

Measuring near stereopsis

Dr Kathleen Vancleef and **Professor Jenny Read** discuss stereopsis and how it is best assessed clinically. (C58138, one distance learning CET point suitable for optometrists and dispensing opticians)



To take part in CET exercises go to cet.opticianonline.net and complete the questions

Stereo vision or stereopsis is the ability to perceive the relative depth of objects based on binocular disparity. Binocular disparity is the small difference in angles between images of objects in left and right eyes (figure 1). Several conditions need to be met for the brain to achieve stereopsis. First, alignment of the eyes is necessary. Second, good visual acuity in each eye is essential. Third, stereoscopic vision can only be achieved by simultaneous perception and superimposition of both images on each retina and fusion of both images into one. Because stereo vision depends upon good vision in both eyes, excellent oculomotor control and the development of binocular brain mechanisms, it is often regarded as the gold standard for binocular visual function.

Stereopsis is not present at birth but develops in the first months of life. That full-term and pre-term children develop stereopsis at the same age post-birth shows that the development depends on visual experience rather than biological maturation of the system.¹ In the early months of life, we develop coarse stereopsis, which operates on high contrast lines and edges and enables us to align our eyes. Alignment then permits the development of fusion and fine stereopsis. Fine stereopsis works over a much shorter range of disparities but enables us to make very fine depth judgments even in densely textured surfaces, such as grass or tree bark, where there are few or no depth cues monocularly.

WHY DO WE NEED STEREOPSIS?

'I am comically bad at sports that involve hand-eye coordination.' 'I was told I ruined the class picture when I was in elementary school.' 'It's a waste of time and money to go to 3D movies.' These quotes by people with amblyopia illustrate the impact poor stereopsis and patching treatment can have on daily life.

Most of our movements are guided by visual information: think about grasping a cup of tea, catching a ball, driving, or walking. Especially for accurate and fast hand movements, stereopsis is highly beneficial. Several studies have shown that children and adults with poor stereopsis have more difficulties in a range of visuomotor tasks than people with normal stereopsis.² For example, researchers in Liverpool asked people with normal stereopsis and people with reduced stereopsis to do a series of tasks. In the first task, the participants had to place as many pins as possible in a pegboard within 30 seconds. In another task, they had to place a fixed number of beads onto a needle as fast as possible.

Participants with poor stereo vision were much slower on both tasks.³ With poor stereopsis, it is much harder to judge where objects are in space and estimate the distance between one's hand and the object. This makes our hand movements less accurate

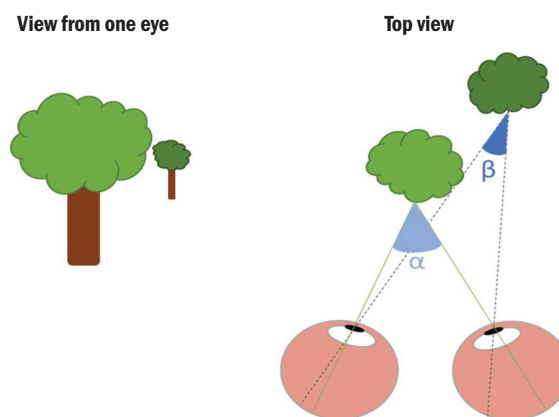


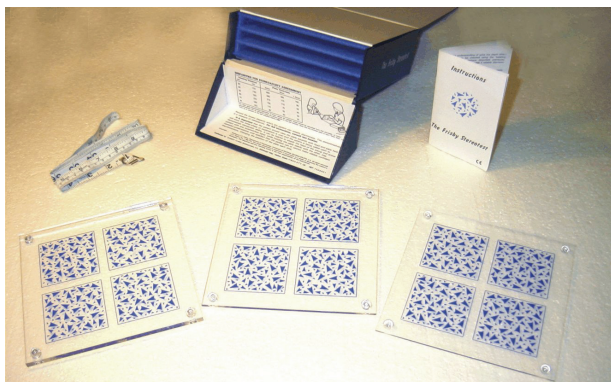
FIGURE 1 The drawing at the left shows the view of two trees from the perspective of the eyes. The light green tree stands in front of the dark green tree. The right drawing shows a top view of the scene. When the eyes are focusing on the light tree, the image is projected on the fovea of the left and right eye. The angle between both projections is angle α . The images of the dark tree are projected on different positions of the peripheral retina in the left and right eye with angle β . Because angle β is smaller than angle α our brain interprets the dark tree as further away than the light tree. The size of the difference between α and β represents the disparity. Large differences in angle indicate large differences in depth

and slower. A more everyday life task like catching a ball is also more difficult with reduced stereopsis.⁴ Children with amblyopia are indeed often clumsy and do not perform as well in sports as their peers. It has also been suggested that professional athletes of football and tennis heavily rely on their stereopsis to judge and predict ball trajectories.⁵ The functional role of stereopsis is not limited to hand movements, also when walking we use stereopsis to detect changes in terrain, like steps or uneven ground. With one eye closed (no stereopsis) people walk 10% slower than with both eyes open (stereopsis available).⁶

Reading abilities and academic performance are also negatively correlated with reduced stereopsis. A research group in Ohio asked children with average intelligence to complete a reading test and a stereotest. They also asked their teachers to give each child a score for mathematics, reading, writing, and spelling. The children with poorer stereopsis presented with lower scores on the reading test and other academic abilities.^{7,8} Reduced stereopsis can also limit a child's career options because for professions

CONTINUING EDUCATION

FIGURE 2 Frisby stereotest



like surgeons, pilots, armed forces, or architects, accurate stereopsis is important.⁵

Last, we should not forget about the psychosocial impact of poor stereopsis or amblyopia in young children. Clumsiness and patching treatment can have a severe impact on a child's social interactions and self-esteem. For instance, a group of researchers in Australia found that children with amblyopia had lower social acceptance than their peers. This was lowest for children who received patching treatment.⁹

Although we have only described a selection of the studies on the topic, the effect of reduced stereopsis on these daily life activities has been shown for all ages, and for amblyopic participants as well as participants with normal vision under monocular viewing conditions. It has been repeatedly demonstrated that the deficits are due to poor stereopsis rather than poor visual acuity, fixation instability or impaired vergence control.⁵

HOW TO MEASURE NEAR STEREOPSIS?

Commonly used stereotests

Frisby

The most commonly used stereotest in the UK is the Frisby stereotest. In this test, four groups of small triangles are printed on a transparent plastic plate (figure 2). In one of the four groups, a circular region is printed on the other side of the plate. This makes the region physically stand out in depth. The child has to indicate which of the four groups has a disk 'popping out'. By using plates with three different thicknesses viewed at nine different distances, 27 disparities between five and 600 seconds of arc can be presented.

This test measures real depth in a detection task without the need for glasses. A disadvantage of the test is that if the plate or the child moves, motion parallax cues can reveal the disk. This means that a child without stereopsis might still perform well on Frisby. For this reason, it is recommended to repeat the test with one eye covered. If the same score is obtained, the result is discarded. Sensitivity for strabismus is estimated at 33% and for amblyopia at 12%,¹⁰ meaning that Frisby detects a very low percentage of children with strabismus or amblyopia and is not suited as a screening method. Also, between two measurements in the same child, stereoacuity on Frisby can almost double due to measurement error (0.78 octave steps).¹¹

Lang

The Lang stereotest I and II are glasses-free stereotests (figure 3). The tests are lenticular-sheet stereograms. The child needs to identify three 3D objects like a cat or a truck. On the Lang II a star is visible with one eye and can be used to see if the child

FIGURE 3 Lang stereotests



FIGURE 4 TNO stereotest



understands the task. The test sheet is to be held at 40cm (16 inches) and measures disparities 550, 600, and 1200 for Lang I, and 200, 400, and 600 for Lang II, one object for each disparity level. As with the Frisby, motion parallax cues are available if the child or test moves. Sensitivity of the Lang II for strabismus (71%) and amblyopia (34%) is better than that of Frisby and similar to other stereotests;¹⁰ no reports on repeatability are available.

TNO

The TNO Stereo test (figure 4) is performed at a distance of 40cm. While wearing red/green anaglyph stereo glasses, the child is first shown Plate I in which two butterflies are presented in a random dot stereogram; one is only visible when both eyes are used, reducing the likelihood of monocular cues. Plate II shows four discs of which two can only be seen stereoscopically. In →

TABLE 1 Characteristics of commercially available near stereotests. Adapted from Read (2015).²

Name	Viewing distance	Available disparities (seconds of arc)	Task	Physical depth?	Monoc contours?	Monoc cue?	Technology	3D glasses needed?
Frisby	30-150cm	5-600	Odd-one out task based on disparity	Yes	No	Yes	Printed pattern on front / back of transparent tiles	No
Lang I and II	40 cm	550-1200 (Lang I) and 200-600 (Lang II)	Detect disparity-defined object	No	No	No	Printed lenticular stereogram	No
TNO	40 cm	15-480	Identify disparity-defined object	No	No	No	Printed anaglyph cards	Yes
Wirt/Titmus/Stereo Fly test	40 cm	40-3000	Report seeing object in depth and Odd-one out task based on disparity	No	Yes	No	Printed polarised vectograph	Yes
Randot	40 cm	20-500	Odd-one out task based on disparity and Identify disparity-defined object	No	Yes	No	Printed polarised vectograph	Yes

FIGURE 5 Titmus stereotest



the last screening plate, Plate III, the child has to identify four geometric shapes. Threshold measures can be obtained using Plates V-VII. In these plates, discs with a sector missing are presented and the child is asked to point to the missing part of the pie or cake. The tested levels of disparity are: 480, 240, 120, 60, 30, and 15 seconds of arc. For each disparity level, two stimuli are shown. In the most recent version of TNO (19th edition of 2016) the disparities 30 and 15 seconds of arc are no longer included. Threshold with TNO tend to be higher than with other stereotests.¹² TNO showed the highest sensitivity for amblyopia (56%) compared with other stereotests and a high sensitivity of 71% for strabismus.¹⁰ However, another study showed that TNO has the poorest test-retest repeatability in children with abnormal binocular vision compared to Frisby, Randot, and Titmus.¹³

Wirt, Titmus and Stereo Fly test

The Stereo Fly Test, Wirt stereotest and the Titmus stereotest all include the same three contour subtests: a fly, circles, and animals (figure 5). For all subtests polarised glasses must be worn. On the right page, a large fly is shown. With polarized glasses the fly will appear in 3D. It is a screening test for gross stereopsis (800-3000 seconds of arc). For some children, the fly can be scary and no

threshold measurement is possible. The optometrist can ask the child to describe the fly or ask whether it stands out. Alternatively, they can encourage the child to touch the wings and see where in space the child grasps (above the page or on the page). Giving the child suggestions as to what they should see is allowed. This non-standardised approach makes an objective judgment of stereopsis difficult.

A graded measurement of stereopsis is available through the Circles subtest. Nine diamonds each have four circles inside them. One of the circles appears to float in front of the page. The child has to detect the circle that is coming closer to them. The nine diamonds each show the floating circle at different disparities between 40 and 800 seconds of arc when held at the recommended distance of 40cm (16 inches). The final score is the last one chosen correctly before two consecutive mistakes. The contours of the circles are visible monocularly especially at high disparities, making it possible for a stereoblind child to solve at least some of the levels.

The Animals subtest has three lines of each five animals. On each line, one animal stands out in depth at 100, 200 or 400 seconds of arc. The child has to detect that animal. Sensitivity of Titmus is estimated at 63% for strabismus and 38% for amblyopia, making the test unsuitable for screening when used without any additional assessments.¹⁰ Also, test-retest repeatability is relatively poor, but better than TNO.¹³

Randot stereo test

The Randot stereotest has three subtests that need to be viewed with polarised glasses. The Circles is similar to the Circles subtest in Titmus/Wirt, but here three circles are printed in a row and on a background of random dots. However, the test is a contour stereotest with the associated monocular cues in the high disparities. The test shows 10 disparities between 400 and 20 seconds of arc when held at the recommended distance of 40cm (16 inches). The final score is the disparity of the last one chosen correctly. The Randot Animals subtest is the same as the animal subtest of the Wirt and Titmus. Like the Wirt and Titmus tests, the Circles and Animals subtests of Randot have monocular cues. Third, a random dot stereotest is included. Simple geometric forms are

FIGURE 6 Thomson stereotest

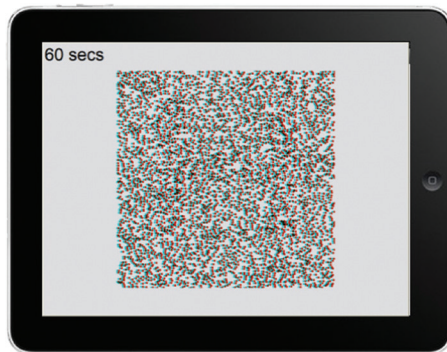


FIGURE 7 Hoya EyeGenius system



shown with a disparity of 250 or 500 seconds of arc. Each disparity level shows three forms and one blank window. The child is asked to detect the shapes, correct identification is not required to pass. This subtest does not have monocular cues if quality of print is good and the child is wearing the glasses and holding their head upright. If the patient tilts their head, the shape can become visible monocularly. Randot has the highest sensitivity for strabismus (75%) of all stereotests, and average sensitivity for amblyopia (36%).¹⁰ Randot stereoacuity improves on average one line (circles test) with repeated testing, but variability within the same child can be up to three lines.¹⁴ Therefore, caution should be taken in interpreting changes of less than three lines.

Table 1 shows a summary of the tests described here.

NEW DEVELOPMENTS

More recently, the first computerised stereotests have become available. Some vision chart software like Thomson Test Chart (figure 6) or Hoya EyeGenius (figure 7) include a stereotest that can be viewed with anaglyph glasses. Paskal 3D makes use of polarised glasses. Vivid Vision combines some stereotesting with amblyopia treatment in a Virtual Reality system. The tasks, disparity levels, and existence of pass/fail criteria in these systems vary greatly and the systems have not yet been evaluated for reliability and sensitivity. Although the costs for combined hardware and software is considerable higher than for the classical stereotests, these systems often include extra features, have higher adaptability and have additional vision tests (eg visual acuity, fusion) or treatment options.

There have also been some developments in the research community for digital stereotests with high accuracy, good repeatability and robustness. The tests provide a wide range of disparities and are not limited to fixed levels like the tests described above. This makes these tests more suited to measure small changes in stereopsis following treatment. Often, great care is taken in assuring high reliability and sensitivity. A staircase algorithm selects the most appropriate disparity level for the child. Children who do well on the first trials will be shown difficult disparities early in the test. Children who struggled with high (easy) disparities will not be challenged or demotivated by being given low (difficult) disparity trials. Two recent examples are the Stereogram Test app developed by McGill University in Canada¹⁵ and ASTEROID developed by Newcastle University (<http://research.ncl.ac.uk/asteroid>).

The Stereogram Test is an iPad app that is to be used with anaglyph stereo glasses at a fixed distance of 50cm. Asteroid is a glasses-free stereotest that runs on a 3D tablet. This test does not require a fixed viewing distance. All instructions are given by game characters. This makes the test fun to play for children and

makes it possible to be administered by an optical assistant. Neither test is (yet) commercially available.

ADVANTAGES AND DISADVANTAGES OF EACH TECHNIQUE

As is evident from the overview above (see also table 1), the commercially available tests differ on a number of domains, each with its own advantages and disadvantages. First, some tests measure coarse or gross stereopsis (contour stereograms like Fly, Circles, Animals) while other tests measure fine stereopsis (random dot stereograms like TNO of the right page in Randot). Since coarse stereopsis can still be intact in children with amblyopia, fine stereopsis is more informative for the screening of amblyopia. Second, some of the tests require polarised glasses (grey glasses), some anaglyph glasses (red/green glasses), while for some glasses are not needed. Glasses can make testing more challenging as not all children are happy to wear them, it may distract them and they can be difficult to fit over other lenses. In addition, the anaglyph glasses used in TNO can promote rivalry, dissociation and suppression which means children perform worse than on other stereotests. A glasses-free stereotest is preferable over a test that requires 3D glasses.

Third, differences in the task can also lead to different results. Simply identifying which region has disparity, as in Frisby or the Circles test, may be a less demanding test of stereopsis than identifying the shape of an object, as in Lang, Randot or TNO.

Fourth, monocular cues are present in many of the stereotest but in different forms. The contours in contour stereotests like Circles, Fly, Animals are often visible monocularly especially for high disparities. This makes it difficult to interpret the results of these tests. Motion parallax can reveal the 'hidden' object in Frisby and Lang stereotest with only one eye. Random dot stereograms (Randot, TNO) suffer less from monocular cues but small dots, dense images and low contrast can make these tests more challenging. In tests with monocular cues, it is advisable to repeat the test with one eye closed. If a child still reaches similar results monocularly, the test cannot be trusted to measure stereopsis with that child.

The levels of disparity also vary greatly between stereotests. One has to be especially careful with interpreting the lowest and highest levels of a stereotest. In healthy people, stereo threshold can be as low as two seconds of arc, but with a stereotest like Titmus it is not possible to measure the difference between a person with a threshold of two seconds of arc and someone with a threshold of 40 seconds of arc. Similarly, the difference between a threshold of 600 and 1000 cannot be detected with a Randot stereotest, because both children will just fail the stereotest.

Other points to take into consideration when deciding which test to use are whether a pass/fail judgement is sufficient or

TABLE 2 Evaluation of the quality of the commercially available near stereotests

Criteria	Frisby	Lang	TNO	Titmus/Wirt + Randot Circles and Animals	Randot Random Dot Stereogram
Measure fine stereopsis	✗	✗	✓	✗	✓
Be glasses-free	✓	✓	✗	✗	✗
Be an easy task	✓	✗	✓	✓	✗
Have no monocular cues	✗	✗	✓	✗	✓
Have large high contrast dots	✓	✗	✗	✗	✗
Measure wide range of disparities with small step sizes	✓	✗	✗	✗	✗
Have good repeatability and sensitivity	✗	✗	✗	✗	✗

whether a stereo threshold is required. For the first, the Fly test or Frisby screening might serve well, for the second a more graded test like the Circles, TNO or Frisby is more suited. Also, repeatability and sensitivity of the test needs to be taken into consideration depending on the aim of the clinical evaluation. Last, one should be aware that the quality of the printed tests decreases over time when the print fades. Replacement of stereotest after a few years is recommended.

CONCLUSION AND RECOMMENDATIONS

An ideal stereotest should:

- 1 Measure fine stereopsis,
- 2 Be glasses-free,
- 3 Be an easy task,
- 4 Have no monocular cues,
- 5 Have large, high contrast dots,
- 6 Measure a wide range of disparities with small step sizes,
- 7 Have good repeatability and sensitivity.

Of all the commercially available tests, none of them meets all these criteria (table 2). Each test has its own advantages and disadvantages. As Frisby stereotest meets four out of seven of the criteria (ticks in table 2), it is probably the best option available at the moment. However, we would recommend using more than one test and combine Frisby with a random dot stereotest like TNO or the right page of Randot to measure fine stereopsis in the absence of monocular cues. Repeating a test with one eye closed is also advisable. The poor sensitivity of all available stereotests (with Frisby being the worst) means that a stereotest should never be used as a sole instrument for screening for amblyopia or strabismus, but should be complemented by other (binocular) vision tests and clinical examinations. Promising new stereotest developments may provide better means for measuring stereopsis in the near future.

Dr Kathleen Vancleef is a Research Associate and Professor Jenny Reed is Professor of Vision Science at Newcastle University.

REFERENCES

- 1 Jandó G *et al.* Early-onset binocularity in preterm infants reveals experience-dependent visual development in humans. *Proc Natl Acad Sci USA*. 109, 11049-52 (2012).
- 2 Read JCA. Stereo vision and strabismus. *Eye* 29, 214-24 (2015).
- 3 O'Connor AR, Birch EE, Anderson S & Draper H. The functional significance of stereopsis. *Investig Ophthalmol Vis Sci* 51, 2019-2023 (2010).

- 4 Mazyn LN, Lenoir M, Montagne G & Savelsbergh GP. The contribution of stereo vision to one-handed catching. *Exp Brain Res*. 157, 383-390 (2004).
- 5 Levi DM, Knill DC & Bavelier D. Stereopsis and amblyopia: A mini-review. *Vision Res*. 114, 17-30 (2015).
- 6 Hayhoe M, Gillam B, Chajka K & Vecellio E. The role of binocular vision in walking. *Vis Neurosci*. 26, 73-80 (2009).
- 7 Kulp MT & Schmidt PP. A pilot study. Depth perception and near stereoacuity: is it related to academic performance in young children? *Binocul Vis Strabismus Q*. 17, 129-34; discussion 133 (2002).
- 8 Kulp MT & Schmidt P P. Visual predictors of reading performance in kindergarten and first grade children. *Optom Vis Sci*. 73, 255-62 (1996).
- 9 Webber AL, Wood JM, Gole GA & Brown B. Effect of Amblyopia on Self-Esteem in Children. *Optom Vis Sci* 85, 1074-1081 (2008).
- 10 Ohlsson J *et al.* Screening merits of the lang II, frisby, randot, titmus, and TNO stereo tests. *J. AAPOS* 5, 316-322 (2001).
- 11 Adams WE, Leske DA, Hatt SR & Holmes JM. Defining Real Change in Measures of Stereoacuity. *Ophthalmology* 116, 281-285 (2009).
- 12 Vancleef K *et al.* Overestimation of stereo thresholds by the TNO stereotest is not due to global stereopsis. *Ophthalmic Physiol Opt*. 37, 507-520 (2017).
- 13 Antona B, Barrio A, Sanchez I, Gonzalez E & Gonzalez G. Intraexaminer repeatability and agreement in stereoacuity measurements made in young adults. *Int J Ophthalmol*. 8, 374-81 (2015).
- 14 Adler P, Scally AJ & Barrett BT. Test-retest variability of Randot stereoacuity measures gathered in an unselected sample of UK primary school children. *Br J Ophthalmol*. 96, 656-661 (2012).
- 15 Hess RF *et al.* A robust and reliable test to measure stereopsis in the clinic. *Investig Ophthalmol Vis Sci*. 57, 798-804 (2016).

ACKNOWLEDGEMENTS

This article presents independent research commissioned by the Health Innovation Challenge Fund (HICF-R8-442, WT102565/z/13/z), a parallel funding partnership between the Wellcome Trust and the Department of Health. The views expressed in this article are those of the authors and not necessarily those of the Wellcome Trust or the Department of Health.

CONFLICT OF INTERESTS

The authors are involved in the development of the ASTEROID stereotest, with Jenny Reed being the principal investigator on the project.