USE OF STEREO FLY COMPARED TO RANDOT STEREOGRAM ON A NORMAL POPULATION

RICHARD M. HALE

The traditionally used clinical method for determining a patients stereo acuity over the past several years has been the Titmus Stereo Fly Test. Some concerns over the dependability and reliability of this test have been expressed by several researches because this test contains lateral displacement clues which might allow a patient to score higher than his actual stereo threshold. A similarly administered test has been developed which tries to overcome this problem in the form of the Randot Stereo Test. This paper will briefly examine the mechanics of these two tests and attempt to compare the results from the two tests on a series of normal clinical patients.

The Titmus Test Booklet consists of three parts which are administered at 16 inches while the patient is wearing polaroid glasses. The first section is a gross stereo test consisting of a large picture of a house fly. If viewed stereoscopically the patient should report that the wings are seen as apparently floating off the page and a stereo acuity of at-least 3000 seconds of arc is present. Suppression checks are also present in the form of an "R" in a circle seen by only the right eye and an "L" in a square which is seen by only the left eye. The second portion of the test consists of three rows of five small animals and the patient is asked to report for each row which animal seems to be closer to the patient. The correct choice in each row scores the following stereo thresholds: first row 400 seconds; second row 200 seconds; and the third row is 100 seconds of arc. The final section of this test is the part most frequently used when testing older children and adults and consists of nine

numbered diamonds. In each dimond there are four circles and the patient is asked to identify which circle in each diamond they visualize as floating off the page toward them. The amount of disparity for each figure is listed below in seconds of arc:

#1	800	#4	140	#7	60
#2	400	#5	100	#8	50
#3	200	#6	80	#9	40

The disparity in the Titmus Circle Test is provided by the lateral shift of one of two symmetrical forms. This shifted image can be detected, particularly in the first three or four figures, without the use of both eyes or even the use of the polaroid glasses. While these circles are not seen as "floating off the page" when viewed monocularly they easily are seen as being different or blurred and on this basis are selected as the correct response. It has been argued that this is an inherent difficulty with contour stereograms and that a test which lacked these monocular clues would be preferable for use in accessing the true stereo acuity threshold levels in patients.

With the use of computers, targets can be generated which present an array of apparently randomly organized dots. These patterns can be generated to contain a central core of laterally shifted dots which when fused through the use of polaroids result in a central stereoscopic form. These stereo grams supposedly lack any monocular clues and because they are unfamiliar to most patients it is the stereo acuity rather than some other function that is being measured.

The Randot Tests makes use of these stereo grams and is presented in much the same manner as the Stereo Fly Test. Again three different testing levels are possible. The gross test consists of six large squares and the patient must identify the shape hidden inside each one. These univerisally

recognized shapes, if identified, measure an acuity of at least 600 seconds of arc. One of these targets, a star, is made up as a demonstration plate and can be clearly seen without the polaroids and even when viewed monocularly. The second part consists of three rows of five smaller squares and the figure within each square needs to be identified. Correct responses in row one measures 400 seconds, row two 200 seconds, and row three 100 seconds of arc. The final section consists of eight numbered diamonds each with four small circles located at the top, bottom, left, and right. Again, the task is to determine for each diamond which circle is seen as floating off the page.

The disparity values in seconds of arc for the Randot Circle Test are as follows when the test is given at 16 inches:

#1	400	#5	50
#2	200	#6	40
#3	100	#7	30
#4	70	#8	20

Suppression checks are also present with this test in the form of a small circle seen only by the right eye and small square seen only by the left eye when viewed through the polaroid glasses.

Because these two tests are not calibrated to measure the same total amount of disparity in the same sequence of steps it can make the direct comparison of the two test results somewhat difficult. For example, if a patient actually had a stereo threshold value of 110 seconds of arc this should allow the patient to pass up to the 140 seconds level with the Fly Test but only the 200 seconds level of the Randot Test. Thus while both tests have under estimated the patients actual stereo threshold, the Randot made the patients threshold appear much higher than the Fly. Likewise, if the patient actually has a threshold of 20 seconds, he can still score no better than 40 seconds with the Fly Test so the Randot in this case would appear more sensitive to the actual threshold of the subject.

This study consisted of a sample of 40 subjects most of whom were general exam patients at either the Ferris State College of Optometry or the Optometric Institute and Clinic of Detroit. A few friends, relatives, and optometry students were also tested. The subjects were tested with their habitual near correction in place unless they did not bring their glasses in with them or had lost or broken them. In these cases they were tested with no near correction in place. The subjects age, near cover test results, near visual acuties, and power of the correction through which the test was run were all recorded along with the actual responses they made for each test item. Both tests were given to each subject and they were encouraged to respond to each item even if they were unsure of the correct response and had to guess. To eliminate one possible way of prejudicing the study results, the order of the test presentation was altered with each trial and essentially identical directions were given to each subject regardless of which test was presented first.

The patient was asked to put on the polaroid glasses (over their own glasses if they were wearing any) and was handed one of the opened test booklets at about 16 inches while being given the test instructions. They were asked to keep it at that same working distance and not to tilt either the test booklet or their head back and forth. If they were unsure of the correct response, they were encouraged to guess. When they had finished with the first test, they were then given the second with no additional instructions except to say that this was the same type of test only a little different.

Only the results of part three were recorded for each patient although the gross stereo portion was used in most cases as a familiarization or demonstration sequence for both tests. The forty subjects that remained in the study consisted only of "normal" patients. All subjects tested that

showed a heterotropia on the near cover test were eliminated as were all of those with subnormal vision or essentially monocular visual acuties. All subjects thus had at least 20/40 near acuity in each eye and a phoria near cover test of under 12 prism diopters. Also any subject that met the above two requirements but performed very poorly on both of the tests was also eliminated. With these patients there was usually some communication or cooperation problems or they were older patients who refused to guess any answers they they were unsure of. The youngest subject left in the study was five years old and the oldest were two patients aged sixty-nine.

Two subjects were marginally kept in the study who possibly should have been totally eliminated. One was the young five year old patient who appeared alert and cooperative, but failed to make any correct response to the Randot Test while getting a perfect score with the Fly Circles. The other was a 43 year old small angle alternating esotrope who seemed to have been intermittent. He made only randomly correct guesses to the Randot but scored 100 seconds of arc with the Fly Test.

After reviewing the test responses, each subject was given a single numerical score for each test.. This was not always an easy clear cut value to determine and alot of subjective interpretation had to enter into these decisions. For example, a subject who gets the first five correct on the Fly and then misses the last four scores a very clear cut 100 seconds, but the patient who scores the first five correct then misses number six and then gets final three correct is another problem. While the rational in determining the threshold scores may not have been completely sound and accurate, it was at least consistent for both tests and for all subjects involved.

Using these asigned values the results were simply compared to see on which of the tests the subjects scored better on. After eliminating the 30 and 20 second level scores from the Randot test so that a 40 second value was the best possible score for each test the following results were found. Of the 40 subjects involved, 25 of them received the same score on both tests. That is for 62.5% of the total group tested it made no difference which method was used in determining there stereo acuity. Of the remaining 15 subjects, ten of these did better on the Fly test and received lower acuity thresholds as compared to the Randot. Thus 25% of those tested scored better by using the Fly. Included within this group were both the five year old and the alternating esotrope. The remaining five subjects, or 15.5% of the total study group, scored better on the Randot than with the Fly. These results are listed at the end of this paper in Table Number One for easier inspection along with the percentages derived by eliminating the two marginal study members.

We next looked at the total number of errors and the percentage of correct responses for each test and derived the following information. On the Fly test there were a total of only 27 errors out of a possible total of 360 responses for the 40 subjects. A 92.5% correct response rate was thus determined for the Fly. On the other hand, the Randot had a total of 52 errors out of 268 possible responses for a correct response rate of only 83.5%. However, over half of the total errors came in the 30 and 20 seconds of arc categories which the Fly was unable to test for. By eliminating these two lower levels and basing the test on a score of 40 seconds, there now remained only 22 errors out of 240 responses for a correct response rate of 90.83% with the Randot. These two values are very comparable with each other and this 1.67% difference in favor of the Fly can be additionally narrowed and reversed if the responses by the five

year old are discounted from each test. This would leave a 92.3% correct rating for the Fly and 93.16% rate for the Randot for the remaining 39 subjects with only a .88% difference between the two, slightly in favor of the Randot. These figures are again presented in Table Number Two for better comparison.

While there appeared to be no significant difference in the total percentage of correct responses between the two tests we next divided the subjects into various age categories to see if age groupings might produce a difference in results between the two tests. The subjects were divided into seven age brackets and the average threshold value for each group was determined for each test and was plotted on a graph. Values for the Randot were determined both with and without the 20 and 30 second values. The two marginal subjects were not included in these considerations. (The graph is included following the tables at the end of the paper.)

While this grouping results in very small samplings, the graph indicates that there was very little difference between the average score for each group on the Fly as compared to the 40 second level based Randot test. The one exception to this was the over age 60 grouping. Here the four subjects averaged better scores on the Fly by just over 22 seconds of arc than with the Randot. This might suggest that the older population may have more difficulty interpreting the random dot patterns than the general population but such conclusions are very tentative when based on such small numbers.

The average differences between the two test results for each age group discounting the over 60 group, was only about 5 seconds of arc. With all groups included there was still only 7.5 seconds of arc difference between the average scores in each group. In four of the seven groups the Randot produced the slightly lower threshold values while in the remaining

three groups it was the Fly which had the lower threshold. The addition of the 20 and 30 second levels presents essentially parallel findings to the 40 second based Randot with an average value of 12 seconds lower threshold for each group.

While the testing sample was admittedly very small several conclusions can be tentatively drawn from these findings. To begin with there appears to be no substantial difference between the stereo acuity levels for the average normal patient as determined by one test as compared to the other. The only exception we found to this was in patients over age 60 where the Fly gave lower thresholds values. This is not to say that there is not a difference between the two tests. The patients almost all uniformily reported, even though they were not asked, that they felt that the Randot test was much more difficult than the Fly. This was true even of those who scored better on the Randot. It took the patients noticeably longer to make there first selection on the Randot than on the Fly. This was true regardless of which test was given first. This may have been because the number one choice for the Fly has twice the disparity of the Randot but another possibility is that responding to the Randot Stereograms is more of a "learned" process. Most patients gave kind of a "light bulb" response such as "Oh yes, now I see it" and were then able to move rapidly through the remaining test targets. This need to "learn" how to interpret randot stereo grams has been expressed in some of the literature and journal articles.

The time and method of presentation and the costs of the two booklets are about identical and they are both readily available through Titmus so there is no major advantage of one test over the other in any of these areas. The difference with and without the polaroid glasses is much more striking and complete for the gross targets with the Randot than the Fly and

this may possibly be an advantage. Also while some small children may be afraid of a "big bug" and be reluctant to touch his wings, few sense any alarm in seeing a "box" or "ball" and should be willing to trace these figures even if they are unable to name them.

In spite of manufacturers claims and theories, however, the Randot figures are not as invisible without the polaroids as they would have you believe. All of the gross targets can be visualized as something being there without the polaroids and the close observer can identify at least the square and the triangle with good guesses. All four circles are also clearly present in diamonds number one and two while in number three, there is only one visible but that is the correct choice. When viewing the test booklet monocularly through the polaroids, however, these defects are not noticeable at all probably due to the overall darkening of the gray dot pattern.

We have suggested that there is no meaning full or consistent difference between the results of these two tests for a normal patient population. By normal we are referring only to patients between the ages of 5 and 60 with at least 20/40 visual acuity in each eye and a heterophoric near cover test of under 12 prisms diopters. We can make no predictions based on this sampling as to what the results might be for patients outside of this group. Very young children, amblyopes, high phroia, microtropes, anisometropes, low vision or elderly patients may all produce substantially different findings that might suggest one test as being superior to the other as a better diagnostic tool. Several of these areas have been discussed in the literature and more work will certinaly continue to be done in these areas to further our understanding as to how visual function may be related to stereo acuity threshold levels.

In the meantime, the Randot Stereo Test appears to be a simple to administer test which will provide at least as valid and consistent results as the traditionally accepted Titmus Fly. With no evidence to suggest rejecting this form of stereo testing and with the theoretical possibility of superior performance for screening high risk cases such as those listed above, the use of the Randot Test would seem to be a very reasonalble and prudent step for the routine measurment of stereo threshold acuity within the clinical setting.

TABLE ONE

	Total Subjects	Scored Same	Better on Fly	Better on Randot
All Subjects #	40	25	10	5
%	(100)	62.5	25	12.5
Excluding # Marginal Subjects %	38	25	8	5
	(100)	65.79	21.05	13.16

TABLE TWO

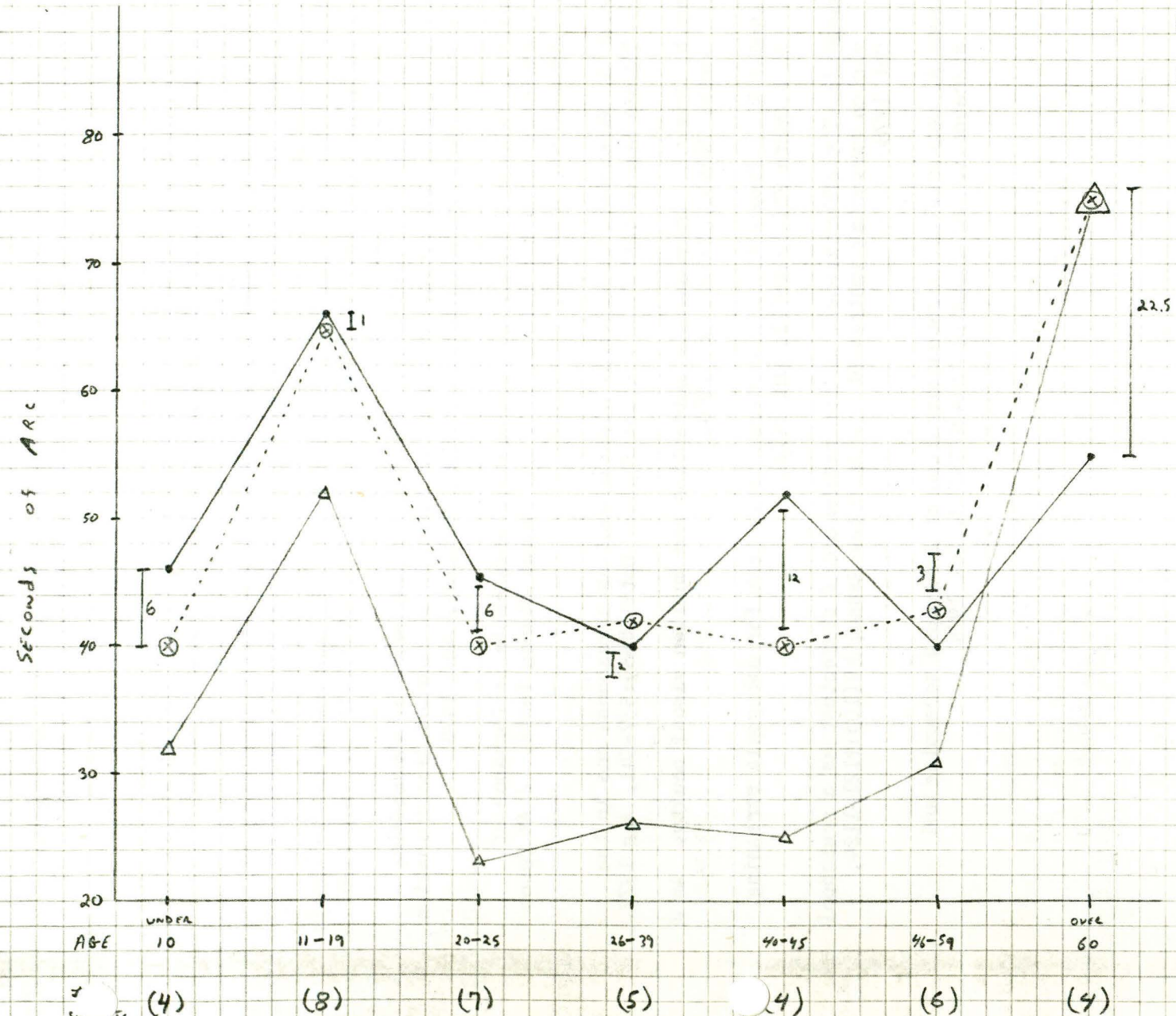
FLY		RANDOT	
Total Errors	27	Total Errors	52
Possible Responses	360	Possible Responses	320
Total Correct	333	Total Correct	268
% Correct	92.5%	% Correct	83.75%

Eliminating 20 & 30 Sec

Total Errors	22
Possible Responses	240
Correct Responses	218
% Correct	90.83%

WITHOUT 5 YEAR OLD

Total Errors	27	Total Errors	16
Possible Responses	351	Possible Responses	234
Total Correct	324	Correct Responses	218
% Correct	92.3%	% Correct	93.16%


Fly •
RanDoT ⊗ (40 sec max)
RanDoT △ (20 sec max)
Seconds of Arc
80
70
60
50
40
30
20
6
6
I1
I2
12
3
22.5
UNDER 10
11-19
20-25
26-39
40-45
46-59
OVER 60
AGE
Subjects
(4)
(8)
(7)
(5)
(4)
(6)
(4)

Partial listing of reference articles which deal with the use of Randot like stereograms which were read and considered in the course of this project:

Cooper J: Clinical stereopsis testing: Contour and random dot stereograms. J. Am. Optom Assoc; 50(1):41-46, Jan 1979.

Cooper J, Feldman J: Assessing the Frisby Stereo Test under monocular viewing conditions. J Am Optom Assoc; 50(7):807-809, July 1979.

Cooper J, Feldman J: Operant conditioning and the assessment of stereopsis in young children. Am J Optom Physiol Opt; 55(8):532-542, Aug 1978.

Cooper J, Feldman J: Random dot stereogram performance by strabismic, amblyopic, and ocular pathology patients in an operant discrimination task. Am J Optom Physiol Opt: 55(9):599-609, Sept 1978.

Cooper J, Feldman J Medlin D: Comparing stereoscopic performance of children using the Titmus, TNO, and Randot Stereo Tests. J AM Optom Assoc; 50(7):821-825, July 1979

Cooper J, Warshowsky J: Lateral displacement as a response cue in the Titmus Stereo Test. Am J Optom Physiol Opt: 54(8):537-541, Aug 1977.

Levy N, Glick E: Stereoscopic perception and Snellen visual acuity. Am J Ophth; 78(4):722-724, Oct 1974

Reinecke R.D, Simons K: A new stereoscopic test for amblyopia screening. Am J Ophth 78(4):714-721, Oct 1974

Rosner J: The effectiveness of the random dot E stereotest as a preschool vision screening instrument. J Am Optom Assoc; 49(10):1121-1124, Oct 1978

Simons K, Reinecke R: A reconsideration of amblyopia screening and stereopsis. Am J Ophth; 78(4):707-713, Oct 1974