Genetic and environmental contributions to strabismus and phoria: Evidence from twins

Jeremy B. Wilmer a,b,*, Benjamin T. Backus b

aWellesley College, Psychology, 106 Central Street, Wellesley, MA 02481, USA
bSUNY State College of Optometry, Vision Sciences, 33 West 42nd Street, New York, NY 10036, USA

Abstract

The causes of manifest (strabismus) and latent (phoria) misalignment of the visual axes are incompletely understood. We calculated genetic and environmental contributions to strabismus based upon a critical review and quantitative meta-analysis of previous strabismus twin studies (n = 3418 twin pairs) and calculated contributions to phoria based upon a new twin study (n = 307 twin pairs). Our results suggest that genetic liability is necessary to develop strabismus, whereas environmental factors are sufficient to cause most phorias. The different etiologies implied by this work suggest that strabismus and phoria should be carefully distinguished in epidemiological work.

1. Introduction

Strabismus, a manifest misalignment of the visual axes, is one of the most common childhood visual disorders, occurring in 3–4% of the population (Robaei, Rose, et al., 2006). The most common cause of amblyopia and an important contributor to childhood visual impairment (Preslan & Novak, 1996), strabismus is also associated with a variety of negative social and employment outcomes (Coats, Paysse, Towler, & Dipboye, 2000; Menon, Saha, Tandon, Mehta, & Khokhar, 2002; Olitsky, Sudesh, & Graziano, 1999; Paysse, Steele, McCreevy, Wilhelmus, & Coats, 2001; Satterfield, Keltner, & Morrison, 1993; Uretmen et al., 2003). Phoria, a latent misalignment of the visual axes, exists to some degree in nearly all individuals. Phorias are normally held in check by the fusion mechanism, which is the binocular system’s drive to fixate the same object with both eyes via vergence eye movements (motor fusion), resulting in sensory fusion. In those with a large near phoria or weak vergence system, the effort to prevent diplopia during sustained near work may cause significant eye strain (Borsting, Rouse, & De Land, 1999; Borsting et al., 2003; Sheedy & Saladin, 1977, 1978).

It has long been known that eye misalignment tends to cluster in families (Jones, 1886), and recent studies have confirmed this fact for both phoria and strabismus (Abrahamsson, Magnusson, & Sjostrand, 1999; Chimonidou, Palimeris, Koliopoulos, & Velissaropoulos, 1977; Francois, 1961; Mash, Hegmann, & Spivey, 1975; Paul & Hardage, 1994; Richter, 1967; Schlossman & Priestley, 1952; von Rotth, 1937). However, either genes or environment can contribute to resemblance among family members; therefore, familial clustering does not clarify etiology. Additional data from twins can discriminate environmental from genetic influences (Falconer & Mackay, 1996; Neale & Cardon, 1992). To quantify the relative contributions of genes and environment to manifest and latent eye misalignment, we therefore embarked upon a critical review and meta-analysis of existing twin strabismus literature, and conducted a new twin study of phoria.

We know of no previous attempts to quantify genetic and environmental contributions to either strabismus or phoria using twins. A large twin strabismus literature has accumulated (Chimonidou et al., 1977; de Decker & Feuerhake, 1978; DeVries & Houtman, 1979; Francois, 1961; Knobloch, Leavenworth, Bouchard, & Eckert, 1985; Kondo, Mori, & Adachi, 1975; Kwapilikova, 1969; Lang, 1990; Matsuo, Hayashi, Fujiwara, Yamane, & Ohtsuki, 2002; Orlebeke & Koole, 1999; Paul & Hardage, 1994; Podgor, Remaley, & Chew, 1996; Reynolds & Wackerhagen, 1986; Richter, 1967; Schlossman & Priestley, 1952; Waardenburg, 1961; Weekers, Moureau, Hacourt, & Andre, 1956; Wei, 1987). However, quantitative estimates of genetic and environmental influence have not yet been derived from these data, nor has this literature been evaluated in light of modern standards for twin research (Neale & Cardon, 1992; Rijndijk & Sham, 2002; Sullivan, Kendler, & Neale, 2003). Our critical review and meta-analysis of the existing twin strabismus literature derives such quantitative esti-
mates while incorporating modern twin methods standards. While one study collected twin phoria data (von Rotth, 1937), it did not report values for individuals and thus did not support quantitative modeling. Our new twin study of phoria supported quantitative modeling.

Using these classic twin study methods, we find evidence that while genetic liability is necessary to develop strabismus, environmental factors alone cause most phorias (see Section 2 for definition of liability). The different etiologies supported by these results suggest that strabismus and phoria differ not only in degree, but also in kind.

2. Methods

2.1. General twin study methods

We identified twin strabismus studies for our critical review and meta-analysis by first conducting inclusive keyword searches of Pubmed, Web of Science, and Google Scholar. We then followed citations forward and backward to identify additional studies and enlisted native speakers to interpret the text of several non-English studies. Our review turned up sixteen primary twin studies (Chimonidou et al., 1977; de Decker & Feuerhake, 1978; DeVries & Houtman, 1979; Francois, 1961; Knobloch et al., 1985; Kondo et al., 1975; Kvpalikova, 1969; Lang, 1990; Matsuo et al., 2002; Orlebeke & Koole, 1999; Podgor et al., 1996; Reynolds & Wackerhagen, 1986; Richter, 1967; Schlossman & Priestley, 1952; Weekers et al., 1956; Wei, 1987), one detailed review of case reports prior to 1961 (Waardenburg, 1961), and one summary of concordance in twin studies prior to Paul and Hardage (1994).

We used the statistical modeling software package Mx (http://www.vcu.edu/mx; Neale, Boker, Xie, & Maes, 2001) to fit standard maximum-likelihood based models of genetic and environmental influence to raw monozygotic (MZ) and dizygotic (DZ) phoria and strabismus data (Neale & Cardon, 1992; Rijsdijk & Sham, 2002; Sullivan et al., 2003). We also used Mx to compute correlations among twins for phoria using the intraclass correlation (ANOVA-based, suitable for continuous data) and for strabismus using the tetrachoric correlation (suitable for dichotomous data; Neale & Cardon, 1992; Rijsdijk & Sham, 2002). Both of these correlations are invariant to arbitrary assignment of twins in a pair to ‘x’ or ‘y’ (Neale & Cardon, 1992; Neale et al., 2001).

2.2. Twin study logic and rationale

Given an environmentally-determined trait, if the crucial environmental effects are those shared between twins – for example, shared prenatal environment, family, school, or societal context – then both MZ and DZ twins should covary perfectly within the precision of measurement. Again, given an environmentally-determined trait, if the crucial environmental effects are not shared (unique) between twins – for example, different teachers or illnesses – then both MZ and DZ twins should have zero covariation (note that unique environmental effects do not contribute to family resemblance). In either case, environmental effects cause covariance between MZ twins to resemble the covariance between DZ twins whereas genetic effects cause MZ twins to covary to a greater extent than DZ twins. Most traits are influenced by some combination of genes and environment.

A simple way to estimate the relative contributions of genes (A, assuming genes have independent additive effects), shared environment (C), and the combination of unique environment and measurement error (E) to variance in a trait of interest is via the following formulas (Falconer & Mackay, 1996; Neale & Cardon, 1992), where $r_{\text{MZ}}$ and $r_{\text{DZ}}$ are the correlation among MZ and DZ twins, respectively:

\[
A = 2(r_{\text{MZ}} - r_{\text{DZ}}) \quad (1)
\]

\[
C = r_{\text{MZ}} - A \quad (2)
\]

\[
E = 1 - r_{\text{MZ}} \quad (3)
\]

(Nota que $A + C + E = 1$).

While these formulas may be intuitive in light of the preceding discussion, they are imprecise in certain situations and do not allow straightforward tests of statistical significance. Therefore, robust maximum-likelihood model-fitting techniques are now standard.

2.3. Assumptions

The twin method relies on several assumptions (Falconer & Mackay, 1996; Neale & Cardon, 1992): (1) MZ and DZ twins share environments to equal degrees (if MZs share their environment to a larger degree than DZ twins, then environmental effects may be attributed to genes). (2) Lack of assortative mating, or lack of tendency for mates to resemble each other genetically on the trait of interest (if assortative mating exists, genetic effects may be attributed to environment due to greater than 50% genetic similarity in fraternal twins). (3) Twins resemble the general population on the trait of interest (if twins differ from the general population, then generalizing the results of twin studies beyond twins requires taking the differences into account). (4) Genes and environment act independently (if a gene–environment interaction is present, it will be attributed to genes). We have no reason to doubt assumption 1 or 2. Assumptions 3 and 4 are addressed in Section 3 (sections “Review of previous strabismus twin data” and “Evidence for strabismus gene–environment interaction”, respectively).

2.4. Dichotomous variables

The standard approach to analyzing twin data for which each individual is classified as either affected or unaffected, as in the case of strabismus, is to apply a liability threshold model (Falconer, 1965; Falconer & Mackay, 1996). Such a model assumes that illness liability is (transformable to) normally distributed in the population, and that manifest illness develops when liability exceeds a certain threshold. The assumption of normally distributed liability is generally considered reasonable except in cases where a single gene or environmental factor exerts a sharply discontinuous effect on liability (Falconer & Mackay, 1996; Neale & Cardon, 1992). Such major effects are rare (Neale & Cardon, 1992) in common disorders such as the concomitant form of strabismus that makes up 95% (Engle, 2007) of strabismus cases, and indeed, current consensus supports a complex, multifactorial etiology for concomitant strabismus (Engle, 2007; Michaelides & Moore, 2004; Parikh et al., 2003). While a small minority of the strabismus cases we reviewed in this paper may have been due to single-gene disorders, these few cases should not have substantially influenced our results. We therefore applied the liability threshold model to our analyses of previous twin studies of strabismus.

The unique step in modeling dichotomous twin data is to infer, given concordance data, overall resemblance between twins. Such resemblance, conceived as correlation in liability between twins, is operationalized as the tetrachoric correlation coefficient, or the correlation needed between two normally distributed variables to produce an observed pattern of concordance and discordance on a dichotomous variable (Neale & Cardon, 1992; Pearson, 1901). In the present case, the underlying correlation among twins in strabismus liability is inferred from observed strabismus concordance and discordance among twin pairs. Tetrachoric correlations...
are then modeled in the same way as covariation on continuous traits. We calculated tetrachoric correlations using Mx (Neale et al., 2001).

Two twin studies of disorders typically recruit only concordant affected and discordant twin pairs, because concordant unaffected pairs (being relatively common) provide less information about correlation in liability. However, if concordant unaffected pairs are not recruited, a separate estimate of prevalence is required to derive the tetrachoric correlation coefficient. If prevalence is overestimated, correlations will be underestimated; if prevalence is underestimated, correlations will be overestimated. Fortunately, the methodologically sound strabismus twin studies we analyzed in depth (Orlebeke & Koole, 1999; Podgor et al., 1996; Richter, 1967) all include concordant unaffected twin pairs, allowing simultaneous estimates of concordance and prevalence in twins.

Estimating genetic and environmental influences on a dichotomous trait requires a systematic, known strategy for ascertaining twins. The two opposite extremes for ascertainment are complete ascertainment and single ascertainment (Rijndijk & Sham, 2002; Sullivan et al., 2003). Under complete ascertainment, twin pairs are recruited from a population without regard to whether they are affected. Conversely, under single ascertainment, probands affected individuals), are recruited from a population without regard to whether they have a twin, and twins of probands are then recruited. Under complete ascertainment, observed pairwise concordance for strabismus [concordant pairs/all pairs] is representative of that in the general population (Neale & Cardon, 1992; Rijndijk & Sham, 2002). However, under single ascertainment, a given concordant pair is twice as likely to be recruited as a given discordant pair; therefore, the ratio of concordant to discordant twin pairs is overestimated by 100% and observed pairwise concordance exceeds that in the general population by a factor of 2/(1 + population concordance) (Rijndijk & Sham, 2002). In the absence of systematic, known ascertainment, accurate estimates of twin concordance are therefore impossible.

2.5. Participants and apparatus for phoria twin study

Twins with normal or corrected to normal vision participated with compensation at the Twin Days Festival in Twinsburg, Ohio during the years 2005−7. Of the 255 monozygotic (MZ) twin pairs, 200 were female and 55 male; of the 52 dizygotic (DZ) twin pairs, 35 were female, 11 male, and 6 opposite-sex. Ages ranged from 18 to 65 years (mean 36.0, SD 14.4). Participants gave written informed consent for this study, which followed the tenets of the Helsinki Declaration and received approval from the Faculty of Arts and Sciences Human Subjects Committee at Harvard University and the Office of Regulatory Affairs at University of Pennsylvania. During a 20-min testing session each participant took a phoria test, completed a questionnaire that assessed zygosity and screened for ocular pathology, and a subset (171 twin pairs) completed the TNO test of stereoeuity. Twin pairs were classified as MZ or DZ based on forced-choice self-report (Heath et al., 2003; Wise, Hansen, Reed, & Breslin, 2007), a simple approach that correctly classified 100% (95% CI 95.7−100%) of 86 twins at the Twin Days Festival in 2003 and 2004 (Wise et al., 2007). We obtained nearly identical twin correlations using three other zygosity questionnaires with known >90% accuracy (Heath et al., 2003; Kasriel & Eaves, 1976; Ooki, Yamada, Asaka, & Hayakawa, 1990).

2.6. Phoria measure

Horizontal nearpoint phoria was measured using a Maddox rod. Note that by ‘phoria,’ we mean here only the magnitude of dissociated deviation of the visual axes, independent of whether this deviation caused symptoms. The Maddox rod test was easy to administer, and because participants adjusted prism power themselves, experimenter-induced bias was minimized. Our testing procedure allowed ample time for slow vergence to decay and included a detailed target for accommodation. In a pilot study, we found that the standard subjective prism-neutralized alternate cover test (Rainey, Schroeder, Goss, & Grosvenor, 1998) produced results that correlated highly with those of the test we used ($r(29) = 0.91$, $p < 0.0001$).

The specific testing procedure was as follows. First, the participant held the Maddox Rod in front of the right eye and fixated a detailed target for accommodation directly under a penlight at 40 cm viewing distance. The Maddox Rod − a red filter combined with high power cylindrical lenses arranged as a grating − smears the view vertically, preventing fusion and making the light appear as a vertical line; if binocular misalignment is present, the line viewed by the right eye and the light viewed by the left eye are offset horizontally. Participants were repeatedly reminded to keep the target in focus. After at least 25 s to allow for asymptotic decay of slow vergence (Schor, 1979), the participant rotated a triangular prism mounted on the Maddox Rod until its horizontal power was such that the point of light and the line were superimposed. Resulting prism power indicated the direction and degree of subjectively experienced binocular misalignment, which corresponds to the phoria in those lacking strabismus and/or anomalous correspondence.

Reliability of the Maddox rod test and other phoria tests is generally high for successive test administrations at short delay (Schroeder, Rainey, Goss, & Grosvenor, 1996). We know of no study of test−retest reliability of a phoria measurement at a delay of weeks to years. We calculated one year test−retest reliability of the Maddox rod test for 91 participants who returned to our research tent in successive years. Phoria was substantially stable over a one year interval ($r(89) = 0.61$, $p < 0.0001$). Note that a difference from year one to year two reflects both genuine change and measurement error.

While the Maddox rod test may fail to produce a measurement in some strabismic patients who suppress, obtaining a successful Maddox rod measurement does not rule out the presence of a strabismus. For this reason, we computed twin correlations excluding participants who either failed to demonstrate stereopsis or reported a history of strabismus (see Section 3). As noted in Section 4, while this analysis should have excluded many strabismic patients, the possibility that some strabismic patients were not excluded does not compromise the conclusions we draw about phoria etiology since we found no evidence for a genetic contribution.

3. Results

3.1. Strabismus

3.1.1. Review of strabismus twin data

Three twin studies of strabismus to date have specified their ascertainment method; each used complete ascertainment with a population-based sample of child twins, aged six (Orlebeke & Koole, 1999), seven (Podgor et al., 1996) and four−seven (Richter, 1967), respectively (see Section 2 for a discussion of the importance of ascertainment methods). We therefore used these three studies to model genetic and environmental influences on strabismus. Concordance and prevalence estimates based on these studies are listed in Table 1. Combined pairwise concordances for MZ and DZ pairs, respectively, are 54% and 14%.

Results from non-systematically ascertainment studies (Chimonidou et al., 1977; de Decker & Feuerhake, 1978; DeVries & Houtman, 1979; Francois, 1961; Knobloch et al., 1985; Kondo...
et al., 1975; Kvaapilikova, 1969; Lang, 1990; Matsuo et al., 2002; Reynolds & Wackerhagen, 1986; Richter, 1967; Schlossman & Priestley, 1952; Weekers et al., 1956; Wei, 1987) are comparable to those from studies with systematic ascertainment (Orlebeke & Koole, 1999; Podgor et al., 1996; Richter, 1967). Combined pairwise concordance for manifest strabismus across all non-systematically ascertained studies is 66% (148/223) for MZ twin pairs and 19% (19/99) for DZ twin pairs (Table 2), with results varying across studies. These figures fall between those obtained from the three studies with known, complete ascertainment (Table 1: 54% and 1%, respectively) and those predicted from those three studies for single ascertainment (70% and 44% respectively; see Section 2 for formula). Results of systematically and non-systematically ascertained studies are therefore consistent, because the concordances obtained for non-systematically ascertained studies fall within the ranges predicted, based on studies with known ascertainment, for varying potential ascertainment scenarios. Concordances in several non-systematically ascertained studies may also have been inflated due to inadequate exclusion of phoria cases (Chimonidou et al., 1977; DeVries & Houtman, 1979; Francois, 1961; Kondo et al., 1975; Kvaapilikova, 1969; Wei, 1987).

In order to estimate twin correlations in liability via the tetrachoric correlation coefficient, strabismus prevalence must be known. If prevalence is overestimated, correlations will be underestimated; if prevalence is underestimated, correlations will be overestimated. Ideally, prevalence and concordance should be estimated from the same data, because then both are based on the same diagnostic procedures in the same population. Fortunately, all three studies with systematic ascertainment allowed such simultaneous estimates of prevalence and concordance (Table 1) (Orlebeke & Koole, 1999; Podgor et al., 1996; Richter, 1967).

Though the twin method standards assume that twins resemble the general population (assumption 3, Section 2), this assumption does not hold for strabismus. Twins have an order of magnitude higher incidence than singletons of low birth weight (<2500 g; 48% vs. 4.8% according to Cohen et al. (1996); 56.6% vs. 6.3% according to Martin et al. (2006)), which is associated with a three- to fourfold heightened risk of strabismus (Bremer et al., 1998; Chew et al., 1994; Robaei et al., 2006; Robaei, Kifley, Cole, & Mitchell, 2006). The presence of this added risk factor coincides with a higher prevalence of strabismus in twins. Specifically, the 6.4% combined prevalence estimate from the three random, population-based samples of twins we reviewed in Table 1 (Orlebeke & Koole, 1999; Podgor et al., 1996; Richter, 1967) (95% CI 5.8–7.0%) was substantially higher (z = 8.655, p < 0.0001) than the 3.9% reported by comparable general population studies (95% CI 3.7–4.2%; based on meta-analysis of 1032 cases among 26366 individuals comparable in age (above 3 years) and racial composition (mostly European descent) to the three studies we review; prevalence in individual studies, summarized in rows 2, 4, 5, 6, 8, 10, 15, and 16 (of 22) in reference 1’s Table 1, ranged from 0.3% to 4.6% (Robaei, Rose, et al., 2006). Assuming greater environmental strabismus risk in twins, twin studies should, if anything, overestimate environmental contribution to strabismus liability. However, as we shall see below (in “Model fit of previous strabismus twin data”), while existing twin studies provided strong evidence for genetic influence on strabismus liability, they provided no evidence for environmental factors that contribute to strabismus liability in the absence of pre-existing genetic liability.

Finally, one study reported strabismus in twins separated shortly after birth and reared in different families (Knobloch et al., 1985). Since this study recruited twins without regard to whether they had strabismus, it was not biased toward recruiting concordant twin pairs. All three MZ twin pairs where at least one twin had strabismus were concordant for esotropia (convergent strabismus), and in every case, this esotropia developed at approximately the same age. Though this finding is consistent with the twin method predictions, it does not necessarily support the twin hypothesis for esotropia, but it does buttress the twin hypothesis for strabismus in general.

### Table 1
Concordance and prevalence of strabismus from published twin studies with systematic ascertainment. MZ = monozygotic; DZ = dizygotic.

<table>
<thead>
<tr>
<th>Study author(s)</th>
<th>Pairwise concordance: MZ twin pairs</th>
<th>Pairwise Concordance: DZ twin pairs</th>
<th>Prevalence: All twin individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Concordant: Total %</td>
<td>Concordant: Total %</td>
<td>Affected: Total %</td>
</tr>
<tr>
<td>Richter (1967)</td>
<td>5: 100</td>
<td>1: 17</td>
<td>17: 234</td>
</tr>
<tr>
<td>Podgor et al. (1996)</td>
<td>4: 44</td>
<td>2: 16</td>
<td>41: 664</td>
</tr>
<tr>
<td>Total</td>
<td>56: 104: 54</td>
<td>16: 115</td>
<td>397: 6192</td>
</tr>
</tbody>
</table>

### Table 2
Concordance data from non-systematically ascertained twin studies of strabismus. Where identifiable, cases of phoria were considered non-affected, and cases of incomitant strabismus as well as higher-order multiple births (triplets, etc.) were excluded. Chimonidou et al. (1977) was not included because its definition of strabismus explicitly included relatively minor phorias and no individual data were given; this study reported one concordant monozygotic (MZ) pair and four concordant dizygotic (DZ) pairs. Matsuo et al. (2002) determined zygosity using chorionicity, a method known to misclassify 30% of MZ twin pairs as DZ (Loos, Derom, Vlietinck, & Derom, 1998). We therefore adjusted DZ concordance from its reported value (8 of 17 pairs concordant) to that expected if 30% of MZ pairs had in fact been misclassified as DZ (4 of 11 pairs concordant), assuming 66% concordance rate for mismatched MZ pairs. Richter (1967) had some systematically ascertained and some non-systematically ascertained twin pairs.

<table>
<thead>
<tr>
<th>Study author(s) and ascertainment type</th>
<th>Pairwise concordance: MZ twin pairs</th>
<th>Pairwise concordance: DZ twin pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schlossman and Priestley (1952)</td>
<td>4: 80</td>
<td>2: 8</td>
</tr>
<tr>
<td>Weekers, Moureau, Hacourt, and Andre (1956)</td>
<td>1: 14</td>
<td>4: 47</td>
</tr>
<tr>
<td>Waardenburg (1961)</td>
<td>58: 76</td>
<td>76: 76</td>
</tr>
<tr>
<td>Francois (1961)</td>
<td>2: 33</td>
<td>6: 33</td>
</tr>
<tr>
<td>Richter (1967)</td>
<td>6: 86</td>
<td>7: 6</td>
</tr>
<tr>
<td>Kvaapilikova (1969)</td>
<td>1: 33</td>
<td>3: 0</td>
</tr>
<tr>
<td>Kondo, Mori, and Adachi (1975)</td>
<td>9: 43</td>
<td>21: 43</td>
</tr>
<tr>
<td>DeVries and Houtman (1979)</td>
<td>8: 47</td>
<td>17: 47</td>
</tr>
<tr>
<td>Reynolds and Wackerhagen (1986)</td>
<td>4: 67</td>
<td>6: 2</td>
</tr>
<tr>
<td>Wei (1987)</td>
<td>13: 76</td>
<td>17: 1</td>
</tr>
<tr>
<td>Lang (1990)</td>
<td>8: 67</td>
<td>12: 4</td>
</tr>
<tr>
<td>Matsuo et al. (2002)</td>
<td>12: 75</td>
<td>16: 4</td>
</tr>
<tr>
<td>Total</td>
<td>148: 664</td>
<td>223: 99</td>
</tr>
</tbody>
</table>
age (preschool, 7–8 years, and 18–19 years, respectively). Since onset of esotropia in the late teenage years is rare, the latter case may not have reflected typical strabismic mechanisms. Nevertheless, the concordance for age of onset in these MZ twin pairs reared in different families demonstrated the importance of genetic factors that—in at least some cases of strabismus—appear to determine a strabismus course relatively immune to environmental differences between families. These cases therefore strengthen the evidence for an important genetic contribution to at least one type of strabismus (that is, esotropia with onset after infancy).

### 3.1.2. Model fit of strabismus twin data

As summarized in Table 1, the three systematically ascertained twin studies of strabismus produced similar concordance and prevalence results. There was no significant difference among the studies in MZ concordance ($\chi^2(2) = 4.7, p = 0.10$), DZ concordance ($\chi^2(2) = 0.07, p = 0.97$), or prevalence ($\chi^2(2) = 0.35, p = 0.64$), though MZ concordance did vary non-trivially among the studies. Therefore, in addition to calculating a combined prevalence estimate (see above), we calculated a combined MZ concordance estimate of 53.8% (95% CI 44.3–63.1) and a combined DZ concordance estimate of 13.8% (95% CI 8.5–21.5).

A strong genetic contribution to strabismus was suggested by the much higher combined concordance rate in MZ than DZ twins (Fisher’s exact test $p < 0.0001$), and the ensuing much higher tetrachoric correlation in MZ ($\rho = 0.92$) than DZ ($\rho = 0.45$) twins (see Table 3; Fisher’s $z = 7.94, p < 0.0001$). As can be seen in Fig. 1a by noting the proximity of the dot labeled ‘strabismus’ to the solid line, these correlations closely matched those expected if additive genetic effects cause all family resemblance in strabismus liability. In fact, the individual results of all three studies that contributed to the combined result matched this expectation (gray dots in Fig. 1a).

Fitting models to systematically ascertained strabismus data produced estimates consistent with family resemblance in strabismus liability being entirely due to additive genetic effects. The results of these model-fitting analyses are shown in Table 3. As the results of two studies (Podgor et al., 1996; Richter, 1967) were not individually large enough to derive stable confidence intervals, we first fit a model to their combined results. For these data, strabismus liability was estimated to result 96% from additive genetic effects (95% CI 36–99%), 0% from shared environment (95% CI 0–58%) and 4% from unique environment (95% CI 1–17%). The estimates derived from the third study (Aurell & Norrissell, 1990; Cantolino & von Noorden, 1969; Crone & Velzeboer, 1956; Ferreira, Oelrich, & Bateman, 2002; Griffin, Asano, Somers, & Anderson, 1979; Podgor et al., 1996; Richter, 1967; Ziakas, Woodruff, Smigh, & Thompson, 2002)]. Assuming the larger correlation in DZ twin data was the tetrachoric correlation coefficient $r(130) = 0.46 (±SE = 0.37–0.54)$. The parallel tetrachoric correlation in first-degree relatives was $r(3525) = 0.29 (±SE = 0.26–0.31) [we calculated this correlation from a strabismus prevalence in the general population of 3.9% (see above; Robaei, Rose, et al., 2006) and a strabismus risk given an affected first-degree relative of 13.4% (95% CI 12.3–14.6%); based on 473 of 3526 affected first-degree relatives of strabismics reported by eight studies: Aurell & Norrissell, 1990; Cantolino & von Noorden, 1969; Crone & Velzeboer, 1956; Ferreira, Oelrich, & Bateman, 2002; Griffin, Asano, Somers, & Anderson, 1979; Podgor et al., 1996; Richter, 1967; Ziakas, Woodruff, Smigh, & Thompson, 2002)]. Assuming the larger correlation in DZ twins than first-degree relatives (Fisher $z = 2.2, p = 0.01$) was not solely due to age-discrepancy-caused discordance among first-degree relatives, these correlations suggest an environmental effect that interacts with pre-existing genetic liability.

### 3.1.3. Evidence for strabismus gene–environment interaction

Twin studies can determine the degree to which genetic liability is necessary—but not the degree to which it is sufficient—to develop strabismus. In other words, twin studies cannot rule out the presence of a gene–environment interaction in which environmental factors exacerbate a genetic liability (assumption 4, Section 2).

Our review of the twin strabismus literature suggested that genetic liability is necessary to develop the common forms of strabismus (i.e. concomitant or non-syndromic), while other research identifies low birth weight, prematurity, maternal smoking, paternal lead exposure, and abnormalities in pregnancy and delivery as strabismus risk factors (Bremer et al., 1998; Chew et al., 1994; Hakim, Stewart, Canner, & Tielsch, 1991; Matsuo, Yamane, & Ohtsuki, 2001; Ponsonby et al., 2007; Robaei, Kifley, et al., 2006; Robaei, Rose, et al., 2006; Taira, Matsuo, Yamane, Hasebe, & Ohtsuki, 2003)]. To the extent that these factors are environmentally caused, a gene–environment interaction is required to account for both necessary genetic liability and environmental influence.

Additionally, since both DZ twins and (non-twin) first-degree relatives share 50% of genes on average, yet first-degree relatives share environment to a lesser degree than DZ twins (especially pre-natally), a greater resemblance among DZ twins than first-degree relatives can support a role for environment. As shown graphically in Fig. 1a, the correlation in liability calculated from extant DZ twin data was the tetrachoric correlation coefficient $r(130) = 0.46 (±SE = 0.37–0.54)$. The parallel tetrachoric correlation in first-degree relatives was $r(3525) = 0.29 (±SE = 0.26–0.31) [we calculated this correlation from a strabismus prevalence in the general population of 3.9% (see above; Robaei, Rose, et al., 2006) and a strabismus risk given an affected first-degree relative of 13.4% (95% CI 12.3–14.6%); based on 473 of 3526 affected first-degree relatives of strabismics reported by eight studies: Aurell & Norrissell, 1990; Cantolino & von Noorden, 1969; Crone & Velzeboer, 1956; Ferreira, Oelrich, & Bateman, 2002; Griffin, Asano, Somers, & Anderson, 1979; Podgor et al., 1996; Richter, 1967; Ziakas, Woodruff, Smigh, & Thompson, 2002)]. Assuming the larger correlation in DZ twins than first-degree relatives (Fisher $z = 2.2, p = 0.01$) was not solely due to age-discrepancy-caused discordance among first-degree relatives, these correlations suggest an environmental effect that interacts with pre-existing genetic liability.

### 3.1.4. Phoria

#### 3.1.4.1. Description of phoria twin data

We measured phorias in 244 monozygotic (MZ) and 51 dizygotic (DZ) twin pairs. Fig. 2 shows scatterplots of phorias for all participants, with each dot representing a pair of twins. MZ twins are shown in Fig. 2a and DZ twins are shown in Fig. 2b.
shown in Fig. 2b. Substantial associations in phoria are evident among both MZ (r(244) = 0.58, p < 0.0001, 95% CI 0.49–0.66) and DZ (r(51) = 0.65, p < 0.0001, 95% CI 0.48–0.80) twins. Ranking of data before computing these correlations changes them little (r ± 0.01), evidence that results are robust to visible bunching due to ceiling and floor effects. Since MZ twins’ phorias are no more similar to each other than DZ twins’, family resemblance in phoria resulted from shared environmental rather than additive genetic influences.

Fig. 1a graphically represents the predicted MZ and DZ associations for the extreme cases where, for a given trait, either shared environment (dotted line, r_{MZ} = \sigma_{sh}) or additive genetic effects (solid line, r_{MZ} = 2\sigma_{ag}) cause all family resemblance (according to the classic twin model defined in Section 2). The correlations for phoria are close to the extreme case where shared environment causes all family resemblance. The point estimate for phoria does not deviate significantly from the line representing 100% environmental contribution.

3.1.5. Model fit of phoria twin data

Fitting models to our phoria data produced estimates consistent with family resemblance in phoria being entirely due to shared and unique environmental effects. The results of these model-fitting analyses are shown in Table 4. We first fit a full ‘ACE’ model to the raw phoria data, including the effects of additive genes (A), shared environment (C), and unique environment (E; this includes measurement error), controlling for gender and age (Table 4, line 1) (Falconer & Mackay, 1996; Neale & Cardon, 1992). There were significant effects of both shared and unique environment (C = 61%, 95% CI 45–67%; E = 39%, 95% CI 33–46%), while the estimated additive genetic effect was zero (A = 0%, 95% CI 0–16%).

To find the most parsimonious model, we fit nested models that lacked either additive genetic or shared environment effects. While dropping the shared environment effect substantially reduced the model’s fit (Table 4, line 2, \chi^2(1) = 17.472, p = 0.001), dropping the additive genetic effect did not reduce the model’s fit (Table 4, line 3, \chi^2(1) = 0, p = 1), again suggesting that phoria is influenced by shared environment but not additive genetic effects. The model including only shared and unique environment gave the most parsimonious fit of all models, as indicated by its low Akaike’s Information Criterion (AIC, Table 4, line 2) (Akaike, 1974). In this best-fitting model, variance in phoria was attributable 61% to shared environment (95% CI 54–67%) and 39% to unique environment (95% CI 33–46%). Our model-fitting analyses therefore suggest that variation in phoria is caused mainly by shared and unique environment.

Fig. 1b shows the percent family resemblance in phoria attributed to additive genetic vs. shared environmental effects, calculated according to the classic twin model defined in Section 2 (Falconer & Mackay, 1996; Neale & Cardon, 1992). The estimated contribution of additive genetic effects to family resemblance in phoria was 0% (95% CI 0–26%; calculated as A/(A+ C)), whereas the estimated contribution of shared environmental effects was 100% (95% CI 74–100%; calculated as C/(A + C)) (Falconer & Mackay, 1996; Neale & Cardon, 1992). This result is opposite to that obtained for strabismus liability (Fig. 1b), evidence that phoria and strabismus have different causes.

3.1.6. Statistical controls for phoria twin data

We performed several control analyses to confirm that our evidence for an environmental contribution to phoria, and against a genetic contribution to phoria, was robust.

Given evidence that being born at low birth weight increases strabismus risk (Bremer et al., 1998; Chew et al., 1994; Robaei, Kifley, et al., 2006; Robaei, Rose, et al., 2006), we obtained self-reported birth weight from a subset of 102 participants to deter-
mine if birth weight predicted phoria. However, birth weight predicted neither raw phoria \((r(102) = 0.10, p = 0.32)\) nor deviation from median phoria \((r(102) = 0.07, p = 0.51)\). We also found no correlation between phoria and age \((r(630) = 0.008, p = 0.84)\), and controlling for age via partial correlation changed neither the MZ \((r_{MZ}(155) = 0.58)\) nor the DZ \((r_{DZ}(52) = 0.66)\) correlation.

Our initial analyses included six opposite-sex DZ twin pairs. It is common practice in twin studies to exclude such pairs because including them could reduce the DZ correlation and lead to an overestimation of genetic influence. Though we found no evidence for a genetic effect, we recomputed the DZ correlation excluding opposite-sex pairs. This correlation \((r_{DZ}(46) = 0.68)\) was similar to that with all individuals included.

Our initial analyses included participants regardless of strabismus or amblyopia history. Given evidence (reviewed above) that genetic liability may be necessary for the development of strabismus, excluding individuals with strabismus might have been expected to decrease a genetic effect had we found one. Though we found no evidence for a genetic effect, we recomputed the MZ and DZ correlations excluding all twin pairs where one or both twins reported a history of strabismus or amblyopia diagnosis. These correlations \((r_{MZ}(211) = 0.62, r_{DZ}(46) = 0.70)\) were similar to those with all individuals included. We recomputed the same correlations again, this time excluding all twin pairs where one or both twins failed the TNO stereoacuity test with a stereoacuity > 480 arcsec (note that the 84 twin pairs who did not complete the TNO test were excluded from this analysis). These correlations too \((r_{MZ}(155) = 0.64, r_{DZ}(33) = 0.60)\) were similar to those with all individuals included.

It is common in clinical practice to consider phoria a dichotomous measure, with phorias larger than a certain critical value being deemed clinically significant, though there is little consensus on how large a phoria – in either the convergent (eso) or divergent (exo) direction – is clinically significant. We conducted several analyses to determine if trichotomizing phoria values into three sequential categories (high eso, normal, and high eso) changed the pattern of MZ and DZ correlations. We took the median phoria (3 prism diptors exophoria) as our reference value, and considered cutoffs for each integral value from one to seven diptors away from this value in both eso and exo directions. In no case was evidence of genetic contribution (i.e. \(r_{MZ} > r_{DZ}\)) obtained.

### 4. Discussion

Quantitative estimates of genetic and environmental contributions to visual dysfunction can guide efforts to identify specific causes, assess their relative importance, and understand their mechanisms of action. We have provided what we believe are the first quantitative estimates of the relative genetic and environmental contributions to both manifest eye misalignment, or strabismus, and latent eye misalignment, or phoria. In a critical review and meta-analysis of previous strabismus twin studies, we found evidence for a strong genetic influence, but no evidence that environmental factors cause strabismus independent of genetic susceptibility. This result suggests that genetic factors are necessary to cause strabismus. In a new twin study of phoria, we found evidence for a strong environmental influence, but no evidence for genetic influence. This result suggests that environmental factors are sufficient to cause most phorias. We believe that future etiological investigations would benefit from distinguishing carefully between strabismus and phoria.

Our phoria study, while excluding stereoblind individuals and those who were aware of having a strabismus, could have included some microstrabismic or intermittent strabismic patients who were unaware of their condition. Importantly, the possibility of some diversity in this sample does not weaken our claims about phoria because the contributions to phoria of any such additional factors appear to have been homogeneous in lacking a genetic basis. There do exist patients whose phorias progress to strabismus later in life and patients with strabismus that results from physical damage to the nerves or eye muscles (von Noorden & Campos, 2002). In the first case, the phoria and strabismus presumably have similar etiology, and in the second, the strabismus was clearly environmental. However, these appear to be exceptions to the general rule. Future studies are needed to determine etiology for increasingly fine distinctions between the different forms of ocular misalignment.

Further work is also needed to determine the functional mechanisms by which genes influence strabismus liability. Previous twin studies have shown a substantial genetic contribution to refractive error (Dirani et al., 2006; Hammond, Snieder, Gilbert, & Spector, 2001). Since uncorrected refractive error is a risk factor for strabismus, it could be that refractive error contributes to genetic strabismus liability. On the other hand, though uncorrected refractive error is also associated with phoria, we failed to detect a genetic contribution to phoria. Perhaps this is because refractive error was corrected in our subjects. Fusional range is also genetically influenced (Kvapilíková, 1969), and we propose that it may contribute to genetic strabismus liability.

The current literature contains an error that should be corrected. The most recent summary of twin strabismus results, by Paul and Hardage (1994), included a misreading of a study by Kvapilíková (1969) that caused a large overestimation of DZ concordance (0/3 concordant DZ pairs and 1/3 concordant MZ pairs reported as 20/34 and 28/34, respectively, due to an error in translation from original Czech). Recent papers on the genetics of strabismus have referenced Paul and Hardage’s figures as important evidence for both genetic (Engle, 2007; Lorenz, 2002; Michaelides & Moore, 2004; Parikh et al., 2003) and environmental (Engle, 2007; Lorenz, 2002; Parikh et al., 2003) contributions to the common forms of strabismus. Our meta-analysis provides clarification, showing that existing twin studies of strabismus support a genetic contribution substantially higher, and an independent environmental contribution substantially lower, than previously recognized.

Evidence that genetic effects contribute importantly to strabismus liability suggests that individual genes may be identified through genetic linkage and association studies. The one strabismus linkage study to date found a candidate genetic locus in one of seven families (Parikh et al., 2003). These results were taken...
as evidence for genetic heterogeneity in strabismus. If strabismus can result from gene–environment interactions, as we have hypothesized, future strabismus linkage studies might improve their statistical power by taking putatively environmental risk factors into account.

Acknowledgments

The authors thank the Twins Days Festival Committee and all the twins who participated; Rachel Kaplan, Jin Lee, Jennifer Liepin, Saad Saadi, Lindsey Sankin, and Kristen Wilmer for assistance with data collection; Zuzana Tothova, Damien Amblard and Xi Haijiang for translating non-English articles; Sarah Medland for providing feedback on methods and analyses; and Ken Ciuffreda, Jeffrey Cooper and Susan Barry for helpful conversations. Author contributions: J.B.W. conceived and designed phoria study in consultation with B.T.B., performed phoria study and strabismus review, and analyzed the data. J.B.W. and B.T.B. wrote the paper. This work was supported by NEI Kirschstein-NRSA postdoctoral fellowship F32 EY16933-02 to JBW and NIH RO1-EY13898 to BTB.

References


Cooper and Susan Barry for helpful conversations. Author contributions: J.B.W. conceived and designed phoria study in consultation with B.T.B., performed phoria study and strabismus review, and analyzed the data. J.B.W. and B.T.B. wrote the paper. This work was supported by NEI Kirschstein-NRSA postdoctoral fellowship F32 EY16933-02 to JBW and NIH RO1-EY13898 to BTB.


Richmond, VA: Virginia Commonwealth University, Department of Psychiatry.


