

THE SIZE-WEIGHT ILLUSION:
A STUDY IN PERCEPTUAL DEVELOPMENT

104
591
THS

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY
Donald David Chezik
1964



ABSTRACT

THE SIZE-WEIGHT ILLUSION: A STUDY IN PERCEPTUAL DEVELOPMENT

by Donald David Chezik

The Up-and-Down Method was used to measure the size-weight illusion with the intent of eliminating constant errors found in other psychophysical methods. Approximately 600 children from kindergarten through sixth grade and 200 college students were tested, the method proving satisfactory in all respects.

Tests of statistical significance revealed that kindergarten children exhibit less illusion than older children and college students, with the latter two types not differing from each other. The difference in means was significant at the .001 level. This finding confirms earlier work that employed quite different measurement technique. There were no significant differences between the sexes.

A second trial on the illusion, with standard and variable switched to opposite hands (from the first) gave results which were similar to the first trial, except that such great variability occurred with kindergarteners that no analysis of their data could be made.

Donald David Chezik

The amount of illusion in the present study was approximately 28 per cent for the kindergarten group and 43 per cent for the other samples.

These values are close to those found in the other study.

THE SIZE-WEIGHT ILLUSION:
A STUDY IN PERCEPTUAL DEVELOPMENT

By

Donald David Chezik

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1964

ACKNOWLEDGMENTS

I would like to express my deep appreciation to Dr. Charles Hanley of the Department of Psychology for his conscientious assistance and guidance throughout this study. Also, I would like to thank Dr. R. E. McMichael and Dr. A. M. Barch, members of my graduate committee, for their wise direction and constructive criticism.

Credit is also due to the school officials in Wheaton and Glen Ellyn, Illinois, and also in scattered parts of Michigan from which school samples were taken. Without their cooperation this thesis could never have been completed.

Finally, I wish to thank my father, whose patience is extraordinary.

Donald David Chezik

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	ii
LIST OF TABLES	iv
LIST OF ILLUSTRATIONS	v
Chapter	
I. INTRODUCTION	1
II. PROCEDURE	6
III. RESULTS	11
IV. DISCUSSION	17
V. SUMMARY	20
BIBLIOGRAPHY	21

LIST OF TABLES

Table		Page
1	Age Ranges and Effective Sample Size (N) of Groups . . .	8
2	Mean, Standard Deviation, and Effective Sample Size for Males and Females, Trial I	11
3	Mean, Standard Deviation, and Effective Sample Size of Males and Females Pooled, Trial I	12
4	Mean, Standard Deviation, and Effective Sample Size with 1st-6th Grade Pooled, Trial I.	14
5	Mean, Standard Deviation, and Effective Sample Size for Males and Females, Trial II	15
6	Mean, Standard Deviation, and Effective Sample Size of Males and Females Pooled, Trial II.	16

LIST OF ILLUSTRATIONS

Figure		Page
1	The Size-Weight Illusion in Percentages	13

CHAPTER I

INTRODUCTION

Inadequate psychophysical methods have long plagued the field of perceptual development. Intrinsic flaws in these methods often render invalid what would otherwise be significant data. Although these psychophysical methods are adequate for many experimental problems with adults, their deficiencies become apparent when they are used in developmental research which cuts across several age levels.

The methods of limits and adjustment, for example, are both prone to starting position effects. Since the experimenter must begin his research at some definite level of intensity, the problem becomes one of handling and correcting the errors produced by the starting position. This is not a cumbersome problem in ordinary perceptual research, but with studies using several age levels its resolution becomes almost impossibly intricate. This entanglement occurs because starting position effects do not remain constant but vary with the age of the subject.

Another damaging effect is present in all methods employing a standard and several values of the variable. Named the "error of the standard," it generally produces an overestimation of the stimulus which is unchanging in the comparisons. Once again this error might be neutralized were it not

that it varies in magnitude with different age groups. In fact, the direction of the error may actually be reversed in some instances depending on the absolute size of the standard being used. (Wohlwill, 1960).

Any study where children are given successive trials is liable to both errors. These considerations leave the field of perceptual development without adequate psychophysical methods to study its phenomena.

The main purpose of this paper is the testing of a method which contains none of the contaminating factors mentioned above. This is the Up-and-Down method (Dixon and Mood, 1948) to be described in detail in a later section of this thesis. Briefly, the method employs a set of variables and a standard, but each subject makes only one judgment. This means that to the subject there is no standard, thereby obviating the previous-mentioned errors.

Another benefit to be derived from this single judgment method is the doing away with of errors which result from practice or fatigue. This should be especially guarded against with children, who are noted for their short interest span and susceptibility to fatigue. Because this method sidesteps these difficulties, it should be excellent for experiments which involve the younger age groups.

Worth considering is the facility of employing the method. Once a set of variables and standard has been obtained, data gathering is quickly completed due to the short length of time required for each subject.

The chief disadvantage of the method is that computation must be based on a sample size which is one-half that of the total number of subjects.

For this reason the method is not amenable to problems for which only a small number of subjects are available.

The particular perceptual phenomenon chosen for this study is the size-weight illusion, which presumably should increase with age. This, of course, is true only if the illusion derives from "set." According to this view (Woodworth & Schlosberg, 1954), the subject "expects" a large object to be heavy and unconsciously adjusts for this when he attempts to lift it. When a smaller object of equal weight is lifted with a larger object, the subject's expectations are upset and the smaller object feels heavier than the larger. The discrepancy between his muscular preparations and the sensory input is chiefly what causes this illusion. Since this expectancy requires experience with objects and their weights, it follows that older children, having more experience, will be more susceptible to the illusion than younger children.

On the other hand it is possible that the origin of the illusion is innate and independent of learning. If this is the case, there should be no differences in magnitude of illusion between age groups.

Another purpose of this paper, then, is to determine if the illusion varies with age. If there are no differences, the learning explanation is obviously unsatisfactory.

Still another reason for this study is the need for a reference point for future researchers who desire to delve more deeply into this illusion. Although the magnitude of illusion may shift when different sized containers are used; nevertheless, these results may at least give some idea of what to expect.

Due to the paucity of literature on this topic, the writer was able to find only one pertinent study. This was by Andre Rey, who sought mainly to verify a lack of illusion in mental defectives which had first been observed by J. Demoor thirty years previously. Rey believed that if this sign were reliable, a new method of diagnosing mental defectives would be made available.

In his study, Rey divided children into two age groups: 7-15 and 5-6 years of age. Then he divided the 7-15 group into normal and retarded, but the latter group is not relevant in the present case.

Rey's technique consisted of presenting two cubes of different volume but equal weight to the child and then noting if the illusion was present. If it was, he introduced additional weights into the cube judged lighter until the illusion was reversed.

Rey's handling of the data leave some doubt as to their meaning. He first counted the frequency of absence or presence of the illusion within each age group and the number of grams needed to reverse it when present. He then found the mode in grams needed for reversal in each group. Although the differences between groups suggest that ages 5-6 experience less illusion than normal 7-15 year olds, Rey made no tests of statistical significance. Instead, he has stated his conclusions in this fashion: "We see that in young children the absence of illusion is found in an important proportion." (1930, p. 293).

While lack of statistical testing is the chief defect of the study, closer scrutiny reveals more subtle flaws. Namely, early trials with the same

subject may subsequently influence his later judgments. In addition to fluctuating errors entailed in cutting across age levels, there is also a factor of loss of attention. Rey himself states: "Experience has shown the least amount of fatigue suppresses the objective value of (the subject's) responses..." (1930, p. 287). The Up-and-Down method escapes this pitfall because it allows only one judgment per subject.

In summary then, the purpose of this thesis is threefold:

1. To provide a testing ground for a psychophysical method which obviates constant errors intrinsic in other methods.
2. To test the hypothesis that young children (5-6 years of age) are less susceptible to the size-weight illusion than older children and adults.
3. To provide a frame of reference for future researchers who wish to examine the size-weight illusion with closer scrutiny.

CHAPTER II

PROCEDURE

The actual mechanics of the Up-and-Down method are quite simple. One subject at a time makes a comparison of two stimuli. In this particular study, the judgment was to determine which of two weights felt heavier when one of them was substantially larger in volume. Following this single judgment, the subject leaves and another comes in to make his judgment.

The experimenter uses a standard and several values of the variable stimulus. This is not known by the subject, who is given only one trial and hence compares just the standard and one value of the variable. If he chooses the variable as being heavier, a "+" is scored at that level of the variable; then a variable stimulus one level lower is shown to the next subject. If he judges the standard to be heavier, an "O" is scored at that level of the variable, and the next subject deals with a variable one level heavier. A data sheet using this method looks something like this:

Levels	Subjects									
1										
2	O			+		O				
3		O		+			O		O	O
4			+				+	+	+	O
5										+

COMPUTATION:

In computing the mean and standard deviation with this method, either

"+" or "O" entries are used depending on which is less frequent. Once this has been determined, a table is set up as follows:

	<u>i</u>	<u>n</u>	<u>in</u>	<u>i²n</u>
	0	-	-	-
	1	-	-	-
	2	-	-	-
	3	-	-	-
	⋮	⋮	⋮	⋮
	<u>K</u>	<u>-</u>	<u>-</u>	<u>-</u>
Total		N	A	B

where $i=0, 1, 2, 3, \dots, k$, with 0 being the value of i given to the lowest level on which the less frequent event occurs, and n is the frequency of the event at a particular level. Then $N=\sum n$, $A=\sum in$, and $B=\sum i^2n$. N is the effective sample size and will always be one-half or fewer of total sample size.

The formula used in computing the mean is:

$$X = y + d\left(\frac{A}{N} \pm \frac{1}{2}\right)$$

where y is the value corresponding to the level on which i is given the value 0, and d is the size of the interval between levels. The plus sign is used when the analysis is based on "O" entries and the minus sign when it is based on "+" entries.

The standard deviation (s) is given by:

$$s = 1.620d \left(\frac{NB - A^2}{N^2} + .029 \right)$$

These formulas are reliable only when d/s is within 0.5 to 2.0. If the interval d is too fine the distribution is spread over too many intervals making the mean unreliable. On the other hand if it is too crude, excessive

lumping occurs with the same effect.

SUBJECTS:

Kindergarten children were from four schools: Morton Elementary School in Marysville, Michigan; Anna Michen Elementary School in Fennville, Michigan; and Wardcliffe and Central Elementary Schools in East Lansing, Michigan. Subjects from the first through sixth grades were obtained from St. Michael's and St. Petronel's Parochial Schools in Wheaton and Glen Ellyn, Illinois, respectively. College-age subjects were students from Michigan State University. Table 1 shows the age ranges and effective sample sizes for each group of students.

Table 1. Age Ranges and Effective Sample Size (N) of Groups

Group	Age Range	N
Kindergarten	6 ⁰ --6 ⁴	92
1st Grade	6 ⁴ --7 ⁵	86
2nd "	7 ³ --8 ¹⁰	94
3rd "	8 ⁴ --10 ⁵	84
4th "	9 ⁵ --10 ⁹	81
5th "	10 ⁴ --11 ⁷	90
6th "	11 ⁰ --13 ²	82
College	18--23	101

METHOD:

A set of variables and a standard were made using common plastic medicinal pill containers cylindrical in shape. The containers used for the variable stimulus series were 5.3 cm. in diameter. The standard was a larger container, 7.4 in height and 3.1 cm. in diameter.

Preliminary examination of the illusion with college students revealed that the optimal difference ($d/s = 1$) was approximately 7 grams when the standard weighed 50 grams.

The variable series began at 10 grams and increased by 7 gram steps up to 66 grams. All containers were filled with copper BB's (with cotton batting to prevent rattling) until the desired weight was attained, then painted cream to prevent the subject from seeing the contents. The caps of the containers were left white.

In testing elementary school subjects, E sat outside the classroom or nearby in an isolated location with the standard and appropriate variable in front of him on a table. The children came in one at a time, and having been told before as a group what they were to do, picked up the two weights and chose the one which felt heavier.

College subjects were tested individually in their rooms and at the university library, and were given the same instructions.

Subjects were given two trials, first with the variable on their right, and second, with the variable on their left. The variable stimuli in the two trials were independent of each other. Since the series of each succession of trials was started one level apart, it was impossible for the variable to

be the same on both trials for any given subject. Results were tabulated separately.

The main reason for giving the second trial was to learn if there would be inordinate differences between the means of the two trials and what direction these differences might take.

The subjects were allowed to handle the stimuli in any way they desired, except they were not permitted to switch the variable and standard to alternate hands.

Because some of the kindergarten children seemed to find the concept of "heavier" nebulous, they were given a test trial after the normal trials. This was done by presenting the child with a variable so much lighter than the standard (10 grams vs. 50 grams) that failing to give a correct discrimination in this case could only mean that the child was guessing or was not judging on the basis of weight. When this occurred, the subject's judgment was not included in the data. Failures did not exceed 1 in 20. Since a guess has a 50-50 chance of being correct, there should be some guessers in the kindergarten sample. Their inclusion with children actually making a real comparison should tend to increase the standard deviation, but not bias the mean.

CHAPTER III

RESULTS

The application of the Up-and-Down method in this study was, in general, quite effective. Virtually no procedural problems were encountered. From the standpoint of actual data collecting, the method leaves little to be desired. For the reasons given earlier (contaminating factors), only the Trial I data test the hypothesis; they are presented first.

TRIAL I:

Table 2 gives the mean and standard deviation for males and females of each age level.

Table 2. Mean, Standard Deviation, and Effective Sample Size for Males and Females, Trial I.

Group	Males			Females		
	$\bar{X}$	s	N	$\bar{X}$	s	N
Kindergarten	36.25	12.35	47	35.28	12.54	45
Grade 1	30.93	9.18	49	31.09	6.84	37
Grade 2	31.29	13.54	48	27.04	5.95	46
Grade 3	28.79	8.05	49	26.70	6.66	35
Grade 4	29.01	10.83	37	28.14	9.00	44
Grade 5	27.83	11.10	42	30.85	8.83	48
Grade 6	29.73	7.62	47	27.50	8.33	35
College	27.36	5.54	50	27.50	7.13	51

Sex differences were tested by comparing the mean for males with the mean for females within each group. T-tests revealed no significant differences and the data were subsequently pooled to obtain Table 3. Figure 1 represents the size of illusion in percentages for the different age groups.

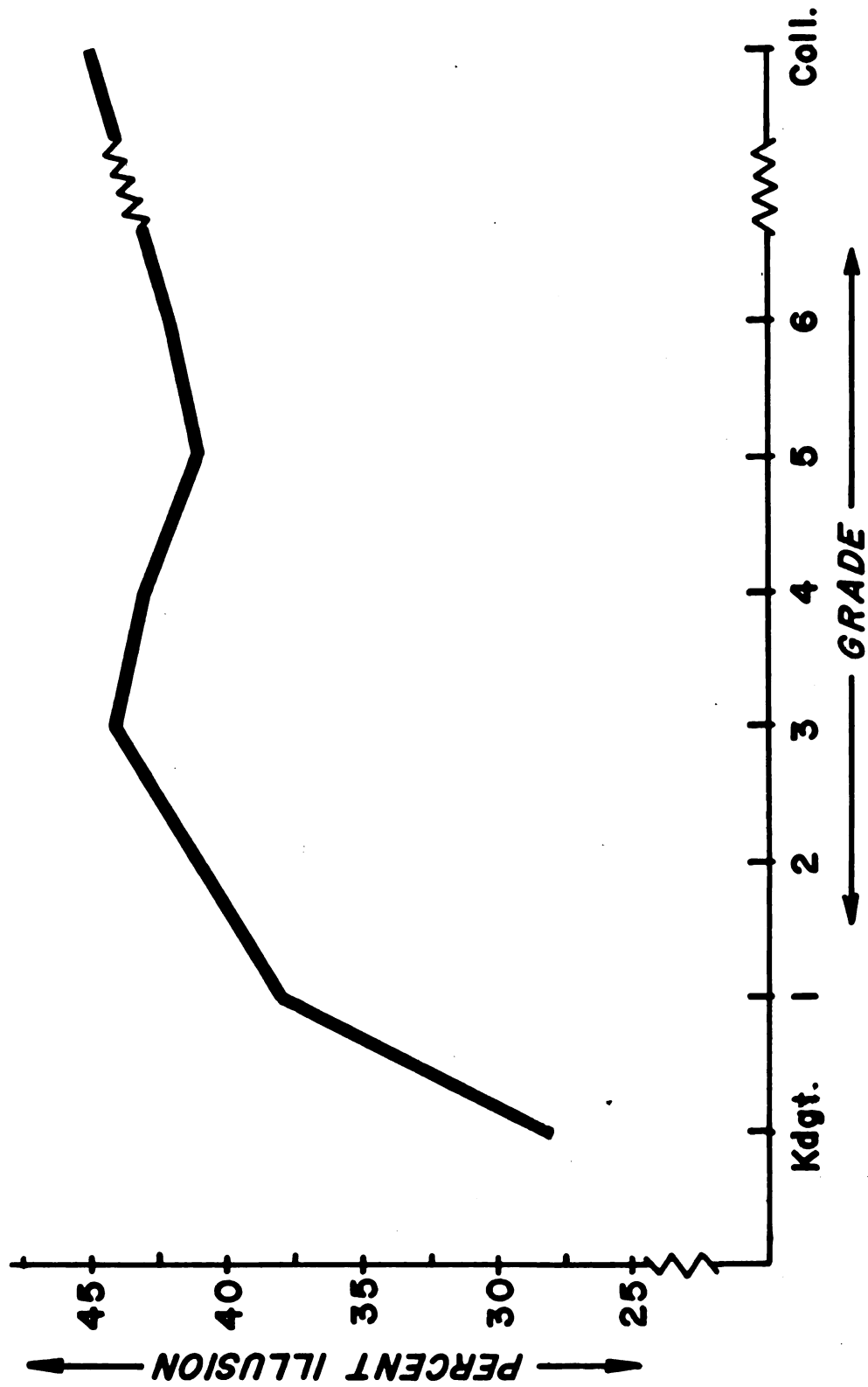
Table 3. Mean, Standard Deviation, and Effective Sample Size of Males and Females Pooled, Trial I.

Group	$\bar{X}$	s	N
Kindergarten	35.95	12.54	92
Grade 1	31.00	8.18	86
Grade 2	29.21	10.86	94
Grade 3	27.92	7.71	84
Grade 4	28.54	9.88	81
Grade 5	29.44	10.42	90
Grade 6	28.78	8.11	81
College	27.50	6.38	101

In Rey's study, children were divided into two groups: ages 5-6 and 7-15. It was Rey's hypothesis that young children (5-6) would not exhibit as much illusion as older ones (7-15). He gave no reason for choosing 6 years as the dividing point between groups, or for assuming that the illusion differential would not continue through the age of 15.

With the data of the present study, it was possible to determine if Rey was correct in dividing the children this way, or assuming that the children

FIGURE 1
THE SIZE-WEIGHT ILLUSION IN PERCENTAGES



older than six experience the same degree of illusion. T-tests between the means in the 1st-6th grade categories (ages 6-4 to 13-2) show only one difference to be significant at the .05 level (Grade 1 vs. Grade 3). With 15 tests of significance, the chance that at least one test will be significant at the .05 level is, of course, much greater than .05. It is therefore possible to assume that these means do not differ significantly and can be pooled to give the material in Table 4. This was done by counting all "+" and "O" entries in the 1st-6th grade group, determining the least frequent of the two, and computing the mean and standard deviation by the procedure described previously.

Table 4. Mean, Standard Deviation,
and Effective Sample Size with
1st-6th Grade Pooled, Trial I.

Group	$\bar{X}$	s	n
Kindergarten	35.95	12.54	92
1-6	29.18	9.42	516
College	27.50	6.38	101

An analysis of these results show the following differences: Kindergarten vs. College, significant at the .001 level. Kindergarten vs. 1st-6th grade pooled, significant at the .01 level. 1st-6th grade vs. College was not significant. These data corroborate Rey's findings that young children exhibit less illusion than older ones, and support his hypothesis.

Table 5 gives the mean and standard deviation for boys and girls on trial II. As in trial I, no statistically significant differences were obtained, thereby permitting the data to be pooled as in Table 6.

Table 5. Mean, Standard Deviation, and Effective Sample Size for Males and Females, Trial II.

Group	Males			Females		
	$\bar{X}$	s	n	$\bar{X}$	s	n
Kindergarten	31.16	20.17	44	29.60	10.42	50
Grade 1	31.07	9.18	51	31.78	11.22	36
Grade 2	29.83	7.10	48	33.31	8.69	47
Grade 3	32.60	9.18	48	27.91	8.96	34
Grade 4	33.92	6.86	36	32.83	14.27	42
Grade 5	30.83	6.94	42	31.15	8.83	48
Grade 6	29.73	7.62	47	31.41	13.13	34
College	29.60	8.15	50	30.16	6.63	50

A cursory comparison between Table 5 and Table 3 shows that with the possible exception of the kindergarten group, there are only minor differences between Trial I and Trial II. The discrepancy in the kindergarten group on the two trials may derive from some of the factors mentioned before, such as loss of attention, practice, etc. Also, since many had not yet formed a stable conception of "heavy," it may be that the second trial confused them somewhat. This could account in part for the sharp increase in the standard deviation.

Table 6. Mean, Standard Deviation, and Effective Sample Size of Males and Females Pooled, Trial II.

Group	$\bar{X}$	s	n
Kindergarten	30.33*	15.12*	94
Grade 1	31.36	9.91	87
Grade 2	31.55	8.58	95
Grade 3	30.66	10.33	82
Grade 4	33.33	10.92	78
Grade 5	31.00	7.95	90
Grade 6	32.54	11.19	81
College	29.86	7.35	100

*d/s < 0.5

Outside of comparison of sexes, no statistical analysis was attempted for Trial II. While it is possible that the kindergarten means on Trial I and II differ, the standard deviation of the second trial was too high (d/s < .5) for reliable computation. Then, too, because of the confounding nature of the variables in Trial II, any interpretation would necessarily be hedged with reservations.

CHAPTER IV

DISCUSSION

There are two limiting features of the Up-and-Down method which should be discussed. First is the fact that one-half of the subjects are discarded in computation. This means with this method a large number of subjects must be readily available, otherwise obtaining subjects becomes so tiresome that the method's other advantages are outweighed. The kindergarten sample in the present study was the only one which presented difficulties in this respect. Here it was necessary to draw from four schools in different parts of Michigan to obtain an adequate sample.

The d/s ratio imposes the other limitation. If this ratio does not lie between 0.5 and 2.0, computation becomes much more involved. Ordinarily a careful selection of d will control this, but in developmental studies, s could vary so much between age levels that no single value of d would keep d/s in bounds. Fortunately the exceeding of these limits occurred only twice in this study, both times on Trial II which was not central in the study.

In the present study, s is largest in the kindergarten sample. The standard deviation can vary because of the nature of young children. First of all, there were some kindergarten children who did not understand exactly what was meant by "heavier." While these children failed the screening test (10 grams vs. 50 grams), there were no doubt others included who had

passed it only by chance. The inclusion of "guessers" does not affect the mean, but does increase the standard deviation. Another explanation for increased variances with younger children is the possibility that some children, while understanding what "heavier" means, do not have as refined a weighing sense as do older subjects. Thus the interval chosen between testing levels may not be great enough to assure a satisfactory measurement of true individual differences in young subjects. What it amounts to is this: E is using an interval which is optimal for some subjects but unsatisfactory for others. The unsatisfactoriness is then manifested by an apparently high standard deviation in those age ranges where the difference threshold is largest.

A comparison between these data and Rey's shows close correspondence. If the illusion is measured in percentages, with the objective difference between the two weights when they are judged equal being divided by the weight of the heavier, we find the following:

	<u>Rey's Data</u>	<u>Present Data</u>
	% illusion	% illusion
7-15 year olds	49%	42%
5-6 year olds	30%	28%

Rey's data should show a higher percentage of illusion, because he increased the weight of the subjectively lighter cube until the illusion was reversed, while the present study gives a mean which represents subjective equality.

The fact that the magnitude of the illusion increases with age suggests the origin lies in a learning explanation. While it is possible that the illusion is due to maturation, this is not as clear-cut or heuristic as an explanation

based on learning. For one thing, it is hard to see how the size-weight illusion could be built into the subject to remain latent until adequate maturation has been reached. Also, innate capacities ought to be useful from an evolutionary standpoint and this human trait has little functional value.

On the other hand, a learning explanation could easily be hypothesized in terms of muscle response and "set" derived from experience. In spite of the credibility of this latter explanation, maturational effects cannot be totally ruled out due to the possibility of the two factors interacting. The only thing we know is that if the data had revealed no differences between age levels, the learning theory would be negated.

Interestingly enough, only statistically insignificant sex differences were found in this study. If maturation were a key factor in this illusion, we would expect that at least in the kindergarten sample girls would experience more illusion since they mature faster.

TRIAL I VS. TRIAL II

There is actually little which can be said about Trial II. It was included in the experiment only because it was easy to give and might have shown some very significant differences from Trial I. Although the data do show a decrease in illusion from Trial I in every group except kindergarten, there are too many possible factors to state anything confidently (such as practice effects, loss of attention, etc.). The discrepancy between Trial I and Trial II with the kindergarteners is also marred by a standard deviation on Trial II which exceeds the maximum limit ($d/s < 0.5$).

CHAPTER V

SUMMARY

The Up-and-Down method was used to measure the size-weight illusion with the intent of eliminating constant errors found in other psychophysical methods. Approximately 600 children from kindergarten through sixth grade and 200 college students were tested, the method proving satisfactory in all respects.

Tests of statistical significance revealed that kindergarten children exhibit less illusion than older children and college students, with the latter two types not differing from each other. The difference in means was significant at the .001 level. This finding confirms earlier work that employed quite different measurement technique. There were no significant differences between the sexes.

A second trial on the illusion, with standard and variable switched to opposite hands (from the first) gave results which were similar to the first trial, except that such great variability occurred with kindergarteners that no analysis of their data could be made.

The amount of illusion in the present study was approximately 28 per cent for the kindergarten group and 43 per cent for the other samples. These values are close to those found in the other study.

BIBLIOGRAPHY

- Wohlwill, J. F. Developmental Studies of Perception. Psychol. Bull. 57, 1960, 249-288.
- Dixon, W. J., and Mood, A. M. "A method for obtaining and analyzing sensitivity data," J. Amer. Statist. Assoc., 1948, 43, 109-126.
- Rey, A. Contribution à l'étude de l'illusion de poids chez les anormaux Arch. de Psycho, 1930, 22, 285-297.
- Woodworth, R. S., and Schlosberg, H. Experimental Psychology. New York: Henry Holt & Co., 1954, p. 227.

ROOM USE ONLY

ROOM USE ONLY



MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03046 3061