# PSYCHOLOGICAL MEASUREMENTS.

MENTAL *Phenomena and Physical Phenomena.* — Experimental psychology, like every empirical science, starts with the facts of our immediate experience. The innumerable phenomena which we experience directly, *i. e.*, of which we are conscious, fall naturally into various classes according to the point of view we happen to take. One of the divisions is that into sights, sounds, tastes, etc., on the one hand, and impulses, emotions, memories, etc., on the other. Just what distinguishing characteristics we find in each class it is impossible to say. Indeed, the division is not at all complete or distinct ;[1] many experiences that we on one occasion put into one class, *e. g.*, hallucinations which we believe to be realities, we at another time place in the other, *e. g.*, hallucinations recognized as such ;[2] in very many cases we are in doubt as to which class a certain experience belongs. The experiences of the former class we regard as belonging to an objective world, as we call it ; those of the latter, and in some degree those of the former, to a subjective world. The objective class consists, by supposition, of phenomena constant in character for constant objective conditions ; the actual variations in these phenomena under constant objective conditions we ascribe to mental elements. For example, we consider colors to be phenomena of the objective world (I am speaking from a purely introspective standpoint); under the most careful physical conditions we can maintain the color in a condition which we know by secondary means to be constant within a very small range of variation. Yet the color actually experienced will be subject to considerable variations ; these we ascribe to the influence of mental phenomena, *e. g.*, attention, fatigue, etc. Suppose we wish to compare two colors together; we so arrange matters that the two classes of variations, the physical and the psychological, are kept as small

[1] Wundt, *Physiol. Psychol.*, 3. ed. II, 2.

[2] Wundt, Zur Frage der Grosshirnfunktionen, *Phil. Stud.*, 1891, VI, 18.

as possible. If we wish to determine that the two colors are physically alike, we seek out the most favorable method for reducing the psychological sources of variation to a minimum. In the spectral photometer the principle of contrast is employed in quite a complicated and unmeasurable way, but the maximum of sensitiveness is obtained and the variations due to psychological influences are quite negligible. On the other hand, if we desire to determine some mental characteristic in regard to the two colors we must obtain sources of light under such conditions that the measurements of their values are carried out with an accuracy to a degree beyond our own sensitiveness ; we must not use complicated and unmeasurable psychological arrangements but the simplest ones possible. Under such conditions the variations measured will be due to psychological influences. In both physical and psychological experiments the same fundamental principles are used. The difference lies in the sources of variation ; in physics we must eliminate psychological influences, in psychology we must make the physical variations comparatively negligible.

Both physical and psychological measurements are concerned directly with the phenomena of immediate experience. In physics we measure certain objective phenomena of consciousness on one another; in psychology we also measure phenomena of consciousness on one another.[1]

It may seem strange that we should treat the objective phenomena of consciousness as physical phenomena. We are accustomed to think of the physical world as something with a separate existence apart from consciousness. By deductions from the objective phenomena of immediate experience physics has built up a system of independent processes subject to the laws of the conservation of matter and of energy, and expressed in the terms of touch and muscular sensations.[2] The attempt is made to reduce the other phenomena of immediate experience (light, heat, etc.) to these terms, or, as the physicist says, to reduce all physical phenomena to the laws of mechanics. This

[1] Wundt, Ueber die Messung psychischer Vorgänge, *Phil. Stud.* 1883, I, 255.
[2] Schwarz, Das Wahrnehmungsproblem, 1. Theil.

attempt has not been successful; it is not possible to transfer the concepts of that most developed portion of physics to the other domains, each of which requires the formation of its own special concepts.[1] The three units, time, space and mass, are not sufficient for the definition of physical quantities except in mechanics; to these a special unit must be added for each domain of phenomena. We have been so long accustomed to attempt to translate all objective phenomena into terms of touch and the muscle sense, and have gotten into such hopeless difficulty that with such a point of view it is quite intelligible that so many should absolutely deny the possibility of psychological measurements. As soon as with Ostwald we look upon the relations of the various classes of objective phenomena as those of equivalence and not of identity, the close interrelation between psychology and physics becomes comprehensible. The two sciences divide the field of immediately experienced facts. Each phenomenon has an objective or physical side and a subjective or mental side, the two being intimately related and sometimes indistinguishable. We can compare physical, or objective, phenomena directly with mental, or subjective, phenomena. Experimental psychology has in great part to do with such comparisons; in a large part of the work the experiments are psychophysical.

*Rejection of Metaphysics as a Basis.* — We have become so accustomed to certain hypotheses that it is difficult for us to look at matters as they are actually given us. From the study of the objective phenomena we have constructed our physical world, in which we find other beings to whom we are inclined to attribute conscious phenomena like our own. By a series of conclusions we suppose that their nervous systems are most closely connected with such phenomena, and it then becomes an object to determine what conscious phenomena are connected with the activities of the nervous system. Thereupon we turn the matter around again and try to look at our own facts of immediate experience as if they were some one else's. When we make a psychological experiment, instead of comparing the

[1] Ostwald, Studien zur Energetik, *Zt. f. physikalische Chemie*, 1892, IX, 565.

resulting subjective phenomenon directly with the physical facts of immediate experience, we take a stand outside of ourselves and imagine the physical phenomenon to be some external affair conducted along our nervous system to a certain place where it is turned into a mental phenomenon. Now, this might all be very well if we had a satisfactory system of concepts for the physical world; our present representation of physical processes as entirely a matter of mechanics with concepts drawn from the muscle sense is not only unsatisfactory in physics, but leads to utter incomprehensibilities in matters of physiological psychology. We cannot represent brain processes in any way that will bring them into harmony with what we know as our facts of consciousness.

If we consider the physical phenomena as something quite outside of and incommensurate with the facts of our immediate experience, then, in the present state of our knowledge, we can readily agree that mental phenomena cannot be measured.[1] We can suppose that mental facts form a world of their own with which physical facts cannot be compared. But we have here broken up the facts as first given us into two classes, drawn hypotheses from the one set, and are now trying to bring the facts of the other class into harmony with hypotheses with which they will not agree.

There is one philosophical theory which goes a step further than this. So far, at least, we are all agreed as to the existence of a mental world governed by its own laws, but even this latter fact is denied by one class of objectors. This school, represented by numerous English and French writers and lately championed by Professor Muensterberg, would claim that no causal relations exist between mental phenomena, that after analyzing them into their elements we should next determine the brain processes to which they belong, and that the co-existence and sequence of mental phenomena find their explanation only in the relations of the brain processes.[2] Of the very naïve

[1] Zeller, Ueber die Messung psychischer Vorgänge, *Philos. u. hist. Abhandl. d. Berliner Akad.*, 1881, 3. März. Berlin, 1882.

[2] Muensterberg, *Ueber Aufgaben und Methoden der Psychologie.* Leipzig, 1891, p. 117.

metaphysical theory laid as the foundation for a science that should be empirical I do not need to speak, although probably few things have done more injury to our science than the general belief that the new psychology finds its justification for existence in such speculations. We can also readily see why even professed psychologists put forward the plea that there is no use in being exact in their measurements. At present we know absolutely nothing of the nature of the brain processes to which mental phenomena correspond; if the only way to trace any connection between mental phenomena is to first get at the brain processes, it is really very hard to see why we should make any experiments at all.

We shall, however, be obliged to reject these and all other metaphysical theories at the outset, and confine ourselves to the facts. In doing so we take up the development of psychology just where the old psychology left it. The material of the new psychology is exactly the same as that of the old psychology, the facts of immediate experience; the only difference lies in the substitution, wherever possible, of exact records and measurements in place of vague descriptions in general terms. The older attempts at applying the method of introspection led to results as often erroneous as true; with the introduction of experimental methods a trustworthy application of the fundamental method of introspection (or reflection, became for the first time possible.[1]

*Measurement in General.* — The fundamental form of measurement is the expression of the judgment that the quantity measured is equal to or unequal to the standard. The primitive method of weighing articles was by balancing them in the two hands; we measure off a yard of cloth by laying a yard-stick down on it and cutting off enough to be equal to the stick; we run the eye over one line, then over the other to see if they are equal; we measure lights by determining that they are alike or unlike in intensity, tones by judging that they are alike in pitch.

[1] Wundt, *Physiol. Psychol.*, 4. ed., I. 4; *Vorlesungen über Menschen- und Thierseele*, 2. ed., p. 14.

All measurements, physical as well as psychological, consist ultimately in the comparison of two phenomena of consciousness, generally two sensations. All physical measures have been developed out of psychological estimates.[1] We measure temperature by noting the agreement of the length of the column of mercury with a certain portion of the scale; we measure the strength of an electric current by noting the angle through which the mirror is deflected or through which the needle passes, and this very angle is in turn measured by some length; we measure time by the agreement of the angle over which the hand has passed with a mark denoting the end of another angle taken as a standard.

Some confusion has been caused by the statement that psychologically we are able to judge only likenesses or difference, with the implication that in physics we are able to do something more, namely, to measure one object as a fraction or a multiple of another. Even if in physics we could directly do this, it would mean nothing more than that we could do so in psychology also, as when we compare two physical lengths we are only performing a psychological process. When we say that one line is apparently three times as long as another, we simply mean that the two mental pictures bear that relation, or that the series of muscular sensations produced by running the eye over the lines bear that relation. The fact is, however, that in the absence of graduated scales we express one quantity as a multiple of another only by estimates directly from our sensations. The graduated scales, by means of which we always work wherever possible, and by means of which we can obtain the accuracy of modern science, are really only records of direct judgments of likeness or difference. The 0 point on the thermometer means that the mercury column occupied that place under certain definite conditions, namely, immersion in the water of melting ice; the 100 mark means that the column was just so long when the thermometer was surrounded by steam at 760 mm. barometric pressure. We usually divide the intermediate space into 100 parts, but these

[1] Wundt, Die Messung psychischer Vorgänge, *Essays*. Leipzig, 1885, p. 158.

divisions of themselves mean nothing. It is only by placing the thermometer in liquids of the intervening temperatures and directly recording the height of the column at each temperature that we could get a definite graduation. As this latter method is too cumbersome, the marks are made at intervals by the dividing machine, and then the actual value of each mark is determined by sending the mercury up to it and noting the temperature required to do so. Thus each mark on the thermometer means that at some previous occasion of a certain character the mercury column reached to that point; when we now make a measurement of temperature we simply *compare the length of the column at present with the record* of its length at some previous time. The same is true of the galvanometer, the clock, and all apparatus in which the graduation is in units of length; exactly similar processes are used to arrive at other scales.

*Exactness in Measurement.* — Since we always measure physical quantities by means of a psychological judgment as to the agreement of two sensations or sets of sensations, we must so arrange matters that in a given case the psychological judgment introduces only a small uncertainty into the measurement. Since all psychological and physical measurements are made by means of apparatus, the error of the apparatus must be sufficiently small in comparison with the quantity measured. For example, in measuring the time between two successive culminations of the same star, the uncertainty introduced into the results by the variations of our judgments in the eye-and-ear or the graphic method are too small to be of importance for most physical purposes, the length of the sidereal day being determinable to within .05 seconds, or $\frac{5}{10000}$ of 1%. In measuring mental times an outside limit of error of $\frac{1}{1000}$ of a second is beyond the needed accuracy; the length of the time measured seldom is less than $\frac{100}{10000}$ of a second; we can thus allow an outside limit of error of 1%. We can therefore use a fork vibrating 100 times per second, whose accuracy has been determined to within 1%, that is, one whose vibrations during a sidereal day amount to 8,616,400 ± 3,200. The

accuracy required for astronomical purposes is something far beyond that for psychological purposes in this case; yet the very thing which we want to measure psychologically, the reaction time, and which we do measure with an accuracy at present beyond our usual needs, is used to determine the unit of measurement, namely, the second; in the latter case we arrange our experiments so that the variations of the psychological quantity are negligible, in the former the inaccuracy of the physical apparatus is negligible.

It is readily seen that if we do not eliminate or render negligible the psychological sources of variation in physical measurements, as was the case in astronomy before the discovery of the personal equation, we are introducing errors into our physical results. Likewise, if we are measuring psychological phenomena, and yet do not know how much of our results and how much of the variations are due to mental influences, and how much to the apparatus, we really do not know what our results mean. It is from the side of psychologists who are not acquainted with the science of measurements that we often hear the remark that it is of no use to be exact in psychological work. They are careless in their methods, careless in their measurements, and careless in their statements. In the published accounts of the work there is often no information as to the elimination or presence of errors. An experimental result whose reliability is unknown to us is nearly worthless. In order to form a judgment on the accuracy of the result, all the necessary data must be given. Any description of a method and result can be criticised as materially incomplete if it does not give all the data needed for such a judgment. Failure to give such data can only be ascribed to the urgent necessity for condensation or to ignorance or neglect on the part of the observer; and either of the latter casts grave doubt on the character of the work.[1]

*Variations in Measurements.* — If we make $n$ independent measurements of the same quantity, physical or psychological,

[1] Holman, *Discussion of the Precision of Measurements.* New York, 1892, p. 36.

we get $n$ different results, provided we make the unit of measurement fine enough. If by $x$ we denote the variations from the arithmetical average, and if $n$ be infinitely large, then the variations will occur with probabilities according to the well-known law,

$$y = ke^{-h^2x^2}dx,$$

provided we make one of two suppositions: (1) the single variations are made up of small elementary independent variations, which are equally likely to be positive or negative; (2) the most probable value is the arithmetical mean. The former is the supposition of Laplace and Hagen; the latter is that of Gauss.

Neither of these suppositions is allowable in psychological measurements, or in physical ones, either, except as furnishing results sufficiently accurate. That they have justified themselves in physics is due to the facts: (1) that in all physical measurements the surrounding conditions are kept in a high degree of constancy; (2) that in all judgments in regard to the accuracy of physical work we presuppose that there were no sources of error comparable in magnitude with the measure of precision. Under such circumstances the occurrence of the elementary errors (or variations) in groups would have comparatively little effect, and we can suppose them to be independent. In psychology the case is different. We cannot yet get our conditions so completely under control as in physics; the state of affairs somewhat resembles that in statistics. We are not justified in supposing that the variations are independent;[1] on the contrary, from the very large and irregular mean variations that we obtain, from our experience in gradually eliminating sources of error, and from our knowledge of varying circumstances that we cannot eliminate or measure, we know that the variations must occur in groups. The variations will therefore not follow the law of probability, and the arithmetical mean may or may not be the most probable value. A critical treatment of the variations, their signs, their successive differences, and the signs of the

[1] Cattell, On Errors of Observation, *Am. Jour. of Psychol.*, 1893, V. 287.

differences[1] will show whether the measurements follow the law of probability or not. If they do not, we have no recourse except to empirical treatment.

*Empirical Treatment of the Results.* — When the usual treatment of our results is not applicable, we are forced to fall back on empirical methods. Let us take our $n$ measurements, say of reaction-time, and lay off on the axis of abscissas values corresponding to the successive results obtained, *e. g.*, $180^{\sigma}$, $181^{\sigma}$, $182^{\sigma}$, . . . and erect ordinates proportional to the number of times each value occurs. If the variations conformed to the suppositions mentioned above we would get a curve resembling the ordinary probability curve. What we actually do get, is a curve with several maxima instead of one; and the curve can be regarded as made up of several probability curves with different mean values and different degrees of steepness. This shows us that our measurements are running in groups, and that the factors going to influence the results are working in combinations. Our measurements were made under conditions that were not controlled so as to give a well defined result. In the measurements of simple reaction-time a curve with two maxima, say one much more prominent than the other, would show that what we had been measuring as simple reaction-time had not been well defined, that there was one form which had predominated and another form not so prominent. If we take the arithmetical mean of all the results we are averaging two different classes of things together. Exactly the same results are obtained in statistical measurements. The arithmetical mean has been found quite unsatisfactory; if we take the mean height of a community composed of part English and part French, we have a mixture of two groups and will get a curve of results with two maxima.

This indication of the grouping of variations leads to a further analysis of the quantity measured till the variations from the probability curve become small in comparison with the desired or the possible accuracy. When this point is

[1] Weinhold, *Physikalische Massbestimmungen*, I, ch. VII. Berlin, 1886.

reached, we have a definite value for each quantity under given constant conditions, namely, the arithmetical mean, and the average of the variations or the probable error will give an index of the accuracy with which that value has been determined.

*Deductions from Results.* — Suppose we have such a value as just mentioned, *i. e.*, the arithmetical mean, what are the conclusions to be drawn? In the first place we can foretell the average value and the probable variations from that value when the conditions of future measurements are exactly the same as those of the set made, or do not differ to a greater degree than is negligible. In the second place, presuming that the same probability relations exist in another set of measurements, we can be sure of obtaining results within a given limit of variation with a definite degree of probability. In the third place, if we have two sets of measurements we can determine within what limits and with what sureness the probability underlying the one is the same as that underlying the other. The formulas for these deductions have been worked out by Poisson (*Recherches sur la probabilité des jugements*) and have been illustrated by Lexis (*Einleitung in die Theorie der Bevölkerungsstatistik*, Ch. V).

*Applications.* — Nothing has been said in regard to how accurate the measurements are to be before we can apply the principles just mentioned. Nothing should be said except that, whenever measurements of any kind are made, the computation of the results must follow the laws laid down by the science of measurement. Whether the accuracy is to 10% or to $\frac{1}{100}$ of 1% is a matter of indifference for the calculations. The claim put forth by some psychologists that the lack of accuracy in the measurements justifies the presentation and lumping of the results without observance of the rules and without a statement of the characteristic variations, enables them to prove anything they please with their figures. One psychologist not long ago made his measurements in groups of twenty-five and then selected twenty of each group from which to compute the result. Concerning the accuracy of the

method we can know nothing without the calculation of some one of the characteristic variations; yet the same experimenter remarks that it was hardly worth while to calculate the mean variations, from which we can draw only one conclusion, that it was hardly worth while to present the results at all. Another psychologist rises superior to the charge of not possessing the faintest idea of accuracy by declaring that psychological phenomena are not measurable quantities, that our measurements are physical, etc., not knowing that the science of measurements has stringent rules for *all* measurements and not seeing that his plea for carelessness simply denies him the right to make any measurements.

No matter how accurate or inaccurate the measurements may be, the amount of trust to be given to the results will be indicated by a proper treatment of the variations and their differences, that is, so far as chance errors and changing systematic errors may have influenced the work. The sources of constant error must unfortunately be left to the experimenter; it is easily seen how fatal the reputation for carelessness must be. There can be no question that the results obtained by many a poor investigator are actually measurements of some error of apparatus or of method and not of a psychological phenomenon at all. One by one we are getting the psychological conditions under control and reducing the amount of error. That some psychologists choose to declare themselves superior to such slow and careful work and prefer to make startling experiments where little or nothing is known of the method or of the complex mass of phenomena measured, is only too unfortunate.

When measurements are made at all, the experimenter must know just how accurate his apparatus, his methods, and his conditions are to be made and are made. Ignorance of the apparatus, laxity in method, and carelessness in work will be shown in the published records, provided a proper account of the apparatus and methods is given and a proper computation of the results is undertaken. In any case where such data are not given, we cannot accept the results.

*Conclusions.* — 1. Experimental psychology differs from the older introspective psychology only in the accuracy and trustworthiness of its results.

2. All measurements involve both physical and psychological elements; in physical measurements the psychological elements are kept at a minimum and *vice versa*.

3. Measurements may be of all degrees of accuracy, but in each case the degree of accuracy must be known and stated.

4. The lower grade of accuracy in psychological measurements is due to the inability to maintain more constant conditions. This furnishes no excuse for still further lowering the accuracy by careless methods.

5. The inference seems justifiable that the main work in psychology should be directed to the attainment of constant conditions and the simplification of methods.

E. W. SCRIPTURE.