1-1-1963

An approach to a psychophysical ratio scale of color saturation.

John Sempowski
University of Windsor

AN APPROACH TO A PSYCHOPHYSICAL
RATIO SCALE OF
COLOR SATURATION

by

JOHN SEMPOWSKI
B.A., Assumption University of Windsor, 1961

A Thesis
Submitted to the Faculty of Graduate Studies through the
Department of Psychology in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts at Assumption
University of Windsor

Windsor, Ontario, Canada
1963

APPROVED BY:

ABSTRACT

This study was an attempt to determine if a satisfactory ratio scale of color saturation could be constructed where an operational definition of color saturation was tentatively proposed.

The experimental procedure was the method of limits. Munsell discs (7.5 R 6/10, 7.5 GY 6/10, 7.5 PB 6/10 and neutral grey 6/) were used as stimulus material and were presented by employing three color mixing wheels.

Analyses of the results showed that a ratio scale, for the three hues, was produced up to Munsell chroma 10, the maximum chroma value used in this experiment. The experimentally produced scales linearly related, over low saturation values, to conventional measures of saturation. However, this relation broke down over higher purity values. This discrepancy was most probably due to the arbitrary definition of the conventional measures at the spectral limit as unity.

PREFACE

This investigation began as a result of Dr. A. A. Smith proposing a new definition of color saturation which appeared to solve some difficulties I had experienced in my attempts to understand color theory.

The author wishes to express his grateful appreciation to Dr. A. A. Smith who initially proposed the problem and whose patient guidance was so helpful in its execution. He is also indebted to Mr. M. Starr whose suggestions were beneficial in organizing and clarifying the text of the thesis. Finally he expresses his thanks to the subjects who gave so generously of their time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

BACKGROUND AND RELATED RESEARCH

The generally accepted object of science is to explore reality. After this exploration the scientist inspects, describes, collects and compares data which he has abstracted from that reality. Utilizing this data the scientist attempts to establish general principles, theories, or laws by means of which the empirical phenomena can be explained, accounted for or predicted. The method scientists employ to symbolize the apparent or assumed isomorphic relation between empirical phenomena and the abstracted principles or laws, is the formal, logical, symbolic system of mathematics. Measurement, then, is a concern of scientific psychology. It would appear to be only a remote aspiration in those applied or clinical areas of the discipline where unmapped complexity, forced introspections and the "mystery of man" are the observed data. Measurement, it seems, is a more immediate goal in the experimental area where the exact identification, isolation and control of relevent variables is the prime objective. Thus, measurement becomes an important preoccupation of psychophysics.

What is the nature or process of measurement? It would appear profitable, before answering this question, to discuss basic assumptions and definitions. An important distinction to be made is one which exists in the intimate relation between two scientific constructs - systems, and properties. Properties are defined as those aspects of objects or events which the scientist observes - such as length, width, color, etc. Whenever the scientist observes or measures a property it always is an aspect of something. This something is called a system - e.g., books, trees, people, etc. Consequently, when properties occur, they occur as characteristics of systems. This distinction between system and property becomes important for it emphasizes the fact that whenever a measurement is performed the operation is comparing the intra and inter-relations of the properties among themselves and is not directly concerned with the systems per se.

A basic assumption made by scientists when measuring is that an isomorphism exists between the empirical relations among properties and the relations within the abstracted, formal mathematical model. Torgerson states: (1958, pp. 14-15)

> In order to represent the property (measure), an isomorphism, i.e., a one to one relation, must obtain between certain characteristics of the number system involved (mathematical model) and the relation between various quantities or instances of the property to be measured empirical relations. Brackets mine.

The reasonableness of this basic assumption, apparently, has been proven by the multitude of fruitful applications of mathematics. i.e., measurement, in almost every scientific discipline and by the consequent greater understanding of reality rendered by such applications.

Since an essential characteristic of measuring is the utilization of numbers a discussion of their properties appears to be required for a better understanding of the problem. Torgerson lists these as: (1958, p. 15)

> (1) Numbers are ordered.
> (2) Differences between numbers are ordered. That is the difference between any pairs of numbers is

> greater than, equal to, or less than the difference between any other pair of numbers.
> (3) The series has a unique origin indicated by the number "zero" . . .

It becomes evident, then, that real numbers possess three fundamental characteristics - order, distance and origin. These three characteristics will serve as one method by which different types of scales can be differentiated. Another type of scale, one which Torgerson does not include but Stevens argues for, is naming. That is, as Stevens argues, whenever properties are put into order they are simultaneously named or labelled.

> The nominal scale is a primitive form, and quite naturally there are many who will urge that it is absurd to attribute to this process of assigning numerals the dignity implied by the term measurement . . . However we christen it, the use of numerals as names for classes is an example of the 'assignment of numerals according to rule'. (Stevens, 1951, p. 26).

Thus, the conclusion that the nature or process of measurement is the assignment of numbers to express the intra-relations of empirical properties among themselves and to symbolize that relation in a formal mathematical model.

> It is possible to establish a one to one relationship between objects possessing this property and those characteristics of numbers. Numbers are then assigned to the objects so that the relations between the numbers reflect the relations between the objects themselves with respect to the property. Having done so, we have measured the property, i.e., established a scale of measurement. (Torgerson, 1958, p. 15).

Both Stevens and Torgerson imply that a scale is an instrument by which measurement is accomplished. The type of scale will be determined on the basis of its characteristics, i.e., order, distance and origin and the number and kinds of mathematical transformations which can be applied to the scale, provided that in these operations the basic essence (order, interval, origin) as determined by the rules under which the scale was constructed are not destroyed or varied. Thus, by using these criteria four basic scales can be differentiated - nominal, ordinal, interval, and ratio. This list, it is noted, is hierarchical, for each possesses the attributes of its predecessor plus its own further refinements.

The nominal is the most primitive scale. This scale represents an unrestricted application of numerals to properties so that the properties are named, where the rule for the construction is, as Stevens states it (1951, p. 20): "Do not assign the same numeral to different classes or different numerals to the same class". A nominal scale equates, according to rule, various heterogeneous properties by grouping those properties relative to a mutual similarity. Also, since these assigned numbers are merely a means of identification, any number can be interchanged for another provided the basic rule is not violated. For example, the similar characteristic of ten men is that they are football players; they are scaled by designating each with a specific number, say from 1 - 20. A one to one substitution of any other football player can be made as long as the basic rule is followed. This transformation can be expressed by the formula $y = f(x)$ where y is any other player and x is any player in the series 1 - 20.

An ordinal scale arises when the rule for construction is that either the natural order or the order resulting from equating properties according to some similarity is serial. Numbers are employed in such a manner that their order corresponds to the order of the properties. This ordinal scale can be manipulated by any mathematical transformation which leaves the rank order intact. This quality of the scale is symbolized by the formula $y = f(x)$ where (x), is any monotonic non-decreasing function which will not change the established rank. For example, ten rocks are arranged according to size so that the smallest of the ten is numbered one and each successive larger rock is numbered 2, 3, 4 . . . and so on until the largest is numbered 10. We can then say that: "rock number two is greater than rock number one", (but not how much greater); and "rock number three is greater than rock number two"; but we cannot as yet say whether the difference in size between rock number three and rock number two is "greater", "less than" or "equal to" the difference in size between rock number two and rock number one.

In the real number system, not only the numbers, but the differences between numbers, are ordered. In particular, certain differences are equal: for example, $3 - 2 = 2 - 1$. Thus when we assign numbers to objects and are able to state that the differences between these objects are equal (that is the difference between rock number 3 and rock number 2 is equal to the difference between rock number 2 and rock number 1) an interval scale is achieved. The only mathematical transformation which can be performed on the interval scale is a linear one. This transformation is represented by the formula $y = ax + b$ where y is any number, and b is a constant relating y and x to a common origin, for example, let y represent any number on the centigrade scale, x any number on the fahrenheit scale and b the distance of both y and x from zero. Thus centigrade can be transferred to fahrenheit or vice versa by putting in any real values of centigrade into the formula and the equivalent fahrenheit value is obtained.

If added to the accumulated requirements of the interval scale the additional requirement of a unique, natural origin or zero point it follows that the linear transformation is reduced to $y = ax$. The constant b is removed for all scales now must proceed from the unique zero point. This refinement creates equal ratios along the scale i.e., the scale demonstrates equal intervals proceeding from zero. This characteristic permits, for example, the statement: "The distance from zero to two is one half the distance from zero to four" or differently stated: $2 = 1/2\ (4)$.

Generally, the reasons for employing any psychophysical method are to determine the acuteness, sensitivity and quality of sense perceptions. Secondly, the particular methods are employed in order to derive the necessary data required for the construction of a scale. How can a psychological scale be constructed: The answer to this question lies in the performance of consecutive steps where the "length" of the sensation continuum is determined by establishing an origin (reiz limen), the limit beyond which the sensation is too "painful" (terminal limen) and a unit which subdivides the interval between the reiz and terminal limens into equal increment steps. The establishment of the reiz and terminal limens is relatively easy. The particular sense modality is merely sounded at various levels and the points where the average person states "I hear it" and "That's painful" determines the "length" of the continuum. However, the determination of the increment steps is more complex. The first problem is to find appropriate units as measures which represent the intervals along the scale. Several have been used but the difference limen seems to be the unit most utilized in the past.

The problem of determining what appropriate unit should be used to represent a man's sense perceptions when the individual is comparing various magnitudes of different sense modalities was first answered by E. H. Weber. Weber stated: "In comparing magnitudes it is not the arithmetical difference, but the ratio of the magnitudes which we perceive", (Woodworth, 1938, p. 430). Implicit in this principle is the acceptance of the "just noticeable difference" (j.n.d.) or differential limen (D.L.) as the basic unit of sense discriminability and that this differential limen is a constant fraction of the physical stimulus. Experimental evidence obtained by Weber tended to support the assumption that the ratio of the stimulus increment to stimulus value for a "j.n.d." of sensation was constant.

This assumption is sometimes formalized as Weber's Law: "A stimulus must be increased by a constant fraction of its value to be just noticeably different". (Woodworth, 1960, p. 194). Weber's Law, as it states, does not provide a measure of sensation. Fechner sought to obtain such a measure by assuming that j.n.d.'s are equal, regardless of the sensation level from which they are just noticeably different and then integrating the Weber function to get $S = k \log I$, or, "the strength of the sensation varies directly as the logarithium of the intensity of the stimulus". (Woodworth, 1960, p. 236). Thus, Fechner formulated a useful unit to be used in constructing scales of sensations.

A renewed interest in the nature and function of psychophysical scales has been noticed in recent years. This interest appears to have been created and strengthened by the ever presence of practical problems which have not been satisfactorily resolved and by the questioned validity of historical assumptions and methods.

Garner (1958) and Stevens (1957) both state that there are two basic classes of experimental techniques which have been used to establish scales. The first includes the "direct response" methods. These methods assume that an individual is able to understand and to use some type of numerical scale in carrying out the required tasks of the experimental situation and that these numerical values are used in the actual scale construction. In one such method, that of "magnitude estimation", S will, for example, be asked to listen to a reference tone of fixed loudness; this, he is told has a loudness of ten units (the modulus). He is then asked to listen to other tones of differing loudness, and to assign numbers to the loudness of the tones so that the ratios of these numbers to the modulus reflect the perceived ratios of the loudness to that of the reference tone.

A second method, that of "fractionation", calls for the S to adjust a comparison tone so that its loudness bears a given ratio to that of a standard. This technique has been used to construct the well known "sone" scale of loudness.

Both these methods are, of course, direct attempts to establish psychologically meaningful ratio scales. A third method, that of "equisection", uses two standard stimuli, with one or more comparison stimuli adjusted so that the interval between the standard is divided into equal parts. This technique establishes initially only an interval scale - unless one of the standard stimuli can meaningfully be said to be at the absolute zero of the property being scaled.

The second class, the "indirect methods", are those methods derived from Fechner which utilize some measure of discriminability (j.n.d.) in the construction of a scale. Historically, the "discriminability" or as Stevens calls it the "discriminal disperson" technique of scale construction took precedence over the direct approach. At present, however, Stevens carries on a continuing argument which challenges the legitimacy of the Fechnerian assumptions and which desires to replace the "indirect methods" with the direct-response methods.

> The lesson of history is that a bold and plausible theory that fills a scientific need is seldom broken by the impact of contracy facts and arguments. Only with an alternative theory can we hope to displace a defective one. The purpose here is to try to do just that - to try to show that there is a general psychophysical law relating subjective magnitude to stimulus magnitude and that this law is simply that equal stimulus ratios produce equal subjective ratios . . . To a fair first-order approximations, the ratio scales constructed by "direct" methods (as opposed to the indirect procedures of Fechner) are related to the stimulus by a power function of one degree or another. (Stevens, 1957, p. 153).

Garner, a firm opponent of the direct methods, points out that when these methods are used the ability of the subject to describe his experiences of sensation in terms of a particular numerical system is assumed.

> We cannot, however, accept simple face validity of the observer's proper use of ratios any more than we would feel free to accept the validity of the observer's proper use of exponents to describe their loudness. (Garner, 1958, p. 1006).

Stevens states in defence of the direct method:

> Contrary to a common assumption, the use of the direct methods does not require knowledge of an underlying, measureable physical continuum. Only a nominal scale is required at the stimulus level, i.e., the stimuli must be identifiable by the experimenter . . . I know from experience that it is more comfortable to take for granted that a direct method is impossible than it is to try to work one out. (Stevens, 1957, pp. 177-178).

Garner suggests that when ratio methods are used the greatest single source of error is large differences in the results of different individuals. Consequently, even though computation of central tendencies is possible they are of little value in predicting the performance of other individuals - "inter-observer differences provide the largest single source of variability in the data". (Garner, 1958, p. 1007). Garner, however, demonstrates that this particular source of error can be minimized by using a bisection method. Garner reaches this sugges-ion in the light of his experimental evidence - "the equisection method should be accepted as more valid than the fractionation method". (Garner, 1958, p. 1007). The author proposes that because of the inconsistent results obtained via the direct methods, indirect methods should be adopted. Garner suggests four variables which tend to vary the results of the direct methods. These are: the questionable face validity that observers can transpose sensations into a numerical system, the demonstrated variability among inter-observers results, inconsistent results of the same individuals and the effect that the context of the particular experiment has upon results. Garner concludes that: "the various ratio scaling procedures produce large differences in the loudness scales". (Garner, 1958, p. 1011).

Stevens, as pointed out above, is an aggressive proponent of the direct methods. He proposes the thesis that the general psychophysical law relating sensation to physical intensity is simply that equal stimulus ratios produce equal subjective sense ratios and that these ratios are related by a power function of one kind or another. Stevens feels that the Weber-Fechner Law is inadequate to serve as a general psychophysical law.

Stevens implies, as does Garner, that perhaps the bisection procedure would be the most efficient method to employ in scale construction for apparently the various sources of error in direct scaling are minimized when the bisection method is used.

There are many examples of attempts to construct scales of sensation. Richardson and Ross (1930); L. B. Hamm and J. A. Parkinson (1932); D. A. Laird, E. Taylor and H. H. Wille, Jr. (1932) and P. H. Geiger and F. A. Firestone (1933) used various techniques in scaling the loudness of sound.

B. G. Churcher appears to summarize the conclusions obtained from the cited experiments when he states:

> . . . the decibel scale (a scale based upon the implications on the Fechner Law) used as a loudness scale, does not yield numerical values proportional to the loudness sensation. (Churcher, 1935, p. 216).

Stevens states essentially the same conclusion:

> The relation between subjective loudness and the physical intensity of the stimulating tone has been investigated by a wide variety of methods yielding a wide variety of results. The decibel change in the stimulus required to produce a 2:1 ratio in apparent loudness has been shown to cover a range of the order of 20 db. (Stevens, 1956, p. 71).

For loudness, then, attempts to construct scales by direct and indirect methods appear to lead to different psychological functions.

A different result obtains when the scaling of visual brightness is considered. R. M. Hanes (1949) set out to determine how easily and reliably individuals could make fractional estimates of brightness and to construct a brilliance scale which would demonstrate the relation, if any, of subjective brightness or brilliance to the physical intensity of the stimulus.

The procedure was the fractionation technique where the subjects were required to adjust the variable stimulus until S was satisfied that the variable was 1½ or 3/4 as bright as the standard. A black disc surrounding the test spot was suggested to function as a reference or zero point when the S's made their fractional estimate. The results show generally a remarkable ability of the O's to make consistent fractional estimates of brightness, especially at low brightness levels. The data obtained from the experiment were utilized to construct a brilliance scale and the technique followed that outlined by Stevens (1936). This scale, it should be noted, agreed very closely with an independent brightness function, based on integrated DL's, obtained earlier by Troland (1929).

The present experiment is concerned with the perception of color, specifically with the perception and scaling of color saturation. Color saturation has been defined as "that attribute of all colors possessing a hue which determines their degree of difference from a grey of the same brilliance", (Jones and Lowry, 1926, p. 25). Saturation is "the attribute of any chromatic color which determines the degree of its difference from the achromatic color most closely resembling it". (Judd, 1940, p. 3).

The Optical Society of America Colorimetry Committee discusses saturation in these terms:

> Saturation also has a quantitative, as well as a qualitative character, and the use of terms high saturation, medium saturation and low saturation are suitable for the descriptions of the saturation of color sensations. When, in a stimulus consisting of a combination of chromatic and achromatic light, the proportion of the chromatic component is increased, the saturation of the color sensation varies correspondingly. (Newhall, Nickerson and Judd, 1943, p. 544).

C. H. Graham states:

> The stimulus terms of the saturation relation have not yet been finally specified in a completely satisfactory manner. One important stimulus variable in saturation discrimination is colorimetric purity, P. Colorimetric purity is defined for a mixture of a spectral color and white as $P = \dfrac{B_2}{B_w + B_t}$, where B2 is the luminance of a

> spectral color and B_W the luminance of the white with which it is mixed. (Graham, 1959, p. 166).

The earliest reported attempt to determine or measure the relationship between various saturation stimulus levels and saturation sensation was performed by A. H. Munsell (1909). Munsell pointed out that it was commonly held that a scale of color intensity (value) follows, at least approximately, the Weber-Fechner Law of the relation between stimulus intensity and sensation. The author suggested that no experimental results had been reported concerning the other two dimensions of color, hue or saturation, and consequently set out to determine the relationship, if any, of saturation sensation to various levels of saturation stimulus.

Munsell's procedure was to construct two homogenous discs (i.e., the discs were taken from the same color sheet) and to cut concentric rings in each disc so that the successive concentric rings, from the center out towards the circumference, contained a smaller proportion of area than its predecessor. Where the amount of area cut from each successive concentric ring for <u>disc A</u>, proceeding from center to circumference, followed the geometric progression of 0, 1/16, 1/8, 1/4, 1/2 and for <u>disc B</u>, followed the arithmetical progression of 1, 1/5, 2/5, 3/5, 4/5.

Secondly, Munsell spun disc A on a white background and found that with most colors a regularly graduated progression of diminishing saturations and an increasing progression of intensity from center to circumference resulted. Munsell spun disc A on a neutral grey background and found an irregular progression of saturation from center to circumference. The author concludes that when there is simultaneous variation of intensity and saturation a regular scale of saturation intensity is found and these scales correspond to the Weber-Fechner Law.

Thirdly, Munsell took Disc B and spun it on a white and neutral grey background. The author's results show that when disc B is spun with white an irregular progression results but with neutral grey, a regular progression of saturation from center to circumference is created.

Munsell concludes that the relationship between various physical levels of saturation stimulus and saturation sensation follows an arithmetical progression rather than a progression corresponding to the Weber-Fechner Law of the relation between stimulus intensity and sensation.

Lloyd A. Jones and E. M. Lowry (1926) determined the size of the difference limen for color saturation and succeeded in constructing j.n.d. scales of saturation for a number of hues. They used lights of pure hues mixed with white light where intensity was held constant. They measured the brightness of the resultant mixture by using a flicker photometer. The first step in the procedure consisted in the illumination of the comparison areas with monachromatic light of equal wave length and intensity. Secondly, a gradual increasing percentage of white light was admitted to one of the comparison areas until a just noticeable difference or a less saturated difference was perceived by the subjects. The authors used the continuous, step by step method and conclude:

> . . . that the number of saturation steps varies for the different color, the larger steps being found in the red and blue, and the smaller in the yellow . . . results seem to offer the possibility of different saturation psychologically, for the different colors, and as a result various numbers of steps are to be expected in their saturation scales. (Jones and Lowry, 1926).

Statement of the Problem

Few other attempts have been made to scale saturation. Part of the difficulty, as pointed out by Graham (1959) lies in the fact that "the stimulus terms . . ., have not yet been finally specified in a completely satisfactory manner". It is, however, generally agreed that when a color of a given hue, brightness and saturation is mixed with an achromatic, neutral grey of equal brightness, the resultant mixture will be a color of the same hue and brightness but of reduced saturation. It would seem plausible, then, to measure the saturation of a given color in terms of the amount of neutral grey required to reduce the saturation by a pre-assigned ratio. The most practical ratio appears to be one-half. Therefore Smith (1962)[1] has proposed that saturation be defined as a function of the amount of neutral grey which must be added to a color so that the mixture is perceived as one-half as saturated as the original.

This definition bears some similarity to the definition of colorimetric purity. It is not, however, the same: in particular, it does not require, as is the case with colorimetric purity, that spectral colors all have the same numerical value.

However, before this suggestion can be proceeded with, it is necessary to demonstrate that the psychophysical operations implicit in it can be satisfactorily carried out with a sufficient degree of reliability. That is, it must be shown that a satisfactory ratio scale of saturation is possible. The investigation of this point is the purpose of the present investigation.

---

[1]Smith, A. A., Personal communication, 1962.

CHAPTER II

METHODOLOGY AND PROCEDURE

Experimental Design

It was decided to attempt the construction of a ratio scale of saturation by the repeated bisection of the interval between a color of a given hue, brightness and saturation, and a neutral grey of equal brightness. Since neutral grey is by definition of zero saturation, such successive bisections should yield a true ratio scale.

In many experiments of this type, it is customary to allow the S to adjust a variable stimulus until, in his judgment, it is midway between the standard stimuli. The apparatus, however, did not allow adjustments by S; all changes in the comparison stimulus had to be made by E. The required bisection points were therefore determined indirectly, using the method of limits, and the point of indifference as the required midpoint.

Subjects

The S's were a group of 10 University students, three males and seven females; ranging in age from eighteen to twenty-one years. No subject with any noticeable or known visual defect was included. All subjects were checked for congenital red-green and blue-yellow color blindness by employing the 1940 Pseudo-Isochromatic Plates of the American Optical Company. The group was experimentally naive with regard to the testing procedure. All subjects were allowed to determine their own appointment schedules. This freedom was permitted in the hope that all subjects, at the time of testing, would be optimally motivated and rested.

Apparatus

Three color mixing wheels (Staelting, models 6112 and Tolboys, model 101) were used. The color wheels were placed inside a wooden box, measuring 29 inches wide, 15 inches high and 19 inches deep, with each wheel positioned directly behind a one by three inch opening cut in the face of the box. Illumination was provided by a six-volt, direct current system which was chosen to control against any stroboscopic effects. Light bulbs (General Electric GM-47) were positioned above and below each stimulus opening in a similar manner.

Color stimuli were provided by four inch discs of Munsell paper of the following Munsell book notations: 7.5 R 6/10, 715 GY 6/10, 7.5 PB 6/10 and N 6/ (neutral grey). All stimuli were therefore of equal brightness, within the limits of error of the lighting system and the Munsell value notations. The discs were placed on the wheels so that stimulus one (the zero standard), to the left, was a neutral grey. Stimulus three (the color standard), to the right, was a mixture of one particular hue with a fixed proposition of grey. Stimulus two (the variable stimulus), located between stimuli one and three, was a varying mixture of a particular hue with a neutral grey. All three color wheels were synchronized by a Dawe Stroboscope (model 1200D) at 4000 r.p.m.'s.

Procedure

Upon entering the room the S was seated and instructed to place his head in a Bausch and Lomb head-rest. The head-rest was situated seven and one-half feet from the apparatus and at a level which allowed S to be seated comfortably but at the same time kept his eyes approximately level with the three stimulus areas. The subjects were instructed as follows:

> The lights in the room will be turned off. We will wait a moment to allow for our eyes to adapt to the darkness. You will be presented with three stimuli; the one to your left, stimulus one, will remain constant throughout the experiment; the one to your right, stimulus three, will remain constant during this first phase of the experiment. However, there are three phases to this particular experiment consequently stimulus three will change three times. You will be instructed as to when these changes occur. Stimulus two, here in the middle, will be constantly changing. Now I would like you to look at stimulus two and tell me which it most resembles; Stimulus one or stimulus three - You can say either right or left.

Any questions the subjects may have asked were answered.

All subjects were required to participate in three experimental sessions (i.e., a session for each hue). These separate sessions were required because, on the average, the subjects needed 40 minutes to bisect the particular hue into the one-half, one-quarter and one-eighth points and it was felt that a 40 minute testing period was long enough. Furthermore each of the sessions were sub-divided into three phases on the basis of each procedure required for the bisection of the one-half, one-quarter and one-eighth points.

Phase I of session one of the experiment consisted in the presentation of Munsell hue 7.5 GY. The zero standard (OS) was neutral grey and was kept constant throughout the session. The variable stimulus (VS) was a mixture of 7.5 GY with neutral grey in amounts dictated by the requirements of the method of limits. The color standard (CS) was 7.5 GY with no neutral grey added. The subjects were given a trial ascending and descending series which was discarded in terms of computations. The trial run, also, served as a cue to the experimenter, aiding him in gaining an idea of the subjects approximate transition point. The acquisition of this approximate point enabled the experimenter to adjust the number of "discrete steps" approaching the estimated bisection point randomly. Three ascending and three descending series were given and the subject's responses were recorded. The light in the room was turned back on and the subject was told to relax. The experimenter, during this time, calculated the individual subject's average point of indifference. This point was taken as the best estimate of the amount of grey required to desaturate the original color standard by one-half.

In Phase II of session one the amount of grey, required by the subject for his bisection point of Phase I, was added to stimulus three. So that the CS changed from a "pure" 7.5 GY to a mixture of 7.5 GY with the experimentally obtained amount of grey from Phase I. Phase II required the subject to judge in a similar manner as he did in Phase I, as to which standard (OS or CS), the VS approximated. Three ascending and three descending series were made and again the point of indifference was calculated.

In Phase III, the CS was changed for the third time to the mixture calculated in Phase II.

Three ascending and three descending series were given and the indifference point was again calculated. This point represented the one-eighth point on the saturation continuum.

Session II and Session III were conducted in the same manner as Session I, except that hue 7.5 R was used in Session II and hue 7.5 PB was used in Session III. All subjects were thanked for co-operation and any problems or questions were answered.

CHAPTER III

PRESENTATION AND ANALYSIS OF RESULTS

The primary data of the present experiment are in terms of the percentages of a given hue which, when mixed with a neutral grey of equal brightness, yields colors of one-half, one-quarter and one-eighth the saturation of the original colors. These data are set forth in Table 1, both for individual subjects and as group averages. Ranges and Standard error of the mean are also given.

Table 1

Percentage Hue at Fractional Saturations for Three Colors

	Munsell Book Notations of Original Colors								
	7.5 R 6/10 Saturation Ratios			7.5 GY 6/9 Saturation Ratios			7.5 PB 6/10 Saturation Ratios		
	1/2	1/4	1/8	1/2	1/4	1/8	1/2	1/4	1/8
Subject # 1	37.0	23.0	11.0	41.5	18.0	10.2	40.0	20.0	-
2	41.0	28.0	18.0	37.0	18.5	9.2	50.0	25.0	-
3	50.0	27.0	12.5	46.0	24.0	15.0	48.0	28.0	-
4	50.0	27.0	12.5	34.0	18.0	11.5	34.5	21.6	7.8
5	35.0	15.0	7.5	50.0	26.5	15.0	46.0	23.0	11.0
6	34.0	14.0	7.0	53.0	28.0	17.4	48.3	26.6	15.8
7	27.0	11.0	4.0	53.0	30.0	17.0	46.0	25.0	19.1
8	47.0	20.0	9.0	34.0	15.0	7.0	56.5	21.5	-
9	42.0	20.0	9.0	50.0	26.3	12.5	51.5	26.0	-
10	30.0	14.0	4.0	53.4	26.0	10.6	44.0	29.0	-
Group Mean	39.3	19.9	9.5	45.0	23.0	12.5	46.4	24.5	13.4
Standard Error of Means	2.5	1.8	1.2	2.4	1.2	1.0	1.8	.7	3.5
Range	23.0	17.0	10.0	19.0	13.0	10.0	22.0	9.0	12.0

One significant aspect of Table 1 is that several individuals were unable to bisect for the one-eighth point in 7.5 PB 6/10. Those individuals who failed to produce a score were rejected either because: 1) They were unable to make a discrimination, 2) They were responding habitually. This was determined by the subjects own verbalizations: "I can't make a judgment. I can only guess". or 3) They were responding randomly in a very inconsistent manner.

The results, as found in Table 1, were transformed into the Munsell system. This transformation was carried out by converting the Munsell book notations of the colors and the neutral grey into C.I.E. tri-stimulus values, x, y, z (data supplied by the Munsell Company); combining the tristimulus values according to the mean percentages found experimentally and reconverting these values back to Munsell renotation values by reference to the renotation charts of the Munsell system. (Newhall, Nickerson and Judd, 1943). The results of these calculations are presented in Row two of Table 3.

The colorimetric purity of the original colors and their desaturated mixtures was determined from the tristimulus values, using the formula:

$$Pc = \frac{(y_{b}/y)\quad(y-y_{w})}{y_{b}\ -\ y_{w}}.$$

In this formula the y's are the trilinear values for, respectively, the spectral limit (y_b), the achromatic point (y_w) and the color whose purity is to be determined (y). y and y_w were determined from the data provided by the Munsell Company, y_b was obtained by interpolation between tabulated data for the spectral limit.. (Committee on Colorimetry (O.S.A.) 1944). Since the spectral limit is very nearly linear in the region of

Table 3

The Munsell Chroma, Colorimetric and Excitation Purity at the Bisection Points

	Munsell - Red 7.5 R 6/10 [1] Wave Length 599 my					Munsell - Green-Yellow 7.5 GY 6/10 [1] Wave Length 562 my					Munsell - Purple-Blue 7.5 PB 6/10 [1] Wave Length 473 my				
Saturation Ratios	1.0	.5	.25	.125	0.0	1.0	.5	.25	.125	0.0	1.0	.5	.25	.125	0.0
Per Cent Hue at Bisection Points	100	39.3	19.9	9.5	0.0	100	45.0	23.0	12.5	0.0	100	46.4	24.5	13.4	0.0
Munsell Chroma at Bisection Points	10	3.9	2	1	0	9	3.8	1.8	1	0	10	5.2	3.0	1.6	0
Relative Colorimetric Purity [2]	"100" (.4557)	36.0 (.1656)	19.7 (.0902)	9.5 (.0459)	0.0 (.0000)	"100" (.8225)	45.0 (.3724)	23.0 (.1932)	9.1 (.0742)	0.0 (.0000)	"100" (.265)	45.0 (.0581)	26.7 (.0336)	13.0 (.0166)	0.0 (.0000)
Relative Excitation Purity	"100" (42.2)	64.0 (27.2)	37.0 (15.7)	25.0 (10.7)	0.0 (0.0)	"100" (61.7)	47.0 (29.1)	21.0 (14.3)	11.0 (7.1)	0.0 (0.0)	"100" (34.8)	58.0 (20.4)	38.0 (13.4)	13.0 (4.7)	0.0 (0.0)

1 - Munsell chromas in column heads are book notations; renotation values are used in body of table.

2 - Numbers in parentheses represent the actual numerical value of Colorimetric and Excitation Purity.

the dominant wave lengths of the values used, a linear interpolation was felt to be sufficiently accurate. The obtained values of colorimetric purity were then transformed to relative colorimetric purity by dividing the colorimetric purity at each fractionation point by the colorimetric purity at the "full" point. The results of these calculations are summarized in row three of Table 3.

The excitation purity of the experimentally obtained saturation ratios was determined directly from tabulated values of dominant wavelengths and excitation purity. (Newhall, Nickerson & Judd, 1943). The excitation purity was also transformed to a relative excitation purity in a similar manner as was employed in colorimetric purity. These results are given in row four of Table 3. (see page 22).

Several of the relationships found in Table 3 could be better understood if they were graphically represented. Consequently the mean per cent of 7.5 R, 7.5 GY and 7.5 PB at the "full", one-half, one-quarter, one-eighth and zero points is graphically depicted in Fig. 1. Fig. 2, 3 and 4 graphically represent the relation of Munsell Chroma, Relative Colorimetric Purity and Relative Excitation Purity, respectfully, of the three hues to the Bisection points. Figure 5 represents the relation between per cent hue and Relative Colorimetric purity. Figure 6 represents the relation of the renotated Munsell values to Relative Colorimetric Purity.

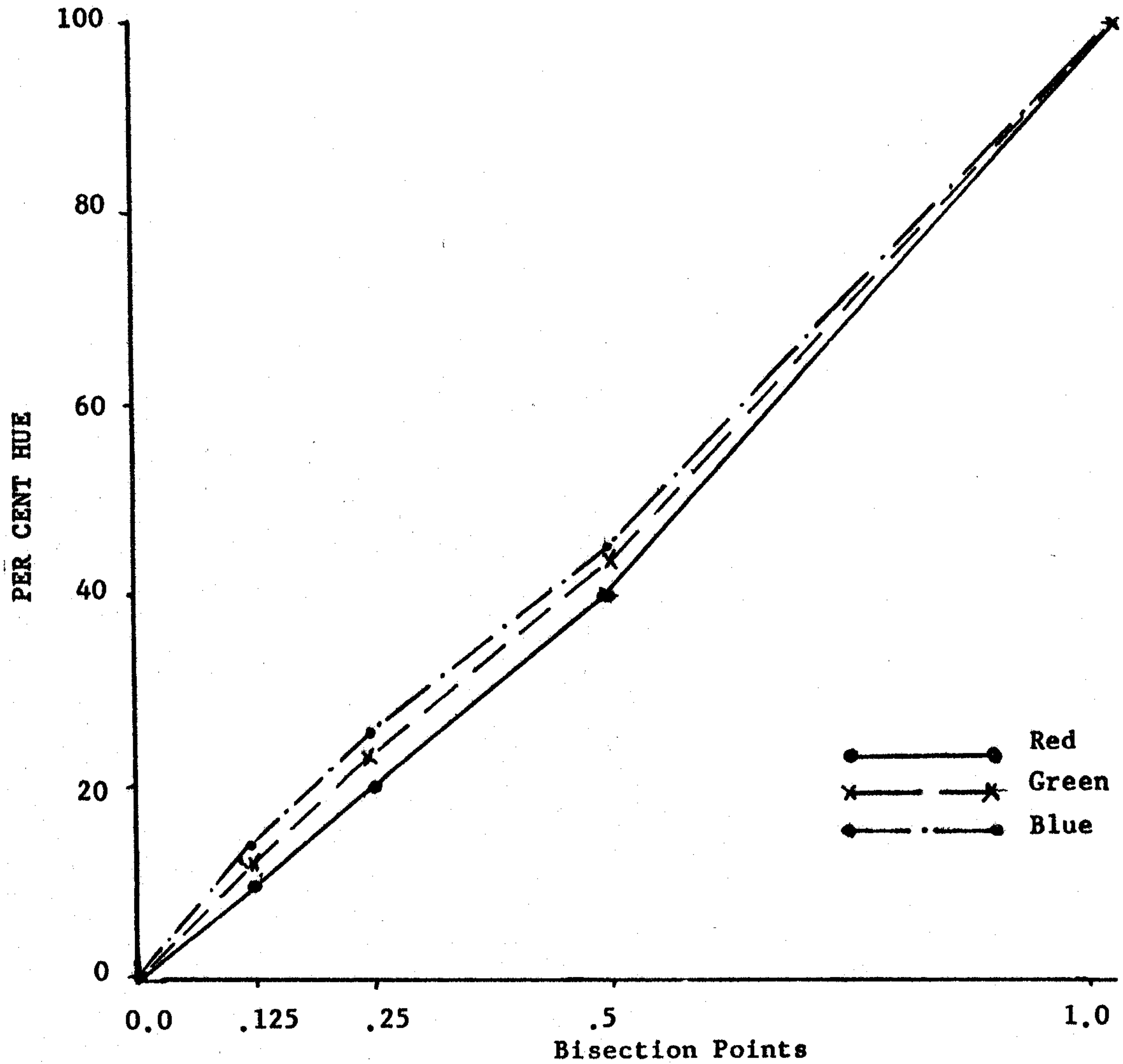

Fig. 1. The per cent hue at the bisection points.

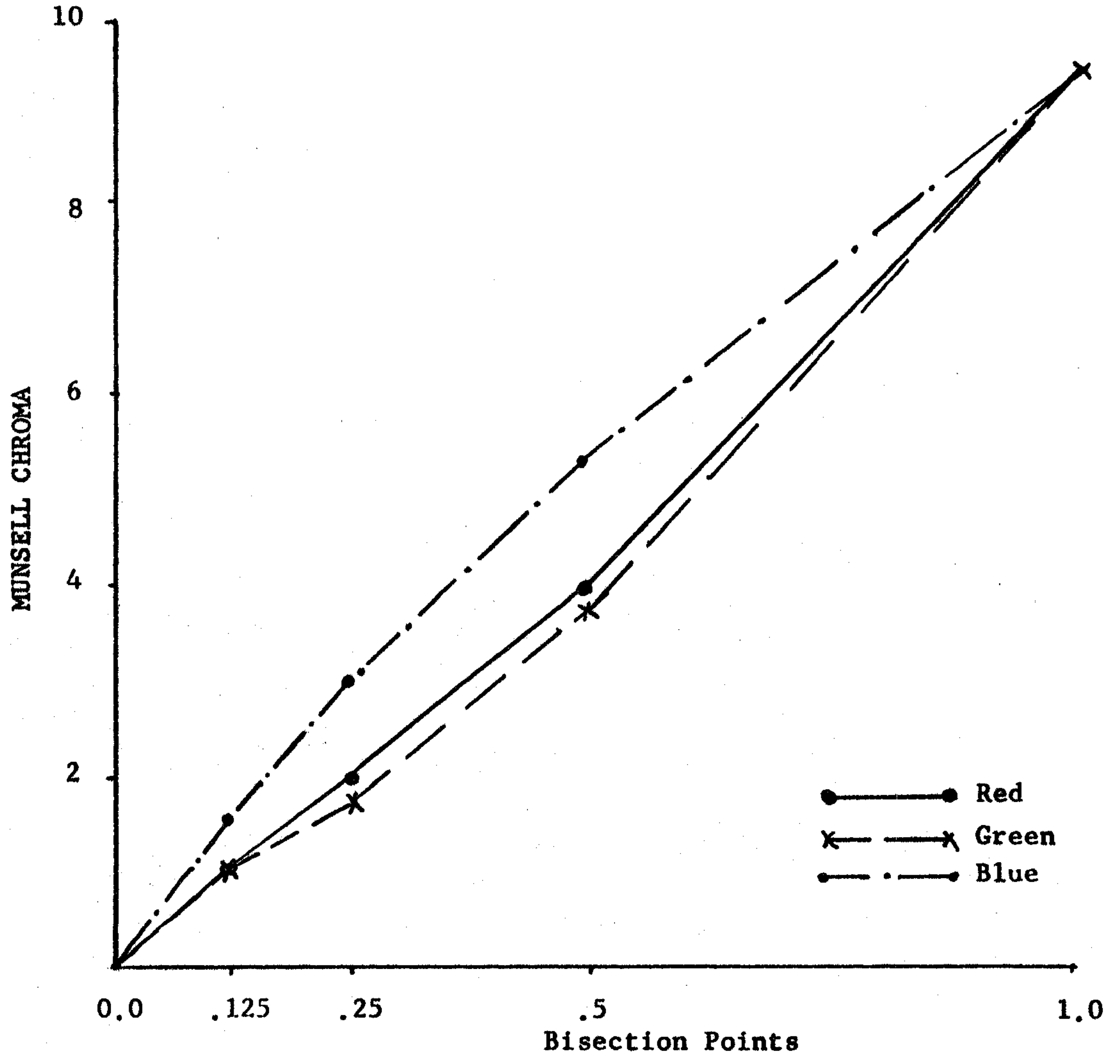

Fig. 2. Munsell chroma at the bisection points.

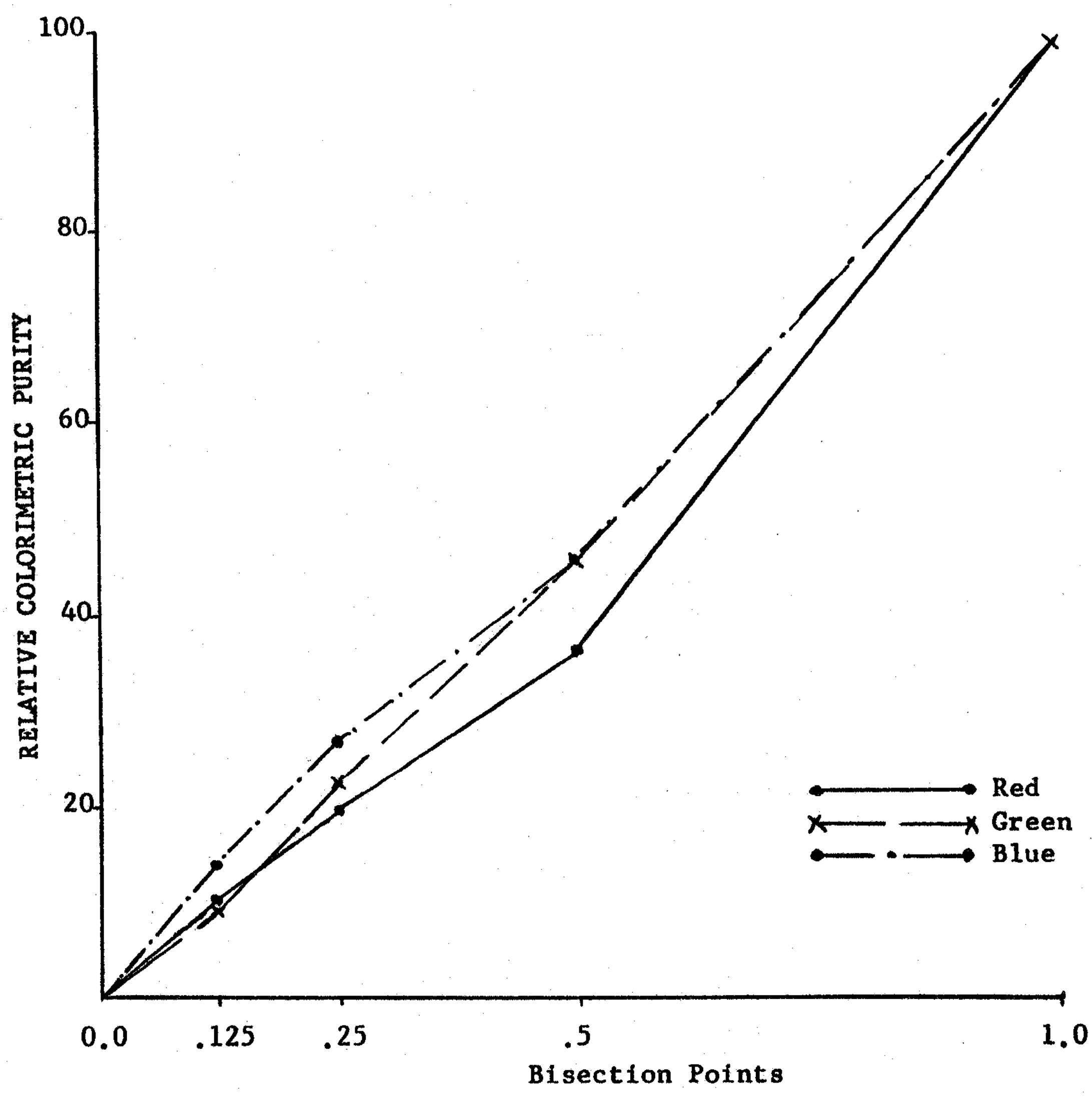

Fig. 3. Relative colorimetric purity at the bisection points.

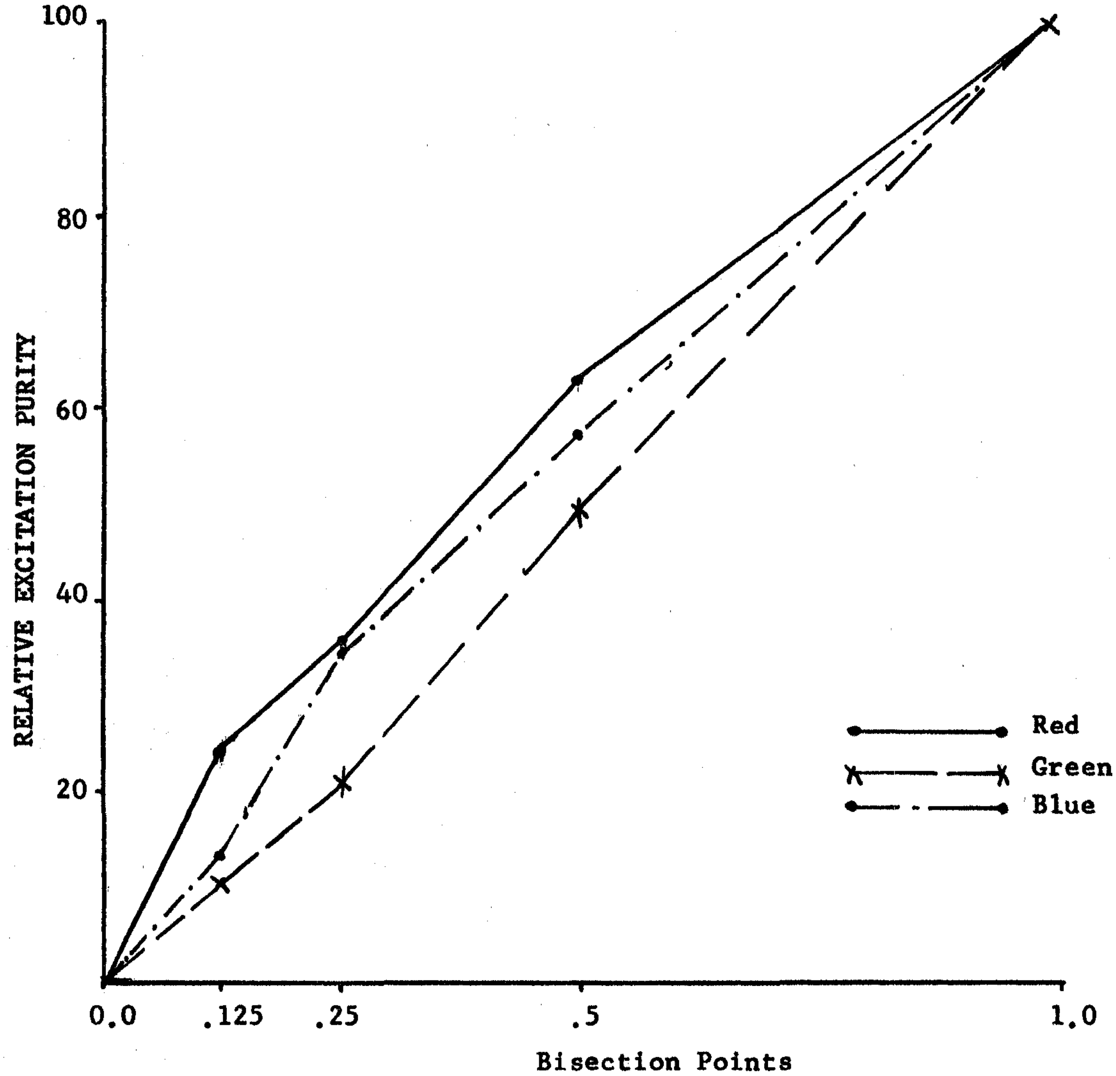

Fig. 4. Relative excitation purity at the bisection points.

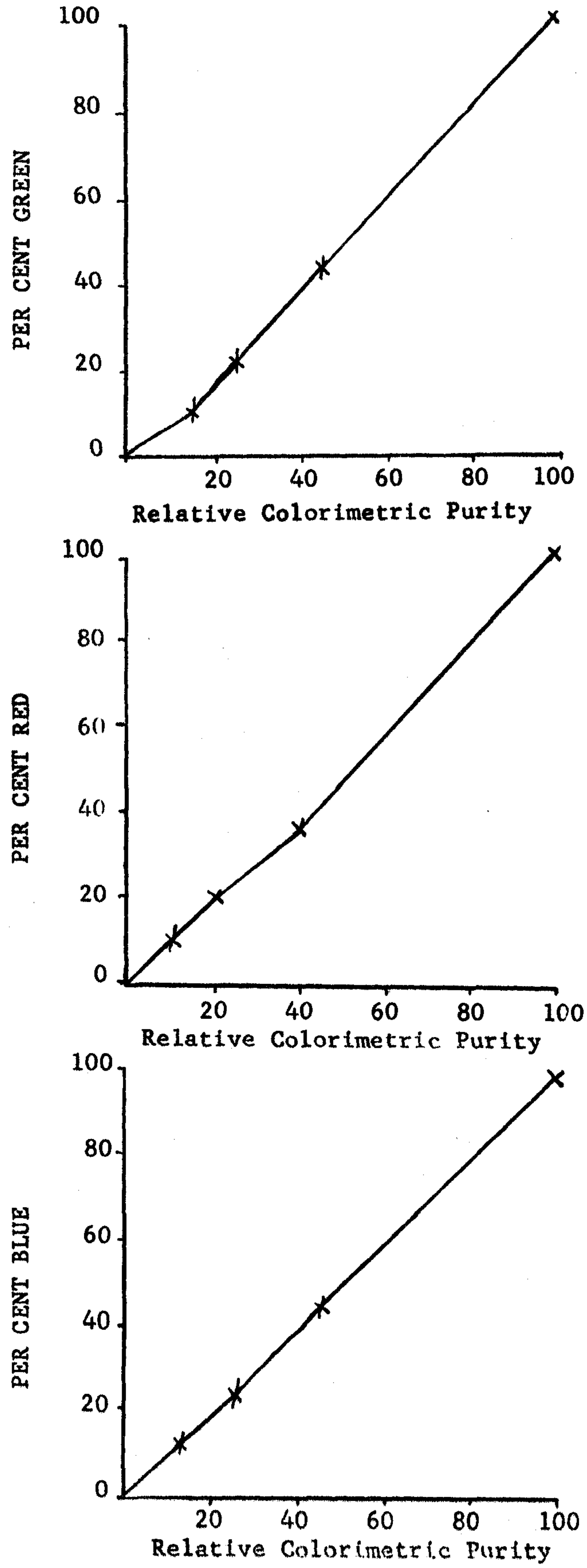

Fig. 5. Per cent green, red and blue to relative colorimetric purity.

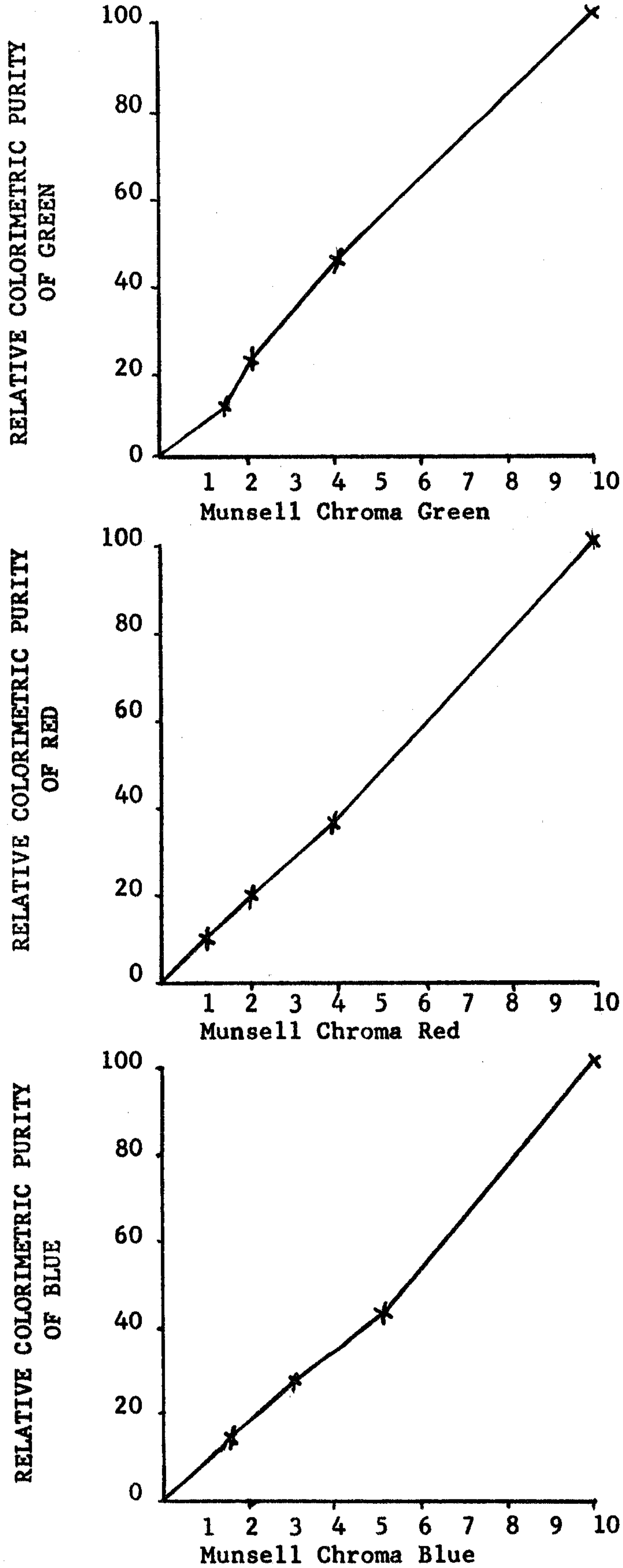

Fig. 6. Relative Colorimetric Purity to Munsell Chroma of green, red and blue.

CHAPTER IV

DISCUSSION OF RESULTS

It can clearly be seen from Figures 1, 2, 3 and 4 that the experimental results produce a ratio scale proceeding from zero to about Munsell Chroma 10. However the algebraically derived scales of Per Cent Hue, Relative Colorimetric Purity, Relative Excitation Purity and Munsell Chroma appear to be ratio scales from zero only to approximately Munsell chroma five. Beyond this point all scales begin to deviate similarly from a ratio scale and follow a curve of undetermined characteristics. The exact point at which this departure begins and the characteristics of the resulting curves cannot be determined from available data. However the change appears to begin in the interval between 4.5 to 5.5 of Munsell Chroma.

From Figures 5 and 6 it appears that the experimental results and the renotated values of the results in the Munsell system are linearly related to relative colorimetric purity. However, colorimetric purity is defined, at the spectral limit for each hue, as 1.0 or unity. Jones and Lowery (1926), on the other hand, demonstrate that along the psychological saturation continuum the particular scales for each hue consist of a varying number of D.L. steps. This implies that scales of colorimetric purity cannot completely represent the psychological continuum. That is almost certainly the reason for the non-linear departure seen in figures 1, 2, 3 and 4.

It was mentioned in Chapter Three that several of the subjects were unable to bisect for the one-eighth point in the blue. Perhaps the significant factor contributing to this inconsistency was a lack of proper control of intensity. C. H. Graham states:

> This condition (constancy of intensity) is a very important requirement. Small differences in the luminance of the two fields can provide an erroneous result such as has, in fact, occurred in some experiments. Thresholds in the blue are particularly susceptible to this sort of error. brackets mine (C. Graham, 1961, page 155).

Secondly, the determination of the tristimulus values, the colorimetric and excitation purity assumed the use of standard illuminant "C" which was only approximately realized in this experiment. Consequently some of the error seen in the curves could be eliminated by a more adequate control of the illumination. In defence of the system used in this experiment it could be stated that the best available lights were employed.

Various extensions of this work seem to be indicated. The experiment could be duplicated where a more adequate experimental control of brightness (intensity) was accomplished. A second extension would be to duplicate this experiment where all ranges, zero to limits, of Munsell hue, value and chroma were used. Once this data was obtained, a third extension would be, to determine more completely the relation between Munsell chroma and an extended psychological scale of color saturation. The establishment of this relation could lead to a correction of the higher Munsell chroma values so that they more accurately reflect the ratios of the psychological scale of color saturation.

CHAPTER V

SUMMARY AND CONCLUSIONS

The present investigation began with a tentative definition of the saturation of a color as a function of the amount of an equally bright grey which, when mixed with the color, will yield a mixture which appears to be half as saturated as the original. Before this proposal could be elaborated, it was necessary to demonstrate that the implied psychophysical operations could be carried out; that a satisfactory ratio scale of saturation was possible.

The construction of such a ratio scale was the immediate purpose of the research. Three hues (red, green and blue) were chosen; the color stimuli were obtained by mixing standard Munsell papers with neutral greys of equal value. The experimental technique was that of the bisection of the interval between grey and the color, using the method of limits to determine the bisection points.

The procedure yielded satisfactory ratio scales up to chroma 10 on the Munsell scale (the maximum saturation used in the experiment). The scales were found to be linearly related to customary measures of saturation (Munsell chroma and colorimetric and excitation purity) up to levels corresponding to Munsell chromas of 4.5 to 5.5. This linear relation broke down for higher saturations. The non-linearity at high levels was tentatively attributed to the fact that the conventional measures of saturation assign the same value to colors at the spectral limit, although there is good evidence that spectral hues are not psychologically equal in saturation.

It was suggested that extensions of the present technique, using mixtures of white and monachromatic light, could lead to simple modifications of the conventional measures, to bring them more in line with the psychological data.

BIBLIOGRAPHY

Churcher, B.G., A loudness scale for industrial noise measurements. J. Acoust. Soc. Amer., 1935 6 216-226.

Committee on Colorimetry, Quantitative data and methods for colorimetry. J. apt. Soc. Amer., 1944, 34, 633-688.

Garner, W.R. Advantages of the discriminability criterion for a loudness scale. J. Acoust. Soc. Amer., 1955, 30, 1005-1012.

Geiger, P.H. and Firestone, F.A. The estimation of fractional loudness. J. Acoust. Soc. Amer., 1933, 5, 25-30.

Graham, C.H. Color theory. In S. Koch (Ed.), Psychology: A Study of a Science. 1959, 1, 145-287.

Ham, L.B. and Parkinson, J.S. Loudness and intensity relations. J. Acoust. Soc. Amer., 1932, 3, 511-534.

Hanes, R.M. A scale of subjective brightness. J. exp. Psychol., 1949, 39, 438-452.

Judd, D.B. A maxwell triangle yielding uniform chromaticity scales. J. apt. Soc. Amer., 1935, 25, 24-35.

Judd, D.B. Hue saturation and lightness of surface colors with chromatic illumination. J. apt. Soc. Amer., 1940, 30, 2-32.

Laird, D.A., Taylor, E. and Wille, H.H. The apparent reduction of loudness. J. Acoust. Soc. Amer., 1932, 3 393-401.

Martin, L.C., Warburton, F.L. and Morgan, W.J. The determination of the sensitiveness of the eye to differences in the saturation of colours. Great Britain Medical Research Council, Special Report Series, 1933, 1-42.

Munsell, A.H. On the relation of the intensity of chromatic stimulus (physical saturation) to chromatic sensation. Psychol. Bull., 1909, 6, 238-239.

Newhall, S.M., Nickerson, D. and Judd, D.B. Final report of the O.S.A. subcommittee on the spacing of the Munsell colors. J. apt. Soc. Amer., 1943, 33, 385-418.

Richardson, H.H., and Ross, J.S. Loudness and telephone current. J. Gen. Psychol., 1930, 3, 288-306.

Stevens, S.S. A scale for the measurement of a psychological magnitude: loudness. Psychol. Rev., 1936, 43, 405-416.

Stevens, S.S. (Ed.) Handbook of Experimental Psychology. New York: John Wiley and Sons, Inc., 1951.

Stevens, S.S. On the psychophysical law. Psychol. Rev., 1957, 64, 153-181.

Stevens, S.S. The estimation of loudness by unpracticed observer. J. Exptl. Psychol., 1956, 51, 71-78.

Torgerson, W.S. Theory and Methods of Scaling. New York: John Wiley and Sons, Inc., 1958.

Troland, L. Psychophysiology II. Sensation. New York: D. Van Nostrand, 1929.

Woodworth, R.S. Experimental Psychology. New York: Henry Holt and Company, 1938.

Woodworth, R.S. Experimental Psychology. New York: Henry Holt and Company, 1960.

VITA AUCTORIS

1939	Born in Cedar Rapids, Iowa to Henry Z. and Jane Francis Sempowski.
1945-57	Educated at St. Thomas School, Villanova Penn., Holy Names School, Birmingham, Michigan and Shrine High School, Royal Oak, Michigan.
1961	Graduated with the degree of B.A., Assumption University of Windsor, Windsor, Ontario. Registered as full-time graduate student at Assumption University of Windsor.