern; ITT = induced tropia test; OD = right eye; OS = left eye

**Table 1**  
Grading Scale for Fixation Preference (FP) Testing

FP Grading Criteria	
Grade	Examiner Observations
A	<ol style="list-style-type: none"> <li>1 Spontaneous alternation between the right and the left eyes</li> <li>2 When switching the prism to the fellow eye causes FP to reverse</li> </ol>
B "Holds Well"	Fixation held with nonpreferred eye under any of the following circumstances before refixation to preferred eye: <ul style="list-style-type: none"> <li>• for <math>\geq 3</math> seconds</li> <li>• during a smooth pursuit</li> <li>• through a blink</li> </ul>
C "Holds momentarily"	Fixation held with nonpreferred eye for 1 to $<3$ seconds
D "Does not hold"	Refixation with preferred eye occurs immediately ( $<1$ sec) when the occluder is removed from the preferred eye

**Table 2**  
 Fixation Preference Grade Compared to Interocular Difference (IOD) in Visual Acuity in Amblyopic Children 30–72 Months of Age

Fixation Preference Grade	Lines			Total
	2	3	≥ 4	
Interocular Difference in Visual Acuity, n				
A	20	6	7	33 (61%)
B	3	1	0	4 (7%)
C	3	1	1	5 (9%)
D	4	3	5	12 (22%)
Total	30	11	13	54

Amblyopia defined as  $\geq 2$ -line IOD in best-measured visual acuity, with  $\leq 20/32$  worse in the worse eye. Forty children had isolated anisometropia, 10 had isolated strabismus, 3 had both strabismus and anisometropia, and 1 had anisometropia with a corneal scar.

**Table 3**  
Fixation Preference Testing for Detecting Amblyopia

Fixation Preference Grade	Amblyopia			Sensitivity (Specificity)
	Yes	No	Total	
Amblyogenic anisometropia				
C/D	9	7	16	20% (94%)
A/B	35	102	137	
Total	44	109	153	
Strabismus				
C/D	9	19	28	69% (79%)
A/B	4	70	74	
Total	13	89	102	
Esotropia				
C/D	7	10	17	70% (72%)
A/B	3	26	29	
Total	10	36	46	
Exotropia				
C/D	2	8	10	67% (84%)
A/B	1	43	44	
Total	3	51	54	