



City Research Online

City St George's, University of London

Citation: van Doorn, L. L. A., Evans, B. J. W., Edgar, D. F & Fortuin, M. (2015). Manufacturer changes lead to clinically important differences between two editions of the TNO stereotest. *Ophthalmic And Physiological Optics*, 34(2), pp. 243-249. doi: 10.1111/opo.12101

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/4867/>

Link to published version: <https://doi.org/10.1111/opo.12101>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Manufacturer changes lead to clinically important differences between two editions of the TNO stereotest

Louise L A van Doorn^{1,2,3}, Bruce J W Evans^{3,4}, David F Edgar³ and Marten F Fortuin²

¹Damme Optometrie, Kesteren, The Netherlands, ²Department of Optometry, Hogeschool Utrecht, University of Applied Science, Utrecht, The Netherlands, ³City University London, London UK, and ⁴Institute of Optometry, London, UK

Citation information: van Doorn LLA, Evans BJW, Edgar DF & Fortuin MF. Manufacturer changes lead to clinically important differences between two editions of the TNO stereotest. *Ophthalmic Physiol Opt* 2013. doi: 10.1111/opo.12101

Keywords: reproducibility, stereoacuity, stereopsis, stereoscopic vision, TNO stereotest

Correspondence: Louise L. A. van Doorn
E-mail address: louise.vandoorn@hu.nl

Received: 27 August 2013; Accepted: 22 October 2013

Abstract

Purpose: Stereoacuity tests used in clinical practice should be repeatable and reproducible. However, it has been observed in a clinical setting that new editions of the TNO stereotest appear to give different values from those obtained using previous versions. The purpose of the present research was to investigate this observation.

Methods: One hundred and twenty-one Dutch subjects, 88 (73%) females and 33 (27%) males, with an average age of 34.0 years (range 18–55) had their stereoacuity measured using two different versions of the TNO stereoacuity test (TNO 13 and TNO 15). The TNO was tested in a counterbalanced order so that consecutive subjects started with alternate editions to avoid bias.

Results: There was a significant difference ($p < 0.001$) between the median value for stereoacuity measured with TNO 13 (30 s of arc) and TNO 15 (60 s of arc). The bias between the two test versions was -0.23 Log arcseconds (95% limits of the differences: 0.15 to -0.60 Log arcseconds).

Conclusion: This study reveals that results obtained with two different editions of a commonplace stereoacuity test are not comparable. New versions come on the market at regular intervals and the assumption that they will give the same results as previous versions may not be valid. Besides the statistically significant difference between the TNO 13 and TNO 15, the Bland-Altman plot also showed a considerable bias and the 95% limits of the differences between the TNO 13 and TNO 15 are more than two steps on the Log arcsecond scale. This difference between two editions of the TNO stereotests is not clinically acceptable and therefore it is inappropriate to use the two versions of the test interchangeably. It is important in both research and clinical records to specify the edition of the TNO test used.

Introduction

Stereopsis is the ability to perceive depth and is important in daily life and an aspect of 'normal' healthy vision. According to Millodot,¹ stereopsis is the 'awareness of relative distances of objects from the observer', and is achieved by means of binocular vision only and based on retinal disparity (the difference between two horizontally displaced retinal images). Stereoacuity measures the threshold (acuity) of stereopsis and is recorded in seconds of arc. It is

routinely measured in hospital eye clinics and also in community optometric practice, most commonly in children. Stereoacuity testing can give the clinician important information from a test of relatively short duration.² It plays a key role in the detection of various binocular vision anomalies,³ and in monitoring the treatment of binocular vision anomalies and certain ocular motility problems.⁴ The appreciation of stereoacuity can indicate the reliability and validity of patient responses during orthoptic exercises. Stereoacuity testing can also give important information

about the visual system, including refractive blur⁵ (decreased monocular or binocular visual acuity) and real-world motor performance.⁶ Stereoacuity is monitored in patients who wear monovision contact lenses^{7–10} and, in addition, the appreciation of stereoacuity is a useful diagnostic tool for detection of malingering or hysteria.¹¹ The widespread use of stereoacuity tests, and in particular their use to monitor the progression and treatment of binocular anomalies, makes it crucial that stereoacuity tests should be reliable and repeatable.

The TNO stereoacuity test was designed by the Institute For Perception, Netherlands Organisation for Applied Scientific Research (TNO) and is distributed by Laméris Ootech BV (<http://www.ootech.nl/>). There are no issue dates on TNO tests, only edition numbers. In 1972 the first edition appeared, and the most recent (17th) edition was published in 2012. Over the years no obvious changes have been made to different editions of the test, instructions, literature or copyright of the manual and no published research has been found regarding changes to the TNO stereotest. The red-green glasses which accompany the TNO test have, according to the manufacturer, never been altered.

However, it has been observed in a clinical setting that new editions of the TNO stereoacuity test appear to give different values than those from previous versions, especially in higher levels of stereoacuity. This observation suggested a need for a scientific investigation of the reproducibility of different TNO stereoacuity test editions.

Materials and methods

This was an observational, balanced cross-sectional method-comparison study. The study was ‘masked’ for the subjects, as both tests were placed in identical book covers. The 13th edition was described as version A and the 15th edition as version B. A literature search revealed one study which assessed test-retest repeatability of a stereoacuity test.¹² Repeatable results were obtained with narrow confidence intervals using 102 participants, aged 2–12 years. Subjects in the current study are older, so our sample is likely to produce more repeatable results. Our target sample size of 100 was in agreement with McAlinden *et al.*,¹³ however recruitment exceeded expectations and the total sample size was 121.

All eligible patients (aged 18–55 years) attending a community optometric practice (Damme Optometrie, Kersteren The Netherlands) had stereoacuity measured using two different versions of the TNO test (data collected between November 2011 and January 2012). All subjects were tested by the same examiner. Written informed consent was obtained and the study was approved by the Research and Ethics Committee of City University London. The study

was conducted according to the tenets of the Declaration of Helsinki.

In the clinical setting it appeared from inspection that in the 15th edition the colour of the plates are subjectively different from previous editions. There are now 16th and 17th editions available, which both appear subjectively to have the same colour differences when compared with editions issued prior to the 15th. Two unused editions were used in the study, one of the 13th edition and one of the 15th, and these are named TNO 13 and TNO 15 throughout this paper.

Inclusion and exclusion criteria

A study informing the design of the current research was carried out by Garnham and Sloper¹⁴ who included the TNO test among the battery of stereotests investigated. Stereoacuity decreased over the age of 55, especially with the TNO. To avoid the possible confounding effect introduced by including over 55 year olds in the current study, this age group has been excluded. Near stereoacuity tests with random dots are more easily degraded by blur¹⁵ because individual dots in the TNO (or any random-dot) test become less visible as blur increases.¹⁶ Accordingly, participants were required to have a monocular distance VA of 0.30 LogMAR (6/12 or 20/40 Snellen) or better. Subjects with an interocular difference in VA were not excluded, provided each of their monocular VAs was better than or equal to 0.3 LogMAR.

A short history and symptoms was taken comprising: chief complaint, refractive history, ocular history, general history, family ocular history and medication. Potential subjects were excluded if they had manifest diplopia, significant ocular disease or a history of squint surgery. Distance VAs (monocular and binocular) were measured according to standard procedures,¹⁷ with habitual correction worn which conforms with common clinical practice and using Snellen optotypes projected by a Magnon CP-670 auto chart projector.

Test procedures

The TNO tests were administered according to the manufacturer’s manual.¹⁸ The tests were carried out at 40 cm with lighting levels set to 500 lux (measured by Vocraft MS-1300 lux meter) and with lighting directed so as to avoid glare affecting the TNO test. Since the TNO manual makes no recommendation about type or level of lighting, this lighting level was chosen based on Richards¹⁹ research finding that 50 fL (=538 lux) is regarded as the standard lighting necessary to perform optimally on visual tasks without a reduction in contrast and fine detail. The 40 cm distance was measured with a cord of 40 cm length attached to a

worktable. The TNO test plates were placed in front of the subject on a durable Plexiglass worktable to ensure a correct reading posture at a 20° angle.²⁰ The book covers of both tests were covered with the same paper to make the study 'masked' for the subjects. Habitual refractive correction was worn. One pair of red-green glasses provided with the TNO 15 was used for all subjects with both tests (the manufacturers confirmed that the red-green glasses had not changed for different editions). The two TNO editions were tested in a counterbalanced manner so that consecutive subjects started with alternate editions (e.g., first subject with 13th edition, 2nd with 15th edition, 3rd with 13th, etc.). The TNO test consists of random dot stereograms viewed through red/green glasses and should be seen in depth. The first three plates (I, II, III) are screening plates with hidden pictures. These quickly establish if the subject has stereoscopic vision of approximately 1980 s of arc.²¹ Plate IV is a suppression test, which can also be used to assess ocular dominance. Plates V, VI, VII are quantitative plates used for more exact determination of stereoacuity (480–15 s of arc). They consist of circles with a 60° sector missing from each one in one of four possible positions. The subject's task is to identify the missing 'piece of the pie'. There are two circles for each disparity (480, 240, 120, 60, 30 and 15 s of arc). To pass at each level the subject must identify the missing 60° sector from both circles correctly (to minimise the possibility of successful guessing). The chance of guessing a correct answer, for one level of stereoacuity, on the random-dot TNO stereotest is 1:16, and the chance of guessing the correct answers on one page of the test (i.e. at two stereoacuity levels) would be 1:256.²¹ All tests took place in one session, lasting approximately 10 min. The first TNO test administered to each patient included both screening and quantitative plates. Eye dominance was tested with plate IV. The remaining TNO tests conducted used only quantitative plates (V, VI and VII).

The TNO stereotests were investigated further at AKZO NOBEL in Sassenheim (The Netherlands) with a photospectrometer, which determined the chromatic properties of the tests through calculations of observed wavelengths. Readings were taken approximately every 5–20 nm in the visible part of the spectrum, and spectral reflectance curves were generated.

The TNO test contains four visible colours: green, red, dark red (comprising the red printed on the green) and white (no print). The suppression plates of both TNO tests were used for the measurement because on this plate there are solid blocks of each colour. On the other plates the same colours are used but the dots are intermixed and too small for easy analysis. The four different colours present on the plates were scanned and spectral reflectance curves generated for the colours red and green. Additionally, the tests were displayed in a 'light chamber' which is a box that

can be illuminated with various sources (daylight, sunrise light and ultra-violet fluorescent light type F6T5/BLB). Both TNO stereotests were also observed and photographed with a standard slit lamp biomicroscope with moderate magnification (40×).

Statistical analysis was performed using SPSS version 20 (www.ibm.com/software/uk/analytics/spss/) and Microsoft Excel 2007. As the steps between the set grades of stereoacuity that can be measured on the TNO are non-linear, the non-parametric Wilcoxon related samples signed rank test was used to check for any significant difference in medians between editions. The agreement between editions was assessed by Bland-Altman difference plots.²²

Results

A total of 121 subjects participated, 88 (73%) females and 33 (27%) males, with an average age of 34.0 years (range 18–55). Mean ages were 34.0 for females and 33.8 for males.

From the short history and symptoms taken, it emerged that one subject wore monovision contact lenses. Another subject had congenital cataract in the right eye. Both of these subjects had a marked difference in visual acuity between the eyes. However, the acuity of each eye was within the acuity limit set by the inclusion criteria, so both were included in the study. One subject had stable keratoconus but was also included as acuities were within the acceptable limit. The mean distance monocular VA was -0.06 LogMAR (~ Snellen 6/5 or 20/15) for the right eye and -0.06 LogMAR for the left eye. The mean binocular VA was -0.11 LogMAR with lowest VA of 0.15 logMAR (~Snellen 6/9 or 20/30) and highest -0.30 logMAR (Snellen 6/3 or 20/10). Most subjects (81%) were equidominant (as measured by TNO test) with most of the remainder (16.5%) being right eye dominant. The stereoacuity values obtained by all 121 subjects from TNO 13 and TNO 15 are displayed in *Figure 1*.

Because the step sizes between grades of stereoacuity recorded in 'seconds of arc' on the TNO test increase exponentially, these values were converted into Log arcseconds, producing equal steps between recorded stereoacuity values.

The median value for stereoacuity for the whole sample for TNO 13 was 30 s of arc, while for TNO 15 it was 60 s of arc. This difference in medians was statistically significant ($p < 0.001$ Wilcoxon signed rank test).

A Bland-Altman difference plot (*Figure 2*) was generated to illustrate the agreement between the two test versions.^{23, 24} Because the differences (between log values) follow an approximately normal distribution, the standard error and confidence intervals can be calculated.²²

The mean difference or bias was -0.23 Log arcseconds (standard deviation = 0.19 Log arcseconds), which

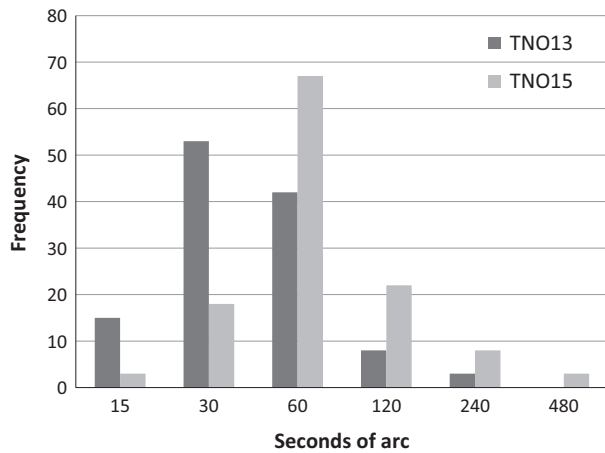


Figure 1. Stereoacuity results (seconds of arc) for all 121 subjects from Editions 13 and 15 of the TNO stereotest.

approaches the 0.30 Log arcseconds difference between any two grades of stereoacuity recordable on the TNO test. This bias on the Log arcsecond scale is also consistent with the finding that the median result was 30 s of arc and 60 s of arc on the TNO 13 and 15 respectively. The 95% limits of the differences are between 0.15 and -0.60 , more than two steps on the Log arcsecond scale (e.g. the difference between TNO stereopsis measurements of 15 and 60 arc seconds is 0.6 Log arcseconds, as is the difference between 60 and 240 arc seconds). Results in *Figure 2* support the view that a clinically significant difference exists between results obtained on the two tests, in addition to the statistically significant difference.

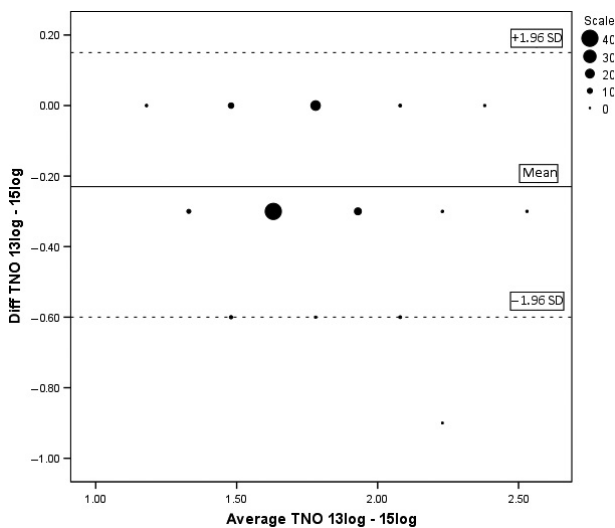


Figure 2. Bland-Altman difference plot for all subjects for both the TNO 13 and TNO 15 tests, with results recorded in Log arcseconds. The log value steps range from 1.18 (15"), 1.48 (30"), 1.78 (60"), 2.00 (120"), 2.38 (240") to 2.68 (480").

Photospectrometry and biomicroscopy of both editions of the TNO test

Figure 3 shows the spectral reflectance of the red and green colours printed in TNO 13 and TNO 15. The red colours are very similar but the greens are clearly different.

With daylight and sunrise illumination, colour differences were observed subjectively between the tests, as measured with the photospectrometer. Under the fluorescent ultra-violet lamp the differences between tests were obvious, with the newer edition scattering back much more short-wavelength radiation than the older edition. This was not measured quantitatively but could be subjectively seen. The UV fluorescence was created by a F6T5/BLB fluorescent light bulb with a narrow spectrum and peak wavelength of 368 nm.

Observation by slit lamp microscope clearly shows differences in image construction between both TNO versions (plate V). TNO 13 (*Figure 4*) shows a much fuller and smoother image compared to TNO 15 (*Figure 5*). Furthermore, it is remarkable that the TNO 15 version gives a pronounced granulated appearance with black pixels included in the image.

Red-green glasses

The standard red-green glasses provided with the TNO tests were analysed with a photospectrometer at HOYA (Uithoorn, The Netherlands). A spectral transmittance curve was generated for both red and green. The transmittance of the red glass was 92% at the peak of 645 nm and for the green glass was 75% at 530 nm.

Discussion

The results support the view that there is a clinically significant difference between the results obtained on the two tests, in addition to the statistically significant difference.

The medians found in the literature from Rosner and Clift²⁵ for adults using the TNO are comparable with the results of the current study. Rosner and Clift²⁵ do not mention which TNO edition they used, though it was clearly an early version of the test. Heron *et al.*²⁶ also investigated the normal values for stereoacuity obtained using different stereotests in young adults (age 18–22), but without stating an edition number for the TNO. They reported a median value for the TNO test of 30 s of arc, which is in agreement with the results of the current study for the TNO 13. Young *et al.*²⁷ found a median value of 60 s of arc based on testing 50 healthy adult subjects. The authors did not state which edition of the TNO was used, but it is possible that a later edition of the TNO was used in this more recent research. This possibility is supported by their finding of median

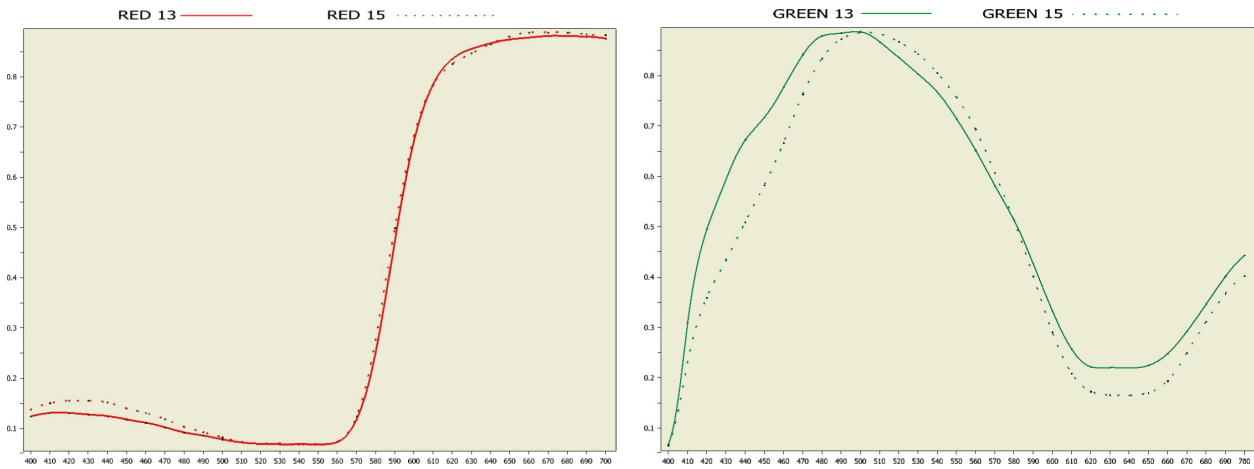


Figure 3. Spectral reflectance curves of the TNO 13 colours red and green (red line/ green line) and TNO 15 colours red and green (red/ green dotted line).

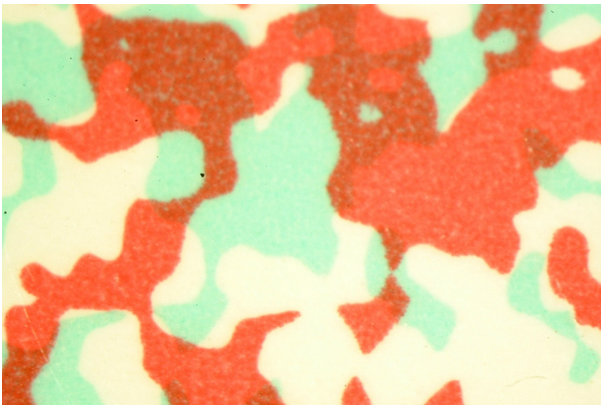


Figure 4. TNO 13 plate V photo graph taken with a slit lamp biomicroscope with 40x magnification. Smooth edges and clear structures are visible.

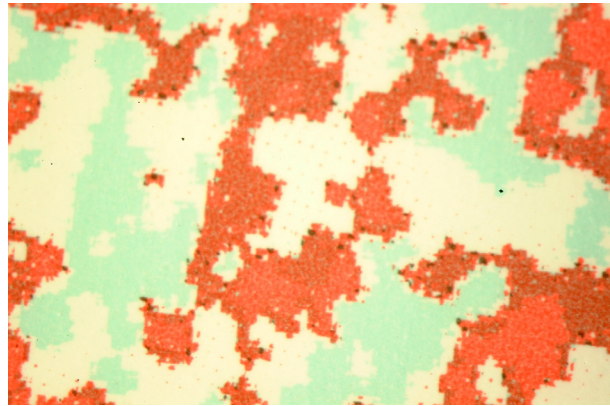


Figure 5. TNO 15 plate V photograph taken with slit lamp biomicroscope with 40x magnification. The edges show a pronounced granulated appearance and a large number of black pixels are visible.

stereoacuity of 60 s of arc which is in agreement with the finding in the current study using the TNO 15. None of the TNO literature reviewed by the authors in the course of the current research states which edition was used. Our results indicate that it is important to specify the edition of the TNO test used in both research studies and in clinical records.

The TNO is based on an anaglyph (red-green) random dot technique. Random dot stereotests are sometimes described as testing global stereopsis (where test features can only be detected binocularly and have no monocularly recognisable form) as opposed to contoured tests (e.g. Wirt circles) where test features can be seen both monocularly and binocularly.²⁸ Anaglyph tests may induce artefacts such as chromatic imbalances that could lead to underestimates of stereoacuity, especially when testing contour stereoacuity.²⁹

Although Larson³⁰ found local stereoacuity with red-green glasses to be higher than global stereoacuity, he concluded that the reduction in stereoacuity was not primarily due to the red-green glasses. Nevertheless, red-green glasses can induce chromatic imbalance because of differences in luminous transmittance.^{31, 32} The red-green glasses used in the current study were provided by the manufacturer of the TNO test (15th edition) and had a luminous transmittance of red 92% at the peak of 645 nm and green 75% at the peak of 530 nm. Bogdanovich *et al.*³¹ and Larson³⁰ found similar transmittance results. Simons and Elhatton³³ and Bogdanovich *et al.*³¹ both reported that image contrast for the red viewing eye was lower than for green. Although luminous transmission of the red lens is higher, the contribution to brightness is much more with the green lens because the green has a higher luminous effect.³¹

If paper manufacturers wish to make paper appear whiter they often add Fluorescent Whitener Additive (FWA). The use of paper with or without FWA can result in colour differences.³⁴ Another possible contributory factor to colour differences is the use of UV-blocking ink/paint, which prevents discolouration with extensive use under UV light. When UV-blocking ink/paint is used in combination with paper with FWA as an additive this can result in even greater colour differences.³⁵ In their research, Andersson and Norberg³⁴ found the greatest colour differences when FWA was added with the colour green. When the TNO 13 and TNO 15 were tested by AKZO NOBEL for the current study, the greatest differences found were also for green.

We speculate that the ink used to print the TNO tests is likely to be a UV-blocking ink, with the new editions perhaps now printed on FWA paper. Additionally, the magnified TNO 15 image appeared granulated and included black pixels in contrast to the TNO 13 (Figures 4 and 5). The resolution of the pictures in TNO 13 appears to be of a much higher quality than in TNO 15. These differences could explain the median differences found in this research but further investigation is needed. According to TNO and Laméris Ootech BV (personal communications 2013) the differences in results between the two editions could be the result of a change in the company printing the plates used in the production of the TNO test.

Limitations of the study

In the current study the researcher could not be masked to the identity of the TNO editions during testing because the colour difference of the paper made each edition's identity obvious. However, the researcher was careful to always use the same verbal instructions for both test versions and not to make any comments or change the tone of her instructions to avoid leading participants.

It would have been desirable to have measured near visual acuity. However, we think it unlikely that this measurement would have explained the findings in this comparison of two editions of a test, using counterbalanced order of presentation, in predominantly pre-presbyopic participants who wore any habitual near vision refractive correction.

Since the subjects were tested within one visit it is possible that the first TNO test could be remembered at the second test, this however seems unlikely because the test has no monocular cues and the quantitative plates consist of 12 circles (two circles for each disparity). Also the counterbalanced test order should minimise systemic bias resulting from memory.

This investigation was limited to two editions of the TNO test. It would be desirable to extend the research to include a comparison of other TNO stereotest editions.

Conclusion

New versions of stereotests become available regularly and caution must be taken not to assume that each new version will give the same results as previous versions. Stereotests should provide accurate results and stereoacuity measurements should be repeatable and reproducible. In particular, there should be minimal variation between results found on the same patient using different editions of a stereotest. However, there was considerable variation in this study between two editions of the same stereotest. This variation between the TNO 13 and TNO 15 is both statistically ($p < 0.001$ Wilcoxon Signed Rank Test) and clinically significant. In clinical terms the Bland-Altman plot showed a mean bias close to the difference between one step and the 95% limits of the differences were more than two steps. These differences in stereoacuity between editions are not clinically acceptable; therefore it is inappropriate to use these two editions of the TNO test interchangeably. This study has demonstrated the importance of comparing new editions of stereotests from the same manufacturer.

Acknowledgement and disclosure

No financial support. No financial interest in TNO stereotest, TNO or Lameris Ootech. Special thanks to Tammo Koster (Akzo Nobel) and Arie van Doorn. Also thanks to TNO & Lameris Ootech for the information provided via personal communication.

References

1. Millodot M. Dictionary of Optometry and Visual Science, 7th rev edn. Elsevier: Oxford, 2009; pp. 11.
2. Lakshmanan Y & George RJ. Stereoacuity in mild, moderate and severe glaucoma. *Ophthalmic Physiol Opt* 2012; 33: 172–178.
3. Von Noorden GK & Campos ES. Examination of the patient V. In: Binocular Vision and Ocular Motility, 6th edn, Mosby inc: St Louis Missouri, 2002; pp. 289–303.
4. Lam AKC, Chau ASY, Leung GYO & Man BSH. Effect of naturally occurring visual acuity differences between two eyes in stereoacuity. *Ophthalmic Physiol Opt* 1996; 16: 189–195.
5. Fricke TR & Siderov J. Stereopsis, stereotests and their relation to vision screening and clinical practice. *Clin Exp Optom* 1997; 5: 165–172.
6. O'Connor AR, Birch EE, Anderson S & Draper H. The functional significance of stereopsis. *Invest Ophthalmol Vis Sci* 2010; 51: 2019–2023.
7. Evans BJW. Monovision: a review. *Ophthalmic Physiol Opt* 2007; 27: 417–439.
8. Jain S, Arora I & Azar DT. Success of monovision in presbyopes: review of the literature and potential applications to refractive surgery. *Surv Ophthalmol* 1996; 40: 491–499.

9. Erickson P & McGill EC. Role of visual acuity, stereoacuity and ocular dominance in monovision patient success. *Optom Vis Sci* 1992; 69: 761–764.
10. Koetting RA. Stereopsis in presbyopes fitted with single vision contact lenses. *Am J Ophthalmol* 1970; 47: 557–561.
11. Levy NS & Glick EB. Stereoscopic perception and Snellen visual acuity. *Am J Ophthalmol* 1974; 78: 722–724.
12. Fawcett SL & Birch EE. Interobserver test-retest reliability of the randot preschool stereoacuity test. *J AAPOS* 2000; 4: 354–358.
13. McAlinden C, Khadka J & Pesudovs K. Statistical methods for conducting agreement (comparison of clinical tests) and precision (repeatability or reproducibility) studies in optometry and ophthalmology. *Ophthalmic Physiol Opt* 2011; 31: 330–338.
14. Garnham L & Sloper JJ. Effect of age on adult stereoacuity as measured by different type of stereotest. *Br J Ophthalmol* 2006; 90: 90–95.
15. Odell NV, Hatt SR, Leske DA, Adams WE & Holmes JM. The effect of induced monocular blur on measures of stereoacuity. *J AAPOS* 2009; 13: 136–141.
16. Larson WL & Bolduc M. Effect of induced blur on visual acuity and stereoacuity. *Optom Vis Sci* 1991; 68: 294–298.
17. Elliott DB & Flanagan J. Assessment of visual function. In: *Clinical Procedures in Primary Eye Care* (Elliott DB, editor), 3th edn, Elsevier: Oxford, 2007; pp. 34–37.
18. Lameris Ootech BV. *Manual TNO test for stereoscopic vision*. 1972.
19. Richards OW. Effects of luminance and contrast on visual acuity, ages 16 to 96 years. *Am J Optom Physiol Opt* 1977; 54: 178–184.
20. Hill SE, Han H & Thorn F. The ergonomics of reading: the relationship of head and eyes to book position. *Invest Ophthalmol Vis Sci* 2006; 47: E-Abstract 1184.
21. Okuda FC, Apt L & Wanter BS. Evaluation of the TNO random-dot stereogram test. *Am Orthopt J* 1977; 27: 124–130.
22. Bland JM & Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8: 135–160.
23. Bland JM & Altman DG. Statistical method for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 8: 307–310.
24. Bland JM & Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 1995; 346: 1085–1087.
25. Rosner J & Clift GD. The validity of the Frisby stereotest as a measure of precise stereoacuity. *J Am Optom Assoc* 1984; 55: 505–506.
26. Heron G, Dholakia S, Collins DE & McLaughlan H. Stereoscopic threshold in children and adults. *Am J Optom Physiol Opt* 1985; 62: 505–515.
27. Young GJ, Sueke H, Wylie JM & Kaye SB. Contribution of a real depth distance stereoacuity test to clinical management. *J Ophthalmol* 2009; 2009: 343827.
28. Evans BJW. *Pickwell's Binocular Vision Anomalies*. 5th edn, Elsevier: Philadelphia, 2007; pp. 33.
29. Yamada T, Scheiman M & Mitchell GL. A comparison of stereopsis testing between red/green targets and polarized targets in children with normal binocular vision. *Optometry* 2008; 79: 138–142.
30. Larson WL. Effect of TNO red-green glasses on local stereoacuity. *Am J Optom Physiol Opt* 1988; 65: 946–950.
31. Bogdanovich G, Roth N & Kohl P. Properties of anaglyphic materials that affect the testing and training of binocular vision. *J Am Optom Ass* 2008; 57: 899–903.
32. Cornforth LL, Johnson BL, Kohl P & Roth N. Chromatic imbalance due to commonly used red-green filters reduces accuracy of stereoscopic depth perception. *Am J Optom Physiol Opt* 1987; 64: 842–845.
33. Simons K & Elhatton K. Artefacts in fusion and stereopsis testing based on red/green dichoptic image separation. *J Pediatr Ophthalmol Strabismus* 1994; 31: 290–297.
34. Andersson M & Norberg O. *Color Measurements on Prints Containing Fluorescent Whitening Agents*, Digital Printing Center, Mid Sweden University: Proc. SPIE 6493. doi: 10.11171/12.704196.
35. Norberg O. Paper whiteness and its effect on the reproduction of colors. *Int Soc Opt Eng* 2007; Proc. SPIE 6492: 64920V.