VOL. XVII. No. 4. July, 1910.

THE PSYCHOLOGICAL REVIEW.

THE METHOD OF CONSTANT STIMULI AND ITS GENERALIZATIONS.[1]

BY F. M. URBAN,

University of Pennsylvania.

We introduced in several previous publications[2] the notion of the probability of a judgment, which is the fundamental notion in the analysis of psychophysical measurement methods. The judgments of a subject who compares two stimuli under well-defined and constant conditions, have the formal and material character of those chance events which are spoken of in the calculus of probabilities. These probabilities are *a priori* entirely unknown and must be determined by observation. Experience shows that if a standard stimulus of given intensity is compared under constant conditions with stimuli of varying intensities, the probabilities of the different judgments vary in a certain way with the intensity of the comparison stimulus. Let the subject be required to express his judgment in either one of the terms smaller, equal or greater. The probability of the judgment greater increases and that of the judgment smaller decreases with increasing intensity of the comparison stimulus, whereas the probability of an equality judgment first increases and then decreases after having reached a certain maximum.

[1] An abstract of this paper, but without the tables, was presented at the spring meeting of the experimental psychologists at Harvard in 1908. A full presentation of this topic with all the necessary demonstrations was given in the treatise on 'Die psychophysischen Massmethoden als Grundlagen empirischer Messungen' in Vols. 15 and 16 of the *Archiv f. d. ges. Psychologie.*

[2] *The Application of Statistical Methods to the Problems of Psychophysics*, Philadelphia, 1908; 'On the Method of Just Perceptible Differences,' PSYCHOLOGICAL REVIEW, 1907, Vol. 14, pp. 244-253; 'Die psychophysischen Massmethoden als Grundlagen empirischer Messungen,' *Archiv f. d. ges. Psychologie*, Vols. 15 and 16; also in the report on 'Die Psychologie in Amerika,' *Archiv f. d. ges. Psychologie*, 1908, Vol. 11, 'Literaturbericht,' pp. 141-143.

This suggests the view that the probabilities of these judgments are functions (in the mathematical sense of the term) of the intensity of the comparison stimulus. A mathematical expression which gives the probability of a judgment as function of the comparison stimulus, is called the psychometric function of this judgment. These expressions are, of course, different for different intensities of the standard stimulus. It is not necessary but perhaps advisable to insist on the fact that such an expression refers only to well-defined experimental conditions, because it is not possible to assign definite values to these probabilities unless one refers to definite physical and psychophysical conditions under which the observations are made. The term psychometric function was chosen in imitation of the term biometric function, which is commonly in use for mathematical expressions which give the so-called probability of dying as function of age. If the psychometric functions of all the judgments are known, one is able to predict the outcome of any set of experiments.

For the practical application of this notion not only the form of these functions but also the values of all their constants must be known or data must be at hand by which they can be determined. Our choice of the expressions which may represent the probabilities of the different judgments as functions of the intensity of the comparison stimulus must be guided by the following consideration: Experience shows that the comparison of a standard with much greater intensities results always, or nearly always, in the judgment greater, whereas the judgment smaller is given exclusively, or almost exclusively, on the comparison with very small intensities. None of the functions, furthermore, must assume values greater than one and smaller than zero, because they represent mathematical probabilities. From this it follows that the psychometric functions of the greater and of the smaller judgments approach the values zero and one asymptotically, the psychometric function of the smaller judgments decreasing and that of the greater judgments increasing with growing intensities of the comparison stimulus. The psychometric function of the equality cases has a certain maximum, on both sides of which it falls off steadily and approaches the value zero asymptotically. Add to this the condition that for any given intensity of the comparison stimulus the sum of all the psychometric functions must be equal to one. These conditions, however, are not sufficient to determine the nature of these functions and one must form a hypothesis about them, which may be based either on experience or on theoretical considerations. The number of psychometric functions in any set of experiments is equal to the number of judgments admitted. There are always two functions, the psychometric functions of which are similar to those of the smaller and greater judgments; judgments of this type are called extreme judgments. Judgments the psychometric function of which is similar in its course to that of the equality judgments may be called middle judgments. The number of middle judgments is always uneven; until now one has not found it necessary to go beyond the number of three middle judgments.

There is no special difficulty in devising algebraic expressions which may serve as hypotheses on the psychometric functions. A criterion of the value of different hypotheses can consist only in the greater or smaller agreement with experience, which is measured by the sum of the squares of the deviations of the observed from the calculated values. A mathematical expression, which may serve as a hypothesis on these functions, depends on a number of constants which must be determined by observation. These observations consist of empirical determinations of unknown probabilities and are, as such, subjected to errors, which prevent us from determining the exact values of the constants. This must cause differences between the true and the calculated values even if our hypothesis on the nature of the psychometric functions is correct. We will speak in this case of a lack of agreement between theory and observation on account of errors of observation. But it also is possible that our hypothesis on the psychometric functions is incorrect and in this case there must be a lack of agreement between theory and observation, even if all the constants were absolutely exact. Errors which are due to an incorrect hypothesis on the psychometric functions may be called errors of theory. It is not easy to decide whether in a certain case we have to deal with errors of observation or with an error of theory, because, errors of observation being inevitable, errors of theory are always intermingled with them. Usually, however, this problem will be put in this way: that it is required to decide which one of a certain group of functions is best suited to represent a given experimental material. The errors of observation in this case remain the same for the different hypotheses and the sums of the squares of the deviations of the calculated from the observed values indicate the greater or smaller agreement of a hypothesis with experience.

One may try to get along without making a definitive hypothesis on the nature of the psychometric functions and, starting from the theorems that every function may be represented by a power series, one may determine from the data of observation as many coefficients as possible. If n comparison stimuli were used in the experiments one may determine n constants, by which all the terms up to that of degree $(n - 1)$ are determined. This representation of the data of observation were absolutely exact, if the psychometric functions could be represented by an algebraic equation of degree $(n - 1)$, but even if this is not the case this method is distinguished in so far as it requires the smallest amount of theoretical additions to make a mathematical treatment of the data possible. The degree of the equations which represent the psychometric functions depends on the number of observations and is different for different sets of experiments, unless it happens that the same number of comparison stimuli had been used. In this sense one may say that the representation of the data by means of an algebraic equation does not involve a definitive hypothesis on the psychometric functions at all. It is easy to see, however, that this method can not possibly give a definitive result. The psychometric functions represent mathematical probabilities and are, as such, restricted to the interval from zero to one, whereas an algebraic function exceeds every limit for sufficiently large values of the independent variable. Lagrange's formula of interpolation and Newton's method of differences are the most convenient ways of treating the data according to this hypothesis and of determining new values of the psychometric functions without actually setting them up. We will illustrate this method by working out a set of results which has served for the illustration of the method of just perceptible differences.[1]

TABLE I.

RELATIVE FREQUENCIES OF THE JUDGMENTS GREATER, SMALLER AND EQUAL.

Comparison Stimulus.	Equal	Greater.	Smaller.
84	0.0444	0.0222	0.9333
88	0.1133	0.0244	0.8622
92	0.1889	0.1111	0.7000
96	0.2578	0.2933	0.4489
100	0 2400	0.5289	0.2311
104	0.0889	0.8156	0.0956
108	0.0800	0.9044	0.0156

Let us suppose we made observations with n comparison stimuli which we call $x_1, x_2, \cdots x_n$ and that these comparison stimuli gave to a certain judgment the probabilities $a_1, a_2, \cdots a_n$. Lagrange's formula then has the form

$$F(x) = \frac{(x - x_2)(x - x_3)\cdots(x - x_n)}{(x_1 - x_2)(x_1 - x_3)\cdots(x_1 - x_n)} a_1 + \frac{(x-x_1)(x-x_3)\cdots(x-x_n)}{(x_2-x_1)(x_2-x_3)\cdots(x_2-x_n)} a_2 + \cdots + \frac{(x-x_1)(x-x_2)\cdots(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)\cdots(x_n-x_{n-1})} a_n.$$

The actual setting up of the equation is laborious but not necessary for interpolating new values of the function. It is a matter of course that one must arrange the computation systematically whenever one has to treat an extended experimental material. The circumstance that the sum of all psychometric functions for any intensity of the comparison stimulus must be equal to one, shortens the work considerably. If three judgments were admitted it is sufficient to calculate by Lagrange's formula the probabilities for two judgments only, that of the third being found by subtracting the sum of the other

[1] PSYCHOLOGICAL REVIEW, Vol. 14, p. 249, 251. These results are taken from subject II, in the experiments on lifted weights, cf., 'The Application of Statistical Methods to the Problems of Psychophysics,' pp. 178, 217, and *Arch. f. d. ges. Psychologie*, Vol. 15, p. 287.

two from the unit. This is the way in which the interpolated values in Table II. were found. From the data of such a table a diagram may be constructed which shows graphically how the probabilities of the judgments vary with the intensity of the comparison stimulus. Such a diagram is shown in Chart 1. The intensities of the comparison stimulus are represented on the abscissa and the corresponding values of the psychometric functions on the ordinate. It is necessary to draw the two axes on different scales; in our drawing the unit of measurement of the y-axis is ten times as large as that of the x-axis.

TABLE II.

COMPARE CHART 1.

Comparison Stimulus.	Equal.	Smaller.	Greater.
84	0.0444	0.9333	0.0222
85	0.0614	0.9101	0.0285
86	0.0786	0.8956	0.0258
87	0.0959	0.8814	0.0227
88	0.1133	0.8623	0.0244
89	0.1310	0.8353	0.0337
90	0.1497	0.7987	0.0516
91	0.1690	0.7533	0.0777
92	0.1889	0.7000	0.1111
93	0.2088	0.6407	0.1505
94	0.2279	0.5775	0.1946
95	0.2447	0.5129	0.2424
96	0.2578	0.4489	0.2933
97	0.2654	0.3875	0.3471
98	0.2659	0.3301	0.4040
99	0.2578	0.2779	0.4643
100	0.2400	0.2311	0.5289
101	0.2125	0.1900	0.5975
102	0.1761	0.1541	0.6698
103	0.1334	0.1230	0.7436
104	0.0889	0.0955	0.8156
105	0.0495	0.0710	0.8795
106	0.0251	0.0494	0.9255
107	0.0294	0.0303	0.9403
108	0.0800	0.0156	0.9044

There are two difficulties in treating the data by Lagrange's formula of interpolation. It happens in many cases that values are obtained which are greater than one or smaller than zero. The occurrence of these impossible values can have symptomatic value only. Their presence is easily seen in a graphic representation of the functions, because in these places the curves rise above a line drawn parallel to the x-axis at the distance unity, or fall beneath the abscissa. The second difficulty is that the course of the functions is irregular near the beginning and the end of the tables, the psychometric functions of the greater judgments not increasing, and those of the smaller judgments not decreasing throughout the whole interval.

In spite of these irregularities, however, there is in the data of all the subjects an interval, which lies in the central parts of the tables, inside of which the psychometric functions behave regularly and are not interrupted in their course by any irregularities. The intensities for which the psychometric functions of the extreme judgments assume the value 1/2 are found inside of these intervals. Between these two stimuli lie all the intensities which do not give a probability equal to or exceeding the value one half to either one of the extreme judgments. For this reason it is called the interval of uncertainty. This interval is determined by the method of just perceptible differences and the comparison of the accuracy of sense perception of different subjects, or of the same subject at different times or under different conditions is based on it. It is easy to determine this quantity by interpolation. In determining the lower limit of the interval of uncertainty one picks out the two stimuli which gave to the judgment smaller probabilities just above and just below one half and interpolates new values in this interval by Lagrange's formula, until the quantity to be determined is included in an interval small enough to admit of an interpolation on a straight line. Owing to the regularity of the course of the psychometric functions in the middle of the table only few interpolations by Lagrange's formula are needed for this calculation. We give here a table of the results of the determination

Subject	Lower Limit of Interval of Uncertainty.	Upper Limit of Interval of Uncertainty.	Length of Interval of Uncertainty.
I.	93.26	100.95	7.69
II.	95.20	99.55	4.33
III.	98.65	100.32	1.57
IV.	95.24	98.26	3.02
V.	93.75	95.83	2.08
VI.	95.82	101.04	5.22
VII.	95.33	100.74	5.44

of the interval of uncertainty by Lagrange's formula for all the persons who served as subjects in our experiments on lifted weights.

The interval of uncertainty is different for different subjects. If we order our subjects by the length of this interval every subject will have a definite place in the series, except when two or more individuals happen to have intervals of uncertainty of the same length. In this latter case the subjects are equally sensitive.

We now turn to the study of the psychometric functions of the equality cases. It is interesting to determine the intensity of the comparison stimulus for which the probability of an equality judgment is greatest. This value can be found easily by means of Table II. We pick out the three greatest values of the table; these values, which we call A, B, C, may correspond to the intensities x_{R-1}, x_R, x_{R+1}, of the comparison stimulus. The maximum of the probability of the equality judgments is found at the point $x_R + \xi$ where the quantity ξ is given by the expression

$$\xi = \frac{A + C}{2\,(A + C - 2\,B)}\,.$$

The maximum probabilities of the equality judgments are found for our seven subjects at the intensities 98.61, 97.57, 100.51, 97.34, 95.89, 99.23 and 96.80. The maximum probability is attained for intensities which are greater than the upper limit of the interval of uncertainty in two cases (subjects III. and V.), from which we conclude that the situation of this maximum has not a definite relation to the interval of uncertainty. It may be remarked that a similar conclusion may be drawn in respect to the arithmetic mean of the equality cases, which represents the abscissa of the center of gravity of the area included between the curve representing the psychometric function of the equality cases and the abscissa.

It is of greater consequence to consider the maximum values of the psychometric functions of the equality judgments. These quantities are found by introducing into Lagrange's formula the values for which the maximum is attained. The results of this calculation for our seven subjects are 0.4860, 0.2667, 0.1508, 0.2111, 0.1969, 0.3771, 0.3797. These numbers show that the probabilities of the equality judgments are not great; in no case do they exceed one-half and only in the case of subject I does this probability come anywhere near this value, while it never exceeds 0.38 for anyone of the other subjects. This shows that we cannot speak of a point of equality in any absolute sense of the word, because we have to require that such a point must give to the equality judgments probabilities greater than one-half.

Let us compare the orders which we obtain when the subjects are ordered by the length of their intervals of uncertainty and by the maximum probabilities which they give to equality judgments. We obtain the following arrangements of our seven subjects.

Subjects Ordered by Maximum Probability of Equality Judgments.	Subjects Ordered by Length of Interval of Uncertainty.
I.	I.
VII.	VII.
VI.	VI.
II.	II.
IV.	IV.
V.	V.
III.	III.

The order of all the subjects is the same in both cases, and we conclude from this fact that the maximum probability of the equality judgments can be used just as well as a measure of the accuracy of sense perception as the interval of uncertainty. It is, therefore, possible to base a comparison of the accuracy of sense perception on the equality cases alone, a fact which may be surprising if one remembers that these judgments are so much of a difficulty in the customary treatment of psychophysical measurement methods that some investigators advised the absolute suppression of these judgments.

The fact that there exist two quantities each one of which may be used for the purpose of comparing the accuracy of sense perception is interesting, because these quantities are independent of one another and are derived from different data. The limits of the interval of uncertainty depend directly only on the probabilities of one of the extreme judgments, on those of the equality judgments and of the other extreme judgment only indirectly through the sum of the probabilities which must be equal to one. The probabilities of the equality judgments have, therefore, no direct influence at all on the length of the interval of uncertainty. The maximum probability of the equality cases, on the other hand, is entirely independent from the probabilities of the extreme judgments, and the methods of calculation, which lead to the determination of the maximum probability of the equality judgments and of the interval of uncertainty, have nothing in common. We will conclude from this fact that there exist certain relations between the different psychometric functions, which must be investigated.

It is not possible to base the definition of the point of subjective equality on the equality judgments. There remains the possibility of basing this definition on the extreme judgments by defining the point of subjective equality as that intensity of the comparison stimulus for which the probabilities of the extreme judgments are equal. If r, s, t are the probabilities of the smaller, equal and greater judgments, this definition of the point of subjective equality implies that $r = t$; no specification is made in regard to the values of either one of these quantities. The approximate determination of this point from a table of the psychometric functions offers no difficulty. We find in the table the stimulus x_1 which gives a higher probability to the smaller than to the greater judgments and which is immediately followed by the stimulus x_2 which gives a higher probability to the greater than to the smaller judgments. Let the corresponding probabilities of the smaller judgments be y_1, y_2 and those of the greater judgments z_1, z_2. If the interval between the points x_1 and x_2 is small enough to admit of an interpolation on a straight line the abscissa of the point of intersection of the curves representing the psychometric functions of the extreme judgments is given by the formula

$$x - x_1 = \frac{(x_1 - y_1)(x_2 - x_1)}{y_2 - y_1 + z_1 - z_2}.$$

We give here the results of this calculation for the seven subjects of our experiments on lifted weights with reference to the limits of the interval of uncertainty.

Subject.	Lower Limit of Interval of Uncertainty.	Point of Subjective Equality.	Upper Limit of Interval of Uncertainty.
I.	93.26	97.20	100.95
II.	95.20	97.36	99.55
III.	98.65	99.46	100.32
IV.	95.24	96.62	98.26
V.	93.75	94.66	95.83
VI.	95.82	98.42	101.04
VII.	95.33	98.59	100.74

These numbers show that the point of subjective equality defined in terms of the probabilities of the extreme judgments lies always inside of the interval of uncertainty and a closer inspection reveals the fact that it coincides very closely with the centre of this interval. Let us divide the distance of the point of subjective equality from the lower limit of the interval of uncertainty by the length of this interval. We obtain the following values for seven subjects: 0.49, 0.50, 0.51, 0.54, 0.56, 0.50 and 0.40, the average of which is 0.50. We conclude from this, that the point of subjective equality coincides with the center of the interval of uncertainty.[1] The difference between the point of subjective equality and the standard gives the constant error which in the case of our experiments is due to the order in which the weights were presented to the subject, *i. e.*, to the so-called time error. We find the following values

[1] This proposition is the justification of all the investigations in which equal appearing stimuli are found by the method of just perceptible differences, as, *e. g.*, in Miss Cook's investigation on the illusion of filled and unfilled tactual spaces. Only the just imperceptible positive and negative differences were determined in her experiments (the just perceptible differences were used only in some series) and the averages of these values were taken. This average determines the center of the interval of uncertainty, *i. e.*, the point of subjective equality which of course has to be determined in a study of an illusion. Miss Cook, however, did not use the method of just perceptible differences in its traditional form but in Sanford's variation, so that one can not make the positive statement that she really determined the stimulus which gives equal probabilities to the extreme judgments. It is not likely that the influence of her variation on the final result is great, but it is inconvenient to have to deal with results which escape an exact interpretation. Cf. Helen Dodd Cook, 'Die taktile Schätzung von ausgefüllten und leeren Strecken,' *Archiv f. d. ges. Psychologie*, 1910, Vol. 16, pp. 451-456.

for this quantity 2.80, 2.65, 0.54, 3.38, 5.34, 1.58 and 1.41. All these differences being negative we clearly have to deal with an over-estimation of the second weight.

An extremely important conclusion may be drawn from this proposition. The similarity of the method of just perceptible differences with the procedure by which we determine unknown empirical quantities by measurement was noticed by several investigators, although one could not agree as to the exact point of similarity. An accurate understanding of the method of just perceptible differences is of the greatest importance for the general theory of empirical measurements. In determining the unknown weight of a body, for example, we start after a preliminary rough determination of the approximate value by comparing the body with known weights which are too great and gradually reduce the difference until no difference is noticed between the two weights. We keep on diminishing the comparison weight until a point is found where the comparison weight is found to be too small. The average of these results is taken as a determination of the unknown weight of the body. The whole process is repeated in the ascending direction, which gives another determination. Several such determinations have to be made for the purpose of obtaining an exact determination and the final result is found by the arithmetic mean of all the individual results.

The fact that we pick out the arithmetic mean as the final determination of the quantity to be measured implies that we regard it as the best value obtainable. This is the meaning of the proposition that the arithmetic mean has to be regarded as the most probable value of a set of measurements of an empirical quantity. Taking this proposition as a principle one may deduce from it the method of least squares, which is a set of rules for finding the most probable values of empirical quantities and the limits of the exactitude of their determination, if the principle of the arithmetic mean is granted. This is the way of reasoning followed by Gauss in his first deduction of the method of least squares. The theory of errors of observation based on this principle has stood the test of practice for more than a hundred years, and it may be granted that the principle of the arithmetic mean seems to be a very obvious one, but it can not be denied that it makes the impression of being artificial. A great number of attempts were made to demonstrate it, but no proof could be given without introducing some other proposition which is equivalent to the principle of the arithmetic mean. As it was known that other definitions of the most probable value of a set of measurements of an empirical quantity lead to entirely different rules of calculation, one began to suspect that the theory of errors of observation required the introduction of a principle of empirical origin, just as much as the use of ordinary geometry in geodesy and astronomy requires the proposition that empirical space is a three-dimensional Euclidean manifoldness.

It is only a step in the same direction of development to prefer an empirical verification of the law of distribution, which follows from the principle of the arithmetic mean, to any *a priori* deduction, because such a demonstration must necessarily start from some other hypothesis. This is the view expressed by H. E. Faye and H. Laurent. For those who maintained this view it became necessary to show that the distribution of the individual results in a set of measurements is such as to warrant the application of the principle of the arithmetic mean. This required observations on actual distributions of errors, which were given in the works of Bessel, C. S. Peirce, Guarducci, Laurent, Helmert, F. Y. Edgeworth and others. In the investigations referred to a satisfactory agreement between theory and practice was observed, but this position became seriously endangered, not to say untenable, by the discovery that most empirical distributions show an essential asymmetry, symmetry being found as an exception only in very rare cases. One might suspect that the symmetry observed in sets of measurements is merely a chance result or the effect of some peculiar conditions, as it very likely is the case of Peirce's observations.[1]

[1] C. S. Peirce, 'On the Theory of Errors of Observation,' Report of the U. S. Coast and Geodetic Survey, 1870, made a very extended series of observations on reaction time using the Hipp chronoscope. His results, curiously enough, show a symmetrical distribution, although it is one of the best established facts that the distribution of reaction times is asymmetrical. It is of course out of question that Peirce did not observe or report correctly and it is all the more

But granting even that there is no flaw in any of these observations, one must ask how it comes that the errors of observations show a symmetrical distribution while all the other empirical distributions are essentially and normally asymmetrical. Neither an empirical nor a mathematical justification of the principle of the arithmetic mean is at hand and the method of least squares seems to hang in the air. We are confronted with the shocking situation that a proposition is triumphantly borne out by an immense indirect experience and that it can be proved neither by mathematical deduction nor by direct experience.

The only case where there does not exist a doubt as to the justification of the principle of the arithmetic mean is that of the empirical determination of unknown probabilities, because the arithmetic mean is the most probable value of the quantity to be determined according to the theorems of the calculus of probabilities. This fact has to be utilized for the theory of errors of observation. A single determination of an empirical quantity is obtained by following a strict rule which determines which intensity has to be put down as the result of an individual measurement. On the basis of any such definition one can set up an expression which has the character of a mathematical expectation and it follows that the most probable value of a set of individual determinations is given by their arithmetic mean. The arithmetic mean, therefore, is the most probable value in all those measurements in which a systematic procedure is strictly followed.

In those cases where the individual determinations were obtained by following the procedure described above we can go a step further. The first result of such a determination is what we call the just imperceptible positive difference, and the point where the balance indicates a difference between the two bodies is the just perceptible negative difference. In the ascending series we determine first the just imperceptible negative, and then the just perceptible positive difference. Since

necessary to explain his result, because Pizzetti, 'I fondamenti matematici per la critica dei risultati sperimentali,' 1892, attributes great weight to Peirce's observations. The symmetry of the distribution in Peirce's results is due to the mixture of different distributions, as will be shown in a paper to be published in the near future.

the final average of the set is not influenced by grouping them and taking the arithmetic mean of the averages of the groups we may combine first the just perceptible and the just imperceptible positive, and then the just perceptible and the just imperceptible negative differences, obtaining in the first case the threshold in the direction of increase and in the second the threshold in the direction of decrease. The arithmetic mean of these two quantities is identical with the average of all the individual results, from which it follows that the final determination of the value of an empirical quantity coincides with the center of the interval of uncertainty. We conclude that we determine by our empirical measurements those intensities for which the probability of the judgment greater is equal to that of the judgment smaller.

We thus obtain the remarkable result that the foundations of the theory of errors of observation are found in the theory of psychophysical measurement. The principle of the arithmetical mean as the most probable value of a set of empirical measurements since more than a hundred years proved refractory to all attempts at a purely mathematical demonstration and empirical demonstrations lack finality because they do not show the cause of the symmetry of this distribution in the face of an indefinite number of asymmetrical distributions. The cause of this failure is to be sought in the notion of the probability of an error of certain size, which is the basis of the theory of errors of observations. The process by which we arrive at assigning a definite value to an empirical quantity is very complicated and requires further analysis. This analysis, however, can be given only by means of the notion of the probability of a judgment, which is entirely foreign to the theory of observations, because this science considers only the result of the process of measuring. The notion of the probability of a judgment belongs to psychology as well as the analysis of the conditions which influence it. It was assumed in most, not to say in all the treatises on the theory of psychophysical measurement that this science has to depend on the theory of errors of observations, which furnishes the data on which we have to build. This is not the case. The relation of these two sciences is just the opposite. Psychology furnishes the notion of the probability of a judgment and thus opens the way to an understanding of the principle of the arithmetical mean, and it also furnishes the notion of the accuracy of sense perception. The theory of errors of observation reciprocates by offering problems and refined observations, which properly belong to the psychology of sense perception.

We now turn to the study of the treatment of the psychometric functions by means of definitive hypotheses. Let $f(x)$, $g(x)$, $h(x)$ be the functions which represent the smaller, equal and greater judgments. Each one of these functions also depends on a number of parameters which must be determined from the data of observation, so that we may write explicitly $f(x;\ a_1, b_1, c_1, \cdots)$, $g(x;\ a_2, b_2, c_2, \cdots)$, $h(x;\ a_3, b_3, c_3, \cdots)$. Generally the number of observations is greater than the number of constants to be determined, so that they must be determined from an overdetermined set of equations. Owing to the fact that empirical determinations of unknown probabilities are not exact, certain discrepancies between the different results will arise which must be eliminated by an adjustment. The procedure to be used for this purpose will become clear by the following considerations. The individual observations consist of empirical determinations of certain unknown probabilities, the probable errors of which are given by the theorem of Bernoulli. Let s_R experiments be made with the comparison stimulus x_R, m_R of which may have resulted in the judgment smaller, o_R in the judgment equal and the rest $n_R = s_R - (m_R + o_R)$ in the judgment greater. The fractions m_R/s_R, o_R/s_R, n_R/s_R are the most probable determinations of the underlying probabilities and the probability that these results are affected by an error of the size γ is given by the expression

$$\frac{h}{\sqrt{\pi}} e^{-h^2\gamma^2},$$

where h, the coefficient of precision, is given for the three probabilities by

$$\sqrt{\frac{s_R^3}{2m_R(s_R - m_R)}}, \quad \sqrt{\frac{s_R^3}{2o_R(s_R - o_R)}}, \quad \sqrt{\frac{s_R^3}{2n_R(s_R - n_R)}}$$

respectively. The results of the experiments with every comparison stimulus, therefore, give three equations of the form

$$f(x_R;\ a_1, b_1, c_1, \cdots) = \frac{m_R}{s_R},$$

$$g(x_R;\ a_2, b_2, c_2, \cdots) = \frac{o_R}{s_R},$$

$$h(x_R;\ a_3, b_3, c_3, \cdots) = \frac{n_R}{s_R},$$

which must satisfy the condition equation

$$f(x) + g(x) + h(x) = 1.$$

One might believe that a condition equation must be entered for every observation, but this is not the case, because the sum of all the psychometric functions must be equal to one for any value of the comparison stimulus, from which it follows that this sum must be identically one. This simplifies the calculation considerably, because one of the psychometric functions may be determined as the difference of one minus the sum of the two other functions, so that the adjustment of the observations need be carried through only for two of them.

Let us suppose that $g(x)$ is determined by $f(x)$ and $h(x)$. We then have only two observations for every stimulus, for which the probabilities of the different judgments were observed. This gives a system of equations for the determination of the quantities $a_1, b_1, c_1, \cdots$ which occur only in the equations originating from $f(x)$, and $a_3, b_3, c_3, \cdots$ occurring only in the equation originating from $h(x)$. The whole system may be solved by treating these two groups separately, and since the constitution of both groups is the same, we may confine the theory to showing how one of them may be solved. Let us take the group which contains the constant of $f(x)$ and refer to them by the letters $a, b, c, \cdots$, omitting the indices.

We introduce here for the sake of simplifying the calculation the assumption that the constants $a, b, c, \cdots$ occur in $f(x)$ in such a way as to form a linear complex

$$\alpha_R a + \beta_R b + \cdots$$

where the quantities $\alpha_R, \beta_R, \cdots$ depend on the intensities of the comparison stimulus and are, therefore, different for different intensities. We also assume that there exists an inverse function $F(x)$ so that

$$\alpha_R a + \beta_R b + \cdots = F\left(\frac{m_R}{s_R}\right).$$

We obtain under these conditions an overdetermined system of linear equations the errors in the determinations of the constants of which follow the exponential law. From this it follows that the most probable values of the unknown quantities must be found by the method of least squares, each equation being put down with the proper weight. Only the fractions $p_R = m_R/s_R$ are directly determined, their coefficient of precision h_R being given by the theorem of Bernoulli. The coefficient of precision in the determination of $F(p_R)$ may be called H_R and is found by the formula

$$\frac{1}{H_R^2} = \frac{\left(\frac{dF}{dp}\right)_{p=p_R}}{h_R^2}.$$

Since the weight of every observation equation is directly proportional to the square of the coefficient of precision in the determinations of $F(p)$, we have all the quantities which we need for the solution of our system which has this form

$$\alpha_1 a + \beta_1 b + \cdots = F_{(p_1)} \text{ with the weight } P_1,$$
$$\alpha_2 a + \beta_2 b + \cdots = F_{(p_2)} \quad \text{" \quad " \quad "} \quad P_2,$$
$$\cdots\cdots\cdots\cdots$$
$$\alpha_l a + \beta_l b + \cdots = F_{(p_l)} \quad \text{" \quad " \quad "} \quad P_l.$$

From these equations the normal equations are derived in the usual way and their solution gives the most probable values of the constants of the function $f(x)$. Introducing these values in the observation equations one obtains the deviations of the calculated from the observed results and the sum of the squares of these deviations permits to state the accuracy obtained in the determination of the constants of the psychometric functions.

The method of calculation explained here is based on two suppositions which must not be overlooked. The first supposition says that the nature of the dependence between the intensity of the comparison stimulus and the probabilities of the different judgments is known, whereas the second supposition specifies this assumption still further by admitting the existence of the function $F(p)$ with all its qualities. This assumption is not necessary, because one can solve the problem without it, but it is extremely convenient and facilitates the calculation considerably. A further argument in favor of this restriction of the problem is the fact that the existence of the function $F(p)$ was tacitly assumed in all previous investigations. No objection can be raised against this supposition, unless one proposes to solve the general problem, an undertaking which is more laborious than difficult.

The question as to the justification of the first hypothesis offers an entirely different aspect. A hypothesis on the psychometric functions is valid only when it expresses the actual dependence of the probabilities of the different judgments on the intensity of the comparison stimulus. The nature of this dependence is not known and cannot possibly be deduced by any considerations *a priori*. It is quite obvious that a function cannot be the object of any immediate experience, but that it must be found by regularities in the results of observation. No such knowledge is at hand at present, so that one hypothesis about the psychometric functions is just as arbitrary as any other. This is the peculiar difficulty of this problem, that one is forced either to make a hypothesis which is not more justified than any other or that one must treat the results by Lagrange's (or some analogous) formula of which one knows positively that it cannot be correct and which takes in errors which are absolutely unknown as to their size and sign.

It seems to be best to test the experimental material at hand by different hypotheses and not to regard any one of them as final, no matter what its success may have been. What we really need is a standard by which we can judge the different hypotheses, because if we have one, we can discard certain mathematical expressions as unsuitable to represent the psychometric functions. The treatment of the same data by different hypotheses furnishes valuable indications. If the constants of the psychometric functions are known we can calculate the probabilities of the judgments according to the different hypotheses and compare them with the results of the observations. The sum of the squares of these deviations is the criterion of the value of a hypothesis. It is not possible to arrive at a final conclusion in this way, because an infinity of hypotheses would have to be gone through, but it may be decided which one of a certain set of mathematical expressions is best suited to represent the psychometric functions. The number of functions which for practical purposes come under consideration as hypotheses on the dependence of the probabilities of the judgments on the intensities of the comparison stimulus is naturally rather limited. The process of calculating the constants of the psychometric functions is rather laborious and it becomes the more so the more complicated the hypotheses are. For this reason one will not be inclined to take up the study of very complicated functions, unless there are strong arguments in favor of them. It may be expected that among these functions there is one which is more suitable as a hypothesis about the psychometric functions than all the others. This hope is as well compatible with the view that the form of the psychometric functions is the same for all subjects, as with the more conservative view that these functions differ for different individuals and perhaps even for the same individual at different times.

We may regard a hypothesis on the psychometric functions merely as a mathematical expression which fits one set of experimental data well, and another, perhaps, less satisfactorily, but we consider it from the start as extremely unlikely that an expression can be found which fits all data equally well. One may support this view by referring to the individual differences between the different subjects, which seem to exclude any such regularity. If, contrary to expectation, such an expression is found we will not have to change our views materially and benefit by this discovery for the facilitation of future work.

The problem which confronts us in the study of the psychometric functions is similar to the problem of determining the probability of dying as a function of age. A mathematical expression which gives this dependence is called the biometric function. These functions are *a priori* just as unknown as the psychometric functions are and the same difficulties are encountered in their *a posteriori* determination, but experience shows that there exists a formula, the so-called formula of Gomperz-Makeham, which has done better service than any other formula tried for this purpose. One therefore expects a similarly satisfactory result for the future and one naturally turns to this formula if new material is to be treated. The modern view about the biometric functions is similar to the one which we gave for the psychometric functions, namely that it is impossible to find an expression which fits all data equally well, a view which is not only supported by past experience, but which also may be backed up by the argument that the conditions under which men live are so different that the existence of any such regularity seems very unlikely. Experience must show whether there is less difference in the psychological make-up of people, but meanwhile we may undertake to find out how different hypotheses on the psychometric functions work out in their application to the results of observation.

We will consider here two hypotheses on the psychometric functions, in which the probabilities of the equality cases are expressed in terms of those of the extreme judgments. The psychometric function of the smaller judgments may be represented by the expression

$$f(x) = \frac{1}{\sqrt{\pi}} \int_{h_1(x-a_1)}^{\infty} e^{-t^2} dt$$

This expression is admissible as a hypothesis on the psychometric function of the smaller judgments, because it decreases with increasing intensity of the comparison stimulus and approaches the limit 1 for $x = -\infty$ and the limit 0 for $x = \infty$. One easily sees that the expression

$$h(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{h_2(x-a_2)} e^{-t^2} dt$$

is admissible as a hypothesis on the psychometric function of of the greater judgments. The probabilities of the equality judgments are given by

$$g(x) = \frac{1}{\sqrt{\pi}} \int_{h_1(x-a_1)}^{h_2(x-a_2)} e^{-t^2} dt.$$

The three functions $f(x)$, $g(x)$, $h(x)$ contain only the constants a_1, a_2, h_1 and h_2 and they are fully determined by them. These constants must be determined in such a way as to fit the data from which they are deduced as well as possible.

This hypothesis may be called the $\Phi(\gamma)$-hypothesis. It is remarkable and well known for the fact that G. E. Mueller uses it in his method of constant stimuli. Mueller starts from the notion of a threshold which is subjected to chance variations, the frequency of which is a function of their size. The mathematical expression for the probabilities of the variations as depending on their size is called their law of distribution. Mueller and his followers assume the exponential law, which frequently but not very appropriately is called the Gaussian law, to hold good for the distribution of the threshold. The ordinate of the maximum of this function is an axis of symmetry, which gave rise to the well-known discussion whether it was admissible to make the assumption that the variations of the threshold follow a symmetrical law of distribution. This objection was strengthened by the fact that all empirical distributions studied until now show an essential asymmetry which is sometimes small but sometimes very considerable indeed. It is not possible to say that the discussion of this problem was very fruitful of important results.

The question as to the symmetry of the law of distribution has the following meaning for the psychometric functions. $f(x)$ assumes the value $\frac{1}{2}$ for $x = a$. Keeping in mind that the probability integral from zero to any positive limit is equal to that from zero to the same negative limit, we see that $f(a-x)$ and $f(a+x)$ are symmetric to the value $1/2$. The curve representing $f(x)$ may be divided into two parts each one of which goes over into the other by being mirrored at the lines $y = \frac{1}{2}$ and $x = a$, the order of this process being indifferent. It is quite obvious that this implies a very special hypothesis on the psychometric functions, but any other hypothesis has to be equally specific, and if one is thoroughly imbued with the conviction that nothing is definitely decided by a provisory acceptance of a hypothesis, one will not attribute too much importance to this question.

It is necessary for the practical application of this method to have a table of the values of this function. One either may construct a table similar to the well-known fundamental table for the method of right and wrong cases or one may use a table of the probability integral. Fechner's table is very convenient for working out sets of 25, 50 or 100 experiments, but in all the other cases it is more convenient to use a table of the proba-

TABLE III.

p	γ	Difference.	p	γ	Difference.
0.50	0.0000	177	0.76	0 4994	230
0.51	0.0177	178	0.77	0.5224	236
0.52	0.0355	177	0.78	0 5460	242
0.53	0.0532	178	0.79	0.5702	249
0.54	0.0710	178	0.80	0.5951	257
0.55	0.0888*	179	0.81	0.6208	265
0.56	0.1067*	180	0.82	0.6473	274
0.57	0.1247	180	0.83	0.6747	284
0.58	0.1427*	182	0.84	0.7031*	298
0.59	0.1609	183	0.85	0.7329	310
0.60	0.1792*	183	0.86	0.7639	326
0.61	0.1975	185	0.87	0.7965	347
0.62	0.2160	186	0 88	0 8308	365
0.63	0 2346*	189	0.89	0.8673	389
0.64	0.2535	190	0 90	0.9062	418
0.65	0.2725	191	0.91	0.9480*	455
0.66	0.2916*	195	0 92	0.9935*	500
0.67	0.3111	196	0 93	1.0435*	558
0.68	0.3307	199	0.94	1.0993*	637
0.69	0.3506	202	0.95	1.1630*	750
0 70	0.3708	205	0.96	1.2380*	920
0.71	0.3913	208	0.97	1.3300*	1220
0.72	0.4121	212	0.98	1.4520*	1930
0.73	0.4333	216	0.99	1.6450	
0.74	0.4549	220	1.00	∞	
0.75	0.4769	225			

bility integral. The tables of Kaempfe[1] and of Bruns[2] are easily accessible to psychologists and they have the great advantage that the interval of the table is very small. The values of the psychometric function calculated from the table of Bruns do not always coincide with the data of the table of Fechner, which is reprinted in most of the treatises on psychophysical measure-

[1] B. Kaempfe, *Psychologische Studien*, 1893, Vol. 9.

[2] H. Bruns, *Wahrscheinlichkeitsrechnung und Kollektivmasslehre*, 1906.

ment methods, and it seems advisable to give here a table of the values calculated from the table of Bruns. The arrangement of this table is identical with that of the fundamental table for the method of right and wrong cases; values marked by an asterisk (*) differ from the corresponding values in Fechner's table. Owing to the symmetry of the function it is sufficient to give the values between 0.50 and 1.00.

The next step consists in finding the formula for the weights of the observation equations. Following the line of argumentation given above we find for the weight of the observation with a comparison stimulus

$$P = \frac{1}{2\pi} \cdot \frac{se^{-2\gamma^2}}{p(1-p)}$$

where s is the number of experiments made. This formula corresponds to the expression which G. E. Mueller gave for the weight of an observation equation, but it differs from it in such a way as to give results which are greater than the corresponding values of Mueller. An analysis of this formula shows (1) that the function representing the weight of the observation equations has a maximum at $p = 1/2$ and that the ordinate of this maximum is an axis of symmetry of the function, and (2) that the function assumes the value zero for $p = 1$ and for $p = 0$, showing that observations which gave the frequency 0 or 1 for one of the extreme judgments are without influence on the determination of the constants of the psychometric function of this judgment, from which it follows that these results simply may be omitted. It is best to use comparison stimuli which give probabilities not differing much from $1/2$, because these observations come down with the greatest weight. It is convenient to have a table of these weights, which owing to the symmetry of the function representing them need cover only one of the intervals from 0 to 0.5 or from 0.5 to 1. Table IV. contains these values for the interval from 0.5 to 1; the arrangement and the use of this table is identical with that of Mueller's table.

The process of setting up and solving the normal equations derived from the observation equations is well known and need

Table IV.

Weights of the Observation Equations According to the $\Phi(\gamma)$-Hypothesis.

p	P	Difference.	p	P	Difference.
0.50	1.000	0	0.76	0 832	14
0.51	1.000	1	0.77	0.818	15
0.52	0.999	1	0.78	0.803	16
0.53	0.998	2	0.79	0.787	17
0.54	0.996	1	0.80	0.770	18
0.55	0.995	3	0.81	0.752	19
0.56	0.992	3	0.82	0.733	20
0.57	0.989	4	0.83	0.713	19
0.58	0.985	4	0.84	0.694	24
0.59	0.981	4	0.85	0.670	24
0.60	0.977	5	0.86	0.646	25
0.61	0.972	5	0.87	0.621	26
0.62	0.967	7	0.88	0 595	28
0.63	0.960	6	0.89	0.567	29
0.64	0.954	7	0.90	0.538	32
0.65	0.947	7	0.91	0.506	34
0.66	0.940	8	0 92	0 472	37
0.67	0.932	9	0.93	0.435	39
0.68	0.923	9	0.94	0.396	44
0.69	0.914	10	0.95	0.352	48
0.70	0.904	10	0.96	0.304	55
0.71	0.894	11	0.97	0.249	62
0.72	0.883	12	0.98	0.187	78
0.73	0.871	12	0.99	0.112	112
0.74	0.859	13	1.00	0.000	
0.75	0.846	14			

not be described here, but a few remarks may prove useful. The opinion is very widespread that this process is difficult or at least very laborious. This is not the case if the computations are arranged properly. With all the necessary checks the calculations ought not to take more than two hours even if one has only little practice in this work. A scheme of calculation which has stood the test of practical application is given in the *Archiv f. d. ges. Psychologie*, Vol. 16.

We give here the results of the computation for the seven subjects in our experiments on lifted weights. The first two columns of Tables V. and VI. give the constants of the psychometric functions and the columns under the headings S_1 and S_2 give the lower and upper limit of the interval of uncertainty. These values may be compared with the results of calculation and observation by the method of just perceptible differences and by Lagrange's formula of interpolation. The results of

TABLE V.

CONSTANTS OF THE PSYCHOMETRIC FUNCTION OF THE SMALLER JUDGMENTS BY THE $\Phi(\gamma)$-HYPOTHESIS. LOWER LIMIT OF THE INTERVAL OF UNCERTAINTY.

Subject.	h_1	c_1	S_1	Method of Just Perceptible Differences.		Interpolated by *Lagrange's* Formula.
				Calculated.	Observed.	
I.	0.125777	11.7398	93.34	93.49	93.30	93.26
II.	0.105252	10.0074	95.08	94.98	94.87	95.20
III.	0.138920	13.6010	97.91	97.88	97.85	98.65
IV.	0.123357	11.7730	95.44	95.56	95 39	95.24
V.	0.127380	12.0011	94.22	94.57	94.47	93.75
VI.	0.110215	10.5023	95.29	95.20	95.31	95.82
VII.	0.114453	10.9650	95.80	96.74	95.79	95.33

these different methods agree so well that no further discussion is needed.

It is a notorious fact that in former investigations the method of constant stimuli did not give the same results as the method of just perceptible differences. This was due to an imperfect understanding of the method of just perceptible differences and one may safely say that the difficulties of the method of constant stimuli were largely due to those of the method of just

TABLE VI.

CONSTANTS OF THE PSYCHOMETRIC FUNCTION OF THE GREATER JUDGMENTS BY THE $\Phi(\gamma)$-HYPOTHESIS. UPPER LIMIT OF THE INTERVAL OF UNCERTAINTY.

Subject.	h_2	c_2	S_2	Method of Just Perceptible Differences.		Interpolated by *Lagrange's* Formula.
				Calculated.	Observed	
I.	0.136113	13.5676	99.68	99 60	99.45	100.95
II.	0.110945	11.0138	99.27	98.71	98.83	99.55
III.	0.145240	14.4346	99.39	99.58	99.28	100.32
IV.	0.117995	11.5917	98.24	98.24	98.08	98.26
V.	0.115708	11.2424	97.16	97.35	97.14	95.83
VI.	0.114995	11.5862	100.75	100.33	99.86	101.04
VII.	0.115465	11.6816	101.17	99.63	99.86	100.74

perceptible differences. As long as one did not see that the result of this latter method could be defined in terms of the probabilities of the different comparison stimuli, one could not possibly use the same material for a test of both methods, and it is very difficult to obtain different sets of results under exactly the same conditions. This is particularly true with reference to experiments made by the method of just perceptible differences in its traditional form and those of the method of constant stimuli, because in the first case one has to present the stimuli in a certain order (ascending or descending), whereas they may be given in random order in the method of constant stimuli. For this reason it is not possible in the method of just perceptible differences to keep the subject in ignorance as to the direction in which the stimuli are varied. This circumstance causes a difference in the attitude of the subject which must influence the judgment, and for this reason it is not very likely that the experimental data obtained by the two methods are strictly comparable.

Differences in the conditions of the experiments may be detected in two ways, by introspection or by a difference in the objective results. We take the view that introspective evidence against a set of experiments makes it suspect, but the absence of any such objection does not put the value of the material beyond doubt. This is the view which is taken in the theory of observations where a set of observations must not be judged off-hand but only on the basis of a minute examination. Differences in the objective results of psychophysical experiments are differences in the values of the observed probabilities and we may call the conditions of two groups of experiments materially different if the judgments have not the same probabilities for the comparison of the same stimuli. Arrangements which are not identical, but which do not interfere with the values of the probabilities of the different judgments, are said to be only formally different. The method of just perceptible difference when looked at as a method of calculation is only formally different from the method of constant stimuli, because, as Tables V. and VI. show, both methods give the same results. The method of just perceptible differences in its traditional form, however, requires results of special kind which are materially different from those obtained and used in the method of constant stimuli. This accounts for the differences thought to exist between these two methods.

The problem of finding the relation between the different psychophysical methods has a well-defined meaning only if one refers to results obtained under conditions which are not materially different. The problem of psychophysics in general is to determine the influence of the conditions of the experiments on our judgment, those conditions being of primary importance which depend on the state of the subject. Various methods have been devised for the purpose of this analysis, the formal character of which has to be perfectly understood before a conclusion can be drawn as to the material difference or identity of the conditions in the different sets of observations. We may express this idea by saying that the purpose of the psychophysical measurement methods is an analysis of the material conditions which determine our judgment on the comparison of stimuli, but for this purpose an understanding of the formal character of these methods is needed.[1]

When the constants of the psychometric functions are known one may calculate the probabilities of the different judgments for all intensities of the comparison stimulus. The results of this calculation are given in Table VII. The numbers under the heading 'observed values' give the difference between the

[1] The distinction between formal and material conditions of an experiment was not favorably criticized in the discussion following the reading of the paper, perhaps because it was not as clearly presented as it might have been. This distinction, however, is absolutely indispensable for a proper understanding of the psychophysical methods. Experience shows that the results of the method of just perceptible differences coincide with those of the method of constant stimuli if the same material is used for this test, and that they are different if different materials are used. There is obviously no reason why the results should agree in one case and not agree in the other, unless there are some similarities between the two methods, which are counteracted by the differences of the conditions under which the material was obtained. In the monograph on 'The Application of Statistical Methods to the Problems of Psychophysics' the emphasis was laid on the formal identity of the method of just perceptible differences with the error methods, because it was a new observation that both methods give the same result if they are tested on the same material. Mr. G. Geiger (*Zeitschrift f. Psychologie*, 1910, Vol. 54, pp. 540–542) in his review of this book and Miss H. D. Cook in her treatise on 'Die taktile Schätzung von ausgefülten und leeren Strecken,' *Archiv f. d. ges. Psychologie*, 1910, Vol. 16, p. 455, justly point out that the differences between the experiments made according to these two methods are not sufficiently dwelt upon. This remark is perfectly to the point and shows the importance of making the distinction between formal and material conditions of an experiment.

Table VII.

Values of the Psychometric Functions.

Compare Chart 2.

Comparison Stimulus	Smaller.		Equal.		Greater.	
	Calculated.	Observed	Calculated	Observed.	Calculated.	Observed.
80	0.9876		0.0011		0.0113	
82	0.9742		0.0030		0.0228	
84	0.9504	—0.0171	0.0075	+0.0147	0.0421	+0.0023
86	0.9117		0.0172		0.0711	
88	0.8540	+0.0082	0.0357	—0.0113	0.1103	+0.0030
90	0.7752		0.0682		0.1566	
92	0.6767	+0.0233	0.1200	—0.0089	0 2033	—0.0144
94	0.5639		0.1946		0.2415	
96	0.4456	+0.0033	0.2920	+0.0013	0.2624	—0.0046
98	0.3319		0.4076		0.2605	
100	0.2320	—0.0009	0.5319	—0.0030	0.2361	+0.0039
102	0.1515		0.6532		0.1953	
104	0.0922	+0.0034	0.7605	+0.0551	0.1473	—0.0584
106	0.0520		0.8465		0.1015	
108	0.0272	—0.0116	0.9091	—0.0047	0.0637	+0.0163
110	0.0132		0.9504		0.0364	
112	0.0059		0.9752		0.0189	

calculated and the observed values of the probabilities, the sign being determined so as to make the arithmetic sum of the terms equal to the observed value. The differences between the calculated and the observed values are very small and some of them are negative and some positive, both signs being distributed quite irregularly throughout the table. The data of this table may be represented graphically, as it is shown in Chart 2. The construction of the curves in this chart is the same as that of Chart 1, so that it need not be explained again. It is seen at a glance that the course of the curves is regular throughout the table and that the ascent and descent of the curves representing the psychometric functions of the extreme judgments is not interrupted by any secondary elevations.

We now turn to the study of another hypothesis on the psychometric functions. One sees immediately that an expression of the form

$$f(x) = \frac{1}{2} - \frac{1}{\pi} \arctan(ax + b)$$

may represent the psychometric function of the smaller judgments, because it approaches the values zero and one asymptotically and decreases for increasing values of the comparison stimulus. For similar reasons is it admissible to suppose that such a function may represent the psychometric function of the greater judgments, if the sign and the constants of the term arctan are determined appropriately. This hypothesis leads to a computation similar to that of the first hypothesis and the necessary formulæ may be easily found by the considerations given above,[1] but for our present purpose only the question is of importance whether this hypothesis agrees better with experience than the $\Phi(\gamma)$-hypothesis.

TABLE VIII.

SUMS OF THE SQUARES OF THE DEVIATIONS OF THE CALCULATED FROM THE OBSERVED VALUES.

Subject.	Arctan-Hypothesis.			Φ(γ)-Hypothesis.		
	Smaller.	Greater.	Equal	Smaller.	Greater.	Equal.
I.	0.019879	0.016931	0.024358	0.001738	0.022826	0.019738
II.	0.010243	0.010667	0 011302	0.001060	0.003492	0.003934
III.	0.019196	0.016839	0.014051	0.008008	0.011209	0 002759
IV.	0.013797	0.020357	0.004427	0.003422	0.004552	0.002023
V.	0.012195	0.036444	0.015876	0.014093	0.011771	0.000743
VI.	0.016717	0.008770	0.017292	0.002878	0.003991	0.003875
VII.	0.022457	0.014181	0.025161	0.002707	0.009569	0.007986
Average	0.016355	0.017741	0.016067	0.004844	0.009630	0.005865

For this purpose it is necessary to calculate the values of the probabilities of the different judgments according to both hypotheses and form the differences between the observed and the calculated values. These deviations are squared and their sums formed. These results are given in Table VIII., which contains the sums of the squares of the deviations for each one of the seven subjects in our experiments on lifted weights. A mere glance at this table shows that the judgment as to the value of the two hypotheses in question cannot be doubtful for a minute. With the only exception of the psychometric function of the smaller judgments for subject V., all the sums of the squares of the deviations calculated by the $\Phi(\gamma)$-hypothesis are smaller than those calculated by the arctan-hypothesis. We,

[1] All the details of the calculation are given in the treatise on 'Die psychophysischen Massmethoden als Grundlagen empirischer Messungen,' *Archiv f. d. ges. Psychologie*, 1909, Vol. 16.

therefore, must say that the first hypothesis fits our results better than the second.

A closer examination of the data of Table VIII. shows that the agreement between the results of calculation and of observation is different not only for different subjects, but also for the different judgments. The averages of these sums for our seven subjects are 0.014767, 0.002829, 0.007325, 0.003332, 0.008869, 0.003581 and 0.006754. It is perhaps worth while noticing that subject II., who had by far the greatest practice in psychological experiments, has the smallest average, whereas subjects III. and V., who were the least reliable, have large averages. From this one might conclude that the psychometric functions approach the $\Phi(\gamma)$-type with increasing practice of the subject, but this conclusion does not agree with the fact that subjects I. and VII. have a large average, although their reliability manifests itself by a small coefficient of divergence. The averages of the sums of the squares of the deviations for the three judgments are given at the bottom of each column in Table VIII. The agreement between theory and observation is best for the smaller judgment, the second place being taken by the equality judgments, the third by the greater judgments. This result agrees very well with the standpoint we have taken before, that the psychometric functions may differ from individual to individual, and that the nature of the dependence of the probabilities on the intensity of the comparison stimulus may be different for the different judgments.