

**PART B**

**Psychophysics, Eye Movements,**

**Physiological Optics**

## THE DEVELOPMENT OF MONOCULAR AND BINOCULAR VEP ACUITY

RUSSELL D. HAMER, ANTHONY M. NORCIA, CHRISTOPHER W. TYLER  
and CHARLENE HSU-WINGES<sup>1</sup>

Smith-Kettlewell Eye Research Foundation, 2232 Webster Street, San Francisco, CA 94115 and  
<sup>1</sup>Department of Ophthalmology, Kaiser Permanente Medical Center, 1200 El Camino Real, South San  
Francisco, CA 94080, U.S.A.

(Received 29 October 1987; in revised form 22 September 1988)

**Abstract**—The development of monocular and binocular grating acuity was measured in 87 infants, 2–52 weeks of age, using the sweep VEP technique. Average monocular and binocular acuity growth functions were nearly identical, with a small (<0.2 octaves) binocular acuity superiority occurring only under 6 months. Interocular acuity differences were small (averaging <1/4 octave, unsigned, with a 95% confidence interval of < ±0.6 octaves) and were not significant at any age. These characteristics make the sweep VEP technique a potentially sensitive tool for the detection of monocular visual losses in the early stages of amblyopia.

Visual development    Acuity    Interocular differences    Visual evoked potentials    Spatial vision  
Binocular vision    Human infants

### INTRODUCTION

Recently, the development of visual acuity under binocular viewing conditions has been measured in human infants from birth to 1 year of age using an efficient electrophysiological paradigm, the sweep VEP technique (Norcia and Tyler, 1985a). From a large sample of infants it was found that grating acuity increased rapidly from about 5 c/deg just after birth to 20 c/deg at 8 months postnatal age.

While these data establish age norms for the development of sweep VEP acuity under binocular viewing conditions, it is important to obtain normative data for monocular measures of acuity development, and for interocular acuity differences, using the same technique and stimulus conditions. Such norms are needed, not only for the purposes of generating cohesive models of normal visual development, but also to provide a database for future clinical studies. The present study, therefore, was designed to measure the development of binocular and monocular sweep VEP acuity, as well as within-subject interocular acuity differences, over the first year of life.

The VEP acuity norms have bearing on at least two theoretical issues. The first involves the hypothesis that the development of spatial

vision, prior to the onset of binocularity, is independent in the two eyes (Atkinson *et al.*, 1982; Birch, 1985). Our monocular data, in which left-eye and right-eye acuities were obtained from the same infants, directly address this issue. The results indicate that development of acuity, up to the level of cortical VEP generators, is identical for the two eyes over the entire first year of normal visual development.

The second theoretical issue involves the basis of previously observed superiority of binocular acuity over monocular acuity in infants (e.g. Birch, 1985; McDonald *et al.*, 1986). Since we have measured VEP acuities under both binocular and monocular viewing conditions in the same infants, we can address the question of binocular acuity superiority in VEP measures.

It has been proposed that the binocular superiority of adult psychophysical acuities can be explained by probability summation between the two eyes (Campbell and Green, 1965). We shall contend that, although the VEP amplitude under binocular conditions often exceeds the amplitude under monocular conditions, the extrapolated acuities in steady-state VEP measurements are not subject to the effects of probability summation. Therefore, monocular and binocular VEP acuities should, in principle, be the same.

## METHODS

### *Apparatus*

Details of the apparatus and analysis techniques may be found in Norcia and Tyler (1985a, b) and in Norcia *et al.* (1985). Briefly, infants were presented with vertical sine-wave luminance gratings of 80% contrast and a space-average luminance of 80 cd/m<sup>2</sup>. Gratings were squarewave modulated at a rate of 12 contrast reversals per sec (rps). During each 10 sec trial, the spatial frequency of a reversing grating was incremented linearly in 19 equally-spaced steps, usually spanning a 30:1 range of spatial frequency.

The EEG was recorded from two bipolar placements. Grass gold-cup electrodes were placed symmetrically 3 cm to the left (L-CH) and 3 cm to the right (R-CH) of a common reference electrode, which was placed 1 cm above the inion on the midline. The analysis was conducted using a 2 sec, 50% Tukey analysis window (Harris, 1978). Adjacent windows overlapped by 1.5 sec, resulting in a 72.7% correlation between adjacent datum points. By means of a Discrete Fourier Transform, the EEG within each window was analyzed to determine the amplitude and phase of the reversal response (12 Hz). In addition, for each window we computed the amplitude of a nearby independent frequency (14 Hz) as an estimate of background noise during the trial (Norcia and Tyler, 1985a). The noise amplitudes were averaged over the 10-sec trial to provide a more reliable estimate of the mean noise level during the trial.

Acuity was estimated from a given sweep record by a linear extrapolation to zero amplitude of the VEP amplitude versus spatial-frequency function. The two primary criteria used to make an acuity estimate (discussed in detail in Norcia and Tyler, 1985a; and Norcia *et al.*, 1985) were that (1) the peak amplitude at the stimulus (signal) frequency had to exceed the mean noise level by a ratio of at least 3:1, and (2) the phase of the response had to be nearly constant or slightly lagging (by no more than 90 deg per bin) over the entire portion of the amplitude response used for acuity extrapolation. Empirically, the amplitude criterion alone yields a false positive rate that is less than 0.3% per bin (Norcia and Tyler, 1985a; Norcia *et al.*, 1985). Combined with requirements for phase consistency and safeguards for local broadband artifact rejection (Norcia and Tyler,

1985a), there is adequate protection level for multiple-testing. From a large sample of adult EEG records (10 sec each) scored using the above criteria, only 2% were spuriously scored.

Single sweep records that do not meet the above criteria are not scored. However, they can contribute to an acuity estimate since they are included in a *vector average* of all the sweep records for a given condition. The SNR in each sampling bin can improve by a factor equal to the square root of the number of trials averaged. This is the maximum possible increase in SNR and requires that the phase be perfectly coherent between trials. Thus, amplitude at the temporal analysis frequency will sum across sweep trials only to the extent that the reversal response is stationary across trials. Linear extrapolation of these vector-average records also yielded acuity estimates and were treated, statistically, the same as single-sweep acuities.

### *Procedure*

Monocular and binocular measurements were made in blocks with the order of left-eye (LE), right-eye (RE) and both-eye (BE) testing randomized. Infants' eyes were occluded with Optclude adhesive patches for monocular testing. Generally, 3-4 sweep trials of 10-sec duration were attempted before moving to the next condition.

During each trial, the infant's attention to the video display was maintained by dangling a small, noisy toy at the plane of the screen. If the infant glanced away or made a large movement, the experimenter suspended the trial by pressing an interrupt button. When the infant's attention was regained, a second press of the button allowed the trial to continue, starting at the spatial frequency value presented 1 sec before the time of the initial interruption. A trial was terminated and discarded if the infant's attention could not be regained, or if any large electrical artifacts were noted by a second experimenter monitoring the EEG. A typical session lasted 30-45 min, including time for electrode attachment, breaks and data collection from LE, RE and BE. A favorable session lasted about 20 min.

### *Infants*

A total of 87 infants between 2 and 52 weeks of age participated in the study. All were born within 2 weeks of their due dates by report of the parents. 44 infants were retested within 2

weeks of the first session, yielding a measure of the intersession reliability.

## RESULTS

### *I. Monocular and binocular acuity development*

*Success rates: binocular testing.* Under binocular viewing conditions, at least one criterion acuity estimate was obtained in 78 out of 87 infants (90%) from the left recording channel (L-CH). For the R-CH, the success rate was 75 out of 87 (86%). However, the use of two recording channels significantly enhanced the likelihood of obtaining a criterion acuity estimate. Thus, 83 of the 87 infants tested (95%) yielded at least one binocular acuity estimate on at least one recording channel.

Of the four infants not yielding a binocular acuity estimate, one was not tested binocularly, and three were tested but failed to yield a criterion VA. All four were under 6 weeks of age. Thus, one or more binocular acuity estimates was obtained on 100% of the infants older than 6 weeks of age.

*Success rates: monocular testing.* For monocular testing, the success rates for the 87 infants tested were: *Left eye*—70.1% (L-CH), 73.6% (R-CH), and 80.5% (channels combined); *Right eye*—66.7% (L-CH), 74.7% (R-CH), and 81.6% (channels combined).

For both monocular and binocular testing, the success rates encountered for channels combined (i.e. success on either channel) were less than expected if the two channels yielded entirely independent acuity estimates ( $\chi^2 > 25$ ,  $P < 0.001$ , d.f. = 2). This same result is found if the binocular data are excluded from the  $\chi^2$  analysis. Independence between the channels predicts combined-channel success rates 92.1%, 91.6% and 98.6% for LE, RE and BE, respectively. If, on the other hand, the two electrode pairs were completely correlated, combining data across channels would simply yield the single-channel performance (about 72% success rate for monocular testing, and about 88% for binocular testing).

To measure an interocular acuity difference, one must obtain at least one acuity estimate from both the LE and RE. We were able to measure an interocular acuity difference in 65 out of the 87 infants (75%). Of the 22 infants for whom we did not obtain a criterion acuity estimate from both LE and RE, most (17) were 8 weeks of age or less. Of these 17, 10 were not tested on either LE or RE (or both) due to

sleepiness or fussiness on the part of the infant. In addition, out of the above 22 infants, 4 infants older than 8 weeks of age were not tested on both the LE and RE conditions. Thus, out of the 73 infants actually tested on both LE and RE conditions, 89% of the tests were successful. Of the 55 infants older than 8 weeks of age, 52 were tested on both LE and RE, and 96% of these were successful.

*Success rates: session 1 vs session 2.* Of the 44 retest infants, 31 were successfully tested on the LE in both sessions, and 30 were successfully tested in both sessions on the RE. Nearly all the retest infants (98%) were successfully tested binocularly on both sessions.

### *Monocular and binocular acuity development*

Acuity growth curves for the LE, RE and BE conditions for each of the two recording channels are shown in Fig. 1 (solid curve: L-CH; dashed curve: R-CH). For each infant we calculated the geometric mean of all criterion acuity estimates obtained for each condition. Data plotted are the group means (solid lines)  $\pm 1$  SEM for each condition. For infants who were tested more than once, only data from the first session are shown, unless more eye-conditions were successfully completed in the second session.

In 43 out of the 87 infants tested, we were able to measure a criterion acuity for both recording channels, in all three eye conditions. Using this sub-group of infants who completed the experimental design, a 3-way analysis of variance of the data was used to assess the effects of *eye* (LE, RE, BE), *channel* (L-CH, R-CH), and *natal age* (0-12, 13-24, >24 weeks). This ANOVA revealed the expected main effect of natal age ( $F = 68.82$ ,  $P < 0.00001$ ) and a main effect of eye ( $F = 5.85$ ,  $P < 0.004$ ), with no significant effects of channel or any interactions.

The main effect of eye is due to a small but significant superiority of binocular over monocular acuities. Averaged across all ages (with data from the two channels combined), mean binocular acuity exceeds monocular acuity by 0.13 octaves. Although the analysis revealed no significant interactions, there is a trend towards age-dependence of the effect; the binocular superiority is 0.17 octaves for the 28 infants  $\leq 24$  weeks of age, and only 0.05 octaves for the 15 infants older than 24 weeks.

Since there were no main effects of channel and no channel  $\times$  natal age or eye  $\times$  channel interactions, we have combined data from the

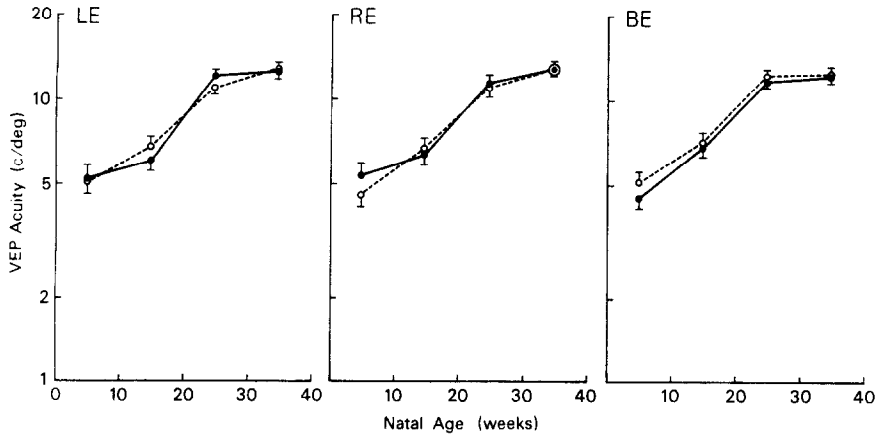


Fig. 1. Left eye, right eye and binocular growth functions for each recording channel. L-CH: ●, solid curves; R-CH: ○, dashed curves. Only four infants between 41 and 52 weeks of age were tested. Their data are combined with the data from the adjacent oldest age group in this and all other figures.

two channels in displaying the data in all subsequent figures. A second ANOVA in which BE data were excluded reveals no main effect of EYE and no interactions, indicating that development in the two eyes is identical (see "II. Interocular Acuity Differences" below). Thus, LE and RE data have been combined to form monocular growth functions in Figs 5 and 6.

Acuity development for the LE, RE and BE conditions are shown in Fig. 2 (○: LE; ●: RE; solid curve without data points: BE). Reported acuity values for each condition are the geometric mean of all acuity scores for that condition regardless of channel.

Figures 1 and 2 show that mean acuity develops from around 6 c/deg in the first month to about 14 c/deg by 8 months of age.\* This upper value is lower than that reported in the previous study (20 c/deg, Norcia and Tyler, 1985a) because we have not followed their use of the *best* significant acuity estimate for each infant. We have chosen to use the *geometric mean* of each infant's acuity estimates for the norms in the present study because they are less variable: the 95% confidence interval around the growth curve based on best-acuities is, on average, 1.35 times larger (0.43 octaves) than the 95% confidence interval around the curve based on

geometric mean acuities. When the binocular acuity data of Norcia and Tyler (1985a) are analyzed similarly, we get the dashed curve in Fig. 2. It is clear that the monocular growth curve is very similar to both binocular curves. Unless otherwise stated, all the following analyses are based on the geometric mean acuities of each infant.

## II. Interocular acuity differences

Next we address the magnitude of interocular differences determined by the sweep VEP method, and the distribution of these differences as a function of age. Figure 3 shows a scatter plot of the (signed) ratio (in octaves) of the LE

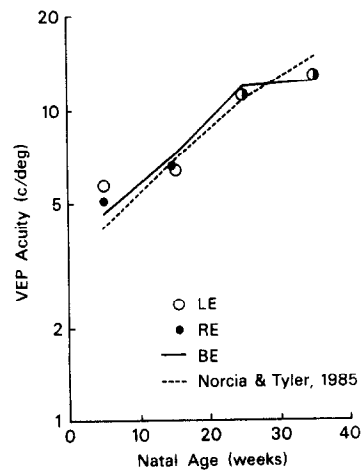


Fig. 2. Left eye (○), right eye (●) and binocular (solid curve) growth functions averaged across recording channel. For comparison, the binocular data from Norcia and Tyler (1985a) have been re-analyzed using the geometric mean acuity from each infant. This growth function is plotted as a dashed curve.

\*Only 4 of the infants tested were 41–52 weeks of age. Their data have been combined with those of the next oldest age group in these and all subsequent figures. Their geometric mean monocular acuity was 16.1 c/deg (SEM = 0.046 octaves). Their geometric mean binocular acuity was 14 c/deg (SEM = 0.117 octaves). Mean interocular differences for these 4 infants were 0.12 octaves (unsigned) with a 95% interocular confidence interval of 0.21 octaves.

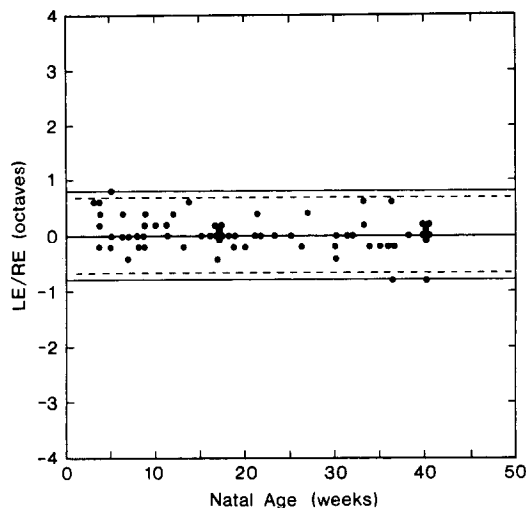


Fig. 3. Scatter plot of interocular acuity differences in octaves [ $\log(\text{LE Acuity}/\text{RE Acuity})/\log(2)$ ]. Points above zero-difference (central horizontal line) indicate that the LE acuity is higher than the RE acuity for those infants. The other pairs of horizontal lines represent the 95% (dashed lines) and 99% (solid lines). These occur at  $\pm 0.6$  and  $0.77$  octaves, respectively.

to the RE acuities as a function of natal age, for the 65 infants who were successfully tested on both conditions. Equality between the LE and RE acuities is indicated by the horizontal line at zero on the ordinate. The dashed and solid pairs of lines above and below this line delineate respectively the 95% confidence intervals ( $\pm 0.6$  octaves) and 99% confidence intervals ( $\pm 0.77$  octaves) for interocular differences. All of the points fall well within  $\pm 1$  octave of zero difference. 58 out of 65 interocular differences (89%) are within  $\leq 0.5$  octaves.

The absolute magnitudes of the interocular differences (regardless of sign) are small, averaging 0.22 octaves across all age groups. The averages of unsigned differences are plotted as a function of age in Fig. 4 (●). They range from a minimum of 0.14 (11–20 weeks) to a maximum of 0.25 octaves ( $\geq 31$  weeks). Also plotted in Fig. 4 are the group means of the individual standard errors for monocular testing, corrected to be comparable with log-acuity differences (■). Since the standard error of a difference between independent sample means (SEM) equals the square root of the sum of squares of the constituent SEMs, we have corrected each group's SEM by multiplying it by 1.414. The latter constitute estimates of the within-session, test-retest differences for each age group.

The test-retest differences represent the inherent variability of the sweep VEP acuity measure.

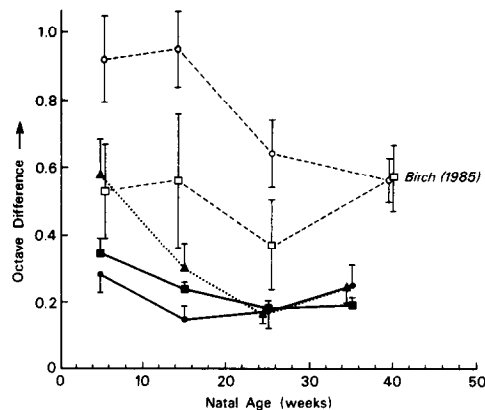


Fig. 4. Interocular and monocular test-retest differences from the present study compared to data from Birch (1985). All points are means  $\pm 1$  SEM. The top two curves are from Birch's study (○: unsigned interocular differences; □: within session test-retest differences). (●) and (■) on the bottom of the graph show the comparable data from the present study. The dotted curve and (▲) depicts the monocular, between-session test-retest data from the present study.

That these differences are nearly identical to the interocular differences (see Fig. 4, and Discussion II below) indicates that LE-RE differences are primarily determined by the test-retest reliability of the sweep VEP measure, and that true interocular differences do not exist in normal development. One would also predict from this that, within subjects, there would be no correlation between the interocular differences measured in one session and those measured in the second session. For 24 infants who completed both monocular conditions in both testing sessions, this correlation was near zero ( $r = -0.0175$ ), using the signed interocular differences, in octaves, for each session.

### III. Between-session test-retest differences

Average values (unsigned) for test-retest differences *between* the two sessions are shown in Fig. 4 as the dotted curve (▲) as a function of age. For each infant, we used the data from whichever eye was successfully tested in both sessions. When both eyes were tested in both sessions ( $N = 24$ ), the octave average of the LE- and RE-retest differences was used. A total of 37 out of the 44 retest infants completed at least one monocular retest condition.

The between-session differences exceed the two within-session (interocular, and test-retest) measures at the youngest age group ( $\leq 10$  weeks). However, for infants older than 10 weeks of age, the magnitude of between-session differences are the same as the magnitude of the

interocular differences and within-session test-retest differences. This indicates that for these older infants, between-session differences are primarily determined by the within-session variability.

The 95% confidence interval for monocular between-session differences is 0.9 octaves, averaged across all ages. For binocular testing, this interval is 0.8 octaves. However, as Fig. 4 illustrates, between-session test-retest differences vary with age. Thus, the 95% confidence interval for monocular retests in young infants ( $\leq 10$  weeks) is 1.3 octaves; for the three older age groups combined it is 0.6 octaves, identical to the 95% confidence interval for interocular differences (0.6 octaves) measured within a single session.

## DISCUSSION

### 1. Properties of monocular and binocular VEP acuity norms

**Robustness of acuity norms as a function of SNR criterion.** As described in Methods, the acuity estimates were obtained from trials in which the signal level (S) exceeded the mean noise for the trial (N) by a criterion ratio (SNR) of 3:1. To verify the robustness of this procedure, we compared acuity growth functions based on the standard criterion ( $\bullet$  in Fig. 5) with an extremely strict SNR criterion of 10:1 ( $\circ$ , dashed curves in Fig. 5). The monocular and binocular comparisons are shown in Figs 5a and 5b, respectively. In Fig. 5a, the LE and RE data have been pooled to form the monocular growth functions displayed since the analysis of variance revealed only a main effect of EYE due to a small ( $< 0.2$  octaves) binocular acuity superiority.

The strict SNR criterion disqualified many records. Of the 76 infants who had at least one successful monocular acuity according to the 3:1 criterion, 36 infants had one or more records which also satisfied the 10:1 criterion. These data are plotted as the open circles, dashed curves in Fig. 5. It is clear that the extrapolated acuities using this extremely strict criterion do not differ substantially from those in which much smaller SNRs are accepted as statistically significant ( $\bullet$ , solid curves in Fig. 5). This result emphasizes the robustness of the technique and the criteria used to score acuities. More importantly, it emphasizes the independence between the signal amplitude and the

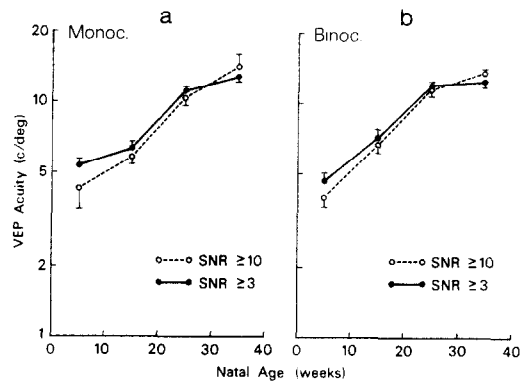


Fig. 5. Growth functions for acuities based on two SNR criteria for the monocular (a) and binocular (b) conditions.  $\text{SNR} \geq 3.0$ :  $\bullet$ , solid curves;  $\text{SNR} \geq 10.0$ :  $\circ$ , dotted curve.

extrapolated acuity value (Tyler *et al.*, 1979, 1981; Norcia *et al.*, 1985).

**Comparison with previous norms.** Previous monocular acuity norms have all come from behavioral studies using variants of the preferential looking paradigm. Figure 6 shows our monocular VEP acuities plotted in comparison with monocular acuity growth curves obtained in behavioral investigations. Two versions of the VEP data are displayed: a growth curve based on the highest criterion acuity estimate obtained from each infant ( $\bullet$  circles) as was used in Norcia and Tyler (1985a), and one based on the geometric mean acuity estimate from each infant (LE and RE data pooled,  $\circ$ ). These two curves parallel each other, separated by about  $1/2$  octave at all ages. All the previous investigations were behavioral studies and are shown as the thin-line curves identified by the symbols indicated in the figure legend (monocular behavioral data not shown in Fig. 6: Atkinson *et al.*, 1982; Maurer *et al.*, 1986; Catalano *et al.*, 1987).

As has been frequently noted for binocular acuity norms, the monocular behavioral acuities are consistently lower than the monocular VEP acuities. The highest of the behavioral norms (Mayer *et al.*, 1982; Birch and Stager, 1985) are 0.5–1.5 octaves below the VEP curve based on mean acuities (open circle curve in Fig. 6). For infants less than 12 weeks of age, the behavioral monocular norms are 1 octave (Birch, 1985) to 4 octaves (Awaya *et al.*, 1983; Yamamoto *et al.*, 1984) below our average VEP curve. Behavioral norms using the recently developed Acuity Card procedure (McDonald *et al.*, 1986) are 1.8 octaves below our mean monocular acuities, and are 2.5 octaves below the best VEP acuities ( $\bullet$ , Fig. 6).

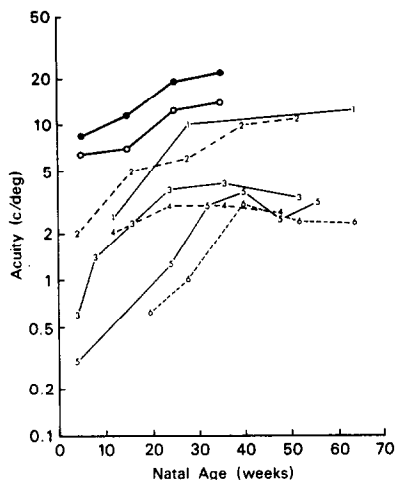


Fig. 6. Comparison of monocular acuity development norms from the present study with previous behavioral norms. The top and bottom heavy solid curve are the current norms based on the best and geometric mean acuity achieved by each infant, respectively. (1) Mayer *et al.* (1982); (2) from Birch (1985) and Birch and Stager (1985); (3) McDonald *et al.* (1986); (4) Dobson (1983); (5) Awaya *et al.* (1983); (6) Yamamoto *et al.* (1984).

Some fraction of this difference between the sweep-VEP and behavioral norms is likely to be due to methodological differences between the studies. For example, all the behavioral norms reported in Fig. 6 were based on tests using stationary gratings. Sokol *et al.* (1987) and Moskowitz *et al.* (1988) have reported that PL acuities are temporally tuned and are higher at 14 reversals per second than at zero rps (stationary gratings). Nevertheless, these authors report

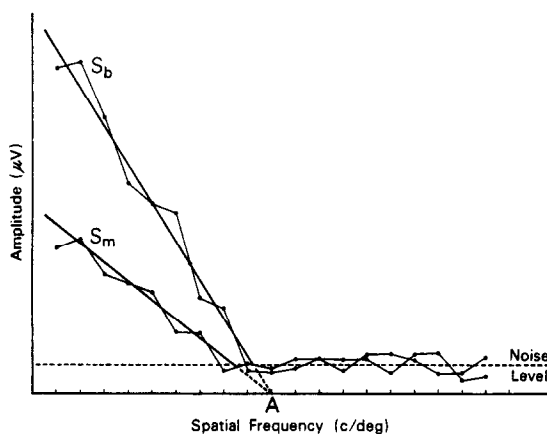


Fig. 7. Schematic representation of monocular ( $S_m$ ) and binocular ( $S_b$ ) VEP responses assuming that the binocular response is comprised of independent monocular responses, and that the monocular responses are equal. These conditions are such as to maximize the influence of probability summation on psychophysical thresholds. Yet extrapolated VEP acuities are the same ( $A$ ) for monocular and binocular viewing conditions.

that PL acuities are still lower than VEP acuities even when both are measured using 14-rps, phase-alternating gratings.

The acuity norms obtained by our methods are about 1.5–4 times less variable than previous values. The 95% (99%) confidence interval around the current norms are  $\pm 1.0$  ( $\pm 1.3$ ) octaves for infants younger than 20 weeks of age, and  $\pm 0.42$  ( $\pm 0.55$ ) octaves for infants between 20 and 52 weeks of age. The 95% (99%) confidence limits for comparable behavioral norms are relatively constant over the first year of life and are considerably larger: i.e. about  $\pm 1.6$  ( $\pm 2.1$ ) to  $\pm 1.8$  ( $\pm 2.4$ ) octaves for infants under a year of age (Birch, 1985; McDonald *et al.*, 1986; Mayer *et al.*, 1982).

*Binocular and monocular VEP acuities.* Our analyses in RESULTS revealed a small but significant main effect of eye. From the cell means in the ANOVA ( $N = 43$ ), it is clear that the effect is due to a slight superiority of binocular over monocular acuities. At most, binocular acuities exceed monocular acuities by 0.17 octaves; and this is in the group of 28 infants under 24 weeks of age. The 15 infants over 24 weeks of age have virtually equal monocular and binocular acuities. This pattern is also evident in Fig. 2 where data from all 87 infants are included. This is in dramatic contrast with the previous behavioral studies of monocular and binocular acuity development which found that binocular acuities were from 0.6 to 1 octave higher than monocular acuities in infants older than 5 months (Birch, 1985; McDonald *et al.*, 1986), and that this binocular superiority was either absent (Birch, 1985) or less pronounced in younger infants (Atkinson *et al.*, 1982).

There are several methodological considerations which could conceivably result in a superiority of binocular over monocular acuities in behavioral testing, but which are less likely to affect sweep VEP acuities. First, infants are less likely to be cooperative when they are wearing an eye patch; this is a factor in both behavioral and VEP testing. This factor may account, in part, for the lower success rate that we achieved in monocular testing. However, a single successful sweep trial can provide a VEP acuity estimate, requiring only 10 sec of sustained looking behavior, whereas behavioral measures require a series of trials to complete one estimate. Thus, for those infants who will submit to patching at all, the VEP test may require substantially less overall cooperation per acuity estimate than is required in behavioral testing.

Second, if infants are less cooperative during monocular testing, this could add noise to the EEG. The effect of added noise would be to eliminate some records that would otherwise have reached the 3:1 SNR criterion. Since each acuity is estimated by a linear extrapolation to  $0 \mu\text{V}$  (and *not* to the mean noise level), records that exceed the 3:1 criterion under the noisier condition should extrapolate to the same acuity value (Tyler *et al.*, 1979; 1981; Norcia and Tyler, 1985a; Norcia *et al.*, 1985). In behavioral testing, added noise due to the monocular test conditions will tend to decrease the acuity estimates by decreasing (for example) the percent correct at a given spatial frequency in a forced-choice preferential looking measure.

Third, behavioral estimates of binocular acuity may be increased by probability summation between the eyes. In psychophysical tests, probability summation has been argued to account for the superiority of acuity and contrast sensitivity in binocular viewing in adults (Campbell and Green, 1965). If the underlying psychometric function is shallow, probability summation across two independent detection mechanisms can greatly increase sensitivity. In fact, if slope of the psychometric function (i.e. "beta", as formulated by Quick, 1974) is unity, addition of a second independent detection mechanism can double sensitivity, mimicking full energy integration (Gorea and Tyler, 1986). In principle, then, probability summation can account for a large superiority of binocular acuity in behavioral tests since infants are known to have shallow psychometric functions (Mayer and Dobson, 1982; Hamer and Schneck, 1984). However, other factors (such as difficulties associated with monocular testing) are likely to be contributing some (unknown) component of the binocular advantage, since it is observed in infants under 6 months of age under some conditions (McDonald *et al.*, 1986).

The possible role of probability summation in VEP recordings has not been explicitly addressed in previous research. Consider the case where binocular responses are comprised of independent monocular responses with the same underlying spatial resolution, in analogy to the psychophysical model. The amplitude response of such a patch of cortical tissue under monocular viewing conditions (either eye) is shown schematically as the curve  $S_m$  in Fig. 7. It has an extrapolated acuity,  $A$ . Under binocular viewing conditions, an independent group of cells will contribute to the overall response measured.

Since the electrode simply records the volume-conducted voltage responses from all sources, addition of a second population of cells, with the same acuity cutoff will increase the response at each spatial frequency by linear summation of the two population responses, yielding sweep amplitude response ( $S_b$ ). However, the extrapolated acuity ( $A$ ) will remain unchanged.

This hypothetical relationship between the binocular and monocular VEP acuities is borne out by the analyses of variance described in Results I (the only binocular acuity superiority found was 0.17 octaves occurring in infants under 6 months of age), and is illustrated by the congruence of the monocular and binocular data presented in Fig. 2. In a separate analysis, we calculated the ratio of binocular to monocular acuities for each of 76 infants who successfully completed at least one of each of these two conditions. Averaged across all age groups, binocular acuity was nearly identical (only 0.014 octaves higher) to monocular acuity. Only in the 11–20 week age group did the binocular superiority occur (0.17 octaves higher than monocular).

Further confirmation of this formulation comes from the data of Orel-Bixler and Norcia (personal communication) who found a small binocular acuity superiority in children of 3–11 yrs of age. These children had no significant refractive errors and normal eye alignment. Thus, they presumably had binocular vision and stereopsis; yet sweep VEP measures of acuity revealed binocular acuities that were, on average, only a factor of 1.09 (0.12 octaves) higher than monocular acuities. The small binocular acuity superiority observed may be due to state variables such as a slight increase in accommodative inaccuracies in monocular viewing, or rivalry.

If this particular invariance of the VEP acuity extrapolation method, i.e. the near equality of monocular and binocular acuities, proves to have generality over other testing conditions, it may simplify the application of the technique. It means that in the absence of monocular acuity norms for a particular testing condition, the binocular acuity norms may provide a reasonable approximation for monocular comparison values.

## II. Test–retest variability is relatively small

The variability inherent in test–retest measures sets a reasonable lower-bound estimate for

interocular acuity differences, since differences between eyes will, on average, tend to be larger than those within an eye. Conversely, measured interocular acuity differences can be considered upper bound estimates of test-retest differences; again, on average, differences within the same eye will tend to be less than or equal to differences between eyes. This, of course, assumes that the within-eye retests and interocular comparisons are performed under conditions that would equate time or practice effects, as well as the effects of adaptation. These are reasonable assumptions for our data.

Thus, we can use the measured (within-session) interocular acuity differences as upper-bound estimates of monocular retest differences within a session. Since the average (unsigned) interocular difference was 0.22 octaves, retests of an eye within a session should differ from an initial acuity estimate by no more than this value. By similar reasoning, the upper bound for the 95 and 99% confidence limits for test-retest within a session should be 0.6 and 0.77 octaves, respectively.

Retesting infants in a separate session adds some variance to the measures. This may be due to a number of factors, such as state changes of the infants from session to session or slight variations in electrode placement. By comparing within-session and between-session interocular differences, we can calculate the contribution to the variance that is due to testing in a separate session. In this calculation, we assume that the variability due to testing in a separate session is independent of the sources of variability within a session; hence the variances deriving from these two sources of variability add. In young infants ( $\leq 12$  weeks), for example, the standard deviation of within-session interocular differences (signed) was 0.33 octaves corresponding to a variance of 0.109 octaves. For between-session interocular differences, the standard deviation was 0.65 octaves (not shown in Fig. 4), i.e. a variance of 0.423 octaves. Subtracting the former variance from the latter, we find that testing in a separate session adds 0.31 octaves to the (within-session) between-eye variance (or 0.56 octaves standard deviation). A similar calculation for infants older than 12 weeks indicates testing in a separate session contributes only a small component (0.017 octaves) of the overall (between-session, interocular) variance (0.091 octaves). Stated another way, for infants older than 12 weeks, the distribution of monocular test-retest differences

was nearly identical to that for interocular differences within a single session ( $\pm 0.6$  octave 95% confidence interval).

Previous studies have found substantially larger test-retest differences than those found in the present study. For example, Birch's (1985) within-session acuity retests (unsigned) averaged about 0.5 octaves across all ages; the comparable value from the present study was 0.23 octaves, averaged across all ages. McDonald *et al.* (1986) reported, using the acuity card procedure, that 47% of their sample (36 infants from 1 to 12 months of age) had between-session acuities that were 1 octave apart under binocular viewing, whereas none of the 43 successful binocular retest acuities in the present study exceeded 0.9 octave.

For monocular testing, McDonald *et al.* (1986) found about 10% of monocular retests yielded acuity differences of 2 octaves. This was the same for interocular and same-eye comparisons. In the current study, we find only 2 out of 61 same-eye comparisons (30 LE and 31 RE) had between-session differences greater than one octave (but  $< 1.3$  octaves). Within-session retest differences from the current study are substantially smaller: e.g. the upper-bound estimate of within-session test-retest differences, determined from the distribution ( $N = 65$ ) of interocular acuity differences, never exceeded 0.9 octaves.

Some other studies that have measured monocular acuity development in infants have not directly measured test-retest acuity differences. However, the interocular differences measured in these studies provide an upper-bound estimate of these differences. Dobson (1983) found 35% of interocular differences in infants 3 to 12 months of age to be  $\geq 1$  octave. On the other hand, Mayer and Fulton (1985) and Mayer *et al.* (1982) found averages interocular differences of only 0.4 octaves, with only 4% of older infants (average age 7.3 months) showing a difference of  $> 1$  octave.

### III. No significant interocular acuity differences

We find that interocular acuity differences measured within a single testing session are small (averaging 0.22 octaves, unsigned, across all age groups), and do not change significantly with age. From the distribution of signed interocular differences, the 99% confidence interval is  $\pm 0.77$  octaves—any interocular difference larger than this should be considered abnormal.

Previous norms for interocular acuity differences have come from behavioral studies. Comparable 99% confidence limits from these studies are 2–4 times larger ( $\geq 1.6$  octaves—Mayer and Fulton, 1985; Dobson, 1983; 1.8 octaves—Birch, 1985; 3.2 octaves—McDonald *et al.*, 1986). Birch (1985) found that infants younger than 3 months postnatal had significant interocular differences (approx.  $\pm 1$  octave, unsigned), i.e. their interocular differences were significantly greater than within-session (within-eye) test–retest differences (0.5 octaves, unsigned). Average interocular differences for older infants were about half the value for the younger infants. For comparison, we have reproduced Birch's (1985) data as the dashed curves in Fig. 4 (○: unsigned interocular acuity differences; □: within-session test–retest differences).

Atkinson *et al.* (1982) also found that interocular differences in a group of 16-week-olds were significantly larger than test–retest differences.

Based on these findings in young infants, both Atkinson *et al.* (1982) and Birch (1985) have made the suggestion that, early in life, the eyes may be at slightly different stages of development, leading to the presence of significant interocular acuity differences found in young infants. Birch further suggests that there may be a causal link between the onset of binocular function, including stereopsis, and the decrease in interocular differences. The implication is that the factors leading to the rapid increase in grating acuity over the first 6 months of development operate independently in the two eyes; at the onset of normal binocular interaction, interocular differences are diminished, bringing the two eyes into equality.

It is possible that our failure to find a significant interocular difference in the younger age groups is due to a kind of sampling bias. The 22 infants from whom we failed to obtain both monocular acuity estimates (and hence failed to measure an interocular difference) may have actually been the infants with real interocular differences who were simply less testable in their "bad" eye. However, data from the group of 44 infants who participated in 2 sessions argues against this possibility. Out of 20 infants who failed to yield an interocular difference in one session, 17 yielded an interocular difference in the other session. The 99% interocular confidence interval for these 17 infants was  $\pm 0.61$  octaves, i.e. comparable to

(actually smaller than) the 99% interocular confidence interval ( $\pm 0.77$  octaves) for the larger group of infants who yielded an interocular difference in one session. Furthermore, for the 6 infants out of the 17 who were under 13 weeks age, interocular differences were nearly identical (99% confidence interval =  $\pm 0.65$  octaves) to those of the larger population of infants. Thus, even quite young infants who fail to yield an interocular difference in one session, show a "normal" (i.e. insignificant) interocular difference in another session.

Our data suggest that the measured interocular differences are primarily determined by the inherent monocular variability of the sweep VEP measure. This view is born out by the near-zero correlation between the interocular differences measured in one session and those measured in another session.

In addition, we have calculated the 95% confidence interval for the *between-session* interocular differences (i.e. LE session 1 vs RE session 2, and vice versa). This confidence interval is 0.9 octaves, identical to the monocular test–retest confidence interval (i.e. LE session 1 vs LE session 2, or RE session 1 vs RE session 2). This is another indication that, in a normal population of infants, interocular acuity differences primarily are determined by the inherent variability of the VEP measure.

Thus, the present data are not consistent with a model of independent monocular acuity development up to the level of the cortical VEP generators. We find no significant interocular acuity differences at any age—VEP interocular acuity differences are small, are nearly constant over age (Figs 3 and 4), and do not differ significantly from the within-session retest error (Fig. 4). In fact, the estimated values for within-session test–retest differences (●, Fig. 4) fall slightly below the interocular data (■, Fig. 4); the opposite result would be required to establish the existence of significant interocular differences.

## CONCLUSIONS

This article presents norms for the development of acuity under monocular viewing conditions for the sweep VEP technique, and replicates the prior norm for binocular viewing obtained using the same technique (Norcia and Tyler, 1985a). For the current stimulus conditions, sweep VEP acuity (based on geometric means) develops from about 6 c/deg at 2–10

weeks of age to 14 c/deg by 20–30 weeks of age, for both monocular and binocular viewing. It is emphasized that these norms are robust and relatively independent of the peak SNR achieved (Tyler *et al.*, 1979; 1981; Norcia *et al.*, 1985). The norms based on the best criterion acuity achieved by each infant are about 1/2 octave higher.

The average acuities measured in the present study are at least 0.5–1 octave higher than the highest of the previous norms, and are as much as 4 octaves higher than some published data. In addition, the confidence limits around the current values are 1.5–4 times smaller than prior estimates. These two factors mean that, for infants under a year of age, much more contrast information is present at the level of the visual cortex at high spatial frequencies than is indicated by behavioral measures of monocular acuity development.

The use of two recording channels increased the likelihood of successfully recording a criterion acuity on any given trial; but the advantage conferred by a second channel was less than that predicted if the two channels were independent. To the extent that the channels are independent, this electrode configuration allows for detection of hemispheric asymmetries and, in monocular testing, asymmetries between crossed and uncrossed pathways.

We find interocular differences to be nearly invariant with age (Fig. 4), and not significant at any age. Thus, our data do not support the concept of independent acuity development in the two eyes (Atkinson *et al.*, 1982; Birch, 1985). The low variability of our measures constrains any developmental differences between eyes prior to the site of the cortical VEP generators to an average value of values less than 1/4 octave.

The absolute magnitude of interocular and test–retest differences are small, averaging about 1/4 octave (unsigned) in infants older than 10 weeks of age. These values are 2–4 times smaller than current behavioral estimates. The confidence intervals for the distribution of interocular and test–retest differences ( $< \pm 0.8$  octaves for 99% confidence) are also 2–4 times smaller than comparable behavioral measures. The 99% behavioral confidence limits for interocular differences range from 1.5 to 3.0 octaves. Interocular differences of this magnitude represent 5–10 standard deviations of the interocular variability in the sweep VEP method. This indicates that the sweep VEP measures of gra-

ting acuity are, in principle, capable of detecting smaller differences between eyes and smaller changes between tests than reported behavioral measures.

We find only a small ( $< 0.2$  octaves) superiority in binocular VEP acuities over monocular VEP acuities occurring only in infants under 6 months of age. This is in striking contrast to results from behavioral studies which measure 0.6–1 octaves of binocular acuity superiority (Birch, 1985; Birch and Stager, 1985; McDonald *et al.*, 1986). We contend that grating acuity at the level of the cortical VEP generators is not subject to the effects of detection probability summation, which has been argued to account for the psychophysical superiority of acuity and contrast sensitivity in binocular viewing (Campbell and Green, 1965). We point out that, since infants have quite shallow psychometric functions (Mayer and Dobson, 1982; Hamer and Schneck, 1984), probability summation across two independent detection mechanisms can, in principle, account for a large magnitude of binocular acuity superiority in behavioral tests, although other factors are almost certainly involved.

A test such as the sweep VEP technique, having relatively small test–retest variability, narrow confidence limits, and yielding high acuity and contrast sensitivity values, may prove to be a sensitive monitor of visual loss in the pediatric population.

*Acknowledgements*—The authors thank the FAMCAP parenting group leaders from Kaiser Hospital, San Francisco, CA, for help in recruiting infant observers for this study. This work was supported by NIH grants EY3622, EY1186 and RR5566, with the cooperation of the Kaiser Research Institute, Oakland, CA.

## REFERENCES

- Atkinson J., Braddick O. and Pimm-Smith E. (1982) "Preferential looking" for monocular and binocular acuity testing of infants. *Br. J. Ophthalmol.* **66**, 264–268.
- Awaya S., Sugawara M., Kodama A., Yagasaki T., Oishi F. and Hirai T. (1983) The preferential looking method for visual acuity assessment in infants with special reference to normal curve, visual acuity difference between the two eyes and success rate in measurement. *Folia Ophthalm. Jpn.* **34**, 1160–1165.
- Birch E. (1985) Infant interocular acuity differences and binocular vision. *Vision Res.* **25**, 571–576.
- Birch E. and Stager D. R. (1985) Monocular acuity and stereopsis in infantile esotropia. *Invest. Ophthalm. visual Sci.* **26**, 1624–1630.
- Campbell F. W. and Green D. G. (1965) Monocular versus binocular visual acuity. *Nature, Lond.* **208**, 191–192.
- Catalano R. A., Simon J. W., Jenkins P. L. and Kandel G. L. (1987) Preferential looking as a guide for amblyopia

- therapy in monocular infantile cataracts. *J. Ped. Ophthalm. Strab.* **24**, 56–63.
- Dobson V. (1983) Clinical applications of preferential looking measures of visual acuity. *Behav. Brain Res.* **10**, 25–38.
- Gorea A. and Tyler C. W. (1986) New look at Bloch's law for contrast. *J. Opt. Soc. Am. A* **3**, 52–61.
- Hamer R. D. and Schneck M. E. (1984) Spatial summation in dark-adapted human infants. *Vision Res.* **24**, 77–85.
- Harris F. J. (1978) On the use of windows for harmonic analysis with the Discrete Fourier Transform. *Proc. IEEE* **66**, 53–59.
- Maurer D., Lewis T. L. and Brent H. P. (1986) The effects of deprivation on human visual development: Studies of children tested for cataracts. In *Applied Developmental Psychology*, Vol. 3 (Edited by Morrison F. J., Lord C. E. and Keating D. P.). Academic Press, N.Y.
- Mayer D. L. and Dobson V. (1982) Visual acuity development in infants and young children, assessed by operant preferential looking. *Vision Res.* **22**, 1141–1151.
- Mayer D. L. and Fulton A. B. (1985) Preferential looking grating acuities of infants at risks of amblyopia. *Trans. Ophthalm. Soc. U.K.* **104**, 903–911.
- Mayer D. L., Fulton A. B. and Hansen R. M. (1982) Preferential looking acuity obtained with a staircase procedure in pediatric patients. *Invest. Ophthalm. visual Sci.* **23**, 538–543.
- McDonald M., Sebris S. L., Mohn G., Teller D. Y. and Dobson V. (1986) Monocular acuity in normal infants: the acuity card procedure. *Am. J. Optom. physiol. Optics* **63**, 127–134.
- Moskowitz A., Sokol S. and McCormack G. (1988) Infant VEP and PL grating acuity for different rates of phase-alternation. *Invest. Ophthalm. visual Sci.* **29**, 25.
- Norcia A. M. and Tyler C. W. (1985a) Spatial frequency sweep VEP: Visual acuity during the first year of life. *Vision Res.* **25**, 1399–1408.
- Norcia A. M. and Tyler C. W. (1985b) Infant VEP acuity measurements: analysis of individual differences and measurement error. *Electroenceph. clin. Neurophysiol.* **61**, 359–369.
- Norcia A. M., Clarke M. and Tyler C. W. (1985) Digital filtering and robust regression techniques for estimating sensory thresholds from the evoked potential. *IEEE Engineering in Medicine and Biology December, 1985*, pp. 26–32.
- Quick R. F. (1974) A vector magnitude model of contrast detection. *Kybernetik* **16**, 65–67.
- Sokol S., Moskowitz A. and Augliere R. (1987) Preferential looking acuity is temporally tuned. *Invest. Ophthalm. visual Sci.* **28**, 5.
- Tyler C. W., Apkarian P. A., Levi D. M. and Nakayama K. (1979) Rapid assessment of visual function: an electronic sweep technique for the pattern VEP. *Invest. Ophthalm. visual Sci.* **18**, 703–713.
- Tyler C. W., Nakayama K., Apkarian P. A. and Levi D. M. (1981) VEP assessment of visual functions. *Vision Res.* **21**, 607–609.
- Yamamoto M., Kanagawa M. and Okuda T. (1984) Preferential looking of infants and younger children with ocular diseases. *Acta Soc. Ophthalm. Jpn* **88**, 885–890.