Test-Retest Reliability of Swept Visual Evoked Potential Measurements of Infant Visual Acuity and Contrast Sensitivity

LOTTE LAURITZEN, MARIANNE HØRBY JØRGENSEN, AND KIM FLEISCHER MICHAELSEN

Center for Advanced Food Studies, Department of Human Nutrition, The Royal Veterinary and Agricultural University, 1958 Frederiksberg, Denmark

ABSTRACT

The aim of the study was to describe variations in swept visual evoked potential (SWEEP-VEP) assessment of visual acuity and contrast sensitivity in infants and to evaluate the best way to estimate visual performance from obtained SWEEP-VEP data. The visual performance of 92 infants (6–40 wk of age) was measured in two separate visits. Results were verified with repeated tests in seven adults. There was a strong association between the two measurements of infant visual acuity ($r = 0.91$, $p < 0.001$), with no constant bias and an inter-assay coefficient of variation of 8.4%. The intra-assay coefficient of variation was 17% and in repeated sessions all obtained acuity measures were normally distributed, indicating that the mean and not the maximum threshold best estimates visual acuity. This estimate of visual acuity also had lower test-retest variability than those calculated from the maximum threshold or threshold from the average EEG signals ($p = 0.001$). Test-retest measures of infant contrast sensitivity had a correlation coefficient of 0.72 ($p < 0.001$) and an inter-assay coefficient of variation of 23%. With the observed test-retest variability, SWEEP-VEP is less valid for estimating the visual performance of individual subjects, but it can give reliable group means. This method was well suited to describe visual development in the infants, which for acuity as well as contrast sensitivity increased by 0.64 octave per doubling in age. However, the variability of the SWEEP-VEP method can be a limiting factor, for example, in the assessment of the potential effect of dietary docosahexaenoic acid in a homogeneous group of infants. (*Pediatr Res* **55: 701–708, 2004**)

Abbreviations
cpd, cycles per degree of the retina
DHA, docosahexaenoic acid
SWEEP-VEP, swept visual evoked potential

Membranes in the retina and brain have a high content of DHA, a long-chain *n*-3 fatty acid derivative of α-linolenic acid. These fatty acids cannot be synthesized by humans and, therefore, it is essential that the diet contains *n*-3 fatty acids in order for DHA to accumulate in the CNS. In the last decade, there has been a great scientific interest in the importance of DHA for the development of visual acuity in human infants (1). Many studies have not been able to show a significant beneficial effect of DHA and one of the explanatory factors may be limitations in the methods used to assess infant visual acuity (1–5). The studies that do find a significant difference tend to use EEG measures of visual evoked potentials (VEP), rather than the Teller acuity cards. Meta-analysis of data from term infant studies found that the acuity difference for EEG-based measures at 4 mo of age in infants consuming a source of DHA compared with infants on a DHA-free diet was 0.26 ± 0.10 octaves (6). Compared with the typical visual acuity in a group of 4-mo-old infants (7–11), this difference is equivalent to approximately 1.2 cpd or approximately 1 SD. We have previously found that milk DHA could explain 9% of the variation in visual acuity in a group of 4-mo-old breast-fed infants (12).

Norcia and Tyler (13) have invented a rapid method of measuring VEP acuity, which, during a 10-s trial, sweeps through gratings with a spatial frequency well below to well above the acuity limits of the subject (SWEEP-VEP). By this method, thresholds in the visual detection are obtained by extrapolating the function relating VEP-amplitude to spatial frequency to noise level. Depending on the number of trials and recording points, this method produces several estimates of the visual threshold of the subject. SWEEP-VEP data have been used in different ways in the estimation of infant visual acuity. In their original setup of the SWEEP-VEP procedure, Norcia and Tyler (13, 14) suggested that the best approximation of infant visual acuity was the highest of all thresholds obtained in any trial or recording point during an assessment. Other investigators have used averages of all EEG signals in a

recording point to extrapolate the visual threshold (the average of which we refer to in this article as "signal-average acuity") (8, 11). The SWEEP-VEP method has been evaluated for the purpose of describing visual development with age in infants (13–15). However, the method has not been thoroughly tested in relation to detecting differences among healthy infants of the same age, as this has only been examined in one small study (15).

SWEEP-VEP has also been used to determine infant contrast sensitivity (16). This property of infant vision has not been studied as much and only a few studies have dealt with whether or not contrast sensitivity can be modified by DHA (17).

To understand the biologic implications of dietary DHA it is necessary to know how much of the obtained variation in infant visual performance is due to variations in the used visual test. This may also be relevant in relation to other areas of visual research. The aim of the present study was to describe the variation of SWEEP-VEP measures of visual acuity and contrast sensitivity in infants. We also wanted to evaluate the importance of how the obtained data could be used in the determination of visual limitations. To do this we examined the visual limitations of a number of infants twice and verified the results with repeated tests in adult subjects.

MATERIALS AND METHODS

Participants and Protocol

A random sample of infants from 5.7 to 39.4 wk of age was identified through the Danish National Birth Register. To be included in the study, the infants had to be born healthy (no admission to a neonatal department), at term, with normal weight for gestation and an Apgar score >7 at 5 min after delivery. The study was approved by the Scientific-Ethical Committees of Copenhagen and Frederiksberg (KF 01–134/95). The mothers of 101 infants agreed to participate and signed an informed consent form; 92 infants completed the study (70% of whom were above 17 wk of age).

Mothers and infants were assessed twice [on average 2.6 ± 2.4 (range, 1–14) days apart] at the Department of Human Nutrition. On the first visit the infant was weighed and measured and the mother was interviewed about her pregnancy, delivery, fish intake, and how she fed her infant. Characteristics of these infants and their mothers are shown in Table 1.

SWEEP-VEP Measurement

The visual acuity and contrast sensitivity of the infants were tested on both visits.

Visual acuity. Binocular visual acuity was assessed by SWEEP-VEP using the NuDiva system (14) equipped with an M2400 high-resolution monochrome monitor (Dotronix, New Brighton, MN, U.S.A.). The subject (an infant on the lap of a parent) was presented with vertical 6-Hz sinewave gratings at 80% contrast and in a field of 35×26 cm^2 with a mean luminance of 47.6 cd. The spatial frequency of the gratings was increased linearly 10 times during 10-s trials. The viewing distance and range of spatial frequencies depended on the age of the subject (Table 2). Visual evoked potentials were re-

Table 1. *Characteristics of study subjects*

No.	92
Maternal age (y)	30.8 ± 4.1
Height (m)	1.69 ± 0.06
BMI (kg/m^2)	22.9 ± 2.8
Frequency of fish intake (pr. mo)	5.7 ± 3.8
Weight gain during pregnancy (kg)	15.8 ± 5.5
Total number of infants	1.7 ± 0.9
Way of birth (% cesareans)	16.3
Infant sex (% female)	54.3
Gestational age (wk)	39.9 ± 1.5
Birth weight (g)	3541 ± 504
Length (cm)	52.4 ± 2.4
Apgar score after 5 min	9.9 ± 0.5
Infant age at time of study [(range) weeks]	21.3 ± 9.3 (5.7–39.4)
Breast-fed at time of study (% yes)	81.5

BMI, body mass index.

Table 2. *Employed viewing distances, spatial frequency (freq.) and contrast sweeps in different age groups of infants*

Age (mo)	View distance (cm)	Spatial freq. at SWEEP start (cpd)	Spatial freq. at SWEEP end (cpd)	Contrast at SWEEP start (%)	Contrast at SWEEP end (%)
1	70	0.5	8	40	1
2	70	0.6	10	40	1
4	100	1.0	16	25	1
6	100	1.3	20	25	1
8	100	1.5	25	20	0.7
Adult	150	2.0	30	—	—

corded with a sampling rate of 450 Hz by gold EEG electrodes attached to the scalp over the visual cortex (9). The attention of the infant was attracted to the screen by small toys or bells and trials were interrupted if the infant looked away from the stimuli.

The EEG was amplified at a gain of 10,000–20,000 in infants and 50,000 in adults. The electrode montage consisted of three channels (O1, Oz, and O2 of the International 10–20 system) each referenced to Oz and two bipolar channels (O1 and O2 each referenced to Cz), as described by Jørgensen *et al.* (9). The amplitude and phase of the VEP were calculated using an adaptive filter technique described by Tang and Norcia (18). Visual acuity was estimated by linear extrapolation of the VEP amplitude *versus* spatial frequency function recorded at 12 Hz to 0 amplitude (19). Individual 10-s sweep records were averaged coherently across trials by addition of the sine and cosine components from each bin as described in detail by Norcia *et al.* (19). The records from the individual trials and vector-average record for each channel were scored automatically using empirically established criteria for distinguishing VEP response from the spontaneous EEG (19). Briefly, the signal-to-noise ratio had to be three or greater for two data points (using noise measured during the trials at 11 and 13 Hz). Additionally, the response phase had to be constant or gradually decreasing lag time between stimulus and VEP as visibility increased. Computer scores were checked manually for errors by one trained observer and manual changes were made if the linear regression was a poor fit to the record. Rescoring most often resulted in more conservative measures of threshold than those produced by the automatic algorithm. Each session could

optimally result in five threshold estimates per trial and an additional five signal average thresholds. We aimed at five trials per session or more if the first five trials did not give 10 successful extrapolations. In some cases, it was not possible to reach these predefined goals. In these cases, we used the number of trials and scores that we could get while the infant was attentive. Infants were given on average 6 ± 2 trials (range, 4–16), which gave on average 20 ± 6 scores per session (range, 1–40, $n = 178$). Unless otherwise stated, visual acuity is given as the mean of all obtained thresholds. One of the infants was inattentive and not testable at both the visits (2.2% of all possible tests) and 89 infants completed the visual acuity test successfully on both visits.

To verify the intra- and interassay variation obtained from double determinations of SWEEP-VEP acuity in infants, we measured the visual acuity of seven female adults (25–45 y of age) on different days. We also tested one of the infants thrice a month (at 20.7, 22.1, and 24.6 wk of age).

Contrast sensitivity. Contrast sensitivity was also assessed by SWEEP-VEP. The system settings were, apart from the following specifications, similar to that of the visual acuity assessments. When we recorded contrast sensitivity we used vertical sinewave gratings with a constant spatial frequency of 1 cpd at a mean luminance of 47.6 cd. During the 10-s trials, the contrast of gratings was increased in 10 logarithmic steps and pattern-reversed at a rate of 6.0 Hz. The view distance and range of contrast depended on the age of the subject (Table 2). Contrast threshold was estimated by linear extrapolation to zero amplitude of the VEP amplitude *versus* log contrasts function in the second harmonic (16). The signals from the individual trials and averages for each channel were scored automatically by the NuDiva system and checked manually for errors. Our goal was five trials and minimum 10 successful extrapolations and, on average, the infants were given 5.1 ± 0.7 trials (range, 3–9; $n = 82$) and had 13 ± 8 scores (range, 0–30). The contrast sensitivity test was of lower priority (always performed after the grating acuity test) and many of the infants were not tested as they were tired after the assessment of visual acuity. Furthermore, the contrast sensitivity test resulted in fewer successful threshold extrapolations and 6.1% of the tests did not result in any successful extrapolations. In 48 infants, we have at least one estimate of the visual contrast sensitivity, and 28 infants completed the contrast sensitivity test successfully at both visits.

Statistical Analysis

All results are given as mean $\pm$ SD. The analysis of measurement error is performed according to Bland and Altman (20–22). Statistical analysis was performed with the Kolmogorov-Smirnov test, linear regression analysis, and *t* test for paired samples using SPSS (version 10.0, SPSS Inc., Chicago, IL, U.S.A.).

RESULTS

Test-retest of infant visual acuity. The test-retest analysis was initially performed on measures of visual acuity determined as the mean of all successful threshold extrapolations in each of the two sessions of assessment. There was a strong association between the visual acuity measures in the two visits (Fig. 1) and no constant bias (the mean difference being 0.05 ± 1.29 cpd, $n = 89$). The intraclass correlation coefficient was found to be 0.88. The mean relative difference, the test-retest variability, was $0 \pm 16\%$, giving a 95% prediction interval (± 1.96 SD) of $\pm 31\%$ (Fig. 1, *inset*). The size of the difference did not depend on the total number of successful extrapolations, age, or visual acuity (data not shown).

To further evaluate the differences in the visual acuity measures relative to the variation among the infants, we calculated the percentile ranking of the infants at the first and second assessment, and the mean difference in ranking was 9.1 ± 7.7 percentiles. The absolute numeric difference between the two visual acuity measures was on average 1.0 ± 0.8 cpd (or 0.19 ± 0.13 octave), but was associated with the visual acuity of the infant (Fig. 2), in contrast to the relative difference (Fig. 1, *inset*). Based on the test-retest differences, one can estimate the SD on the infant visual acuity measure [SD $= \sqrt{(\Sigma\text{difference}^2)/2 \cdot n}$]. Because only the relative differences were constant over the entire range of visual acuities, we used the log-transformed visual acuity (logMAR) in this calculation, as suggested by Bland and Altman (22). The obtained interassay coefficient of variations [CV% $= $ (SD/mean visual acuity) $\cdot 100$] was 8.4%. The variation coefficient on all estimates of infant visual acuities obtained within a session (*i.e.* the intraassay variation) was on average $17 \pm 5\%$ ($n = 89$).

Estimation of visual acuity from SWEEP-VEP data. We compared the visual acuity calculated as the mean threshold of
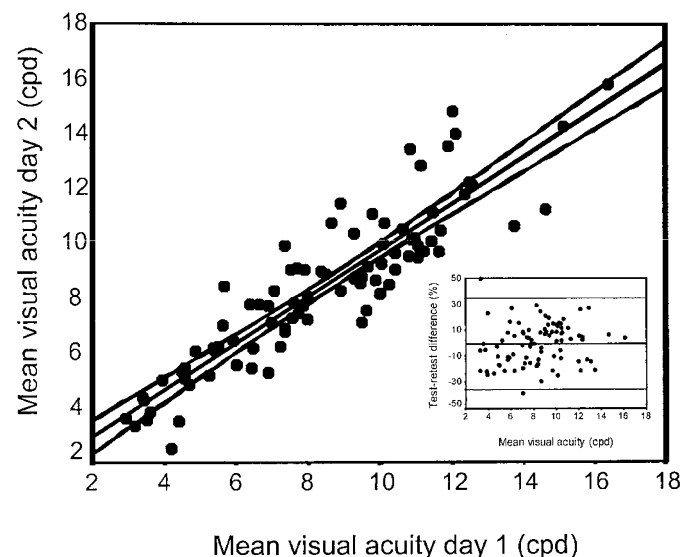

Figure 1. Visual acuity of 89 infants at the second assessment *vs* that determined at the first session (2.6 ± 2.4 d before). The data points represent visual acuity given as the mean threshold of all successful determinations. Linear regression gives a regression coefficient for the correlation between the two assessments of 0.91 ($p < 0.001$). The linear regression line is drawn in the plot with 95% confidence intervals. The ratio between these two visual acuity measures was on average 1.01 ± 0.17 ($n = 89$). (*Inset*) The individual discrepancies between the two visual acuity measures are expressed relative to the measured mean visual acuity of the subject. The center line depicts the average relative difference (constant bias) and the outer lines depict the measurement error 95% prediction interval (1.96 SD).
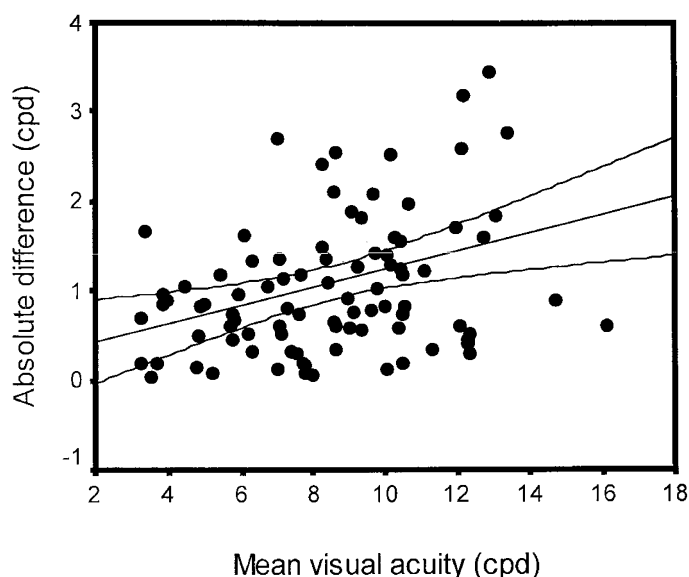
Figure 2. Relationship between the individual differences between visual acuity from the two assessments *vs* the mean visual acuity of the subject. The differences are expressed as absolute numeric differences in the measured mean acuities. The linear regression line (with 95% confidence intervals) has a coefficient of 0.32 ($p = 0.002$, $n = 89$).

all successful extrapolations (mean acuity) with visual acuity assigned by the best of the obtained threshold values (max acuity) or by the mean of the 3, 5, or 10 best values and with the mean threshold obtained by extrapolation of the average EEG signal in each of the five channels (signal-average acuity). Table 3 gives visual acuities obtained by each of these methods (mean for all infants) and also for each measure the absolute numerical and relative difference between the two measurements. Visual acuity calculated as mean acuity had the lowest relative difference. The difference in acuity calculated as means of the 10 best scores was also low, but this way of calculating visual acuity results in fewer successful assessments as not all infants obtained 10 scores. The relative difference on the mean acuity assessment was significantly lower than that on the signal-average acuity estimate ($p = 0.001$), and also tended to be lower when compared with max acuity ($p = 0.13$).

Visual acuity measurements in adults. The SWEEP-VEP visual acuity assessments in seven adult subjects also showed large differences both within and between days (Table 4). The interassay coefficient of variation, being $8 \pm 4\%$, on visual acuities determined by the mean acuity method was lower than

Table 3. *Comparison of test-retest and population variation in infant visual acuity determined by SWEEP-VEP; the visual acuity of the infants is assessed in six different ways from obtained VEP thresholds*

	No.	Range	Visual acuity mean $\pm$ SD (cpd)	Difference between days (Δ, cpd)	Relative difference (Δ/mean * 100)
Max. threshold	89	4.4–19.1	11.4 $\pm$ 3.7	1.7 $\pm$ 1.5	15.0 $\pm$ 12.1
Mean of 3	88	4.0–19.0	10.9 $\pm$ 3.6	1.5 $\pm$ 1.2	14.8 $\pm$ 10.5
Mean of 5	88	3.8–18.7	10.5 $\pm$ 3.5	1.4 $\pm$ 1.1	13.9 $\pm$ 10.3
Mean of 10	79	3.5–17.9	10.1 $\pm$ 3.1	1.2 $\pm$ 1.0	12.9 $\pm$ 10.7
Signal average	89	3.2–17.3	9.1 $\pm$ 3.2	1.4 $\pm$ 1.1	16.3 $\pm$ 12.6
Mean of all	89	3.3–16.1	8.5 $\pm$ 2.9	1.0 $\pm$ 0.8	12.9 $\pm$ 9.1

Table 4. *Comparison of three ways of assessing the SWEEP-VEP measurements of visual acuity in seven female adult subjects (25–45 y of age)*

No. of determination	4.4 $\pm$ 2.9(3–11)
Average no. of scores per session	13.7 $\pm$ 7.5
Average visual acuity (mean acuity, cpd)	18.0 $\pm$ 1.4
CV% on visual acuity from max. thresholds	13.4 $\pm$ 7.6
CV% on visual acuity from signal-average thresholds	11.1 $\pm$ 3.1
CV% on visual acuity from mean of all thresholds	8.1 $\pm$ 3.5
Mean interassay CV% on mean acuity	15.6 $\pm$ 7.0

the variation on max acuity and signal-average acuity method. The intra-assay coefficient of variation on mean acuity within the individual day was on average $16 \pm 7\%$ ($n = 7$). The inter- and intra-assay variation on visual acuity of an infant measured in three separated recording sessions was $17 \pm 2\%$ and 4.8%, respectively (data not shown) and, thus, comparable to adult values.

The distribution of all the single measures of visual acuity of the adult subject with most data were in agreement with a normal distribution ($p > 0.2$ in Kolmogorov-Smirnov test, Fig. 3). The peak of the distribution for this person was 17.8 ± 4.1 cpd ($n = 289$), giving a CV% of 23. The peak of all visual acuity measures on the individual days (the mean acuities) varied around the overall distribution peak (17.9 ± 1.8 cpd, $n = 11$; range, 15.3–20.4 cpd). Eight of the 11 individual session data sets agreed with a normal distribution. The total numbers of single estimates of visual acuity for the other subjects were lower (37–184), and for half of the subjects the obtained results followed a normal distribution. All obtained measures of visual acuity in the infant, who was tested thrice, were also normally distributed ($p > 0.20$ in Kolmogorov-Smirnov test) (Fig. 3, *inset*).
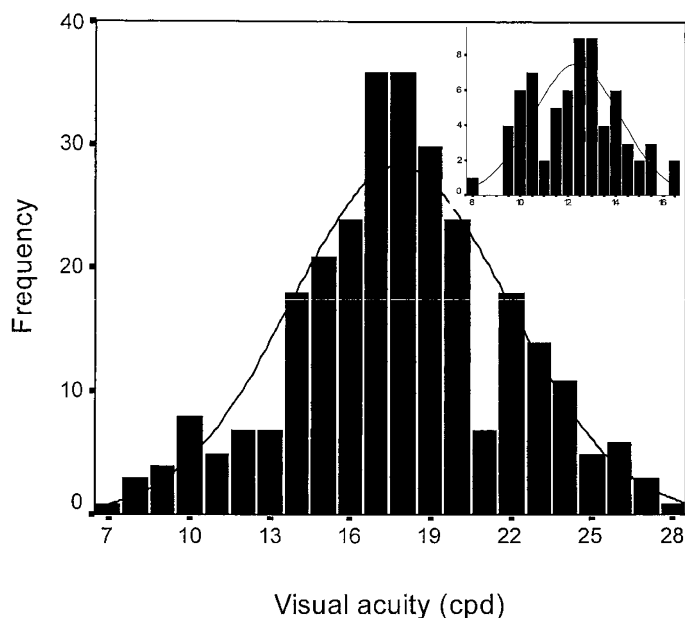

Figure 3. Frequency histogram of all single measures of visual acuity in an adult subject. Total number of observations was 289, the peak of the distribution occurred at 17.8 cpd, and the SD was 4.1 cpd. (*Inset*) A similar histogram for a 5-mo-old infant. Here, the total number of observations was 69, the peak of the distribution occurred at 12.3 cpd, and the SD was 1.8 cpd.
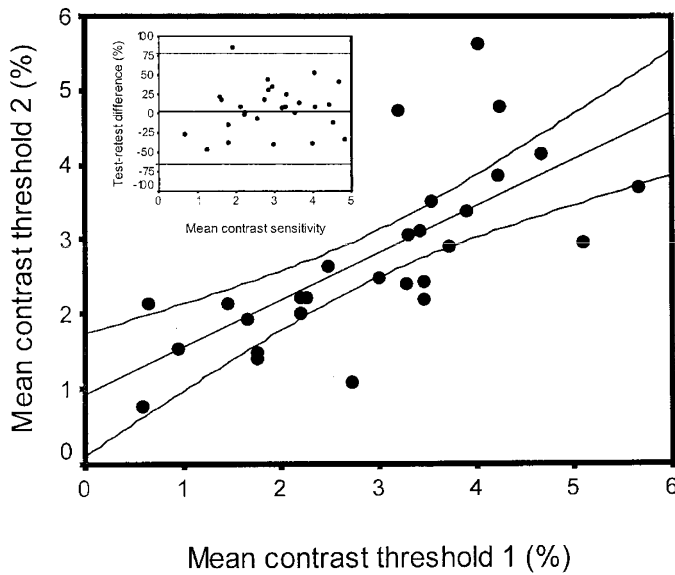
Figure 4. The contrast sensitivity of 28 infants at the second assessment *vs* that determined at the first session (2.7 ± 2.1 d before). The data points represent contrast sensitivity given as mean percentage contrast at threshold in the all successful VEP-contrast plots within the two sessions. Linear regression gives a regression coefficient for the correlation between the two assessments of 0.72 ($p < 0.001$). The ratio between the two measures was 1.1 ± 0.4 ($n = 28$). (*Inset*) The relative variation was constant for all visual acuities. The plot shows the individual test-retest differences relative to measured contrast threshold as a function of mean threshold.

Contrast sensitivity. There was also a significant association between the contrast sensitivity measures in the two sessions (Fig. 4). There was no constant bias [mean difference in VEP threshold was 0.2 ± 1.0% contrast (or 0.05 ± 0.56 octave, $n = 28$)], but the test-retest difference was associated with the measured contrast threshold ($r = 0.37$, $p = 0.05$). The percentile ranking of the infants at the first and second assessment was 16 ± 13% points. The examination was most successful (most scores and the best signal-to-noise ratios) in 10- to 25-wk-old infants (data not shown). The variation coefficient of infant contrast thresholds obtained within a session (*i.e.* the intra-assay variation) was on average 54 ± 19% ($n = 45$). Although the "mean" assessment of contrast sensitivity had the lowest relative difference between days, this was not significantly different from that of the absolute lowest of all obtained threshold (equivalent to the max acuity method; paired *t* test, $p = 0.086$; Table 5). The mean relative difference, the test-retest variability, was 2 ± 37%, giving a 95% prediction interval of approximately ±75%, as shown in Figure 4 (*inset*). Based on the 28 double assessments of infant contrast sensitivity, the threshold had an inter-assay coefficient of variation of 22.9%, equal to 0.39 octave.

Table 5. *SWEEP-VEP contrast sensitivity measurements of infants*

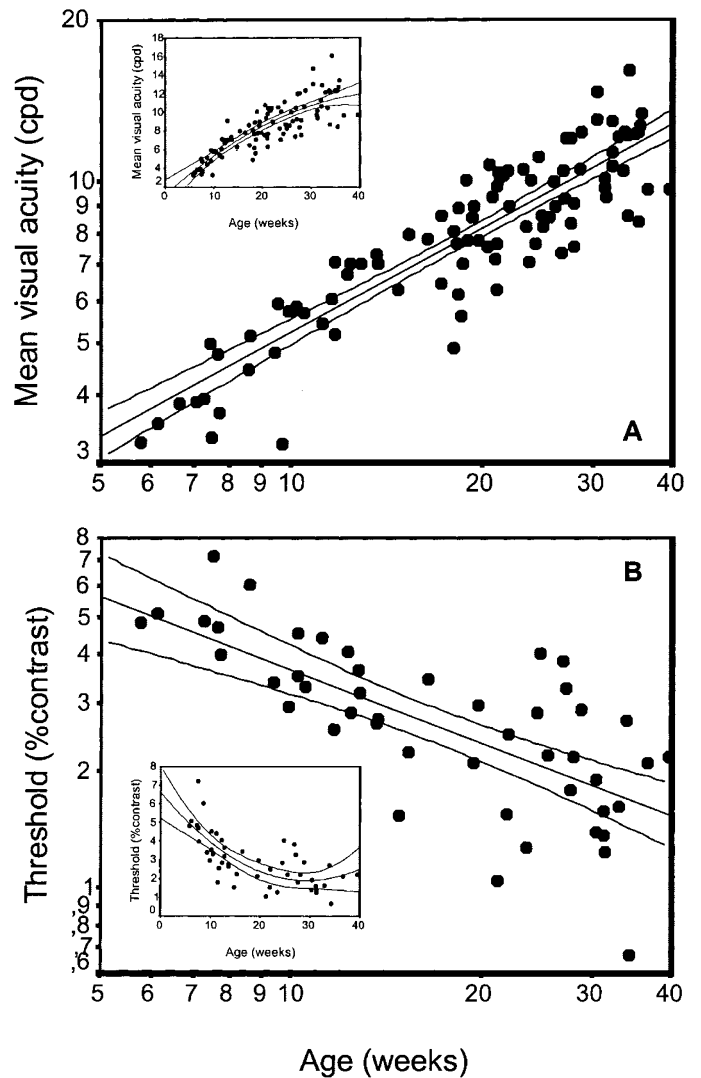|	No.	Range	Contrast threshold mean ± SD (%)	Difference between days (Δ, %)	Relative difference (Δ/mean * 100)
Max.	28	0.5–3.3	1.2 ± 0.8	0.8 ± 1.4	42 ± 38
Mean of all	28	0.7–4.8	2.9 ± 1.1	0.7 ± 0.6	27 ± 25

Figure 5. Visual acuity (*A*) and contrast sensitivity (*B*) as a function of infant age. The data points represent mean values for individual subjects. The relationship between visual acuity and infant age in a double log-plot is described by the line: log cpd = 0.08 + (0.64 · log age) ($n = 100$, $r = 0.88$, $p < 0.001$). The contrast sensitivity has fewer observations ($n = 48$), but the relationship with infant age in a double log-plot is still significant and described by the line: log %contrast = 1.19 − (0.63 · log age) ($r = 0.70$, $p < 0.001$).

Visual development. Figure 5 shows development with age in both of the measured visual functions. For visual acuity as well as contrast sensitivity there was a log–log relationship with infant age. A direct plot of threshold in visual acuity (in cpd) or contrast (% contrast) *versus* infant age, showed a asymptotic development toward the fully developed function, which was approached around 30−40 wk of age (Fig. 5, *A* and *B* insets). Based on the linear regression analysis of the log–log development curve, visual acuity increased by 0.64 octave per doubling in age [octave = (log acuity age 2 − log acuity age 1)/log 2]. Contrast sensitivity also increased by age from a 2.8% contrast detection limit at 15 wk of age to 1.8% contrast at 30 wk of age, giving a constant rate of 0.64 octaves per doubling of age for this visual function as well.

DISCUSSION

Method of assessing visual performance from SWEEP-VEP results. Norcia and Tyler (13, 14) suggested that the best approximation of infant visual acuity was the highest of all thresholds obtained during the session (here called the max acuity). The rationale for this procedure was that variability in infant attentiveness, muscle activity, and accommodative state all conspire to lower the measured activity, but cannot raise it if an adequate statistical criterion is applied. This was experimentally supported by showing that the best acuities obtained in sweeps at different temporal frequencies were significantly correlated and that the within-subject variance was reduced with this method compared with when all the single trial data were used (14). When measuring SWEEP-VEP acuity in an adult subject, one must assume that the subject is attentive throughout the assessment and does not move around while being tested. In this study, we found large variations in SWEEP-VEP acuity assessment in adults, both within the single session and between days (Table 4). The relative difference between two visual acuities determined as mean acuity tended to be lower than that of max acuity (Tables 3 and 4). Furthermore, we showed that all obtained thresholds from one adult and one of the infants were normally distributed (Fig. 3), indicating that they were all meaningful and that the most accurate visual acuity estimate is the peak of the Gauss curve. A good estimate of the peak of the distribution is easier to obtain than the absolute maximum, as session means vary around the overall peak whereas session maximums are all somewhat below the overall maximum.

Other investigators (8, 11) have used the threshold of the average EEG signal from a minimum of three trials (here called signal-average acuity). An estimate of visual acuity based on the average EEG signals should, in theory, be good, as this estimate accumulates all VEP from the tested subject and usually has a much better signal-to-noise ratio than the single trial determinations. However, we found that the relative difference between two SWEEP-VEP measures of visual acuity based on the mean of all thresholds was significantly lower than those based on averages of EEG signals (Tables 3 and 4). The obtained estimates of visual acuity by the mean acuity method and the signal-average acuity method deviate by 0.6 ± 0.9 cpd ($p < 0.001$, $n = 188$). A similar nonsignificant difference was seen in the SWEEP-VEP contrast sensitivity measures (Table 5). Therefore, our results seem to indicate that the most precise estimate of visual detection limits is obtained by averaging all obtained VEP thresholds and that this should preferably be based on at least 10 scores for each infant.

Using the best-threshold-procedure, Prager *et al.* (15) found a poor correlation among estimates of infant visual acuity assessed by SWEEP-VEP and Teller acuity cards ($r^2 = 4\%$), SWEEP-VEP acuity being on average 3.5 cpd better than the Teller card acuity. SWEEP-VEP acuity, obtained as signal-average acuity, has also been shown to be better than Teller acuity (23). The absolute difference decreased with age and reached a stable minimum at 6 mo of age, where SWEEP-VEP acuity was 1.5 times better than Teller acuity (23). This could indicate a general overestimation of visual acuity in the SWEEP-VEP method. However, the two methods might not measure the same type of visual performance, as the Teller method depends on other aspects of infant behavior and observer judgment.

Test-retest variation in SWEEP-VEP acuity. In the present study, we showed that the coefficient of variation in the SWEEP-VEP visual acuity assessment of infants was 17% within each session and 8.4% between sessions. These coefficients of variation were of similar size in adults. The intra-assay variation of SWEEP-VEP was of similar order of magnitude as the interassay variation (test-retest variability). A previous study found a test-retest coefficient of variation of 15% in SWEEP-VEP acuity of 20 4-mo-old infants tested 5 d apart (15). The alternative method, Teller card acuity, has been shown to be two to four times more variable than SWEEP-VEP (15, 24). We found that the SWEEP-VEP contrast sensitivity assessment in infants had an intra-assay coefficient of variation of 54% and an interassay coefficient of variation of 23%, which is in agreement with a previous report of test-retest difference (25).

Whether or not a test-retest variability of $10-40\%$ should be considered large depends on the biologic variation between subjects that are being studied (*e.g.* age, diet, gender). Norcia and Tyler (14) found that between- and within-subject variance each accounted for approximately 50% of the overall variance in single-trial SWEEP-VEP acuities in a group of 17- to 25-wk-old infants. The age span in our group of infants was larger ($6-40$ wk). Consequently, the within-subject variation contributed less to the population SD on visual acuity (about 33%). In spite of the relative low precision of the SWEEP-VEP assessment of infant visual acuity, there was a very strong association with infant age (Fig. 5*A*). Also, in a longitudinal study we saw a pronounced increase in SWEEP-VEP acuity with age (from 2 to 4 mo of age and 4 to 9 mo) (unpublished experiments). Therefore, we conclude, like Prager *et al.* (15), that the method is well suited to detect visual development in a group of infants, as originally shown for visual acuity by Norcia and Tyler (13). The method has been shown to be sensitive enough to detect monocular loss of visual capacity in patients with amblyopia and to determine the optimal time for operation to regain visual capacity (26, 27).

In this study, we found that visual acuity increased by 0.64 octave/2-fold increase in age, which is equivalent to previously observed developmental rates (13, 28–30). Individual rates of development vary, but the average developmental curve from longitudinal data has been shown to be in agreement with cross-sectional data (29). Prager *et al.* (15) found that visual acuity as assessed by SWEEP-VEP increased by 0.8 cpd (0.11 octave) between 4 and 8 mo of age in 103 infants, whereas their Teller acuity increased by 7.8 cpd (1.29 octave). This apparent difference in rate of development measured by the two methods is very unusual. Other longitudinal studies found that the increase in visual acuity measured by SWEEP-VEP and Teller cards was comparable (31), although some have reported a 2-fold higher rate with Teller acuity cards (23).

In spite of large SWEEP-VEP variation and fewer infants, we also found that the contrast detection limit was significantly associated with age and decreased with 0.64 octaves for each

doubling in age (Fig. 5*B*). Previous studies have shown a similar development in infant contrast sensitivity (25, 29, 30). Some of these studies have found that the developmental curve had a steeper slope between 0 and 10 wk of age and a more pronounced plateau thereafter (25, 29).

Our results, like those of Prager *et al.* (15), show that the SWEEP-VEP assessment is less valid for estimating visual acuity in individual subjects, due to the large coefficient of variation. This is in accordance with recent studies (32, 33), which have found a low predictive validity (0.2–0.6) of infant acuity measured by Teller acuity cards. Infant acuity measured as the relative position within the normal range at 1 y of age was not predictive of the relative position later in childhood (32, 33), and, as such, there was no tracking. Prager *et al.* (15) conclude that SWEEP-VEP as well as Teller acuity can give reliable group means. We agree with this conclusion. However, the reliability of the method in relation to detection of differences in between-group means will depend on the size of the difference, the number of infants/group, and how homogeneous the group is (*e.g.* age or degree of breast-feeding, which will affect the SD). The SD of the mean visual acuity in the infants in this study was 2.9 cpd (equal to 34% of the mean visual acuity). In a group of 4-mo-old infants, we found a mean SWEEP-VEP visual acuity of 7.2 $\pm$ 1.4 cpd ($n = 148$) (unpublished experiments), giving a coefficient of variation on the group mean measure of 19%. According to Norcia and Tyler (13), the cross-sectional variability remains roughly constant across age at an average SD of 0.44 octaves. However, the coefficient of variation in a group decreases as the group becomes more homogeneous. As such, in our group of 4-mo-old infants, variation became even smaller (0.5 cpd or 8%) if only fully breast-fed infants were included. The variability of the SWEEP-VEP method is similar in size to the SD in an age-homogeneous group of infants. This means that a very large percentage of the observed variation was due to the method and that the biologic variation constituted only a minor part. It is possible to estimate the relative contribution of the assessment method and biologic factors (*e.g.* genetic dispositions and DHA) to the overall variation in visual acuity by comparing the method variation with the overall variation in the population [hypothetical maximun r^2 adjusted $= 1 - (s^2$ method/total $s^2)$]. A rough estimate shows that the method variation in a very homogenous group of infants (*e.g.* fully breast-fed, 4-mo-olds) can explain a very large part (almost all) of the overall variation. The less homogeneous the group of infants is (with respect to age and other factors that affect visual function), the less important the variation of the method becomes. However, the variability of the SWEEP-VEP method can be a seriously limiting factor when one wants to investigate the importance of dietary DHA for infant visual development. In a previous study, we found that the DHA content of breast-milk accounted for 9% of the variation in visual acuity in a group of 4-mo-old infants (12). In light of the observed SWEEP-VEP variability, this may turn out to constitute a substantial amount of the biologic variation.

The recent meta-analysis by SanGiovanni *et al.* (6) estimated that the acuity difference at 4 mo of age in infants consuming a source of DHA compared with infants on a DHA-free diet was 0.26 $\pm$ 0.10 octaves, which in 4-mo-old infants would be equiv-

alent to approximately 1 SD (7–11). A potential effect of dietary DHA of this size is not insignificant when compared with the rate of visual acuity development (0.64 octaves from 2 to 4 mo of age). The visual reduction in infants who are fed standard infant formulas without DHA equals an approximately 3-wk set back in acuity development (or 18%) compared with infants who are breast-fed or supplied with DHA.

REFERENCES

1. Lauritzen L, Hansen HS, Jørgensen MH, Michaelsen KF 2001 The essentiality of long chain n-3 fatty acids in relation to development and function of the brain and retina. Prog Lipid Res 40:1–94
2. Heird WC, Prager TC, Anderson RE 1997 Docosahexaenoic acid and the development and function of the infant retina. Curr Opin Lipidol 8:12–16
3. Gibson RA, Makrides M 1998 The role of long-chain polyunsaturated fatty acids (LCPUFA) in neonatal nutrition. Acta Paediatr 87:1017–1022
4. Simmer K 1998 Long chain polyunsaturated fatty acid supplementation of terms infants. The Cochrane Library, Issue 4. Oxford, U.K., Update Software Ltd., pp 1–16
5. Neuringer M 2000 Infant vision and retinal function in studies of dietary long-chain polyunsaturated fatty acids: methods, results, and implications. Am J Clin Nutr 71:256S–267S
6. SanGiovanni JP, Berkey CS, Dwyer JT, Colditz GA 2000 Dietary essential fatty acids, long-chain polyunsaturated fatty acids, and visual resolution acuity in healthy fullterm infants: a systematic review. Early Hum Dev 57:165–188
7. Makrides M, Neumann M, Simmer K, Pater J, Gibson RA 1995 Are long-chain polyunsaturated fatty acids essential nutrients in infancy? Lancet 345:1463–1468
8. Birch EE, Hoffman DR, Uauy R, Birch DG, Prestidge C 1998 Visual acuity and the essentiality of docosahexaenoic acid and arachidonic acid in the diet of term infants. Pediatr Res 44:201–209
9. Jørgensen MH, Hølmer G, Lund P, Hernell O, Michaelsen KF 1998 Effect of formula supplemented with docosahexaenoic acid and gamma-linolenic acid on fatty acid status and visual acuity in term infants. J Pediatr Gastroenterol Nutr 26:412–421
10. Makrides M, Neumann MA, Simmer K, Gibson RA 2000 A critical appraisal of the role of dietary long-chain polyunsaturated fatty acids on neural indices of term infants: a randomized, controlled trial. Pediatrics 105:32–38
11. Auestad N, Montalto MB, Hall RT, Fitzgerald KM, Wheeler RE, Connor WE, Neuringer M, Connor SL, Taylor JA, Hartmann EE 1997 Visual acuity, erythrocyte fatty acid composition, and growth in term infants fed formulas with long chain polyunsaturated fatty acids for one year. Ross Pediatric Lipid Study. Pediatr Res 41:1–10
12. Jørgensen MH, Hernell O, Hughes E, Michaelsen KF 2001 Is there a relation between docosahexaenoic acid concentration in mother's milk and visual development in term infants? J Pediatr Gastroenterol Nutr 32:293–296
13. Norcia AM, Tyler CW 1985 Spatial frequency SWEEP VEP: visual acuity during the first year of life. Vision Res 25:1399–1408
14. Norcia AM, Tyler CW 1985 Infant VEP measurements: analysis of individual differences and measurement error. Electroencephalogr Clin Neurophysiol 61:359–369
15. Prager TC, Zou YL, Jensen CL, Fraley JK, Anderson RE, Heird WC 1999 Evaluation of methods for assessing visual function of infants. J AAPOS 3:275–282
16. Norcia AM, Tyler CW, Hamer RD 1989 Measurement of spatial contrast sensitivity with the SWEPT contrast VEP. Vision Res 29:627–637
17. Fitzgerald KM, Hartmann EE, Montalto MB, Auestad N 1996 Longitudinal VEP measures of contrast sensitivity in full-term human infants fed different dietary fatty acids. Invest Ophthalmol Vis Sci 37:4904
18. Tang Y, Norcia AM 1995 An adaptive filter for steady-state evoked responses. Electroencephalogr Clin Neurophysiol 96:268–277
19. Norcia AM, Clarke M, Tyler CW 1985 Digital filtering and robust regression techniques for estimating sensory thresholds from the evoked potential. IEEE Eng Med Biol Mag 4:26–32
20. Bland JM, Altman DG 1996 Measurement error and correlation coefficients. BMJ 313:41–42
21. Bland JM, Altman DG 1996 Measurement error. BMJ 312:1654
22. Bland JM, Altman DG 1996 Measurement error proportional to the mean. BMJ 313:106
23. Riddell PM, Ladenheim B, Mast J, Catalano T, Nobile R, Hainline L 1997 Comparison of measures of visual acuity in infants: Teller acuity cards and sweep visual evoked potentials. Optom Vis Sci 74:702–707
24. Hamer RD, Norcia AM, Tyler CW, Hsu-Winges C 1989 The development of monocular and binocular VEP acuity. Vision Res 29:397–408

25. Kelly JP, Borchert K, Teller DY 1997 The development of chromatic and achromat-iccontrast sensitivity in infancy as tested with the sweep VEP. Vision Res 37:2057–2072
26. Polat U, Sagi D, Norcia AM 1997 Abnormal long-range spatial interactions in amblyopia. Vision Res 37:737–744
27. Norcia AM, Hamer RD, Jampolsky A, Orelbixler D 1995 Plasticity of human motion processing mechanisms following surgery in infantile esotropia. Vision Res 35:3279–3296
28. Skoczenski AM, Norcia AM 1999 Development of VEP Vernier acuity and grating acuity in human infants. Invest Ophthalmol Vis Sci 40:2411–2417
29. Norcia AM, Tyler CW, Hamer RD 1990 Development of contrast sensitivity in the human infant. Vision Res 30:1475–1486
30. Allen D, Tyler CW, Norcia AM 1996 Development of grating acuity and contrast sensitivity in the central and peripheral visual field of the human infant. Vision Res 36:1945–1953
31. Salomao SR, Birch EE 1996 Individual growth curves for infant visual acuity measured by sweep-VEP and FPL. Invest Ophthalmol Vis Sci 37:4902
32. Mash C, Dobson V 1998 Long-term reliability and predictive validity of the teller acuity card procedure. Vision Res 38:619–626
33. Dobson V, Quinn GE, Siatkowski RM, Baker JD, Hardy RJ, Reynolds JD, Trese MT, Tung B 1999 Agreement between grating acuity at age 1 year and Snellen acuity at age 5.5 years in the preterm child. Invest Ophthalmol Vis Sci 40:496–503