THE INFLUENCE OF COLOR ON APPARENT WEIGHT. A PRELIMINARY STUDY

BY J. E. DE CAMP

Stanford University

CONTENTS

I. STATEMENT OF THE PROBLEM

The apparent weight of an object is a function of its size. Two objects of like material, appearance, etc., and having the same weight, but of different size, will appear unequal in weight not only to the unsuspecting subject, but as well, though to a less marked extent, to the experienced individual who is thoroughly familiar with the nature of the illusion involved. A second psychological factor influencing the weight of an object has been pointed out by Seashore.[1] This factor is the material out of which the object is thought to be constructed. The apparent weight of an object is influenced by many other factors than its size and material. It depends upon such factors as the neuro-muscular adjustment brought about by lifting one or more weights, the time elapsing between this adjustment and the lifting of the weight to be judged, the velocity with which the weight is raised from its support, the time occupied in the lifting process, whether it

[1] Univ. of Iowa Studies in Psychol., 1899, 2, 36–46.

is lifted before or after the weight with which it is to be compared, bodily conditions such as fatigue, etc.[1]

The question arose in the mind of the writer whether the color of an object might not influence its apparent weight. Would objects alike in all respects save that of color appear equal in weight? Is color in any way a determinant of the apparent[2] weight of an object? If the color of an object does influence its apparent weight we should like to know how the latter varies with a variation in hue, tint, and chroma. This paper and the experimental work herein reported will serve only as an approach to an answer to these questions. Its main object is to suggest, with experimental evidence for the suggestion, an answer to the particular question: Is color in any way a determinant of the apparent weight of an object? In other words, is there a color-weight illusion just as there is a size-weight illusion and a material-weight illusion?

II. Literature

So far as the writer is aware no experimenter has given a careful consideration to the problem with which we are now concerned. Seashore,[3] in connection with his experiments on the material-weight illusion, made a few tests on children to determine the presence or absence of a color-weight illusion. His method was that of selecting from a graded series a block thought to be equal to a comparison block. Only black and white were tested. His results indicated the absence of any color-weight illusion. It is obvious that the increment used

[1] Compare in this connection C. E. Seashore, 'Psychology in Daily Life,' N. Y., 1913, 175 ff.; G. E. Müller and F. Schumann, *Arch. f. d. gesamm. Physiol.*, 1899, 45, 37–112; L. J. Martin and G. E. Müller, 'Zur Analyse der Unterschiedsempfindlichkeit,' Leipsig, 1899.

[2] *Psychological* weight might be substituted for *apparent* weight. The qualification is necessary to distinguish the apparent weight when lifted by an experiencing person from the physical weight of the object. I have often noticed that the beginner in psychology not infrequently has difficulty in immediately realizing that the psychological weight of an object is just as real as its physical weight. The fact that reference is always made to a physical standard leads to the use of the term illusion. The pound of lead is really psychologically heavier than the pound of feathers. Scripture's law or formula (*Science*, N. S., 1897, 5, 227), of the size-weight illusion is only a means of passing from one real to another.

[3] Univ. of Iowa Studies in Psychol., 1899, 2, 45.

(5 grams) in his standard series of blocks was too large to separate out any factor whose influence is very slight. A second approach is some work done by E. Bullough[1] on 'The Apparent Heaviness of Colors,' which is, as the writer of the article points out, a contribution to the æsthetics of color. He is concerned in particular with the æsthetic effect of placing one color above another. It may be said that he attempts to investigate scientifically the rule "that dark colors should stand below light ones."[2] His method is that of choices or preferences. It is concluded that dark colors are heavier than light ones. Accoiding to him "it is the luminosity which forms the principal factor in the preferences,"[3] *i. e.*, the relative luminosity of a color determines whether it will be preferred above or below another color. Seemingly the empathic sense in which the term weight is here used differs quite markedly from that in which we use the term weight in speaking of a lifted object, but possibly the significance of the term in the two cases is quite the same. Our ignorance of the factors involved in the æsthetic arrangement of two or more colors at present precludes a definite answer to this question.

There are various articles dealing with the many factors influencing the apparent weight of a lifted object, such as the size of the object, the rate at which it is lifted, etc., but since we have endeavored to make such factors constants in our tests they need not claim our attention here, and it is not necessary to refer to the literature dealing with them.

III. Experimental Work

Section I

The experimental work reported in this section was carried on by the writer and two of his advanced students[4] in

[1] *British J. of Psychol.*, 1907, 2, 111–152.

[2] *Loc. cit.*, 113.

[3] *Loc. cit.*, 133; *cf.* also p. 152 where he says: "Since the saturated colors are not all of the same luminosity, but yellow and green are of higher luminosity than blue and violet, yellow and green are also of apparently lighter weight than blue or violet when seen singly."

[4] Misses Irene Hollenbeck and Laura Stephan. Sixty-five per cent. of the tests were made by the writer.

the psychological laboratory of the University of California during the second semester of the year 1915–1916. Of the sixty-three subjects tested nineteen were men and forty-four were women. The subjects were drawn from Professor Stratton's class in applied psychology. Seven of the subjects were seniors, thirty-five were juniors, nineteen were sophomores, and two were specials. The test was always given in the forenoon some time between 8 A. M. and 12 M. Due to a cloudy or foggy morning there were a few instances where the light conditions were slightly below normal (ordinary diffuse daylight). The test usually occupied about forty-five minutes, and was given to each subject individually. Each test fell into two main parts: for convenience we will call the first part Part 1, and the second part Part 2. Part 1 always preceded Part 2.

A. Apparatus.—The apparatus used was twenty-five cubical ($2\frac{1}{2} \times 2\frac{1}{2} \times 2\frac{1}{2}$ in.) soft pine blocks. These blocks were smoothed and made the desired weight by boring out part of the wood from the bottom of the block. Nine of the twenty-five were used in Part 1, all were used in Part 2. The nine blocks for Part 1 were all made exactly the same weight, seventy-six grams. Each one of the nine was covered with colored paper by glueing the paper evenly and closely to five sides of the block. The bottom was left uncovered. In this way the nine blocks were alike in size and weight; in fact they were alike in every respect save that of color. Each of the nine differed from every other one in respect to color. The following colors were represented: red, orange, yellow, green, blue, violet, purple,[1] black, and white. It will be seen later that the present work does not demand a refined photometric method such as is described by Sir Wm. Abney,[2] or like that given by Yerkes and Watson.[3] An attempt was made to obtain the approximate luminosity value of each color by the method suggested by Rood.[4] The following values were received for the luminosity of the different

[1] Probably more correctly designated by the commercial term *cerise*.

[2] 'Researches in Color Vision,' London, 1908.

[3] 'Behavior Monographs,' 1911, No. 2.

[4] *Amer. J. of Sci. and Arts*, 1878, 15, February.

colors: red, 5.2; orange, 27.2; yellow, 75.7; green, 44.3; blue, 14.7; violet, 8.5; purple, 6.6; black, 0; white, 100; gray, 16.9. The figures given represent the percentage of white contained in the colored discs. No pretense is made to absolute accuracy. There can be little doubt, however, that the relative luminosity of the different colors is correctly shown.

The sixteen additional blocks used in Part 2 were of the same material and size as those used in Part 1. These were covered on five sides with gray paper (No. 16[1] of Hering's series). They all differed in weight. The sixteen made up a graded series from fifty-five to one hundred grams, the increment in each case being three grams. The set of sixteen blocks were alike, then, in every respect save that of weight.

Additional apparatus consisted of (*a*) a stand supporting a horizontal rod 16 cm. above the base of the stand, (*b*) a second stand supporting a black cardboard screen, and (*c*) a large piece of soft black velvet cloth. The horizontal rod mentioned under (*a*) was wrapped with black cloth.

B. Method: Part 1.—The general method used in Part 1 was that of Paired Comparisons. This method is so well known that it needs no discussion here.[2] The ranks and files of the table constructed for the thirty-six (*i. e.*, between nine objects; $\frac{9 \times 8}{2} = 36$) possible comparisons were headed thus:

	R	O	Y	G	Bl	V	P	Bk
O	1							
Y	2	3						
G		4	etc.					
Bl								
V								
P								
Bk								
W								

The observer[3] was seated at one side of a fairly large table, the experimenter at the other. The table was covered with the large piece of black velvet cloth. The stand supporting

[1] W. Brown has shown (Psychol. Rev., 1915, 22, 520 f.) that this particular gray is the seventeenth (instead of the sixteenth as numbered) in Hering's set.

[2] See E. B. Titchener, 'Experimental Psychology,' Vol. 1, Pt. 1, 92 f.

[3] Throughout this paper the terms *subject* and *observer* are used interchangeably.

the horizontal rod was placed near the observer. The purpose of the rod was to give uniformity to the height of each lift made by the observer. The height of the rod above the table was 16 cm. The observer was definitely shown how the lift was to be made. The elbow was to be used as a fulcrum; no other part of the arm or hand was to be bent. The subject[1] was told to have as nearly as possible a uniform method of grasping the blocks. The blocks were placed in position by the experimenter for the subject. Only two blocks, those being compared, were allowed on the table at the same time; the others were kept in the table drawers on the experimenter's side of the table. The cardboard screen was placed in such a way that the protocol of the experimenter was concealed from the subject's view.

The first comparison was made between red and orange, the second between red and yellow, the third between orange and yellow, etc. Care was taken that in the eight comparisons belonging to any color the block representing the color in question should be presented four times before and four times after the blocks with which it was compared. This served to eliminate any temporal error. All blocks were lifted from the same position, in that way avoiding any positional error. Before the test began the observer was impressed with the following four points:

1. The experiment is a comparison of weights.

2. The weights are all different; no two have the same weight.[2]

[1] *Vide* note 3, p. 351.

[2] It might be thought that the proper procedure would be to omit any reference to the weight of the blocks. In such a case should the observer be presented in the first few trials of the experiment with blocks whose color has little relative influence upon their weight he would very likely get the impression that the blocks were all equal in weight. The influence of this idea would invalidate the results obtained so far as the factor under investigation is concerned. Suppose that the different colors have little relative influence upon the apparent weight of objects, the introduction of this common factor will greatly aid in the determination of the little and will not materially affect its validity. The suggestion will aid in fixing in the mind of the observer the idea that the aim of the experiment is a comparison of weights and will tend to exclude from him any knowledge of the factor under consideration. This is highly desirable since the untrained observer can not avoid distorting actual occurrences; his results are biassed by what he thinks the results ought to be under such conditions.

3. The second weight in each comparison is to be judged in terms of the first.

4. The eyes are to be kept fixed upon the block while it is being lifted.

After these instructions had been given a few preliminary trials were usually made with some of the blocks with the graded gray series for the purpose of (1) familiarizing the observer with the experimenter's method, and (2) setting up a kind of muscular adjustment in each individual case for the perception of differences in weight. At first fairly large differences were given, then smaller ones. In these preliminary tests the subject was informed whether his judgments were right or wrong.

Part 1 was then conducted in accordance with the method of paired comparisons. The observer was in no case informed whether his judgment in any particular instance was correct. Sometimes when the observer showed too much anxiety about the correctness of his judgments the experimenter made such casual remarks as, "You are doing very well," or, "You are getting along very nicely." Immediately upon each individual judgment the experimenter entered the judgment in the table previously prepared for that purpose. Unless the observer showed signs of fatigue the test was carried directly through. This part of the test required approximately fifteen minutes. In case of fatigue, and this very seldom happened, the observer was allowed a brief respite.

Method: Part 2.—In this part the complete set of twenty-five blocks was used. Only two blocks were on the table at a time. Beginning with the red block and taking the colored blocks in spectral order each one served as a standard with which a number of the sixteen gray blocks were compared. In each case the colored block was lifted first, then the gray block. The gray blocks were taken in the order of their weight, *i. e.*, 55, 58, 61, ..., 100. Before beginning this test the observer was told that in this case some of the blocks compared might be just alike in weight, in this way indicating to the observer that in some of the comparisons 'the same' might occur as a correct judgment as well as 'lighter' or 'heavier' in other comparisons. The observer was not informed that the gray blocks constituted a graded series.[1] Record was made by the experimenter of the subject's judgments. In the first few experiments the comparison of any colored block with the graded series of grays was stopped upon the subject's first judgment of 'heavier.' In the later tests the comparison was stopped only when one judgment of 'heavier' was followed in the next comparison by a second judgment of 'heavier.' It was assumed that a continuation of the series would lead to a judgment of the remaining heavier weights as 'heavier.'

After the completion of Part 2 four of the colored blocks, red, yellow, green, and blue, were placed before the observer ranging from his left to right in the order named. He was told to lift these blocks separately as many times as he desired, no horizontal rod being used to restrict the height of each lift, and to place them in the order of their weight, the heaviest one to his left, the next heaviest one place to the right, and so on. By means of a stop-watch was recorded the time that it took the observer to arrange these four blocks in the manner described.

All nine of the colored blocks were then placed in chance order before the subject. He was told to select that block which was most pleasing, *i. e.*, whose color was most pleasing, also to select the most pleasing block of the remaining eight, the most and the next most displeasing blocks.

In no case, not even at the close of the test, was the subject told the purpose of the experiment or the point under investigation. No information was given the subject about the results obtained in the experiment.

[1] This varies slightly from the method recommended by Whipple, 'Manual of Mental and Physical Tests,' Baltimore, 1915, Pt. 2, 226, for a similar effect in the case of the size-weight illusion. In his case the graduation while not distinctly mentioned is open to easy discovery. In Part 2 of the present experiment, following upon Part 1, in which no apparent (to the observer) order is followed, no impression is conveyed that the grays are given in any order other than a promiscuous one. Doubtless the repeated giving of the graded series with the colored blocks led to an idea, on the part of the observer, of the gradation, yet in no case, to the knowledge of the writer, was any indication given that the gradation was noticed. With the results of this part of the experiment at hand it seems probable that more reliable results would have been obtained had the subject been informed that the series was a graded one.

The subject's name, sex, year in the university, and the condition of the light were indicated upon the record sheet.

C. Data and Results.—The results obtained by the method of paired comparisons (Part 1) are graphically shown in Fig. 1.

Arranging the colors in the order of their apparent heaviness we have red, white, orange, violet, green, purple, blue, yellow, black. In this decreasing series three characteristic step effects may be clearly seen. The highest is formed by the two colors red and white; the next step, whose height corresponds fairly closely to the line of chance, may be thought of as made up of orange, violet, green, and purple; the third step is made up of blue, yellow, and black. Stated in another way, the results seem to indicate that red and white have a positive; blue, yellow, and black a negative; orange, violet, green, and purple an indifferent; effect upon the apparent weight of the objects lifted.

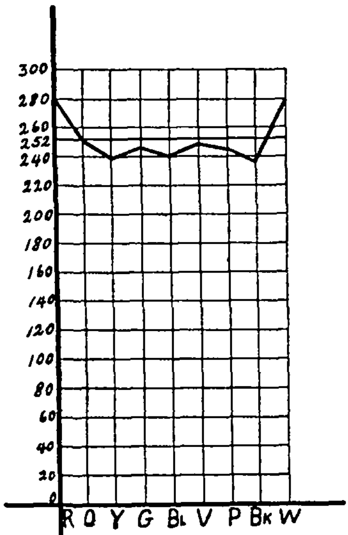

Fig. 1

Fig. 1. Method of paired comparisons. The results for 63 subjects are shown. The ordinates represent the total number of judgments of 'heavier' given the different colored blocks.[1] The different colors are represented along the line of abscissas. The height to which a color should rise, assuming the absence of all factors save that of chance, is shown by the horizontal line at ordinate 252.

The average value,[2] the average value in terms of per cent.,

[1] Though no encouragement was given to the subject to reply that the blocks compared were of the same weight such judgments occasionally occurred. In cases of this kind, also in those cases where the subject seemed quite doubtful, the comparison was repeated, but in the curve here shown these cases were disposed of by giving to each color involved one half of one unit credit, it being the plan to give to the color judged 'heavier' one unit credit.

[2] The 'average value' for each color is obtained by dividing the total number of times for all subjects it was judged 'heavier' by the number of subjects. If any color had always received a judgment of 'heavier' regardless of the color with which it was compared its average value would be 8. Assuming no influence other than chance the average value for any color would be 4. The M. V. as given in the table is computed from the average.

and the mean variation for each color are given in the following table, Table I.:

TABLE I

Color	Av.	%	M. V.
Red	4.452	55.6	1.083
Orange	4.008	50.0	0.899
Yellow	3.794	47.4	1.203
Green	3.905	48.8	1.011
Blue	3.810	47.6	1.180
Violet	3.960	49.5	0.943
Purple	3.897	48.7	1.194
Black	3.754	46.9	1.218
White	4.421	55.3	1.194

Table II. is presented to show the gross distribution about the chance line of the individual results obtained. The results, for the men, for the women, and for the men and women are shown in separate columns.

TABLE II

Color	Men			Women			Men and Women		
	>C	<C	=C	>C	<C	=C	>C	<C	=C
Red	12	4	3	22	13	9	34	17	12
Orange	5	9	5	17	15	12	22	24	17
Yellow	4	10	5	15	22	7	19	32	22
Green	8	6	5	13	22	9	21	28	14
Blue	7	8	4	12	20	12	19	28	16
Violet	7	9	3	16	11	17	23	20	20
Purple	8	7	4	14	21	9	22	28	13
Black	5	8	6	14	25	5	19	33	11
White	10	5	4	23	16	5	33	21	9

In the columns headed > C is given the number of individuals whose credit value for the particular color was greater than chance would have allowed. In column < C (= C) is given the number of individuals whose credit value for the different colors was less than (equal to) that allowed by chance.

Considering the totals we note the same step-effect as was pointed out above, and each step is made up of exactly the same colors as before: red and white make up the first; orange, green, violet, and purple, the second; yellow, blue, and black, the third. Even though the number of men tested is quite too small to merit serious consideration of their results apart from the combined results of both men and women it will be seen that save for orange and blue their distribution is quite in accordance with the statements made above.

If we now note the results obtained from the auxiliary test[1] with the four colors, red, yellow, green, and blue, we have the following table, Table III.:

TABLE III

	Red		Yellow		Green		Blue	
First place	25	13	11	13	10	13	6	13
Second place	16	13	11	13	12	13	13	13
Third place	7	13	17	13	19	13	9	13
Fourth place	4	13	13	13	11	13	24	13

First place (second place, etc.) is used to denote the fact that a block falling therein was designated by the observer as heaviest (next heaviest, etc.) of the four. The first column under each of the four colors indicates the number of individuals giving to the various colors the first, second, third, and fourth place. The second column gives the distribution according to chance.

If we adopt an arbitrary system of units in which the occurrence of a colored block in the first place is given a value of 4; its occurrence in the second place, a value of 3; in the third place, a value of 2; in the fourth place, a value of 1; and compute upon this basis the aggregate value for each of the four colors, we have: red, 166; yellow, 124; green, 125; blue, 105. The figure of chance is 130. These results, save those for yellow, quite accord with our previous statement (see p. 355). To yellow is attached a slightly larger figure than we should have anticipated on the basis of our previous results. To agree with the previous results it should have attained a figure near that of blue, about 105. This slight discrepancy may be due to the presence in this part of the test of complicating factors arising out of the increased freedom in method allowed the subject.

The average time spent in arranging the four blocks in

[1] The method of this test is given on page 354 *supra*.

the order of their apparent weight was 1 min. 59 secs. This result is obtained by averaging the results of forty-seven subjects. The experimenters began to take the time with the sixteenth subject in the series, and failed to get the time in a single subsequent case.

The test requiring a choice of the two most pleasant and the two most unpleasant of the nine blocks led to the following results, Table IV.:

TABLE IV

	Pleasant		Unpleasant	
Red	2	4	1	4
Orange	4	8	6	9
Yellow	4	4	3	5
Green	7	14	3	4
Blue	14	11	1	4
Violet	7	2	4	9
Purple	6	5	20	11
Black	0	1	12	2
White	6	1	0	2

In the first column under Pleasant (Unpleasant) appears the number of persons choosing red (orange, yellow, etc.) as the most pleasant (unpleasant) color of the nine. In the second column under Pleasant (Unpleasant) is given the number of persons choosing red (orange, yellow, etc.) as the next most pleasant (unpleasant) color. From the table it readily appears that the most generally pleasant color was blue, with green next; the most generally unpleasant color was purple, with black next. Three of these colors, black being the one excepted, showed no particular influence upon the block so colored. Red and white called out little affective quality, and it was these two colors[1] that seemingly had the most marked effect upon apparent weight. Again, of the three colors, yellow, blue, black, which appeared to influence apparent weight negatively, one (blue) is generally pleasing, one (black) is rather displeasing, the other (yellow) is neither decidedly pleasant nor decidedly unpleasant.

The following table, Table V., shows the results obtained from the second part of the experiment (designated above, Part 2):

[1] *Supra*, p. 355.

TABLE V

Red	68.27	Violet	70.78
Orange	68.32	Purple	70.07
Yellow	67.46	Black	69.57
Green	68.22	White	70.56
Blue	68.30		

The figures opposite the various colors are obtained in the following manner: for each individual test an interpolated value for each color is computed by averaging the value of the last judgment of 'lighter' and that of the first judgment of 'heavier'; the sum of these individual interpolated values is then averaged, and it is the resulting figure that appears in the table. Probably little of a definite nature is to be gained from a consideration of this table. It seems that the method was too crude for satisfactorily disengaging the factor under consideration from the many others which complicate the problem of the apparent weight of a lifted object. Doubtless one of the main complicating factors was the fact that the colored blocks towards the end of the series would not be approached with the same naïveté as were those of the first part of the series. It will be noted that there is a decided tendency to overestimate the weight of the gray block. This overestimation is due, I take it, to the serial arrangement of the grays.[1] It is the result, often noted in the case of unsophisticated observers, and clinging even to the trained observer, of the premature equalization, when there is a problem of equalization, of members of a series of diverse stimuli with a standard stimulus. Seemingly in all such cases there is a tendency to judge that the goal is reached *before* it is reached rather than to judge that the goal is reached *after* it has been reached. This tendency is quite noticeable in the various experiments connected with the problem of the limen. The following is sufficient evidence that this factor was really an effective one in the present experiment. Though each gray block was always heavier than the preceding one very frequently the observer would

[1] Subsequent experiments (Section II.) clearly show that this overestimation is not due to the color (gray) of the block. It may be that the grayness has some influence, but it is relatively small. Compare results given on p. 363 *infra*.

judge a very light block (say one of 64 grams) to be equal to or heavier than the colored block (weight, 76 grams), then judgments of 'lighter,' 'the same,' and 'heavier' would follow. Only one of the sixty-three individuals tested failed to definitely show this tendency, and it was not an uncommon occurrence for this tendency to appear more than once during any particular test. In the opinion of the writer anticipatory attention has entered as a complicating factor in the averages presented. If so, its influence should be stronger for those colors appearing first in the series, since numerous comparisons would in all probability tend to lessen or altogether eliminate it in the case of those colors coming towards the end of the series. Were this factor alone present it seems that the series of values from red to white should show a gradual increase. It is perfectly clear that in the figures this increasing tendency is present. A separate average for the first, second, and third group of three values gives 68.02, 69.10, and 70.07 respectively. But it is quite probable that anticipatory attention is not the only influence present, and the results of our test by the method of paired comparisons, which seems to be the most adequate method of attacking the problem in hand, would lead us to assume an influence of the different colors. The low value (lowest of the nine) of yellow accords with the results previously obtained.

Section II

This section contains a report of additional experiments carried out with the hope of obtaining more satisfactory results relative to the existence of a color-weight illusion. The work was done at Stanford University during the year 1916–1917. The subjects used, sixty-eight men and thirty-one women, were drawn from various classes in psychology, but mostly from my class in mental hygiene. First-, second-, third-, and fourth-year students were represented.

While the apparatus and method employed were similar to those used in the first part of Section I., a brief description is necessary for a clear understanding of the results to follow.

Apparatus.—Six cubes ($2\frac{1}{2} \times 2\frac{1}{2} \times 2\frac{1}{2}$ in.) were constructed from ordinary Bristol board; the bottom of each was left open; the remaining five sides were covered over with a colored paper, the color varying with the cube. The six colors represented were red, yellow, blue, black, white, and gray. These colors were the same as those used in Section I. with the exception of the gray, which in the present case was No. 17 of Hering's series. (No. 16, the gray previously used, was not available.) Each cube was made to weigh seventeen grams. Adjustment of weight was accomplished by glueing a cork to the bottom of the upper side of the cube and adding shot or paring away cork. The choice of a small absolute weight for the cubes was dependent upon the consideration that if the color of an object has an influence upon its apparent weight that influence should be relatively greater the lighter the object, since the lighter the object the smaller the appreciable difference in weight.

Method.—The subject was seated at one side of a table, the experimenter at the other. The following experiments were carried out:

1. The four cubes, red, yellow, blue, and white, were placed on a large sheet of dark cardboard lying upon the table. The right-to-left arrangement of the cubes varied from subject to subject. The following four orders of presentation were used, beginning at the subject's right: (1) red, yellow, blue, white, (2) yellow, blue, white, red, (3) blue, white, red, yellow, and (4) white, red, yellow, blue. The first (1) represents the order in which the cubes were placed before the first subject; for the next subject the order was changed to that shown in (2); etc.

The subject was told that the cubes were all different in weight and that he should arrange them in the order of their weight, placing the heaviest at his left. He was definitely instructed to look at the weights while lifting them. An occasional prompting was sometimes found necessary to make the subject conform to this requirement. The subject was allowed to rearrange the order[1] as often as desired until

[1] A change in the order often occurred. One of the most frequent means of bringing about such a change was the following: the subject would lift the various

the proper order was obtained. The subject's final arrangement was noted and recorded.[1]

2. A test similar to the one just described was given using the three cubes, white, gray, and black.

3. The subject was tested with cubes red, yellow, blue, white, and black, using the method of paired comparisons. This method has already been referred to under Section I. (p. 351). The same general precautions were taken here as there.

4. In some of the later tests the subject was asked to designate from 'looks' alone the order of the apparent heaviness of the cubes. After the subject had given his order he was asked to explain why the cube so designated looked heaviest and another seemed lightest.

The results obtained for Test 1 appear in Tables VI. (*A*) and VI. (*B*).

TABLE VI. (*A*)

Color	Men (68)				Women (31)				Men and Women (99)			
	1	2	3	4	1	2	3	4	1	2	3	4
Red	18	16	19	15	6	5	13	7	24	21	32	22
Yellow	15	19	22	12	8	10	8	5	23	29	30	17
Blue	17	20	17	14	9	10	5	7	26	30	22	21
White	15	13	10	30	8	6	5	12	23	19	15	42

Under 1, 2, 3, and 4 appears the number of times the different colors were ranked as heaviest, second heaviest, third heaviest, and lightest. The numbers in parentheses following Men, Women, and Men and Women indicate the number of subjects.

Assigning the following values to the different judgments, heaviest, 4; second heaviest, 3; third heaviest, 2; lightest, 1; and computing the aggregate value for each color we have Table VI. (*B*).

The significance of these results is not very clear. They

cubes one or more times and place them in a certain order, then he would compare the extremes, seemingly as a confirmatory test, and that would frequently lead to a change in the established order.

[1] This method will be referred to later as the 'arrangement' method.

TABLE VI. (*B*)

Color	Men (68)	Women (31)	Men and Women (99)
Red	173	72	245
Yellow	173	83	256
Blue	176	83	259
White	149	72	221

do not support the interpretation given to the results of Section I., neither do they accord with the results obtained by the method of paired comparisons referred to later in this section (see p. 364). There seems to be a tendency to judge white and red as relatively light, blue is heaviest, while yellow follows as next heaviest. It is to be noted that there is not complete agreement with the rule: light colors—light weight, dark colors—heavy weight.

The results for Test 2 are given in Tables VII. (*A*) and VII. (*B*).

TABLE VII. (*A*)

Color	Men (68)			Women (31)			Men and Women (99)		
	1	2	3	1	2	3	1	2	3
White	23	16	29	10	5	16	33	21	45
Gray	23	30	15	9	12	20	32	42	35
Black	23	22	23	12	14	15	35	36	38

Under 1, 2, 3 appears the number of times the different colors (grays) were ranked as heaviest, next heaviest, and lightest.

Assigning the following values to the judgments, heaviest, 3; next heaviest, 2; lightest, 1; and computing the aggregate value for each color we have Table VII. (*B*).

TABLE VII. (*B*)

Color	Men (68)	Women (31)	Men and Women (99)
White	186	56	242
Gray	205	69	274
Black	215	71	286

From these results for three blocks differing only in tint it is quite apparent that the tendency is to judge the darker block as heavier. This is true for both men and women.

The results for Test 3 appear in Table VIII.

TABLE VIII

Color	Men (42)	Women (8)	Men and Women (50)
Red	97	23	120
Yellow	90	21	111
Blue	96.5	29	125.5
Black	96.5	30	126.5
White	101	23.5	124.5
Gray	101	25	126

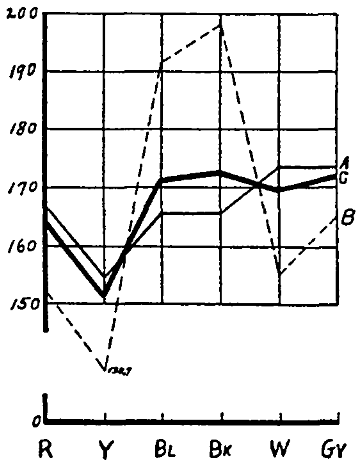

FIG. 2. Curve *A* indicates the relative difference in the apparent weight of six different colored cubes of the same weight, as judged by 42 men. Curve *B*, the same for 8 women. Curve *C* is obtained by combining the results for both men and women. The colors are represented along the line of abscissas, the relative number of judgments of 'heavier' along the line of ordinates.[1] Method of paired comparisons.

The numbers represent the relative number of judgments of 'heavier' given in a comparison of different colored cubes. A graphic representation of these results is shown in Fig. 2.

For the men white, gray, and red are the heavier colors, yellow is the lightest; for the women[2] black, blue, and gray are the heavier, as with the men yellow is the lightest. White is heaviest for the men while for the women it is slightly heavier than black and red.

Since the method of paired comparisons was used in the experiments of both Sections I. and II. it was deemed advisable, for the sake of increasing the number of comparable results from like tests, to combine the values received in the two. The colors common in the two groups were red, yellow, blue, white, and black. The judgments resulting from a

[1] In this figure, as well as in all subsequent ones, the absolute ordinate values appertaining to any individual curve have been transformed so that their summation equals 1,000.

[2] The number of women tested is too small to make sure that their results are typical.

comparison of all possible combinations of these five colors were separated out from the two wider groups of results belonging to Sections I. and II. Table IX. presents the results obtained.

TABLE IX

Color	Section I			Section II			Sections I and II		
	Men	Women	Men and W.	Men	Women	Men and W.	Men	Women	Men and W.
	(19)	(44)	(63)	(34)	(8)	(42)	(53)	(52)	(105)
Red	44.5	109	153.5	69	16	85	113.5	125	238.5
Yellow	34	82	116	67	15	82	101	97	198
Blue	32	78.5	110.5	65.5	17	82.5	97.5	95.5	193
Black	39.5	88.5	128	70	18.5	88.5	109.5	107	216.5
White	40	82	122	68.5	13.5	82	108.5	95.5	204

The table brings together the results, method of paired comparisons, of Section I., Section II., and Sections I. and II. for red, yellow, blue, black, and white. In the different columns appear the values given by men, women, and men

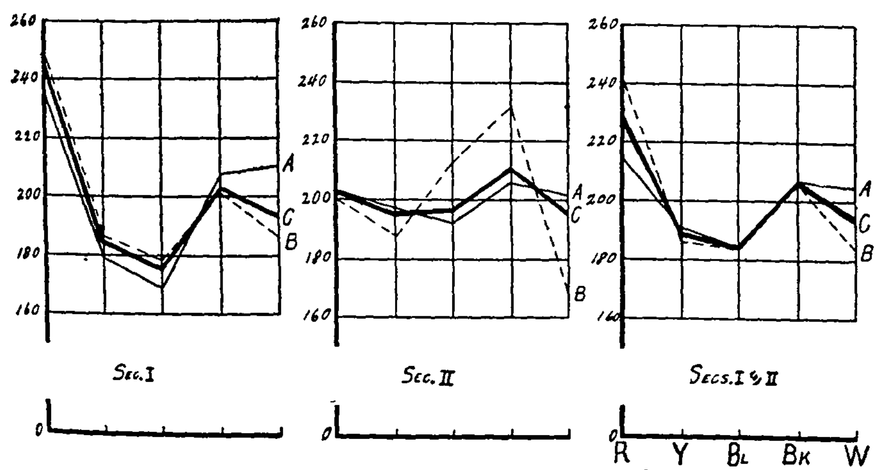

FIG. 3. Method of paired comparisons. The results of Sections I., II., I. and II. combined, are shown graphically. In each case Curve *A* represents the apparent relative weight of the five colored cubes, red, yellow, blue, black, and white as judged by the men. Curve *B*, the same for the women. Curve *C* represents the combined results for men and women. The colors are shown along the line of abscissas; the relative number of judgments of 'heavier' along the line of ordinates.

and women for each color. The numbers in parentheses designate the number of subjects.

For the purpose of facilitating comparison the results of Table IX. are graphically represented in Fig. 3.

It will be seen that there is great similarity in the general direction of the various curves. The results of both Sections I. and II. seem to indicate that there is a general tendency to judge red and black as relatively heavy, white fairly heavy, and yellow and blue light. The main departure from this tendency occurs in the women's judgment of white. For them white occupies a low position similar to that taken by yellow and blue, red and black being the heavy colors. Attention may be called to the fact that for the women red is noticeably the heaviest of all.

The results for Test 4 appear in Table X.

TABLE X

Color	Men (12)	Women (4)	Men and Women (16)
Red	44	15	59
Yellow	23	9	32
Blue	50	16	66
Black	56	23	79
White	25	8	33
Gray	44	13	57

The numbers represent the relative influence of 'looks,' without actual lifting, upon the different colored cubes. These numbers are obtained by assigning the following values to the different judgments, the block judged heaviest, 6; next heaviest, 5; ...; lightest, 1; and computing the total credit due each color.

A graphic representation of the results appearing in Table X. is shown in Fig. 4.

The results for Test 4 are much more uniform than those received in the other tests. The uniformity from subject to subject rendered the testing of a larger number of subjects unnecessary. The relative position for each of the six colors used is the same for both men and women, save in the case of yellow and white which show a small difference. These results favor the rule: light-colored objects appear light in weight; dark-colored objects appear heavy in weight.

The introspections taken in connection with the judgments of Test 4 indicated quite clearly that the judgments made were due in most cases to associations of a general and not of a specific nature. For example, the connection between the black cube and heavy weight was immediate, because dark objects are usually thought of as heavy, and did not depend upon the arousal of some specific association, like the idea of coal or iron.

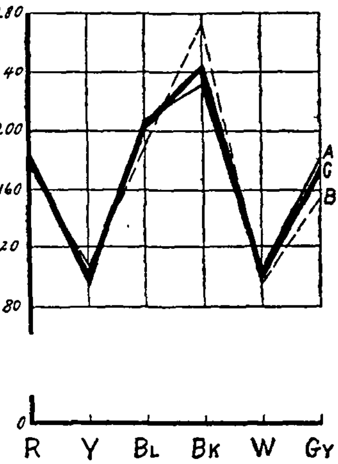

FIG. 4. Curve *A* indicates the relative difference in the apparent weight, judged by 'looks' alone, of six different-colored cubes as judged by 12 men. Curve *B*, the same for 4 women. Curve *C* represents the combined results of both men and women. The colors are shown along the line of abscissas. The ordinates represent the relative weight of the different colors.

We raised an objection to accepting at their face value the results of Part 2 of Section I. because of the probable influence of anticipatory attention. The values for the colors, red, yellow, blue, black, and white (the colors for which we have extended results with the method of paired comparisons), do not vary widely in their distribution from the distribution obtained by the method of paired comparisons. If the value for red were slightly greater (and if we allow for anticipatory attention it would be relatively larger) the curve of distribution for these values would be quite similar to those obtained by the method of paired comparisons. The low position of yellow is to be noted. If anticipatory attention is not admitted the results for blue in the two cases do not accord.

IV. Discussion

The results accruing from the problem herein attacked probably need little further discussion. They indicate that when the method of paired comparisons is used there is a general tendency to judge the apparent weight of red and black as heavier than that of yellow and blue. The fact that red and black always lie above,[1] yellow and blue below, the line of chance (at ordinate 200) may be interpreted as due to the overestimation of the weight (judged after lifting) of the colors red and black and the underestimation of yellow and blue.

The results received when the arrangement method was used (p. 362 f.) are not in harmony[2] with the above interpretation. They seem to indicate quite clearly that white is judged lighter than red, yellow, or blue.

There is no doubt in the writer's mind that the results obtained by the method of paired comparisons must be considered as more reliable than those obtained by the arrangement method. The former appears to be the more scientific method. A number of distinct judgments is involved and the situation immediately preceding each judgment is comparatively simple. When the arrangement method is used the placing of all the blocks before the subject renders the situation preceding his judgments rather complex. That the method of arrangement permits of rearrangements is probably an argument against it. Often the cause of the rearrangement seems to be extrinsic and not dependent upon the appearance of the cubes (see note, p. 361 f.).

The fact that the results of the two methods, that of paired comparisons and that of arrangement, the cubes being actually lifted, show no definite correlation[3] with the results received when the judgments were made by 'looks' alone seems to indicate that the lifting of the cubes may have introduced an additional factor. One might be disposed to draw the inference that this factor is not entirely dependent upon the appearance of the cubes.

If the claim that the results obtained by the method of paired comparisons demands prior consideration is admitted, the question arises, To what is the differential influence indicated due? The fact that the effect produced is quite analogous to that present in the size-weight illusion suggests the

[1] In Curve *B* of Fig. 3 (Section II.) the ordinate for red is exactly 200.

[2] That is, if they are taken as at all indicative.

[3] With the exception of the results for blue there is a fair degree of correspondence between the curves of Figs. 3 and 4.

possibility of a similar explanation. The size-weight illusion is usually explained on the basis of a close association between large objects and heavy weight, and between small objects and light weight. Making use of a similar argument we should expect yellow and white, since they *appear* light, to be overestimated after lifting. It may be remarked that in the case of white this occurred in the experiments of Section I. But the results for yellow, the most consistent throughout, indicate just the opposite. Further, we should expect the weight of black, blue, and red to be underestimated after lifting, but this is opposite to the general tendency as shown by the results (see curves in Fig. 3). The results, then, militate against an associative explanation so simply conceived.

Since white objects are known, other things being equal, to appear larger than black objects it might be thought that a slight difference in apparent size would exist with our blocks and consequently influence their apparent weight. If the white (black) block is judged larger (smaller) than the black (white) block, and no other factor is involved, it in consequence ought to be judged lighter (heavier) than the black (white) block. This is in accordance with the general tendency of the results for white and black, but does not hold for blue. Even with white the results of Section I. indicate that it was often judged relatively heavy. Again, it has been found that 'there is a small, but decided influence of color'[1] upon the apparent size of an object. According to Quantz[2] red, orange, yellow, and purple increase, blue and violet decrease, the apparent size of an object, green is near the borderline of influence. According to these results the red object should appear larger than the blue or violet object and consequently be judged, when lifted, lighter than they, but this was not found to be the case. A similar argument with respect to some of the other colors shows the inadequacy of apparent size as a basic explanation of the results obtained. Hence we conclude that some factor other than the apparent size of the blocks is involved.

[1] J. O. Quantz, *Amer. J. of Psychol.*, 1895, 7, 40.
[2] *Loc. cit.*, 39.

There is a possibility that the different colors exert different dynamic influence upon the observer such that red and black, say, call out but little neuro-muscular force, and blue and yellow call out more of such force. This idea is quite too conceptual to be introduced into the present discussion, which already has involved too much theorizing.

V. Conclusions

The writer is fully aware that the conclusions to be drawn from the results herein presented must necessarily be tentative. It is perfectly clear that for a final solution of the present problem much careful and accurate work must be done. At the present time I only wish to suggest the following conclusions as sufficiently justified by the results obtained:

1. The influence of the color of an object upon its apparent weight is relatively slight.
2. There is a tendency in many cases, at least when the method of paired comparisons is used, to judge a red or black object to be slightly heavier than a yellow or blue object of the same weight.
3. Apparently the influence of these colors is not due to their tint value alone; it seems necessary to consider their hue as a minor factor.
4. The results obtained do not lend themselves to a simple associative explanation.
5. Seemingly there is no simple correlation between the affective quality of a color and its influence upon apparent weight.

The writer wishes to express his indebtedness to Professor Warner Brown for a helpful criticism of an early draft of this paper.