

A New Visual Evoked Potential System for Vision Screening in Infants and Young Children

John W. Simon, MD,^a John B. Siegfried, PhD,^b Monte D. Mills, MD,^c Joseph H. Calhoun, MD,^d and Judith E. Gurland, MD^e

Introduction: With a prevalence of 3-5%, amblyopia represents a major public health problem. Effective treatment depends on early detection, and a broad consensus of professional opinion supports vision screening of infants and young children. No single method of screening has been demonstrated to be superior in detecting amblyopia and all methods have significant limitations. **Methods:** We assessed a new, "child-friendly" visual evoked potential (VEP) system (ENFANT™ II, Diopsys Corp., Metuchen, NJ) for use in screening. We studied 122 children, aged 6 months to 5 years, comparing test results in a masked fashion to results of standard ophthalmologic examinations. A statistical program analyzed VEP differences between fellow eyes to determine a "pass" or "fail" for each child. For verbal patients, clinical amblyopia was defined as an interocular difference of two or more lines in best-corrected visual acuity. For preverbal patients, clinical amblyopia was defined by the clinician's decision to treat with occlusion or atropine penalization. Preverbal children with significant refractive errors or structural eye pathology were also considered clinically abnormal. **Results:** The test was completed by 94% of the study group, each child requiring an average of 10 minutes to complete testing of both eyes. The sensitivity was 0.973, the specificity 0.808, the positive predictive value 0.706, and the negative predictive value 0.984. **Conclusion:** With its easy electrode placement and rapid, attractive stimulus, the new system overcomes technical difficulties which were associated with older VEP techniques. The test shows promise as a screening tool for detecting amblyopia and other visual deficits in young children. (J AAPOS 2004;8:549-554)

Amblyopia is a major public health problem. With a prevalence of 3-5%, it accounts for more cases of monocular reduced vision in children than all other causes combined. It is also the most common cause of monocular reduced vision in young adults.^{1,2} Because early detection allows for more effective treatment, a broad consensus supports universal vision screening in infants and young children. However, there has been no consensus on the appropriate screening methods.^{3,4} All too often, screenings during the preschool years are simply not performed.⁵

In addition to screening for amblyopia itself, the ideal method would detect amblyogenic factors, such as refractive errors, strabismus, and media opacities, particularly in infants and preverbal children. Some have argued that the most effective screening method is a complete ophthalmologic examination, including cycloplegic retinoscopy.⁶ Such examinations would require investment of substantial effort on the part of highly trained examiners, however, and most ophthalmologists consider this approach impractical for screening.

Photorefractometry provides information on the ocular alignment and the refractive state, both associated with amblyopia, but does not measure amblyopia directly. Although it continues to be used in screening programs, variable sensitivity and specificity results have led several authors to question its reliability for vision screening.⁷⁻¹² Autorefractometry without cycloplegia and Bruckner reflex testing have also proved unreliable.^{13,14}

Preferential looking (eg, Teller acuity cards) is not objective, in that the examiner must interpret the infant's gaze to one side (grating) or the other (blank field). Largely because of the training required for test personnel and the time required to complete testing, the method has not been used in large screening programs. Stereopsis screening is difficult in preverbal children and has shown inconsistent results in verbal children.^{15,16} Optotype rec-

From the Department of Ophthalmology/Lions Eye Institute,^a Albany Medical College, Albany, NY; Department of Basic Sciences,^b Pennsylvania College of Optometry, Elkins Park, PA; Division of Ophthalmology,^c Children's Hospital of Philadelphia, and Department of Ophthalmology, University of Pennsylvania Medical School, Philadelphia, PA; Department of Pediatric Ophthalmology,^d Wills Eye Hospital, Philadelphia, PA; Department of Ophthalmology,^e Bronx Lebanon Hospital Center, Bronx, NY.

Disclosure: Dr. Siegfried serves as a consultant to the Diopsys Corporation. The other authors performed the research described under contract between Diopsys and their offices or academic institutions, but have no other financial interest in the Corporation.

Submitted March 9, 2004.

Revision accepted August 30, 2004.

Address reprint requests to John W. Simon, MD, Department of Ophthalmology/Lions Eye Institute, 35 Hackett Blvd., Albany, NY 12208. E-mail: Simonj@mail.amc.edu

Copyright © 2004 by the American Association for Pediatric Ophthalmology and Strabismus.

1091-8531/2004/\$35.00 + 0

doi:10.1016/j.jaapos.2004.08.009

ognition acuity testing, the traditional mechanism of screening, is limited in younger children and is often plagued by high non-completion rates and low specificities.^{17,18}

Visual evoked potentials (VEPs) have been used to assess the vision of infants and young children since the 1970's.^{19,20} Although it is a sensitive test which has been useful in research, the use of VEP in screening for amblyopia has been problematic. Among factors that have limited its usefulness are the cumbersome process of attaching and standardizing electrodes, the time required for testing, the relatively monotonous stimuli, and the complexity of the generated waveforms.

We were interested in assessing the efficacy of a new VEP device, the ENFANT™ II system (Diopsys Corp., Metuchen, NJ), which was designed to circumvent these problems. The device employs a sweep stimulus in a simplified operation that removes the subjective component in waveform interpretation.²¹ We compared the results of testing with standard ophthalmologic examinations, which served as the "gold standard." We were especially interested in the ease of testing infants and young children and the sensitivity, specificity, and negative and positive predictive values of the test in detecting interocular differences.

PATIENTS AND METHODS

Study Patients

After approval by the local Institutional Review Boards, a total of 122 patients were recruited at four test sites: Albany Medical Center/Lions Eye Institute, Children's Hospital of Philadelphia, Wills Eye Hospital, and Bronx Lebanon Hospital Center. Children varied in age from 6 months to 5 years (mean, 3.3 years). Both active patients and normal volunteers were included. Recruitment was completed during a 6-month period, during which parents of eligible children were offered the opportunity to participate. Some parents declined, and this therefore does not constitute a consecutive series. Normal controls included siblings and active patients who had been examined and determined to have no visual system pathology. All underwent standard pediatric ophthalmologic examinations, including motility testing, cycloplegic refraction, and (for verbal patients) measurement of best-corrected visual acuity. Fixation behavior was used in place of recognition visual acuity testing in preverbal children. Children born more than 1 month prematurely were excluded, as were those with known neurologic and developmental disorders.

The four sites were chosen to represent a reasonable diversity of geographic, socio-economic, and ethnic populations. Testing was masked so that the four technicians performing the tests were unaware of clinical findings and diagnoses, which were recorded on a data sheet by the physician at each site. Similarly, the four physicians were unaware of the VEP test results. The technicians were all



FIG. 1. Headband being applied for VEP testing.

trained by Diopsys Corporation using a protocol for VEP testing provided by the International Society for Clinical Electrophysiology of Vision. This training required approximately 16 hours.

Clinical amblyopia was defined as an interocular difference of two or more lines in best-corrected visual acuity. For preverbal children, clinical amblyopia was defined by the clinician's decision to treat with patching or atropine penalization. Preverbal children were considered clinically abnormal if they had monocular strabismus, significant structural eye pathology, greater than 3.00 diopters of uncorrected myopia or astigmatism, or hyperopia greater than 5.00 diopters. Included in this group were patients with optic nerve hypoplasia, glaucoma, Leber congenital amaurosis, and nystagmus. Preverbal children with various conditions deemed clinically to have no impact on visual acuity (eg, alternating strabismus, conjunctivitis, mild ptosis, iris nevus, juvenile rheumatoid arthritis without ocular involvement, and orbital fracture) were considered normal. Teller acuities were not used to determine the presence of amblyopia for this study.

VEP Testing Methods

A proprietary disposable headband with integrated electrodes (Fig. 1) was used for recordings. The headband aligned the occiput (O_2), the mid-forehead (Fp_2), and the temple (ground). Skin contact with the pre-gelled electrodes was enhanced with a small amount of EEG conductive paste. Infants were positioned on a parent's lap and children were seated in a comfortable chair at a measured distance of 57 cm from a 17-inch (43-cm) display monitor, so that the stimulus subtended a total visual angle of 20°. The room was darkened except for the light from the testing equipment. Testing was performed monocularly, with spectacle correction if prescribed, using an adhesive occluder over the fellow eye.

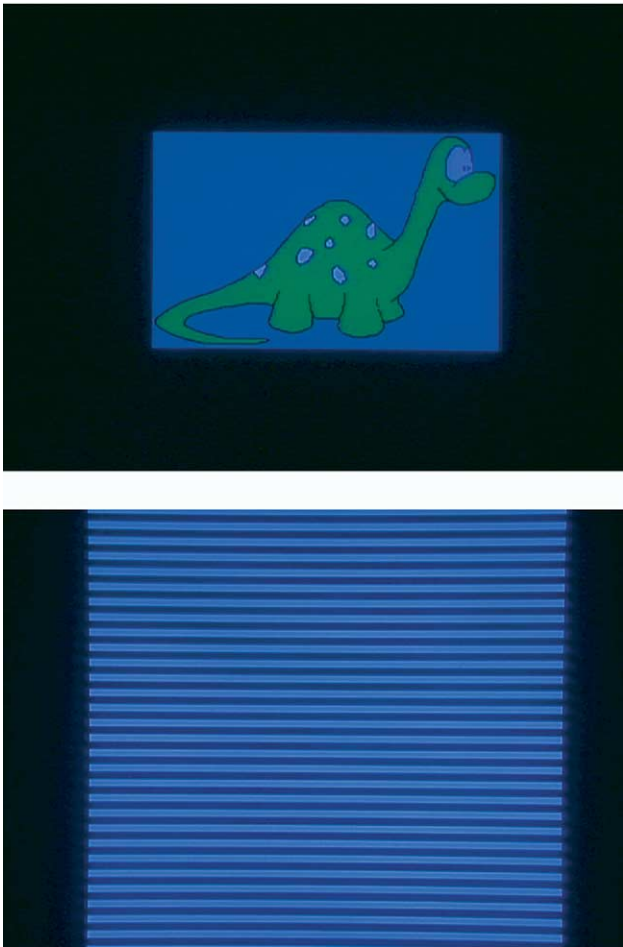


FIG. 2. Test stimulus, showing a cartoon figure (top), which appears before the sequence of gratings (sample at bottom).

The stimulus was chosen to maximize attention and accommodation (Fig. 2). Six horizontal square-wave grating patterns were presented in one-octave steps, at spatial frequencies from 0.4 cycles/degree (Snellen equivalent 20/1500) to 12.8 cycles/degree (Snellen equivalent 20/45). The largest grating pattern was not used for analysis, but only to orient the child to the stimulus. The stimulus field was 20×20 cm and the contrast was fixed at 95% (Michelson contrast). Each pattern was temporally modulated in counter phase at 7.5 Hz, yielding 15 contrast reversals per second. The sweep format presented each spatial frequency, contrast reversing, for 1 second.²¹ Therefore, each sweep required only 6 seconds for completion.

At the initiation of each sweep, a different children's song was played to attract the child's attention. A large color cartoon image of an animal (elephant, tiger, giraffe, etc.) then appeared on the display. When the test was about to commence, the animal's sound was presented and its image became smaller in three rapid steps. After each sweep, a different animal image appeared on the screen, and the sequence was repeated until five usable sweeps were recorded for each eye. The patient's fixation was

monitored continuously using a small infrared video camera to permit the technician to exclude recordings taken during periods when the child was not attending to the screen. Although the actual recording required as little as 30 seconds per eye, the entire testing sequence, including electrode placement and automated analysis of responses for the two eyes, required approximately 10 minutes. Data acquisition was synchronous with stimulus presentation to minimize contamination of the response. An automatic fault detection feature was incorporated which eliminated recordings contaminated by excessive noise or artifacts.

Monocular test data were automatically compared for the two eyes at the completion of testing. The actual statistical criterion used to determine asymmetry in the monocular responses was based on the empirical results of testing to optimize the sensitivity and specificity.²² First, a monocular T_{circ}^2 statistic was calculated, using both amplitude and phase data, to determine the variability associated with the VEP estimates at each spatial frequency.²³ Recordings were repeated until five sweeps fell into a 90% confidence interval. Then, a binocular T_{circ}^2 statistic was calculated on the differences between the two eyes at each spatial frequency, with a critical F statistic value determined at the 96% level of confidence. A Receiver Operating Characteristic (ROC) analysis was then performed using the number of failed spatial frequencies to determine an optimal criterion to separate normal and abnormal patients. The criterion selected for test failure was a statistical difference at four or more consecutive spatial frequencies.

RESULTS

Figures 3 and 4 depict representative waveforms from both eyes of normal and abnormal subjects. Figure 5 illustrates the ROC analysis.

Of the 122 subjects qualifying for the study, 115 (94%) completed monocular VEP testing of both eyes. Of these, 78 were considered clinically normal. A statistical assessment of screening accuracy for the 122 children is presented in Table 1. Using the clinical examination as the "gold standard," the VEP test's sensitivity (probability of correctly identifying an abnormal patient) was 0.973. The test's specificity (probability of correctly identifying a normal patient) was 0.808. The positive predictive value (probability that a child with an abnormal VEP proved to be abnormal on clinical examination) was 0.706. The negative predictive value (probability that a child with a normal VEP proved to be normal on clinical examination) was 0.984. Table 1 indicates these statistics separately for verbal and preverbal children and for the combined group. Only one false negative was identified. This 6-year-old passed the VEP test but was found to have corrected Snellen acuities of 20/40 in the right eye and 20/60+ in the left eye. There were 15 false positives. Of these, six had strabismus, which may have been associated with subclinical amblyopia in one eye.

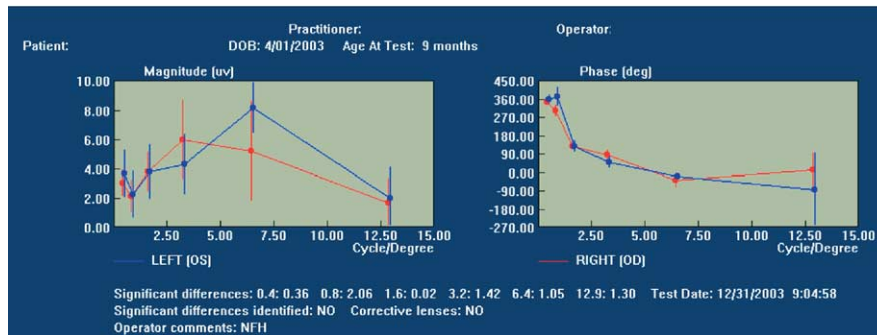


FIG. 3. Example of normal patient recording. Both magnitude (left) and phase (right) of VEP responses for the two eyes are statistically similar through all spatial frequencies.

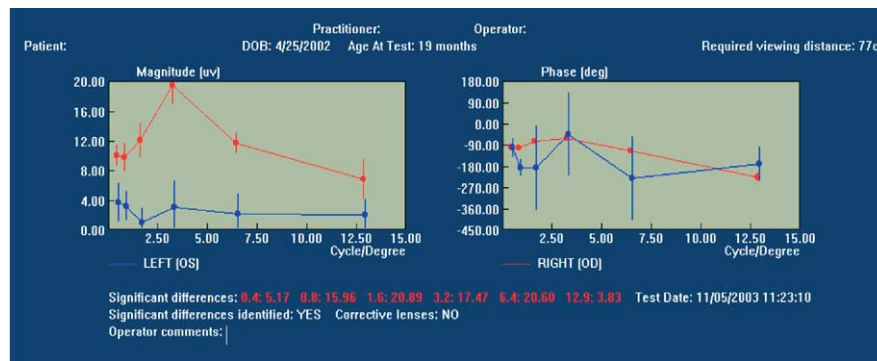


FIG. 4. Example of abnormal patient recording. Magnitude of VEP responses (diagram on left) is higher on the right (red tracing) than in the left (blue tracing) eye at all spatial frequencies, indicating poor vision in the left eye. Phase difference (diagram on right) is also significant at 0.8 cycles/degree.

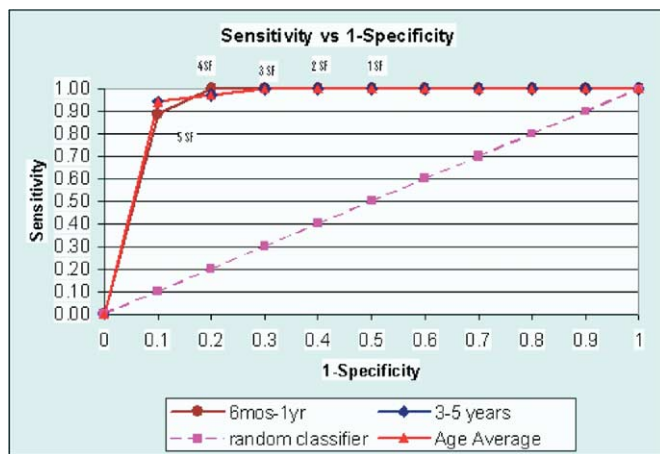


FIG. 5. ROC analysis with the criterion for test failure. A significant difference on four consecutive spatial frequencies of five was selected as the optimized criterion. Gratings of 0.4 cycles/deg were used only for orientation and were excluded in the analysis.

DISCUSSION

Although VEPs have long been used to measure the visual function of infants in clinical studies, the test has yielded

TABLE 1. Screening accuracy in verbal and preverbal children.

| Group | Yield Rate | Sn | Sp | PPV | NPV |
|----------------------------|------------|-------|-------|-------|-------|
| Combined (<i>n</i> = 122) | 94% | 0.973 | 0.808 | 0.706 | 0.984 |
| Preverbal (<i>n</i> = 42) | 91% | 1.00 | 0.808 | 0.706 | 1.00 |
| Verbal (<i>n</i> = 80) | 98% | 0.960 | 0.792 | 0.705 | 0.977 |

Yield Rate = (Number patients who completed the screening)/(Number patients enrolled in the screening) × 100.

Sn (Sensitivity) = TP/(TP + FN) × 100. Sensitivity is the probability of correctly identifying an abnormal patient. TP = true positives; FN = false negatives.

Sp (Specificity) = TN/(TN + FP) × 100. TN = true negatives; FP = false positives. Specificity is the probability of correctly identifying a normal patient.

NPV (Negative Predictive Value) = TN/(TN + FN) × 100. NPV is the probability that a child with a normal VEP proved to be normal on clinical examination.

PPV (Positive Predictive Value) = TP/(TP + FP) × 100. PPV is the probability that a child with an abnormal VEP proved to be abnormal on clinical examination.

variable results when compared to clinical examinations and preferential looking assessments.²⁴⁻²⁶ Similar to preferential looking, the test has proved too cumbersome to apply to large-scale vision screening. Typically, a young child will begin to lose interest during the application of electrodes and standardization of equipment and will not attend to the stimulus for prolonged testing.

The method we describe was designed to circumvent these difficulties. The headband with the electrodes embedded is attached quickly. The stimulus, with its songs, cartoons, and animal noises, is intrinsically interesting to young children, and the sweep program presents all necessary spatial frequencies in seconds. The infrared video monitoring and automatic fault detection systems allow exclusion of recordings taken during periods when the child is not looking at the stimulus screen. The automated statistical analysis of waveforms allows a pass/fail determination to be made without subjective input, so that the test can be performed by technicians without extensive training.

The 94% testability rate attests to the ease and attractiveness of the described method. Although we anticipated that testing would be difficult in children between the ages of 12 and 36 months, these children had a similar test completion rate and similar reliability to younger and older children. The test seems equivalent in accuracy during both the preverbal and the young verbal age ranges.

Our analysis of VEP results has focused primarily on interocular differences. Although several children were identified to have significant but clinically symmetric visual pathology in the two eyes (eg, glaucoma and Leber congenital amaurosis), all had sufficiently asymmetric VEP recordings to be counted as positives. We acknowledge that strictly symmetrical visual loss might escape detection using the protocol we describe. We are currently investigating methods to relate VEP results more directly to visual acuity to address this problem. We recognize that ours was not a typical screening population, as children were recruited through pediatric ophthalmology practices. Technicians had undergone more training than might apply in primary care settings.

On the other hand, our experience suggests that the *Enfant*TM II test has promise as a screening tool. Its sensitivity, specificity, and positive and negative predictive values all compared favorably with those of other screening methods. As indicated above, 6 of the 15 false positives may have had subclinical amblyopia associated with strabismus. Similar observations have previously been made in VEP studies.^{27,28} There was only one false negative among the 125 children we tested. Further investigations in screening populations are necessary.

References

1. Donahue SP, Arnold RW, Ruben JB, and the AAPOS Vision Screening Committee. Preschool vision screening: what should we be detecting and how should we report it? *J AAPOS* 2003;7:314-6.
2. Attebo K, Mitchell P, Cumming R, Smith W, Jolly N, Sparkes R. Prevalence and causes of amblyopia in an adult population. *Ophthalmology* 1998;105:154-9.
3. Simmons K. Preschool vision screening: rationale, methodology and outcome. *Surv Ophthalmol* 1996;41:3-30.
4. Prager TC, Zou YL, Jensen CL, Fraley JK, Anderson RE, Heird WC. Evaluation of methods for assessing visual function of infants. *J AAPOS* 1999;3:275-82.
5. Castanes MS. Major review: the underutilization of vision screening (for amblyopia, optical anomalies and strabismus) among preschool age children. *Binocul Vis Strabismus Q* 2003;18:217-32.
6. Gandjour A, Schlichtherle S, Neugebauer A, Russmann W, Lauterbach KW. A cost-effectiveness model of screening strategies for amblyopia and risk factors and its application in a German setting. *Optom Vis Sci* 2003;80:259-69.
7. Tong PY, Bassin RE, Enke-Miyazaki E, Macke JP, Tielsch JM, Stager DR Sr, Beauchamp GR, Parks MM. Screening for amblyopia in preverbal children with photoscreening photographs: II. Sensitivity and specificity of the MTI photoscreener. *Ophthalmology* 2000;107:1623-9.
8. Tong PY, Enke-Miyazaki E, Bassin RE, Tielsch JM, Stager DR Sr, Beauchamp GR, Parks MM. Screening for amblyopia in preverbal children with photoscreening photographs. National Children's Eye Care Foundation Vision Screening Study Group. *Ophthalmology* 1998;105:856-63.
9. Simons BD, Siatkowski RM, Schiffman JC, Berry BE, Flynn JT. Pediatric photoscreening for strabismus and refractive errors in a high-risk population. *Ophthalmology* 1999;106:1073-80.
10. Weinand F, Graf M, Demming K. Sensitivity of the MTI photoscreener for amblyogenic factors in infancy and early childhood. *Graefes Arch Clin Exp Ophthalmol* 1998;236:801-5.
11. Hatch S, Tibbles CD, Mestito IR, Read R, Traveis L, Richman J. Validity and reliability of the MTI photoscreener. *Optom Vis Sci* 1997;74:859-64.
12. Berry BE, Simons BD, Siatkowski RM, Schiffman JC, Flynn JT, Duthie MJ. Preschool vision screening using the MTI-Photoscreener. *Pediatr Nurs* 2001;27:27-34.
13. Barry JC, Konig HH. Non-cycloplegic screening for amblyopia via refractive findings with the Nikon Retinomax hand held autorefractor in 3 year-old kindergarten children. *Br J Ophthalmol* 2001;85:1179-82.
14. Paysse EA, Williams GC, Coats DK, Williams EA. Detection of red reflex asymmetry by pediatric residents using the Bruckner reflex versus the MTI photoscreener. *Pediatrics* 2001;108:E74.
15. Ohlsson J, Villarreal G, Sjoström A, Abrahamsson M, Sjostrand J. Screening for amblyopia and strabismus with the Lang II stereo card. *Acta Ophthalmol Scand* 2002;80:163-6.
16. Ohlsson J, Villarreal G, Abrahamsson M, Cavazos H, Sjoström A, Sjostrand J. Screening merits of the Lang II, Frisby, Randot, Titmus, and TNO stereo tests. *J AAPOS* 2001;5:316-22.
17. Cyert L, Schmidt P, Maguire M, Moore B, Dobson V, Quinn G, and Vision in Preschoolers Study Group. Threshold visual acuity testing of preschool children using the crowded HOTV and Lea Symbols acuity tests. *J AAPOS* 2003;7:396-9.
18. Newman DK, East MM. Preschool vision screening: negative predictive value for amblyopia. *Br J Ophthalmol* 1999;83:676-9.
19. Weiss AH, Kelly JP. Acuity, ophthalmoscopy, and visually evoked potentials in the prediction of visual outcome in infants with bilateral optic nerve hypoplasia. *J AAPOS* 2003;7:108-15.
20. Norcia AM, Tyler CW. Spatial frequency sweep VEP: visual acuity during the first year of life. *Vision Res* 1985;25:1399-408.
21. Zemon V, Hartmann EE, Gordon J, Prunte-Glowacki A. An electrophysiological technique for assessment of the development of spatial vision. *Optom Vis Sci* 1997;74:708-16.
22. Harper R, Reeves B. Compliance with methodological standards when evaluating ophthalmic diagnostic tests. *Invest Ophthalmol Vis Sci* 1999;40:1650-7.

23. Victor JD, Mast J. A new statistic for steady-state evoked potentials. *Electroencephalogr Clin Neurophysiol* 1991;78:375-88.
24. Westall CA, Ainsworth JR, Buncic JR. Which ocular and neurologic conditions cause disparate results in visual acuity scores recorded with visually evoked potential and Teller acuity cards? *J AAPOS* 2000;4:295-301.
25. Katsumi O, Denno S, Arai M, De Lopes Faria J, Hirose T. Comparison of preferential looking acuity and pattern reversal visual evoked response acuity in pediatric patients. *Graefes Arch Clin Exp Ophthalmol* 1997;235:684-90.
26. Lauritzen L, Jorgensen MH, Michaelsen KF. Test-retest reliability of swept visual evoked potential measurements of infant visual acuity and contrast sensitivity. *Pediatr Res* 2004;55:701-8.
27. Watts PO, Neveu MM, Holder GE, Sloper JJ. Visual evoked potentials in successfully treated strabismic amblyopes and normal subjects. *J AAPOS* 2002;6:389-92.
28. Shawkat FS, Kriss A, Timms C, Taylor DS. Comparison of pattern-onset, -reversal and -offset VEPs in treated amblyopia. *Eye* 1998;12:863-9.



An Eye on the Arts – The Arts on the Eye

I see everything.

That is why I don't like new places. If I am in a place I know, like home, or school, or the bus, or the shop, or the street, I have seen almost everything in it beforehand and all I have to do is to look at the things that have changed or moved. For example, one week the Shakespeare's Globe poster had fallen down in the classroom at school and you could tell because it had been put back slightly to the right and there were three little circles of Blu-Tack stain on the wall down the left-hand side of the poster. And the next day someone had graffitied CROW APTOK to lamppost 437 in our street, which is the one outside number 35.

But most people are lazy. They never look at everything. They do what is called *glancing*, which is the same word for bumping off something and carrying on in almost the same direction, e.g., when a snooker ball glances off another snooker ball. And the information in their head is really simple. For example, if they are in the countryside, it might be

1. I am standing in a field that is full of grass.
2. There are some cows in the fields.
3. It is sunny with a few clouds.
4. There are some flowers in the grass.
5. There is a village in the distance.
6. There is a fence at the edge of the field and it has a gate in it.

And then they would stop noticing anything because they would be thinking something else like, "Oh, it is very beautiful here," or "I'm worried that I might have left the gas cooker on," or "I wonder if Julie has given birth yet."

But if I am standing in a field in the countryside I notice everything.

—Mark Haddon (from *The Curious Incident of the Dog in the Night-Time*)