5-1993

Clinical trials of the Tomlinson Chromatic Flicker color vision test

John P. Lowery
Pacific University

Debby McClure
Pacific University

Clinical trials of the Tomlinson Chromatic Flicker color vision test

Abstract

Pacific University College of Optometry was contracted to evaluate a new color vision test based on the principle of chromatic flicker fusion. A clinical trials was constructed to compare the efficiency and diagnostic capability of this new test to four standard tests commonly found in clinical settings. The Tomlinson test proved to be quick and easy to administer. The the results show that it detects protans with accuracy and sensitivity, but it misdiagnoses many normal and most deutan subjects as tritans. The fact that it is purely a test of foveal color vision is hypothesized to be the reason for the test's tendency to produce tritan false positives. The underlying assumption that all color deficiencies can be gauged on the same scale of luminosity appears to be at the root of the Tomlinson test's failure to clearly reveal deutans.

Degree Type

Thesis

Degree Name

Master of Science in Vision Science

Committee Chair

Robert Yolton

Subject Categories

Optometry

CLINICAL TRIALS OF THE TOMLINSON CHROMATIC FLICKER COLOR VISION TEST

by

JOHN P. LOWERY AND DEBBY McCLURE

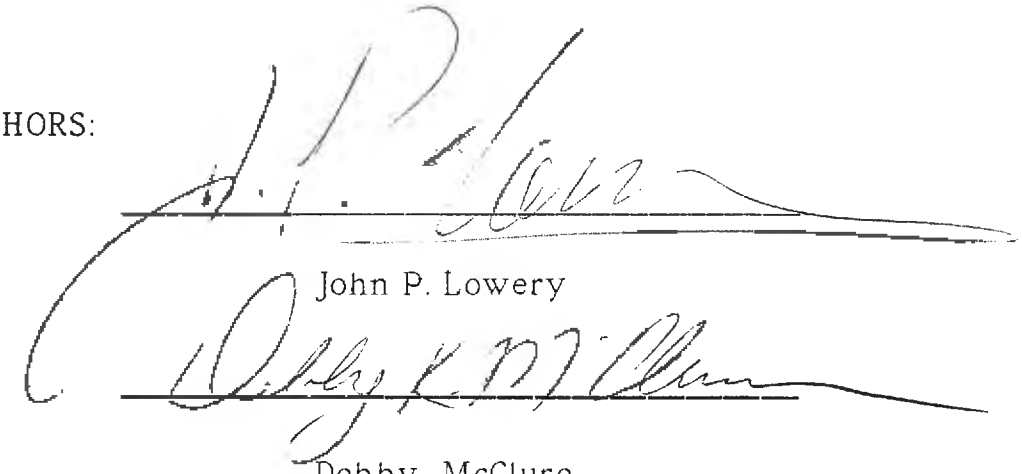
A thesis submitted to the faculty of the
College of Optometry
Pacific University
Forest Grove, Oregon
for the degree of
Doctor of Optometry
May, 1993,94

Advisor:

Robert Yolton O.D., PhD.

Signatures

AUTHORS:



John P. Lowery

Debby McClure

ADVISOR:

Robert Yolton O.D., PhD.

Biographies

John P. Lowery was born and raised in Orinda, California. He received a Bachelor's Degree in Biochemistry and Cell Biology from the University of California, San Diego in 1988. He will receive the Doctor of Optometry Degree from Pacific University, College of Optometry in May, 1993. His academic and professional affiliations include the American Optometric Student Association, College of Optometrists in Vision Development, Amigos, and Beta Sigma Kappa International Honor Society. His primary professional area of interest lies in pediatrics and visual learning disorders. After graduation he plans to complete a residency in pediatrics or seek an associateship in a optometric practice that will allow him to specialize in his area of interest. In his free time, John enjoys astronomy, photography, painting, hiking, and all outdoor sports.

Debby K. McClure was born in Nuremburg, Germany but was raised in the United States. Having obtained a Bachelor's Degree in Psychology from Willamette University in Salem, Oregon in May1987, she continued on and worked in psychiatric research for three years. She will receive the Doctor of Optometry Degree from Pacific University, College of Optometry in May, 1994. Future plans for Debby include an associateship in a general practice with possibly an emphasis on contact lens, although this is undecided at this time. Debby spends most of her free time with outdoor activities.

Clinical Trials of the Tomlinson Chromatic Flicker Color Vision Test

John Lowery and Debby McClure

ABSTRACT

Pacific University College of Optometry was contracted to evaluate a new color vision test based on the principle of chromatic flicker fusion. A clinical trials was constructed to compare the efficiency and diagnostic capability of this new test to four standard tests commonly found in clinical settings. The Tomlinson test proved to be quick and easy to administer. The the results show that it detects protans with accuracy and sensitivity, but it misdiagnoses many normal and most deutan subjects as tritans. The fact that it is purely a test of foveal color vision is hypothesized to be the reason for the test's tendency to produce tritan false positives. The underlying assumption that all color deficiencies can be gauged on the same scale of luminosity appears to be at the root of the Tomlinson test's failure to clearly reveal deutans.

ACKNOWLEDGEMENTS

We would like to thank Dr. Richard Reinke for providing us with his personal copy of the H-RR which has been out of print for many years. Its fine diagnostic capability was extremely valuable to this project.

We would also like to thank Lester Tomlinson for his time and input during the course of the research. His contribution to color vision research and the development of new tests to meet the increasing demands of the clinical setting is commendable.

We would like to extend our appreciation to Beta Sigma Kappa for providing funding for this project.

Finally, we would like to thank Dr. Yolton for all his help, wisdom, and patience.

INTRODUCTION

Normal color vision is based on three receptors, each containing a unique photopigment. Each of the three photopigments have a peak wavelength sensitivity: blue cone at 440 nm, green cone at 540 nm, red cone at 570 nm. The spectral sensitivity of all three photopigments overlap to some degree which allows us to perceive the entire spectrum from approximately 400 nm to 700 nm. (see Figure A) Few sources in the natural world are truly monochromatic. It is the neurological processing within the retina that sums up the cone input within a receptive field to produce the perception of a single color. Thus, we perceive a single color from a polychromatic source.

Color vision deficiency can affect an individual's performance in everyday tasks, especially tasks that are based on color coding. According to Steward and Cole,[1] nearly one-half of dichromats, and one-fifth of anomalous trichromats reported difficulty with traffic lights, as well as having trouble with color related tasks associated with their career. In the same study, many color deficient individuals reported that their color vision had affected their career choice, and some reported it had actually excluded them from a chosen occupation. Careers in the military, transportation, and electrical fields, to name a few, are occupations from which color defectives are usually excluded. Because of these occupational limitations, it is important that color defective individuals be informed of their deficiency at an early age. Color vision testing should be a consistent part of every pediatric examination, especially males. Baseline color vision information may also prove to be vital in the early detection of disease of the optic nerve or retina.[2] Fortunately, these disease states usually present unilaterally while genetic color loss is almost always bilateral. However, in those rare cases with bilateral optic nerve or retinal disease, the color loss may be mistaken for genetic in the absence of previous color test records and objective signs of nerve damage.

INTRODUCTION

Normal color vision is based on three receptors, each containing a unique photopigment. Each of the three photopigments have a peak wavelength sensitivity: blue cone at 440 nm, green cone at 540 nm, red cone at 570 nm. The spectral sensitivity of all three photopigments overlap to some degree which allows us to perceive the entire spectrum from approximately 400 nm to 700 nm. (see Figure A) Few sources in the natural world are truly monochromatic. It is the neurological processing within the retina that sums up the cone input within a receptive field to produce the perception of a single color. Thus, we perceive a single color from a polychromatic source.

Color vision deficiency can affect an individual's performance in everyday tasks, especially tasks that are based on color coding. According to Steward and Cole,[1] nearly one-half of dichromats, and one-fifth of anomalous trichromats reported difficulty with traffic lights, as well as having trouble with color related tasks associated with their career. In the same study, many color deficient individuals reported that their color vision had affected their career choice, and some reported it had actually excluded them from a chosen occupation. Careers in the military, transportation, and electrical fields, to name a few, are occupations from which color defectives are usually excluded. Because of these occupational limitations, it is important that color defective individuals be informed of their deficiency at an early age. Color vision testing should be a consistent part of every pediatric examination, especially males. Baseline color vision information may also prove to be vital in the early detection of disease of the optic nerve or retina.[2] Fortunately, these disease states usually present unilaterally while genetic color loss is almost always bilateral. However, in those rare cases with bilateral optic nerve or retinal disease, the color loss may be mistaken for genetic in the absence of previous color test records and objective signs of nerve damage.

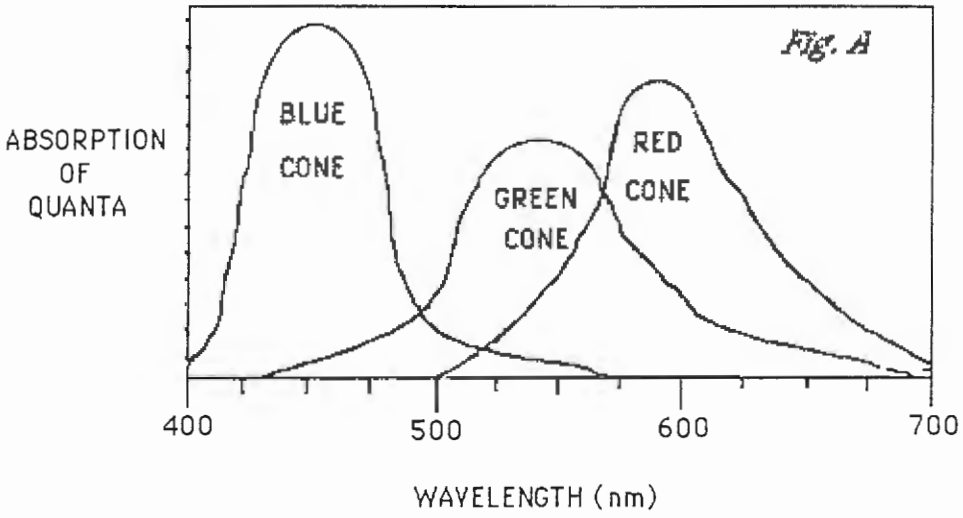
Hereditary color deficient individuals are either missing a pigment or have an anomalous pigment. Table 1 summarizes the categories and distribution of genetic and pathological color deficiencies in the human population. It should be noted that color "blindness" is a misnomer because true monochromats are very rare, rather, the majority of color deficient individuals have decreased sensitivity to portions of the visible spectrum.

As opposed to normal trichromatic vision, dichromats have only two cones, and thus their color discrimination is reduced according to the pigment that they are missing. For instance, protanopes have considerable difficulty discerning reds because their green (540 nm) cone is used to perceive the red wavelengths which it is considerably less efficient at capturing. Anomalous dichromats are thought to have a cone with a displaced peak sensitivity. For instance, a protanomalous individual may have a red cone which peaks at 560 nm instead of 570 nm. It is also possible that the anomalous cone has reduced sensitivity over it's normal range of function.

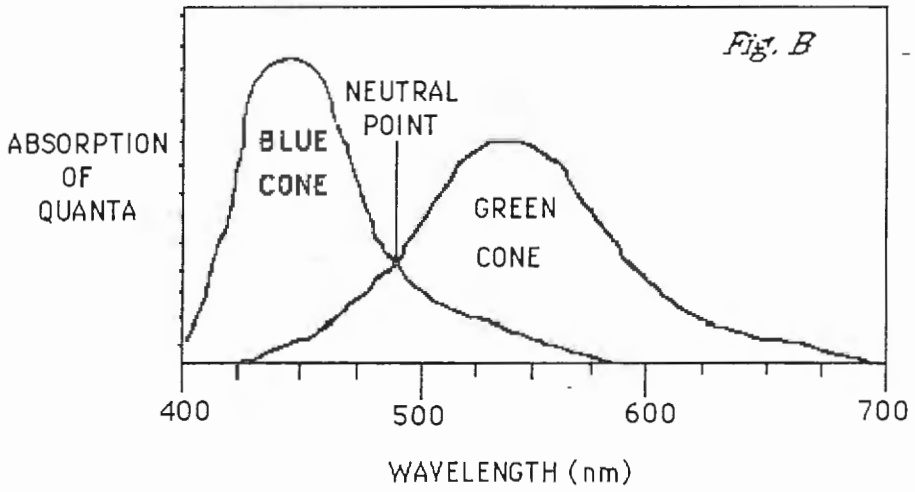
Table 1

COLOR VISION TYPE	PHYSIOLOGICAL BASIS	PREVALENCE IN U.S. POP.
Trichromats	3 Cone Receptors Plus Rods	90% of males
Anomalous Trichromats		7% of males
Protanomalous	570 nm receptor absorption spectrum shifted or reduced	1% of males
Deuteranomalous	540 nm receptor absorption spectrum shifted or reduced	5% of males
Tritanomalous	440nm receptor absorption spectrum shifted or reduced	very rare
Dichromats		2-3% of males
Protanopes	570 nm Receptor Absent	1% of males
Deuteranopes	540 nm Receptor Absent	1% of males
Tritanopes	440 nm Receptor Absent	.005% of males
Monochromats	2 or More Color Receptors Absent	very rare
Acquired Color Loss		
Red-Green	Associated With Ganglion Cell/Optic Nerve Disease	rare
Blue-Yellow	Associated With Intraretinal Nerve Disease	rare
Blue	Age Related Protein Deposition in the Lens and Retina	> 90% of people over 50

SPECTRAL ABSORPTION OF THE THREE HUMAN PHOTOPIGMENTS



SPECTRAL ABSORPTION OF THE PROTANOPE



In testing color vision, there are three component variables of colored stimuli that must be considered; hue, luminosity, and saturation.[3] Hue corresponds to the wavelengths that comprise a color while luminosity is a measure of the number of quanta absorbed of a given wavelength. Saturation is a measure of the depth of a color and is typically defined as the ratio of pure hue to white light in a color stimulus.

Color tests found in clinical settings use hue as the variable to assess color discrimination, and strive to maintain luminosity and saturation as constants. It is essential to the validity of these tests that these latter parameters be kept constant, not only within the framework of the test stimulus, but also in the illumination source that is used. The light source must present equal quanta of wavelengths across the visible spectrum in order to avoid a bias in the reflection spectrum coming off the test stimulus. The established illumination for plate and arrangement tests is known as Standard Illuminant C which approximates natural daylight of the northern hemisphere afternoon sky.[4] The most common source used is the Macbeth Easle Lamp which incorporates a filter over a standard Tungsten source.

Current accepted color vision testing includes plate tests, arrangement tests, and anomaloscopes. The most common plate tests are pseudoisochromatic in which the observer must identify a colored symbol imbedded in a multi-colored background. The only difference between the figure and the background dots is hue. Arrangement tests require a patient to arrange color samples in sequential order according to hue. The most commonly used arrangement test is the Farnsworth D-15 which evaluates color confusion. The D-15 differentiates protan, deutan, and tritan defects from normal by the axes along which color confusions are made. The most accurate color vision test is the anomaloscope. It is an optical instrument in which the subject must alter control knobs to create a match between two colored fields in the display.[4] The anomaloscope is the standard testing device against which all other color tests are compared.

The plate and arrangement tests have limited specificity in testing color discrimination because they present color stimuli that represent ranges of wavelengths. These tests do not allow the clinician to differentiate anomalous trichromats from dichromats. At best, they indicate the kind of defect, (red-green vs. blue-yellow), and the relative severity. Plate and arrangement tests trade off sensitivity in diagnosis for clinical efficiency. The anomaloscope presents individual wavelengths which allows for a very precise

determination of the wavelength ranges that the subject fails to discriminate and, accordingly, a very accurate diagnosis of the color deficiency. Dichromats have a neutral point or specific wavelength that lies at the midpoint between the peak sensitivities of the two cones present.[3] At this neutral point, no color can be perceived, since there is an equal input from both cones. (see Figure B) The resultant perception is white light which is the same as presenting an evenly distributed whole spectrum to a trichromat. Only the anomaloscope can determine the existence of a neutral point and differentiate the dichromat from the anomalous trichromat. Unfortunately, contemporary anomaloscopes are too complex and time consuming to be clinically useful, and are used mainly for research purposes.[4]

Color vision testing is becoming more clinically important and probably will continue to do so. Well supported research has indicated that subtle color vision loss may precede all other signs of retinal and optic nerve damage in certain diseases like Glaucoma and Diabetes.[2,5,6] For the purpose of disease detection, sensitivity and accuracy in color testing is paramount. Pathological color loss also does not present in the predictable patterns of genetic loss. The ideal test that would serve all clinical needs would accurately determine the specific type of color vision loss as well as reliably quantify the depth of the defect.[7]

Perhaps in response to these demands, a different approach to color testing has been studied. The basic premise of this approach is to use luminosity as the test variable, keeping hue and saturation constant. To test color vision in this manner, you could simply present two bordering stimuli of different hues and vary the luminosity between the two until the color deficient subject cannot distinguish the two. In reality, there are only limited ranges of hue in which a color deficient will perceive a match, and these ranges of hue are very specific to the type of deficiency. In addition, the extent of the hue range would have to be varied according to the severity of the color deficiency. A dichromat would perceive matches over a much larger range than an anomalous trichromat. Thus, to assess color vision in this manner, many different opposing hue ranges would have to be presented sequentially until the exact matches of the observer were known. Even then, the results would need interpretation. This process is very much like the analysis performed with an anomaloscope, thus no purpose would be served by its development..

One way of overcoming the need to present variable hue ranges to different observers is to change the conditions of the stimulus making it easier for a match to be made. This has been done

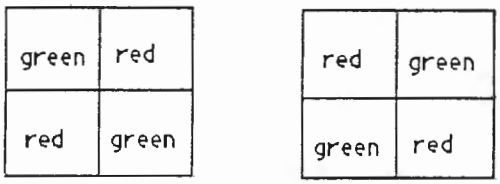
in recent years using the phenomenon of flicker fusion.[8] If two colored stimuli are presented in a test grid each flickering off/on in an alternating square wave pattern, then adjusting the relative luminosity of the two stimuli as close to unity as possible will reduce the perception of flicker to a minimum. In other words, by decreasing the luminosity contrast between the two stimuli to a minimum, the degree of flicker perceived between the stimuli will also be at a minimum. Therefore, a color testing device can be created out of this system by simply measuring the relative luminosities of the flickering source at the subject's perceived minimum flicker point.

The first test using this principle, the Oscar (Objective Screening of Color Vision Anomalies), was developed by Estevez et al. in 1983.[9] The test screens out red/green deficiencies by presenting alternating red and green lights in the same position. No blue stimulus is presented in this test; therefore, it does not evaluate tritans. The normal, deutan, and protan observer will each adjust the red and green stimuli to different relative intensities to create the test endpoint of a minimum flicker. For example, the protan will make the red flash brighter compared to the green in order to perceive the least flicker, the deutan just the opposite. Highly anomalous trichromats and dichromats may reach a point of no flicker perception because the contrast between the sources may be reduced nearly to zero for them. The relative luminosities of the two sources at the test endpoint not only screen out color defectives, but also determine the severity of the defect. For example, the more deficient a protan's red cone is, the more intensity he will add to the red source to reach the test endpoint.

The test apparatus consists of a small hand held box that the subject looks into to see the flickering field. The subject adjusts a dial that controls the relative intensities of the two sources to find the minimum flicker point. The test is easy and quick to administer to all ages, and is not dependent upon visual acuity or form perception. The developers of the test found that its results correlated well with the results of other color vision tests, especially the Nagel Anomaloscope.[7,9]

Using the same principles as the OSCAR, Professor M.L. Tomlinson of the University of Christchurch, New Zealand has developed a computer based test that screens protans, deutans, and tritans. The test is a program that runs on an IBM compatible computer with MS DOS and a VGA color monitor. The test works essentially the same way as the OSCAR except that it consists of four

squares with two colors presented diagonally and alternating position to create the flickering stimulus.



cycle one cycle two

The 4-square grid is approximately 3 cm square, and the test distance is 1 meter making visual acuity an insignificant factor. The test grid subtends an angle of approximately 1.91^0 on the retina (or covers an area of .675 mm^2). The human fovea measures about 1.5mm in diameter,[11] so this test basically analyzes foveal color vision only. This may be a very important parameter that we will discuss later.

The intensity of each test color is determined by the drive to the VGA monitor gun which is on a scale from 1-63, 1 being nearly black, 63 being brightest. The test is subdivided into three color pair matchings for which a minimum flicker point must be found (red/green, blue/red, & blue/green). As the drive to one gun is increased, the other is decreased and they start at opposite ends of the scale so that a minimum contrast or *minimum flicker* point is found somewhere in the middle. The computer program takes the ratios of the drives to the color guns for all three tests and computes a score as a percentage of the norm. The norm is established by simply having a known color normal subject complete the test, and then a normalization key is pressed to lock in those ratios. At the upper left-hand corner of the screen is a table that shows the values (1-63) of the drive to the gun of the first color of each pair. These values represent the raw data of the test, and it is these that are used when the computer is normed at the beginning of each test session.[10]

The physiological optics lab under the direction of Dr. Robert Yolton at Pacific University College of Optometry was asked to evaluate the Tomlinson Color Test in the spring of 1990. We constructed a clinical trials that would establish the usefulness of the Tomlinson Test by comparing its efficiency in terms of time and labor and its diagnostic capability to standard tests in clinical use.

METHODS

The subject population consisted of 122 normal and color deficient men and women ranging in age from 6 to 76 years old. The mean for the normals was 25.1 years old with 6.3 years standard deviation. The mean for the color deficients was 34.3 years old with a standard deviation of 16.5 years. The majority of the subjects were Pacific University College of Optometry students. In order to reach a more diverse age group and obtain more color deficient subjects, an article was placed in the local newspaper describing the project and asking for people who suspected they had color deficiencies to participate in the project in exchange for a complete vision examination. Notices were also placed around campus.

For our standard clinical tests that we would use to compare the results from the Tomlinson test we chose the Farnsworth D-15 arrangement test and three pseudoisochromatic plate tests, the AO-Hardy-Rand-Rittler, Ishihara, and the standard AO Pseudoisochromatic test. The D-15 screens gross color defects across the visual spectrum and provides a classification of the type of defect by the axis along which the color confusion is made. It also indicates the depth of the defect by the number of axis crossovers. The AO-HRR consists of 6 screening plates and 14 diagnostic plates that allow determination of both the type and severity of the color deficiency. The Ishihara is the most sensitive standard that we used. It screens red/green defects with 15 plates and indicates the the type (protan or deutan) by two diagnostic plates. The AO consists of 18 plates that screen red/green deficiencies without indicating the specific type of defect. For more specific information regarding these color tests see Procedures for Testing Color Vision.[4]

All subjects were assigned a number in order to maintain confidentiality. They were then tested using a standard Snellen nearpoint card to determine the 20/200 minimum visual requirement for this project. 20/200 is the suggested visual acuity requirement for the plate and arrangement tests to be reliable. Those requiring strongly tinted lenses were excluded from our trials. Light tints like contact lens handling tints were allowed, but these cases were recorded. The HRR, AO, D15, and the Tomlinson computer test were administered in random order. All tests were performed monocularly at a 40 cm working distance for the plate tests and D15, and a 1m distance for the Tomlinson computer test. The HRR and AO

plate tests were performed under a MacBeth easel lamp, as was the D15. The plate tests were administered and scored according to established clinical procedure. The D15 was rearranged in a random order prior to the subject's arrival. The subject had the suggested 2 minutes to respond, although no one seemed to need more. For the younger subjects that seemed hesitant or had trouble calling out the numbers in the plate tests, they were given a small soft bristled brush to trace the numbers they saw.

Throughout the course of the study one researcher administered the standard tests (Ishihara, HRR, D15, and the AO), while the other administered the experimental computer test so that for each subject the results of the comparison tests remained unknown to the researchers.

Administration of the Tomlinson test runs as follows. The computer must be turned on a few minutes prior to test administration in order to give the color phosphors time to warm up. The program must be normalized by inputting the predetermined values from the standard normal subject and pressing key 2. Then, pressing the initialization key 1 puts all three raw values back to 32 which is the starting point. The subject is seated so that his face is 1 meter from the test screen. The test begins with the red/green subtest and the subject is asked to adjust the flicker he sees to a minimum using the left-right arrow keys on the keyboard. The subject is encouraged to bracket the endpoint in order to attain greater accuracy. Occasionally a subject would be hesitant or uncertain as to the manipulation of the arrow keys and the researcher would then demonstrate how the arrow keys caused the flicker on the screen to change without making reference as to his perception of the flicker. Pressing the up or down key moves the program to the next subtest. When all three subtests are completed, the examiner presses the down arrow and the evaluation key 3. The computer now comes up with sensitivities to red, green, and blue as a % of the norms. Also, a confidence figure appears at the bottom of the screen. This confidence figure represents the product of all three raw score ratios and should be within .1 of 1.00 in order for the test to be valid. Basically, this confidence figure tells us the reliability of the subject.[10]

Scoring the test is simple according to Tomlinson. Normals should be near 100% of the norm for all three colors. Protans should be slightly reduced in green perception and markedly reduced in ability to perceive red. Deutans should be most reduced in green perception and slightly reduced in red. Tritans should be most reduced in blue perception with some reduction in the green as

well.[10] No absolute values were prescribed for cut-offs between normal and defectives at the time we began our evaluation. During the data compilation, we chose a break of 90% between normal and color defectives.

The raw ratio scores from each test are converted to % scores by the computer program using the following criteria:[10]

BLUE = 100% if: green < blue < red
red < blue < green
red < green < blue
green < red < blue

So, the only way blue sensitivity can be established at less than 100% is if it is the least luminous color to the observer.

Reliability was established by repeating the Tomlinson test on 15 normals and 12 color defectives. The same color normal person was used as the standard throughout the experiment, and was retested at the beginning of each testing sequence to insure the reliability of the normative data and against possible instability of the computer monitor. All data analysis and comparisons were performed after completion of all subject testing. Subjects were scored pass or fail and deficiencies were categorized on all of the color tests according to procedure instructions by test authors.

RESULTS

Out of the total subject population for our clinical trials, 31 were determined by one or more tests to be color defective. Table A summarizes the total color defectives identified by each test.

TABLE A

TOTAL COLOR DEFECTIVES SCREENED OUT BY:	
D-15	22
HRR	29
ISHIHARA	30
AO	27
TOMLINSON	26

Next we broke down the color defective population by comparing the Tomlinson Test classification to that of the D-15 and HRR. The AO only determines if the subject has a red/green defect and is only useful for comparing relative severity in terms of how many plates are missed. The Ishihara, although it claims to differentiate protans and deutans by two diagnostic plates, was found to be inadequate for this purpose. Basically, it determined all color defectives to be deutans. The following tables B-D summarize this comparison data.

TABLE B

COLOR DEFECTIVES CLASSIFIED BY THE TOMLINSON TEST AS PROTANS

Criteria: RED $\leq$ 90% Total = 16

CLASSIFICATION	D-15	HRR
PROTAN	8	10
DEUTAN	1	0
TRITAN	0	0
NORMAL	5	2
UNCLASSIFIED	2	4

TABLE C

COLOR DEFECTIVES CLASSIFIED BY THE TOMLINSON TEST AS DEUTANS

Criteria: GREEN $\leq$ 90% Total = 0

CLASSIFICATION	D-15	HRR
DEUTAN	11	14
PROTAN	0	0
TRITAN	0	0
NORMAL	4	0
UNCLASSIFIED	0	1

TABLE D

COLOR DEFECTIVES CLASSIFIED BY THE TOMLINSON TEST AS TRITAN

Criteria: BLUE $\leq$ 90% Total = 35

CLASSIFICATION	D-15	HRR
TRITAN	0	0
DEUTAN	8	10
PROTAN	0	0
NORMAL	27	24
UNCLASSIFIED	0	1

As we can see from the preceding tables, the Tomlinson test is very good at detecting Protans and may be more sensitive than the other tests in this regard indicated by the few undiagnosed (normal) and unclassified subjects that it diagnosed as mild Protans. In detecting deutans, the Tomlinson test fails based on the Tomlinson classification scheme because the test misdiagnoses deutans as tritans (blue-loss only). Note that there are no color defectives with a green % $\leq$ 90, yet the other tests found 15 deutans combined. (see table C) Looking at Table D we can see that the color defectives that the Tomlinson test calls tritans are really deutans as determined by the other tests.

The Tomlinson test appears to be quite consistent in terms of values produced for the same subject on retest. Table E shows the raw scores and percentage scores on two consecutive trials for 15 normal subjects and 12 color defectives. Note that among the normals there was very little variation between scores on consecutive trials, but two out of the color defective group scored values that differed enough to yield different diagnoses. Especially notable is subject #58 who was diagnosed as protan on the first run and tritan on retest. From the subjective observation logs of the researchers, it was noted that this particular subject was very tentative and unsure of her ability to find the test endpoint of minimum flicker. No other subjects in the study required as much

direction and reassurance as subject #58. Other than this one subject, the data indicates that the Tomlinson test is very reliable.

TABLE E

	#	COMP R/G	COMP R%	COMP G/B	COMP G%	COMP B/R	COMP B%	COMP CONF	TYPE (P,D,N)
	45	43	98	21	103	33	100	1.06	N
		43	98	21	103	33	100	1.06	N
	47	41	97	21	100	34	94	0.99	N
		43	95	22	95	34	100	1.22	N
	50	43	95	20	100	34	94	0.93	N
		43	95	20	100	34	94	0.93	N
	51	41	99	20	100	36	92	1.04	N
		41	99	20	100	36	92	1.04	N
	52	43	97	20	100	35	96	1.12	N
		43	97	19	100	36	92	1.11	N
	54	41	99	21	100	35	95	1.05	N
		41	99	21	100	35	95	1.05	N
	56	41	96	43	100	31	100	0.94	N
		42	95	23	97	32	100	1.07	N
	57	43	95	20	100	33	95	0.99	N
		43	98	21	100	33	100	1.06	N
	59	43	94	22	96	32	100	1.07	N
		44	95	21	97	33	100	1.14	N
	61	42	95	19	100	34	90	0.91	N
		42	97	20	100	34	94	0.98	N
	66	44	92	22	91	33	100	1.23	N
		43	97	21	101	32	100	1	N
	67	43	95	20	100	33	95	0.99	N
		43	95	19	100	34	92	0.98	N
	68	44	93	21	100	31	100	1.01	N
		44	91	22	95	31	100	1.08	N
	69	42	97	20	100	34	94	0.98	N
		43	100	20	100	35	98	1.12	N
	70	42	99	19	100	36	92	1.04	N
		42	100	19	99	37	93	1.1	N
MEAN VARIANCE OF NORMALS			1.33		1.66		1.8		
	72	45	86	22	95	28	100	0.96	P
		46	88	21	95	30	100	1.1	P
	85	46	70	26	77	24	100	1.05	P
		46	70	26	77	24	100	1.05	P
	87	44	89	22	93	31	100	1.08	P
		42	87	21	97	29	100	0.97	P
	46	41	97	19	100	36	88	0.97	BL
		42	90	16	100	36	78	0.82	BL
	48	40	99	18	100	38	84	0.95	BL
		41	96	18	100	37	85	0.96	BL
	49	41	96	18	100	37	85	0.96	BL
		41	96	18	100	37	85	0.96	BL
	53	40	100	17	100	40	83	1	BL
		40	98	17	100	39	82	0.94	BL
	55	47	72	25	81	25	100	1.14	P
		47	75	24	84	26	100	1.13	P
	58	44	80	25	86	26	100	0.96	P
		41	97	19	100	35	89	0.91	BL
	60	41	98	16	100	39	81	0.93	BL
		41	100	17	100	39	85	1.01	BL
	62	42	95	18	100	35	87	0.9	BL
		42	99	19	100	36	92	1.04	N
	64	40	99	18	100	37	86	0.89	BL
		40	97	17	100	37	82	0.83	BL
MEAN VARIANCE OF COLOR DEFECTIVES			3.66		1.75		3		
MEAN VARIANCE ALL REPEATED SUBJECTS			2.37		1.70		2.33		

As an additional measure of the consistency of the Tomlinson program over time, we recorded the normative data that was used to begin each test session. It shows that the Tomlinson program, our computer monitor as well as the color perception of our standard (Debby) were very consistent over the course of our study. (see table F below)

DEBBIE	RED/GREEN	GREEN/BLUE	BLUE/RED	CONFIDENCE
4/8/91	40	23	33	1
4/13/91	40	23	34	1.06
4/15/91	41	22	34	1.06
4/17/91	40	23	33	1
4/18/91	40	23	33	1
4/23/91	41	22	33	0.99
4/25/91	41	22	33	0.99
4/29/91	41	22	33	0.99
4/30/91	40	22	34	0.99
5/1/91	40	22	33	0.93
MEAN	40.4	22.4	33.3	1.001
SD	0.52	0.52	0.48	0.04

The Tomlinson test proved to be quick and easy to administer taking approximately 2.5 minutes to complete testing on both eyes. The vast majority of the subjects were able to find the test endpoints easily. An important attribute of this test is its objectivity; there is no manipulation of data by the administrator. Also, it is no more difficult for the color defective to perform than the normal subject. There are no incorrect perceptions. In fact, it may be easier for the color defective to find the test endpoints because the contrast between the flickering colors is less for them, effectively dampening the perception of flicker. Many of our color defective subjects reported difficulty or frustration in taking the plate and arrangement tests, but were pleased with how easily they completed the Tomlinson test.

Even though the Tomlinson test incorrectly diagnosed deutans as tritans, manipulation of the data indicates that it has some capability to detect deutans. Looking at our population of 15 deutans as classified by the D-15 and/or H-RR, it is apparent that not only are they reduced in blue perception according to the Tomlinson test, but 10/15 were also reduced in green color perception as well. The majority (74.7%) of our normal subjects showed some reduction in blue perception, but only 18.7% were determined by the Tomlinson test to have a reduction in green color perception. Altering the criteria for classifying deutans by the Tomlinson test, it yields some interesting results. After all, no clear cut-offs between normals and color defectives had been established for this new test prior to this study. It is was our comparison data that would determine these values. So, if we establish the criteria for classifying deutans by the Tomlinson test as Blue $\leq$ 91% and green $<$ 100, this yields 10/15 (66.7%) of the actual deutans as determined by the other tests. Unfortunately, this still leaves 1/3 of the deutans misdiagnosed and also produces a false positive rate of 3/91 (3.3%).

A question that remains to be answered is why so many of the normal subjects subjects were determined by the Tomlinson test to have reduced blue color perception. If we maintain our criteria for establishing tritans at 90%, this classifies 24/91 normals as tritans or a false positive rate of 26.3%. Unfortunately, we had no true tritans in our study, so we cannot determine if the Tomlinson test would screen them. Maybe, it would show tritans as *extremely* reduced in blue perception, but we just don't know.

DISCUSSION

After considerable analysis of the data, we are still puzzled by some of the trends that were uncovered. The Tomlinson test does not appear to test color perception accurately and reliable across the visible spectrum. Clearly, it does assess red perception with at least the accuracy of the HRR and D15 and may be more sensitive than these standards at picking up red color deficiency. Also, our analysis seems compelling enough to suggest that the Tomlinson program could be fine tuned to weed out deutans from the normals that show blue color loss as we have done. However, our data pool of deutans was not large enough for us to draw conclusions about the Tomlinson Test's deutan detection capability. We can conclude that the Tomlinson test in its present form does not possess the sensitivity to reliably detect green color loss. This fact seriously limits its clinical usefulness.

So why does the Tomlinson test fail to detect deutans when the OSCAR which works on the same principles of human perception was shown to be quite reliable in this respect? There are a few obvious and subtle differences between the two tests that may be contributing to their unequal diagnostic capabilities. The most obvious difference is that the Tomlinson test assesses blue color perception in addition to red and green. This design is commendable in that it potentially makes the test useful over a broader range of clinical applications. However, blue color perception is highly variable in the population due to many preretinal factors.[12] For a test that is supposed to assess photoreceptor physiology and retinal integrity, this presents a problem.

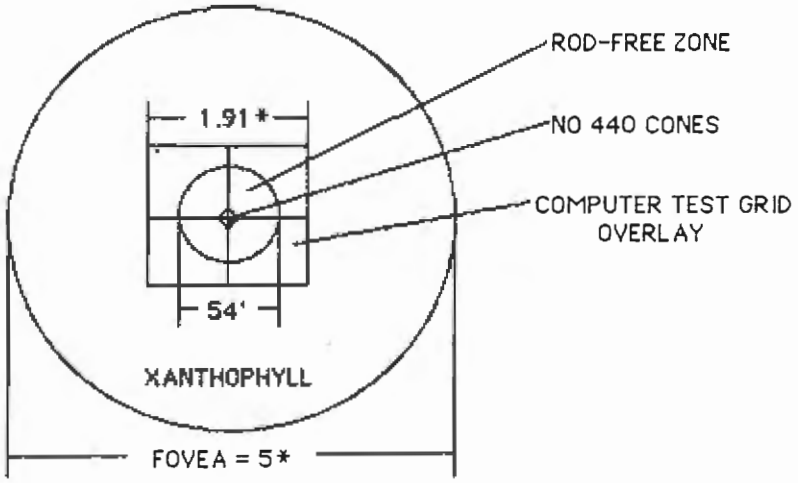
We did find that blue color perception was the most variable and reduced in our normal population. This variable blue color loss seems especially surprising in light of the Tomlinson programs bias to set blue perception at 100% (see page 9). Perhaps the author designed this intentionally to mask the variable blue perception in the human population and reveal blue loss only if it was abnormal. Regardless, the results of the Tomlinson test in terms of blue color perception at the fovea are most likely valid. However, human color perception is based on extrafoveal receptor pathways as well, so the blue subtest results on the Tomlinson test are probably skewed to show blue deficiencies that are not necessarily present.

Under normal conditions, the primary variable in blue color perception is the attenuation of blue light by the various media of the eye, mainly the crystalline lens. It is well known that this absorption of blue wavelengths by the lens increases with age, however there was no correlation between blue color perception loss and age in our subject population. This leads us to look at more specific conditions of the Tomlinson test to determine why our subjects had such variable blue color perception.

At the root of this variability may be the fact that the Tomlinson test stimulus falls entirely within the fovea (see figure C). The human fovea contains variable amounts of the blue absorbing pigment xanthophyll.[13] It is this foveal pigment that is most likely contributing to the blue color loss on the Tomlinson test. In fact, the Tomlinson test may be a direct way of measuring the relative amounts of xanthophyll in an individual's fovea, but we have no way of verifying this from our study.

Fig. C

THE FOVEA



AREA OF XANTHOPHYLL VARIES UP TO 5* FROM THE CENTER OF OF THE FOVEA. IT IS LIKELY THAT FOR THE MAJORITY OF OUR SUBJECTS, THE RETINAL IMAGE OF THE TEST GRID FELL COMPLETELY WITHIN THIS AREA.

Another factor that may play a small role in the variable blue color perception as determined by the Tomlinson test is the lack of 440 nm (blue absorbing) cones at the very center of the fovea.[3,11] The retinal image size of the Tomlinson test stimulus is quite a bit larger than the area void of 440 nm cones, but it is possible that this physiological phenomenon plays a small role in the decreased blue perception of subjects on the Tomlinson test. Lending support to this theory is the subjective perception of differential flicker between the center and perimeter of the Tomlinson test stimulus that is most pronounced in the blue-green and blue-red subtests. In fact, many subjects asked if they should use the outside or the center of the test grid in determining the minimum flicker point for those two subtests. We consistently told them to use the center to make the judgement. Thus, the subjects were most likely using the center of their foveas where there are few if any 440 nm cones to perceive blue light. Consequently, they would adjust the blue brighter to reach the minimum flicker point. This adjustment between where the minimum flicker was perceived in the center versus the perimeter of the stimulus was only one unit on the scale from 1 to 63. One unit in the raw ratio score of the Tomlinson test corresponds to a 4 to 6% difference in the final assigned color perception. Added into this scenario is the variable nature of human fixation and the likely variation in the size of the 440 nm cone-free zone in the human population. It is likely that the lack of 440 nm cones in the central fovea is playing some role in the variable blue perception of our subjects on the Tomlinson test. Just how much of a role this plays is very uncertain.

Regardless of the exact mechanism of the variable blue color vision results on the Tomlinson test, we believe that the addition of the blue stimulus contributed to the test's inability to clearly detect deutans. The fact that the test grid falls within the fovea certainly confounds the results of the blue subtest and, because of the way the Tomlinson test is constructed, this weakness in the blue portion has ramifications for the validity of the entire test.

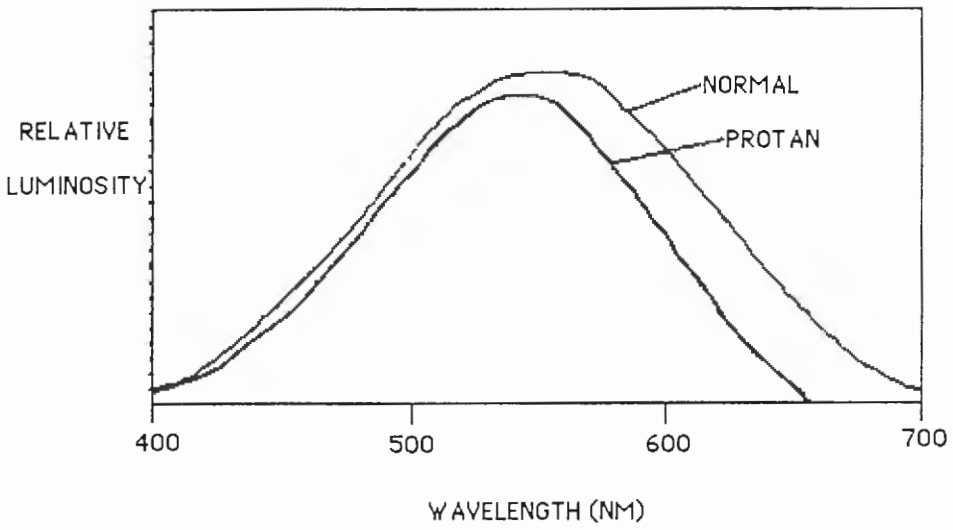
The real problem with the Tomlinson test is that it uses the ratios of the three colors from the paired comparisons to determine the relative color loss in any one part of the three parts of the spectrum that it tests. These ratios are then multiplied together and must yield a product of unity (1.00) for the test to have a reliable confidence figure. Thus, if there is a decrease in perception of any one part of the spectrum, it correspond to a relative increase in the other parts. So, if there is a decrease in the perception of the blue

stimulus, this will translate to an increase in the green and red. The problem is that a deutan has decreased sensitivity to both green and red as well. So, the Tomlinson test masks the real reduction in green and red perception of the deutan because the blue is also reduced.

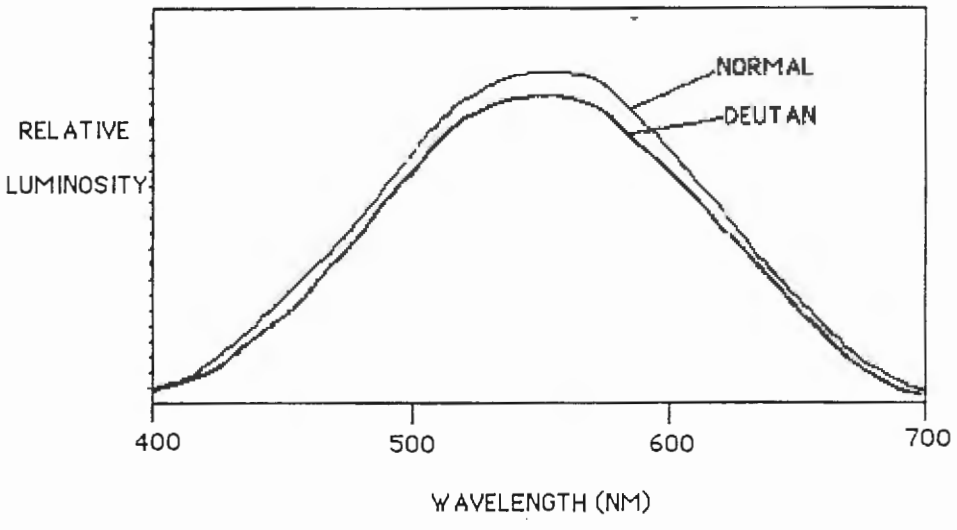
This artificial masking of green and red color loss is probably just part of the underlying problem with the Tomlinson test construct. It is well documented that the overall spectral sensitivity of the deutan is not significantly reduced below that of the color normal subject, whereas the protans spectral sensitivity curve is clearly
(see Fig. D) This fact is probably why luminosity was not considered as a variable for testing color vision until the OSCAR was developed. The authors of the OSCAR must have employed some means of the green luminosity difference between normal and deutan subjects. It is not clear from their paper describing the OSCAR how they increased the relative sensitivity of the test over the green portion of the spectrum, but the results of their comparison study proves their success. One thing they do mention is that they used a relatively broad band LED for their green source and a nearly monochromatic LED red source. This may have allowed the OSCAR to pick up decreases in green color perception over a broader spectral range and sum this to increase the sensitivity of the test for deutans.

Fig. D

OVERALL SPECTRAL SENSITIVITY
OF THE PROTAN



OVERALL SPECTRAL SENSITIVITY
OF THE DEUTAN



In concept, the Tomlinson test is precisely what is needed to meet the demands of all clinical color vision testing. However, based on the evidence that we have presented, the Tomlinson test falls short of its expectations because it assumes that color loss in different parts of the spectrum can be placed on the same linear scale. A certain amount of color perception reduction (measured in units of luminosity) from the normal in the red or blue is not equal to the same amount of reduction in the green part of the spectrum. In other words, one unit of luminosity does not produce the same perceptual difference across the visible spectrum. A color test based on luminosity must take this into account.

REFERENCES

1. Steward JM, Cole BL. What do color vision defectives say about everyday tasks? Optometry and Vision Science 1989; 66, No. 5: 288-295.
2. Ichikawa k, Ichikawa H, Tanabe S. Detection of acquired color vision defects by standard isochromatic plates, part 2. InVerriest G Ed. Colour vision deficiencies VIII, Dordrecht: Morfinus.Nijhorf/Junk, 1987; 133-40.
3. Yolton R, Student's guide to psychophysiology of vision, Class lecture notes, 1992.
4. National Research Council. Committee on Vision. Procedures for testing color vision: Report of working group 41. Washington, DC: National Academy Press, 1981.
5. Mantyjarvi M. Screening of colour vision defects in diabetic patients. Acta Ophthalmol. 1987; 65:178-84.
6. Falcao-reis FM, O'Sullivan F, Spileers W, Hogg C, Arden GB. Macular colour contrast sensitivity in ocular hypertension and glaucoma: evidence for two types of defect. Br. J. Ophthalmol. 1991 Oct; 75(10):598-602.
7. Fischer, ML. Clinical color vision testing as a diagnostic tool. Problems in Optometry 1991;3, no.1:94-119.
8. Brown JL. Flicker and Intermittent Stimulation, Ch 3 in Vision and Vision Perception. Graham CH, Ed. NY: Wiley, 1965.
9. Estevez D, Spekreijse H, Van Dalen JTW, Verduyn Lunel HFE. The Oscar Color Vision Test: Theory and Evaluation (Objective Screening of Color Anomalies and Reductions), Amer. Jour. of Opt. and Phys. Optics. 1983;60, no.11:892-901.
10. Tomlinson. Introductory paper describing the Tomlinson test.
11. Cohen AI: The Retina, in Hart WM, ed: Adler's Physiology of the Eye: Clinical application, 9th ed, St. Louis, CV Mosby, 1992.
12. Hart WM: Color Vision, in Hart WM, ed: Adler's Physiology of the Eye: Clinical application, 9th ed, St. Louis, CV Mosby, 1992.
13. Hart WM: Radiometry and Photometry, in Hart WM, ed: Adler's Physiology of the Eye: Clinical application, 9th ed, St. Louis, CV Mosby, 1992.
14. LeGrand I. Light, Colour and Vision, 2nd ed. London, Chapman and Hall LTD, 1956;338,9.
15. Hsia Y, Graham CH. Color Blindness, Ch 14 in Vision and Vision Perception. Graham CH, Ed. NY: Wiley, 1965.