Fusional stamina: An alternative to Sheard's criterion

Arnulf Myklebust[1], **Patricia Riddell**[2]

[1] Statped, Deptartment of Visual Impairment, Oslo, Norway.
[2] University of Reading, School of Psychology and Clinical Language Sciences, Reading, United Kingdom.

Received April 14, 2016, accepted September 26, 2016.
Keywords: Vision test: Fusional stamina, Sheard's criterion, norms, repeatability, children
Correspondence: arnulf.myklebust@statped.no

Abstract

Sheard postulated that symptoms from heterophoria can be avoided if the fusional reserves in the opposite direction are at least twice the size of the phoria itself; this is known as "Sheard's criterion". Although later studies have provided supporting evidence for this postulate, dichotomous criteria of this type are subject to a lack of sensitivity since small differences in measurements can result in a change in classification. The main purpose of this study was to provide a reliable continuous alternative measure to Sheard's criterion in children; we have called this "Fusional Stamina".

Heterophorias and positive fusional vergence (PFV) were measured in a group of 82 typical children aged 5–11 years. The suggested new variable, fusional stamina, was calculated by dividing PFV by two and then subtracting the (hetero-) phoria. Repeatability at near was checked after one year for a subgroup of 40 children assumed to be at greater risk for binocular vision deficits. Mean fusional stamina was 8.1 PD (Prism Dioptres base out) (SD = ±5.9 PD) for distance and 12.3 PD (SD = ±8.2 PD) for near. Repeatability at near was $r = 0.62, p < 0.001$, compared to $r = 0.38$, $p = 0.02$ for PFV break-value. Five children passed Sheard's criterion on only one of two repeated tests, but showed low fusional stamina on both tests.

Normative values of fusional stamina for a group of typical children are presented. Repeatability for near is high and better than for PFV. Results demonstrate that there is a risk of missing binocular problems based only on passing Sheard's criterion. The continuous alternative of fusional stamina can be useful for researchers as well as clinicians in quantifying binocular vision problems and to monitor the effects of treatment.

Sammendrag

Sheard postulerte at symptomer fra en heterofori kan unngås hvis fusjonsreservene i motsatt retning er minst det dobbelte av foriens størrelse. Dette er kjent som «Sheards kriterium». Selv om nyere studier har gitt støtte for dette postulatet, kan slike dikotome (todelte) kriterier lide under mangel på sensitivitet, siden små forskjeller i målinger kan medføre en endring i klassifisering. Hovedformålet med denne studien var å gi et pålitelig kontinuerlig alternativ til Sheards kriterium for barn. Vi har gitt dette benevnelsen «Fusjonell Stamina».

Heteroforier og positiv fusjonell vergens (PFV) ble målt i en gruppe bestående av 82 representative barn i alderen 5–11 år. Den foreslåtte nye variabelen, fusjonell stamina («motstandskraft»), ble beregnet ved å dele PFV på to og deretter trekke fra (hetero-) forien. Repeterbarhet på nært hold ble undersøkt etter ett år for en undergruppe av 40 barn som var antatt å ha større risiko for binokulære vansker.

Gjennomsnittlig fusjonell stamina var 8,1 PD (prismedioptier basis ut) (SD = ±5,9 PD) for avstand og 12,3 PD (SD = ±8,2 PD) for nær. Repeterbarheten på nær var $r = 0,62, p < 0,001$, sammenlignet med $r = 0,38, p = 0,02$ for PFV «break-verdi». Fem barn innfridde Sheards kriterium på bare en av to gjentatte tester, men viste lav fusjonell stamina på begge. Normative verdier for fusjonell stamina for representative (norske) barn er presentert. Repeterbarheten for nært hold er høy, og bedre enn for PFV. Resultatene demonstrerer at det er en risiko for å overse binokulære vansker basert kun på Sheards kriterium. Det kontinuerlige alternativet, fusjonell stamina, kan være nyttig for både forskere og klinikere for å kvantifisere binokulære synsproblemer og å evaluere effektene av behandling.

Introduction

While there is little dispute that disorders of binocular vision can cause symptoms, what is less clear is which measures can be used to assess these difficulties. Since 1886, when Landolt introduced measurements of positive and negative relative vergence, or fusional vergence reserves (Landolt, 1886), several attempts to establish criteria for acceptable values of this and other binocular measures have been made (Daum, Rutstein, Houston, Clore, & Corliss, 1989; Hofstetter, 1945; Lie & Opheim, 1985; Percival, 1928; Pestalozzi, 1975; Sheard, 1930; Sheedy & Saladin, 1978). One rationale for doing this is to enable verification of a symptomatic binocular problem, and to treat patients with values outside an acceptable range.

In the United Kingdom, use of the Mallett fixation disparity unit is widespread (Karania & Evans, 2006). Fixation disparity, which normally represents a fraction of the measured phoria in the same direction, can be corrected with aligning prisms, and the patient will see the effect while watching the Mallet unit. Conway, Thomas, and Subramanian (2012) compared results from this unit with measures of fusional vergence reserves in 500 adults, and found a strong inverse correlation between positive fusional vergence (PFV) and reported (exo) aligning prism from the Mallett unit for near evaluations (40 cm). At near, 299/500 (≈60%) participants required no aligning prism, however, of these, 107/299 (≈36%) complained of visual symptoms, and 89/299 (≈30%) failed to meet Sheard's criterion (see below). It is therefore possible that the Mallet unit is not sensitive enough to detect subtle binocular problems.

One commonly used method of evaluating binocular vision is Sheard's criterion (Dalziel, 1981; Daum et al., 1989; Evans, 2000; Mitchell Scheiman & Wick, 2008; Sheedy, 1983). This criterion establishes a norm for typical binocular vision such that the relative vergence is at least double the fusional demand, i.e. the phoria in the direction opposite to the vergence response. If, for example an exophoria of 7 PD (Prism Dioptres, base in) is measured, the relative (con)vergence measured at the same distance needs to be at least 14 PD (base out) for the criterion to be fulfilled. If, for instance, only 11 PD base out is measured, a prism-correction of 1 PD base in may be prescribed. This alters the measured phoria to 6 PD base in and the convergence to 12 PD base out, thus fulfilling Sheard's criterion. Treatment, in the form of spherical lenses, surgery, orthoptics or vision therapy, are other options (Atzmon, Nemet, Ishay, & Karni, 1993; Berard & Reydy, 1984; Dalziel, 1981; Daum, 1986; Griffin & Grisham, 2002).

Since Sheard's criterion is dichotomous (met or not met), it has its limitations for both clinicians and researchers. The difference in relative vergence before and after treatment, or between two patients, might be very small but span the criterion for typicality. Thus, despite only a very small change in absolute measurement for an individual, the measure taken before treatment might be atypical (Sheard's criterion not met) while the measure taken after treatment is typical (Sheard's criterion met). Sheard's criterion gives no indication of whether the change before and after treatment is large or small. Additionally, there

doi:10.5384/SJOVS.vol9i2p1 - ISSN: 1891-0890

is no evidence that symptoms from a phoria increase at the moment the criterion is not fulfilled or that all symptoms disappear when it is barely met. Attempts to convert Sheard's criterion to a continuous variable have been made previously, but primarily for statistical reasons (Conway et al., 2012; Evans, Busby, Jeanes, & Wilkins, 1995). Moreover, norms and repeatability of measurements have not been published. The main objective of this study was to provide a simple measure that can convert Sheard's criterion into a continuous variable which can be used by clinicians and researchers alike. We have named this measure "fusional stamina". A second aim was to determine normative values for, and test the repeatability of, this measure.

Material and methods

Design

A combination of descriptive cross-sectional and repeated-measures designs was used as part of a single masked randomised case control study.

Participants

For the purpose of the study reported here, 90 typical children aged five to eleven years old were recruited from two kindergartens and a public school in the western part of Oslo. From the kindergartens, fifteen 5-year-old children were recruited. The remaining group consisted of 15 children from each school grade 1 to 5, i.e. from 6 to 11 years of age. Three children were initially excluded; two children due to premature birth and one child due to a diagnosis of autism. For the purpose of this particular study, five more children were excluded due to intermittent or manifest strabismus. Thus, data from 82 typical children collected as part of a more comprehensive visual examination were used to establish normative values. All children were refracted (range –1.75 DS to +4.25 DS mean spherical equivalent on retinoscopy) and wore their habitual optical corrections during the assessments.

Procedure/Clinical tests

Tests used for this study were conducted early in the examination process after measurement of acuities and objective refraction. During the first testing session, cover test and fusional reserves base out were measured at 6 metres and 40 cm. For the second testing session, only measures at 40 cm were conducted. All testing was completed by the same examiner (AM), who was blind to initial results.

The alternating cover test (between eyes) was conducted at distance (6 m) and near (40 cm) to reveal any phoria and to measure the size of deviation. In the case of no visible movement, the child was asked if the fixation object appeared to be moving or jumping; "subjective cover test". Any movement was neutralised by a Behrens prism bar (Guilden Ophthalmic), and the size of the deviation was scored in prism dioptres (PD). A positive value was used for exodeviations, and negative for esodeviations. None of the children examined were found to have clinically significant vertical deviations.

Positive fusional vergences (reserves) for each target distance were measured with the Behrens prism bar. This was moved downwards at the speed of about one step per two seconds until the fixation object became double or appeared to start moving. When the "break (or suppression) point" was reached, the child was asked to make an extra effort to fuse the objects (or stop it from moving). The first prism value at which the child was unable to fuse the target was registered as the break point. By moving the prism bar in the opposite direction, a recovery point was registered when the child was able to fuse the images again. If a child was able to fuse the target with the largest prism on the bar (40 PD), 15 PD of loose prisms were placed in front of the child's eyes and the prism bar was again used from the 20 PD step until fusion was lost. There were nine children (10.3%) who were able to fuse the 55 PD (40 + 15 PD) in the first study. These children were given a score of 60 PD for break and 50 PD for recovery for statistical purposes.

Some children found it hard to report the break and recovery points, but these could be identified objectively since, at the break point, children used version movements to alternate between the two objects seen. Version movements were replaced by convergence when the recovery point was reached. Thus, break- and recovery-values could be recorded for all children.

Fusional Stamina is calculated by subtracting the heterophoria measured with the cover test, from half of the fusional reserve value measured in the opposite direction. If, for example the measured phoria is 4 PD exo, and the PFV at the same distance is measured as 10 PD, the fusional stamina will be (10/2) – 4 = 1 PD. Fusional stamina will be positive when Sheard's criterion is met, and negative when Sheard's criterion is not met.

Repeatability

Measures of repeatability are based on the 12 children from the typical population with lowest binocular values established by near point of convergence (NPC) and fusional stamina (see below) along with 28 premature children tested at the same time. To be considered eligible for this part of the study, children were included if they had normal distance acuities ($\leq$ logMAR 0.0 or $\geq$ 20/20 Snellen acuity) and a positive result on the Lang II stereopsis test. Ten children from each of the populations had NPC greater than 9 cm (range 9–31 cm). The two remaining typical children had a fusional stamina of (+)3 PD and −6 PD respectively. The remaining 18 premature children all had NPC below 9 cm and positive fusional stamina values.

This group of 40 children was retested after approximately one year. No child received any additional care from the Centre in this period.

Informed consent was obtained from all families involved after they received an explanation of the nature and possible consequences of the study in writing. The study was approved by the regional ethics committee for medical research at the University of Oslo and followed the tenets of the Declaration of Helsinki.

Data analysis

Means and standard deviations for phorias, PFV break and recovery, and fusional stamina, at 6 m and 40 cm are reported for the 82 children. This study also allowed us to calculate test-retest reliability for 40 children with Pearson's correlation coefficient. According to Field (2009), a correlation coefficient (r) of 0.1 is commonly viewed as a small sized effect, 0.3 a medium effect and 0.5 as a large effect. Repeatability was also assessed with a Bland-Altman analysis. Fusional stamina means and standard deviation were calculated for subgroups who either passed or failed Sheard's criterion. The bootstrap facility of the SPSS 21 program was used to ensure normality of data.

A multivariate analysis of variance (MANOVA) was run with all clinical measures as dependent variables and gender (male and female) and school grade (5 levels) as between-subject factors. The analysis showed that neither school grade nor gender had any significant impact on the measures. A repeated measures MANOVA showed the same result for all variables used in the second part of the study (n = 40). All analyses have therefore been collapsed across age and gender.

Results

Table 1 shows the descriptive data for phorias, fusional vergences and fusional stamina for all children tested and after excluding children with esophorias. Differences between the whole group and after excluding children with esophoria are small with a change in fusional stamina at distance of only 0.8 PD (All: 8.9 PD; Excluding Esophoria: 8.1 PD) and at near of

1.8 PD (All: 14.1 PD; Excluding Esophoria: 12.3 PD). Figure 1 shows the distribution of fusional stamina by age at distance and near for all children (including the 12 children with esophoria for each distance). There was no age related change in fusional stamina over the range tested here.

Table 1: *Comparison of measured values for all participants vs. when participants with esophoria are excluded.*

Mean values for measures	all participants $N = 82$		excluding esophorias $n = 70$	
	Mean values in PD (SD)	95% Confidence Interval (PD)	Mean values in PD (SD)	95% Confidence Interval (PD)
Phoria (exo) at 6m	0.3 (2.8)	0.3 eso–1.0 exo	0.9 (2.0)	0.5–1.4 exo
Phoria (exo) at 40 cm	2.5 (3.9)	1.7–3.4 exo	3.6 (3.1)	2.9–4.3 exo
PFV (break) at 6 m	18.5 (11.1)	16.1–20.8	18.0 (10.7)	15.5–20.6
PFV (recovery) at 6 m	11.3 (9.5)	9.4–13.4		
PFV (break) at 40 cm	33.4 (14.5)	30.2–36.7	31.9 (14.0)	28.6–35.2
PFV (recovery) at 40 cm	24.7 (13.6)	21.8–27.8		
Fusional Stamina at 6 m	**8.9 (6.6)**	**7.5–10.4**	**8.1 (5.9)**	**6.7–9.5**
Fusional Stamina at 40 cm	**14.1 (9.2)**	**12.3–16.2**	**12.3 (8.2)**	**10.4–14.3**

Note: PD = Prism Dioptres; PFV= Positive Fusional Vergence.

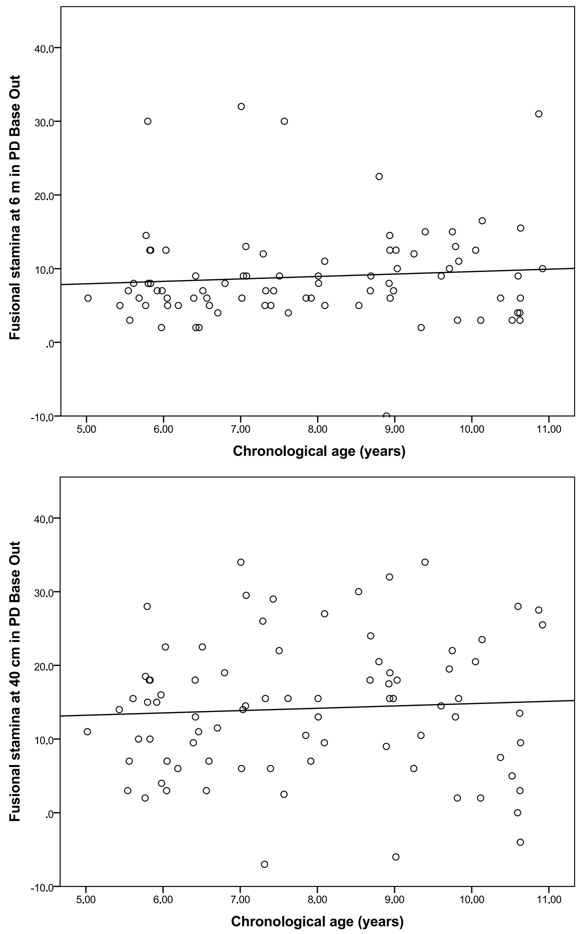

Figure 1: *Distribution of fusional stamina as a function of age for all $N = 82$ subjects. Top: At 6 m. Bottom: At 40 cm. There is no significant association between the two variables for distance or near. Fusional stamina below zero PD (Prism Dioptres) is equivalent to failing Sheard's criterion. This is the case for one child at distance and three children at near.*

In order to assess the test-retest relationship, correlations between values at two testing times were calculated for the group of 40 children believed to be more at risk of binocular vision problems. There were significant correlations between the two time points for near PFV break ($r = 0.38, p = 0.02$), PFV recovery ($r = 0.50, p = 0.01$) and fusional stamina ($r = 0.62, p < 0.001$). Additionally, there was no statistical difference between the fusional stamina means measured at each time point for near; 9.5 PD (±9.7 PD) for time 1 and 9.4 PD (±7.9 PD) for time 2; $F(1, 39) = 0.01, p = 0.92$.

Figure 2 shows the individual data for fusional stamina for all ($N = 40$) subjects. Each vertical line represents the difference between the two measurements for a participant. A negative value for fusional stamina (y-axis) is equivalent to failure on Sheard's criterion. Note that there were five children (12.5%) for whom values were negative on one visit (failed Sheard's criterion) and positive (passed Sheard's criterion) on the other. This suggests that while Sheard's criterion is a stable measure for most children, there is a potential for misdiagnosis when fusional stamina is close to zero.

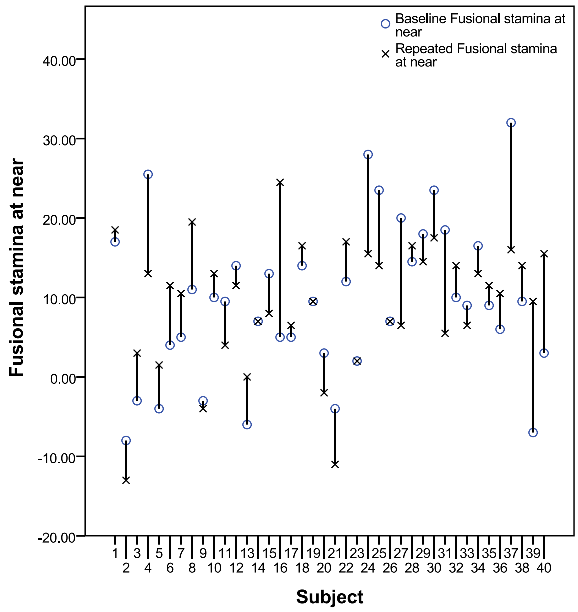

Figure 2: *Distribution of fusional stamina at near repeated after one year for all $N = 40$ subjects. Each vertical line represents the difference between the two measurements for each subject. A negative value on the Y-axis means that Sheard's criterion has not been passed. Note that n = 5 subjects have positive values in one trial and negative in the other.*

To further evaluate test-retest reliability, a Bland-Altman analysis was conducted (Figure 3). A one sample t-test showed no significant difference between the mean expected value of 0 D difference between measurements and the calculated mean difference of 0.125 D, $p = 0.92$. In addition, a linear regression with the mean fusional stamina dioptre value from the two testing times as the independent variable and the difference between them as the dependent variable confirmed that there is no proportional bias; $F(1, 39) = 2.65, p = 0.11$. However, a standard deviation of 7.9 PD for the differences between the two testing times, results in a rather large 95% confidence intervals as shown in the figure.

The children were then divided by whether they passed or failed Sheard's criterion. The fusional stamina data for each group is shown in Table 2. The mean fusional stamina value for children passing Sheard's criterion on both occasions is above 12 PD, while children who failed on both occasions have a value which is −5 PD or less. For the five children who only passed on one of the two testing times, mean fusional stamina for the occasion they did pass was 3.4 PD (±3.6 PD). This value is low, and lies more than one SD below the mean for typical children (Table 1).

Table 2: *Mean Fusional Stamina (SD) in prism dioptres when Sheard's criterion @ 40cm was either passed both times, failed both times or passed and failed on different times*

Participants, N = 40	FS Time 1	FS Time 2
Pass both times (mean), n = 32	12.9 (7.6)	12.2 (5.0)
Fail both times (mean), n = 3	–5.0 (2.7)	–9.3 (4.7)
Pass-fail, n = 1	3.0	–2.0
Fail-pass, n = 4	–5.0 (1.8)	3.5 (4.2)

Note: FS = Fusional Stamina.

Discussion

Fusional stamina is a continuous variable based on Sheard's criterion which provides a measure of the relationship between PFV and the exophoria at a given distance. The concept of fusional reserves implies that PFV represents resources to overcome any exophoria. Several studies, although without placebo control groups, have shown that symptoms are significantly reduced or eliminated by increasing PFV and / or reducing the exophoria (e.g. Atzmon et al., 1993; Aziz, Cleary, Stewart, & Weir, 2006; Lie & Opheim, 1990). It is therefore indicated that the more positive the value for fusional stamina (i.e. more resources), the lower the likelihood that the exophoria will be symptomatic for the child / patient, while the more negative the value (i.e. less resources) the greater the likelihood of a symptomatic exophoria that requires treatment. This study provides normative values for fusional stamina in a group of typical children. When children with esophorias were excluded, fusional stamina values were 8.1 PD for distance and 12.3 PD for near.

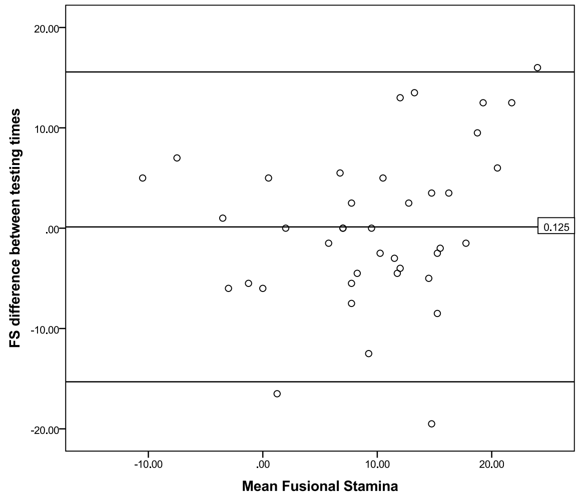

Figure 3: *Bland-Altman plot for repeated fusional stamina (FS) at near. The middle horizontal line represents the mean difference between the two measures (0.13 dioptres). The upper and lower lines represent the upper (15.57 D) and lower (−15.31 D) confidence-levels of the mean difference respectively.*

The concept of converting Sheard's criterion into a continuous variable is not new. Conway et al. (2012) defined "Sheard's value" as "the fusional reserve opposing the heterophoria to blur point (or, if no blur to break point) divided by the heterophoria" (p. 2). This implies that Sheard's criterion is unmet if the value is below two. Yet, this method has three limitations compared to fusional stamina. Firstly, it is somewhat harder to calculate in most cases, secondly it will be lees useful in cases of "Restricted relative convergence" (see next paragraph) and thirdly, it will not work at all if there is no heterophoria. Apart from these minor limitations, the idea of Sheard's value is almost identical to that of fusional stamina. Our research adds to this by introducing a value which is more clinically useful and by providing repeatability-measures and values for a typical population of children.

One example of the benefit of measuring fusional stamina can be illustrated by considering "Restricted relative convergence (base out)". This has been used as a means of diagnosing binocular problems that can result in symptoms, e.g. after acquired brain injury (Ciuffreda et al., 2008). Thus, low PFV is assumed to implicate potential binocular problems, independent of the phoria size. This would not be indicated using Sheard's criterion. For example, a child who has 6 PD of PFV and a 2 PD exophoria, has PFV that is three times higher than the phoria value, which is well above Sheard's criterion. In contrast, fusional stamina will be low in this case ($6/2 - 2 = 1$ PD) and therefore would indicate potential binocular problems. Fusional Stamina is also easier to evaluate since it avoids the need to either compare two parameters that are estimated separately, or use a set criterion.

Especially in American studies of binocular vision, meeting Sheard's criterion (often in addition to low PFV) is commonly used as an inclusion criterion. For example, the CITT (Convergence Insufficiency Treatment Trial) studies did not include children who passed Sheard's criterion unless PFV at near was 15 PD or less (CITT-Group, 2008). We believe it would be better to simply state that children with a fusional stamina of 7.5 PD or less could be included, and to use this measure for evaluations as well. This would provide an additional means of determining the size of change after any intervention.

Comparisons with previous findings

Phoria Measurement

Table 3 compares phoria at 6 m and 40 cm measured in this study with those from previous studies. Despite differences in method and age of participants, values are close to zero (orthophoria) at 6 m, while there is a small (< 6 PD) exophoria at 40 cm. Our results are consistent with the values reported in this table.

Table 3: *Mean value of heterophoria in prism dioptres at 6 m and at 40 cm comparing values from previous research to those found in this study.*

Author/Year	Criteria	Age	N	6 m	SD	40 cm	SD
Morgan Jr (1944)	Von Graefe	Adult	800	1.0 x	2.0	3.0 x	5.0
Jackson and Goss (1991)	Von Graefe	8–16	244	1.0 x	2.0	3.0 x	4.0
Rouse, Borsting, and Deland (2002)	Von Graefe	10–11.5	20[a]			4.3 x[b]	6.5
Letourneau and Giroux (1991)	Maddox	6–13	2035	0.6 s	2.5	0.8 x	4.5
Jimenez, Perez, Garcia, and Gonzalez (2004)	Maddox	6–12	1015	0.6 s	1.9	0.4 x	3.0
Daum et al. (1989)	Cover test	Adult	100	0.8 x	1.8	1.7 x	10.0
Conway et al. (2012)	Cover test	Adult	500	1.6 x	3.3	5.9 x	3.2
This study	Cover test	5–11	82	0.3 x	2.8	2.5 x	3.9

Note: x = Exophoria; s = Esophoria.
[a] Examined three consecutive times by two different observers at two occasions (= 240 measures).
[b] Performed at 30 cm.

While within- and between-session reliability of the von Graefe method for measuring near phoria has been shown to be high in sixth graders (Rouse et al., 2002), the inter-examiner reliability was lower for this measure. A previous study which directly compared methods in adults found that prism-neutralised cover test with subjective reporting had greater reliability compared to the von Graefe method (Rainey, Schroeder, Goss, & Grosvenor, 1998).

doi:10.5384/SJOVS.vol9i2p1 - ISSN: 1891-0890

Positive Fusional Vergence

The correlation between heterophoria and subjective symptoms is generally low (Goss, Reynolds, & Todd, 2010; Sheedy & Saladin, 1978), and has not been found to be significant for exophoria. However, whether a given child is symptomatic is likely to be affected by individual differences in the ability to compensate for heterophoria. This is why a measure of positive fusional vergence is needed over and above phoria measurement. Table 4 summarises results from studies that have measured PFV break value using various methods in samples of different ages. Values measured using the synoptophore method are higher than those using prisms. If these values are excluded, the range for PVF measured in previous studies is 17–20.8 PD at 6 m and 18.0–27.1 PD at 40 cm. Our results for PFV break are within this range at 6 m (18.0 PD) but above the range at 40 cm (31.9 PD). Possible explanations for our high PFV break for near include instructing children to make an extra effort to fuse the images and the fact that our age group is the youngest listed in Table 4. It is also likely that there is a ceiling effect in other studies using prism bars and Risley prisms, as these normally are limited to 40 PD.

Table 4: *Mean value of positive fusional vergence (break-value) in prism dioptres at 6 m and at 40 cm comparing values from previous research to those found in this study.*

Author/Year	Criteria	Age	N	6 m	SD	40 cm	SD
Daum et al. (1989)	Synoptophore	Adult	100	29.0	19.0	33.0	19.0
Morgan Jr (1944)	Risley prisms	Adult	800	19.0	8.0	21.0	6.0
Jackson and Goss (1991)	Risley prisms	8–16	244			27.0	8.0
Rouse et al. (2002)	Risley prisms	10–11.5	20[a]			22.0[b]	17.0
Wesson (1982)	Prism bar	7–12	79			19.0	11.0
Scheiman, Herzberg, Frantz, and Margolies (1989)	Prism bar	6–12	386			23.0	8.0
Chen and Abidin (2002)	Prism bar	7–12	60			19.4	9.4
Jimenez et al. (2004)	Prism bar	6–12	1016	17.0	7.0	18.0	8.0
Conway et al. (2012)	Prism bar	Adult	500	20.8	6.4	27.1	8.2
This study	Prism bar	5–11	82	18.5	11.1	33.4	14.5

[a] Examined three consecutive times by two different observers on two occasions (= 240 measures).
[b] Performed at 30cm.

Several studies have shown moderate to low repeatability for PFV-measures (Goss & Becker, 2011; Jackson & Goss, 1991; Rouse et al., 2002). Sources of error include instructions used (e.g. whether children are encouraged to try to keep one single image), speed of introducing prisms, and time allowed for the child/patient to try to re-fuse the images if they become double. The limited reliability of PFV-measures implies that these numbers should be used with caution.

Fusional Stamina

Calculations of fusional stamina for the studies appearing in Tables 3 and 4 are shown in Table 5. After excluding the measure made using the synoptophore method, the range of values for fusional stamina is lower than for PFV (6 m: range 8.5–9.1 PD; 40 cm: range 6.7–14.1). As for PFV, the value measured in this study is close to the range for previous studies for the 6 m measure but higher than previous studies for the 40 cm measure. Indeed, almost all values for the 40 cm measure using prism methods are outside the 95% confidence intervals for this study (10.4–14.3 PD). This is the result of the high PFV break value used to calculate fusional stamina in this study.

Table 5: *Calculated Fusional Stamina in prism dioptres at 6 m and at 40 cm comparing values from previous research to those found in this study.*

Author/Year	Criteria	Age	N	6 m	40 cm
Daum et al. (1989)	Synoptophore	Adults	100	13.7	14.8
Morgan Jr (1944)	Risley prisms	Adults	800	8.5	7.5
Jackson and Goss (1991)	Risley prisms	8–16	244		10.5
Rouse et al. (2002)	Risley prisms	10–11.5	20[a]		6.7[b]
Jimenez et al. (2004)	Prism bar	6–12	1015	9.1	8.6
Conway et al. (2012)	Prism bar	Adults	500	8.8	7.7
This study	Prism bar	5-11	82	8.9	14.1

Note: distribution (SD) cannot be calculated from information provided in former published articles.
[a] Examined three consecutive times by two different observers at two occasions (= 240 measures).
[b] Performed at 30cm.

Reliability of Fusional Stamina

Fusional stamina measured one year apart was not significantly different across two testing times with means of 9.5 PD and 9.4 PD, respectively. Indeed, 35 children (87.5%) had either both positive or both negative values each time. While measures of PFV were significantly correlated with a medium sized effect when repeated after one year, the effect size for fusional stamina was large, indicating a higher level of repeatability. This suggests that fusional stamina is a reliable measure. The Bland-Altman plot (Figure 3) shows that the variance is large, however.

Sensitivity of Fusional Stamina

The five children (12.5%) who did not provide consistently signed fusional stamina values, had lower positive values than expected for typical children on both occasions. Values of fusional stamina for these children were more than one SD below the group means (Table 1) for the trial they passed. Sheard's criterion shifts when fusional stamina has a value of zero, and this corresponds to 1.5× SD subtracted from the fusional stamina means. This indicates that typical non-strabismic children who do not pass Sheard's criterion, by definition have fusional stamina values within the lowest 7% in this population. Thus, our data support the use of fusional stamina as a more sensitive measure than Sheard's criterion for the categorisation of the relationship between PFV and exophoria in an "at-risk population". However, this does not necessarily mean that patients with low fusional stamina require treatment.

Limitations

When calculating Sheard's criterion it is commonly recommended that the "blur-point" is used as the PFV value. When this is visible to the patient, it is always lower than the value for "break" as used in this study (Sheard, 1930). Using the blur-point when calculating fusional stamina will therefore produce lower values than those reported here. However, this would have required that every child was fully refracted, and that fixation objects were suitable for each child's near acuity. Moreover, blur is harder to define than "double", cannot be seen by the examiner, and would therefore require more attention from each child. For these reasons, we have used the break point here.

There is also a potential source of error in refraction. The typical children wore their habitual lenses (if any) but were not fitted with "optimal lenses" before the measurements were performed. For the second part of the study where repeated measures were taken, the prerequisite of normal distance acuities makes uncorrected hypermetropia the most likely refractive error. According to the known relationship between accommodation and convergence (AC/A), an increase in exophoria can result from applying + lenses (Fry & Haines, 1940). If the children in this part of the study were corrected one could therefore expect a reduction in fusional stamina measured with base out prisms. Since acuities and refraction were not repeated after one year, changes in refraction might have influenced the results and therefore this will be an added source of noise in our binocular measures. It is therefore possible that repeatability of these measures would have increased if these changes had been taken into account.

It might be argued that the same population should have been tested in both studies reported here. We used different populations in the two parts of this study because typical children are required when calculating reliable norms, and reliability of measure in a clinical population is of greater clinical interest. We chose not to measure base in fusional reserves because esodeviations combined with low base in fusional reserves are not very common, and would require several thousand participants from a typical population to yield significant results (Conway et al., 2012).

Further studies are required to relate fusional stamina to symptoms of poor binocular vision and to expand the age range used here. Moreover, it would be of interest to evaluate fusional stamina repeatability at distance. Clinically, it is also important to remember that no single test for binocularity is perfect, and a combination of different tests should be used for diagnostic and treatment purposes (Evans, 2008).

Conclusions

This study has been used to introduce fusional stamina, a continuous measure of binocular function which is easy to calculate and more sensitive than Sheard's criterion: Instead of simply indicating if the criterion is met or not, fusional stamina provides a tool to state how far from passing the criterion a subject or patient is. This can be of importance for the researcher as well as the clinician concerned with treating binocular vision problems. Similarly to Sheard's criterion but with greater precision, fusional stamina can be used as a reliable baseline-measure, for between-group comparisons and for progress evaluations as treatment progresses.

We have also shown that there is a risk of missing potential binocular problems if the clinician relies only on Sheard's criterion, since 12.5% of the children in our group did not pass the criterion in one of the two trials. Low levels of fusional stamina therefore seem to be better predictors of binocular vision problems in symptomatic patients. Norms for a typical group of non-strabismic schoolchildren have been provided, and we have shown that the repeatability is high and better than for fusional vergence.

Acknowledgement

The work leading to this research was carried out at Huseby Resource Centre in Oslo, Norway. This research has been supported by grants from the Norwegian Extra Foundation for Health and Rehabilitation through EXTRA funds and the Premature Association of Norway, C-Optikk and The Norwegian Optometric Association. We thank the employees of the two kindergartens and the school involved, and all participating families and children. Thanks also to Prof Bruce JW Evans for his discussions with us about this topic. Finally, we would like to thank the anonymous reviewers in helping us to improve the original manuscript for this article. The authors have no declared conflict of interest or financial disclosure.

References

Atzmon, D., Nemet, P., Ishay, A., & Karni, E. (1993). A randomized prospective masked and matched comparative study of orthoptic treatment versus conventional reading tutoring treatment for reading disabilities in 62 children. *Bin Vis Eye Muscle Surg Q*, *8*(2), 91–106.

Aziz, S., Cleary, M., Stewart, H. K., & Weir, C. R. (2006). Are orthoptic exercises an effective treatment for convergence and fusion deficiencies? *Strabismus*, *14*(4), 183–189.

Berard, P. V. & Reydy, R. (1984). Treatment of heterophorias. *Documenta Ophthalmologica*, *58*(1), 35–43. doi:10.1007/BF00140896

Chen, A. H. & Abidin, A. H. Z. (2002). Vergence and accommodation system in Malay primary school children. *The Malaysian journal of medical sciences: MJMS*, *9*(1), 9–15.

CITT-Group. (2008). Randomized clinical trial of treatments for symptomatic convergence insufficiency in children. *Arch Ophthalmol*, *126*(10), 1336–49. doi:126/10/1336[pii]10.1001/archopht.126.10.1336

Ciuffreda, K. J., Rutner, D., Kapoor, N., Suchoff, I. B., Craig, S., & Han, M. (2008). Vision therapy for oculomotor dysfunctions in acquired brain injury: a retrospective analysis. *Optometry-Journal of the American Optometric Association*, *79*(1), 18–22. doi:10.1016/j.optm.2007.10.004

Conway, M. L., Thomas, J., & Subramanian, A. (2012). Is the aligning prism measured with the Mallett unit correlated with fusional vergence reserves? *PloS one*, *7*(8), e42832. doi:10.1371/journal.pone.0042832

Dalziel, C. C. (1981). Effect of vision training on patients who fail Sheard's criterion. *American journal of optometry and physiological optics*, *58*(1), 21–23. doi:10.1097/00006324-198101000-00006

Daum, K. (1986). Double-blind placebo-controlled examination of timing effects in the training of positive vergences. *Am J Optom Physiol Opt*, *63*(10), 807–812.

Daum, K., Rutstein, R. P., Houston, G. I., Clore, K. A., & Corliss, D. A. (1989). Evaluation of a new criterion of binocularity. *Optometry & Vision Science*, *66*(4), 218–228. doi:10.1097/00006324-198904000-00008

Evans, B. J. (2000). An open trial of the Institute Free-space Stereogram (IFS) exercises. *British Journal of Optometry and Dispensing*, *8*(1), 5–14.

Evans, B. J. (2008). Optometric prescribing for decompensated heterophoria, oip, volume 9, issue 2, 2008, b j evans. *Optometry in Practice*, *9*(2), 63–78.

Evans, B. J., Busby, A., Jeanes, R., & Wilkins, A. J. (1995). Optometric correlates of Meares–Irlen syndrome: a matched group study. *Ophthalmic and Physiological Optics*, *15*(5), 481–487. doi:10.1016/0275-5408(95)00063-J

Field, A. P. (2009). *Discovering statistics using spss : (and sex and drugs and rock 'n' roll)* (3rd). Los Angeles ; London: SAGE.

Fry, G. A. & Haines, H. F. (1940). Tait's analysis of the accommodation-convergence relationship. *Optometry & Vision Science*, *17*(9), 393–401. doi:10.1097/00006324-194009000-00001

Goss, D. A. & Becker, E. (2011). Comparison of near fusional vergence ranges with rotary prisms and with prism bars. *Optometry*, *82*(2), 104–107. doi:10.1016/j.optm.2010.09.011

Goss, D. A., Reynolds, J. L., & Todd, R. E. (2010). Comparison of four dissociated phoria tests: reliability & correlation with symptom survey scores. *J. Behavioral Optom*, *21*, 99–104.

Griffin, J. R. & Grisham, J. D. (2002). *Binocular anomalies: diagnosis and vision therapy* (4th). Amsterdam ; London: Butterworth-Heinemann.

Hofstetter, H. W. (1945). *The zone of clear single binocular vision*. American Academy of Optometry.

Jackson, T. W. & Goss, D. A. (1991). Variation and correlation of standard clinical phoropter tests of phorias, vergence ranges, and relative accommodation in a sample of school-age children. *J Am Optom Assoc*, *62*(7), 540–7.

Jimenez, R., Perez, M. A., Garcia, J. A., & Gonzalez, M. D. (2004). Statistical normal values of visual parameters that characterize binocular function in children. *Ophthalmic and Physiological Optics*, *24*(6), 528–542. doi:10.1111/j.1475-1313.2004.00234.x

Karania, R. & Evans, B. J. (2006). The Mallett Fixation Disparity Test: influence of test instructions and relationship with symptoms. *Ophthalmic and Physiological Optics*, *26*(5), 507–522. doi:10.1111/j.1475-1313.2006.00385.x

Landolt, E. (1886). *The refraction and accommodation of the eye and their anomalies*. Edinburgh: Young J. Pentland.

Letourneau, J. E. & Giroux, R. (1991). Nongaussian distribution curve of heterophorias among children. *Optom Vis Sci*, *68*(2), 132–7. doi:10.1097/00006324-199102000-00008

Lie, I. & Opheim, A. (1985). Long-term acceptance of prisms by heterophorics. *J Am Optom Assoc*, *56*(4), 272–278.

Lie, I. & Opheim, A. (1990). Long-term stability of prism correction of heterophorics and heterotropics; a 5 year follow-up. Part I: Heterophorics. *J Am Optom Assoc*, *61*(6), 491–498.

Morgan Jr, M. W. (1944). The clinical aspects of accommodation and convergence. *Optometry & Vision Science*, *21*(8), 301–313. doi:10.1097/00006324-194408000-00001

Percival, A. S. (1928). *The prescribing of spectacles*. Bristol, England: J. Wright and Sons.

Pestalozzi, D. (1975). Erfahrungen mit der kombination von polatest und prismenvollkorrektion in der behandlung von störungen des binokularsehens. *Ophthalmologica*, *170*(2-3), 274–279. doi:10.1159/000307219

Rainey, B. B., Schroeder, T. L., Goss, D. A., & Grosvenor, T. P. (1998). Inter-examiner repeatability of heterophoria tests. *Optometry & Vision Science*, *75*(10), 719–726. doi:10.1097/00006324-199810000-00016

Rouse, M. W., Borsting, E., & Deland, P. N. (2002). Reliability of binocular vision measurements used in the classification of convergence insufficiency. *Optom Vis Sci*, *79*(4), 254–264. doi:10.1097/00006324-200204000-00012

Scheiman, M., Herzberg, H., Frantz, K., & Margolies, M. (1989). A normative study of step vergence in elementary schoolchildren. *J Am Optom Assoc*, *60*(4), 276–80.

Scheiman, M. [Mitchell] & Wick, B. (2008). *Clinical management of binocular vision : heterophoric, accommodative, and eye movement disorders*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Sheard, C. (1930). Zones of ocular comfort. *American Journal of Optometry*, (7), 9–25. doi:10.1097/00006324-193001000-00001

Sheedy, J. (1983). Validity of diagnostic criteria and case analysis in binocular vision disorders. *Am J Optom 1980a*, *57*(9), 618–631.

Sheedy, J. & Saladin, J. (1978). Association of symptoms with measures of oculomotor deficiencies. *American journal of optometry and physiological optics*, *55*(10), 670–676. doi:10.1097/00006324-197810000-00002

Wesson, M. D. (1982). Normalization of prism bar vergences. *Am J Optom Physiol Opt*, *59*(8), 628–34. doi:10.1097/00006324-198208000-00002