Pseudo-Mathematics in the Mental and Social Sciences
Author(s): H. M. Johnson
Source: *The American Journal of Psychology*, Vol. 48, No. 2 (Apr., 1936), pp. 342-351
Published by: University of Illinois Press
Stable URL: http://www.jstor.org/stable/1415754
Accessed: 24/01/2010 10:24

NOTES AND DISCUSSIONS

PSEUDO-MATHEMATICS IN THE MENTAL AND SOCIAL SCIENCES

The admirers of a certain very famous psychologist often credit him with having said: "If anything exists, it exists in some degree, and therefore can be measured." "Moreover," say these disciples, "it should be measured. No fact of observation can become a scientific fact until it be measured, and its measure expressed by a number. Only quantitative facts are scientific facts, and we intend to be scientific."

So they say; and so they have set out to measure many kinds of *personal traits,* which they call 'intelligence,' 'emotionality,' 'self-reliance,' 'dominance,' 'introversion,' etc., as well as persistent *social habits* of the individual, which they call 'aptitudes,' 'opinions,' 'attitudes,' 'systematic prejudices,' 'fears,' 'superstititions,' etc. The list of these so-called measurables is long, and it becomes longer whenever a pychological journal issues a new number.

Unfortunately, many psychologists dislike to examine and analyze the fact-finding methods which they employ, fearing, perhaps, that if they should do so, then some one might suspect them of being logicians and philosophers. Hence, some of them have published many studies in a form which they believe to be mathematical, and therefore scientific. Many of their colleagues, being lazy, preoccupied in their own work, fearful of authority, or desirous of peace, have allowed this kind of activity to go unprotested until its results became a nuisance. It is now time to examine them.

Given a set of observational data, it is anti-scientific to subject them to illicit mathematical treatment, just as it would be non-scientific not to apply mathematical treatment if the latter is appropriate. But some kinds of psychological, educational and social data lend themselves to mathematical expression, while other kinds do not. Hence we should always begin by determining what kind of data we have. Thus we must apply a set of universal rules, which belong to elementary logic and elementary algebra, and make sure whether the mathematical treatment that is proposed satisfies *all* these rules or not. If it fails *any* of them it is invalid. Fortunately, these rules are very few and also very simple.

If a collection of observational facts is to be treated mathematically, it is *necessary,* though not *sufficient,* that each fact can be perfectly denoted by a number. But three kinds of numbers lend themselves to denoting, and only one kind to counting: hence, before we perform any operations on the fact-denoting numbers we should determine their kind.

Nominal numbers. We use these numbers each to denote a particular member of the collection, and thus as substitutes for ordinary *names,* as of football players, factory employees, prisoners, railroad cars. If two nominal numbers stand in some particular relation to each other, the fact does not imply that the objects which they denote stand in any corresponding relation to each other. Thus, as Cohen and Nagel remark,[1] from the fact that one prisoner is called "500" and another prisoner

* Read at the St. Louis meetings of the A.A.A.S., before Section I, on December 30, 1935.

[1] M. R. Cohen and E. Nagel, *Introduction to Logic and Scientific Method,* 1934, 294 f.

"100" one could not infer that the first prisoner is "five times as dangerous or wicked" as the second prisoner. Neither could one infer that the first prisoner was sentenced to a longer term than the second prisoner, or even that "500" entered the prison later than "100," since if a convict leaves the prison his number can be reassigned to a newcomer.

Sometimes nominal numbers denote sentences. Thus, in the flag-code of a merchant marine the number 373 might denote, "Yellow fever on board; keep off." By convention, an instructor might write on a student's theme the number 95 to denote, "If you hand in many other papers like this one, I shall recommend you for honors;" or the number 55 to denote, "Unless you go to work effectively I shall exclude you from a continuation-course;" or the number 19 to denote, "You ought to be ashamed to hand in this paper for criticism." Certainly this code does not imply that the first paper exhibited 19/11 times as much information as the second and 5 times as much information as the third; any other set of three numbers could be made to denote the same three sentences as these three numbers respectively denote, if only the code were agreed to. Obviously, no meaning attends the result of any operation performed on nominal numbers, such, for example, as taking their mean.

Ordinal numbers. These numbers, such as "the first," "the second," . . . "the *n*th," are assigned to individual objects, not to denote them as individuals, but to denote the *places* which they severally occupy in some *ordered series.* Note that if two objects in the ordered series should exchange places, they would have to exchange ordinal numbers. To arrange the objects in a collection into an ordered series, one first chooses some one property which all the objects have in common. Thus, Mohs chose the property of *hardness,* which is shared by all the members of his collection of minerals. In general, call the common property x. Next, one has to choose some *operation* which defines the relation "x-er than." Thus, to define the relation 'harder than,' Mohs chose the operation of *scratching.* If any mineral A scratches another mineral B, and if also B does not scratch A, then by definition A is 'harder than' B. Thus by the defining operation the relation 'harder than' is *asymmetrical.* Moreover, experiment is required to show that the relation "x-er than" is *transitive: i.e.* if $A > B$ and $B > C$, then $A > C$. (Here the symbol $>$ means "x-er than.") Thus, Mohs showed that if A scratches B and B scratches C, then A scratches C.

If any relation is asymmetrical and also transitive, it permits the objects which it connects to be arranged in an ordered series, otherwise it does not. Thus, Mohs arranged his minerals in the order of their hardnesses, placing diamond, which scratches all the others, at one end, and talc, which is scratched by all the others, at the other end. To diamond he assigned the number 10, to sapphire 9, to topaz 8, . . . and to talc the number 1, assigning to each mineral a number larger than the number of any mineral that it scratches, and smaller than the number of any mineral that scratches it. Thus, carborundum, which scratches sapphire, but which is scratched by diamond, would receive a number greater than 9 but less than 10. These numbers are ordinal numbers: they denote nothing except the *places* in which the several minerals stand in a series that is ordered with respect to the relation *harder than.*

Note that Mohs chose these numbers *arbitrarily.* If he had assigned to diamond the number 3 and to talc the number 1; or if he had assigned to diamond the number 10 and to talc the number 7, he could still have denoted perfectly the place that each mineral occupies in the hardness-series, simply by following the general rule that we mentioned above. Hence it is meaningless to say that diamond is 10 times as hard as talc, or three times as hard as talc, or 10/7 times as hard as talc, or that its hardness bears any other ratio to the hardness of talc. The relation *harder than,* which is identical with *scratches,* is perfectly denoted by *any* of an infinite number of series of ordinal numbers; and operations on ordinal numbers give meaningless results.

Similarly, if we were to subject say 100,000 persons to the so-called National Intelligence-Test, count the conventional answers that each individual returns, and then ascertain that 70% of the population answered fewer of these questions conventionally than John Walker answered conventionally, and that 35% of the same population answered fewer of these questions conventionally than did William Carpenter, then we might assign to Walker the number 70 and to Carpenter the number 35 to denote their respective centile ratings in the population of tested individuals. But then these numbers would be *ordinal* numbers, and it would be absurd to say that because Walker's place in the series is denoted by a number which is twice as great as the number which denotes Carpenter's place, therefore Walker has twice as much National Test Intelligence as Carpenter has, or that the relations between these two ordinal numbers denote any relation between the 'amounts of' National Test Intelligence which these individuals may be thought of as possessing.

In general, if a property can be *adequately* denoted by nominal numbers or by ordinal numbers, then it is non-additive and non-measurable. Among such properties are hardness, shape, structure, generosity, dominance, radicalness, intelligence and the like.

Cardinal numbers. These and only these numbers express the result of *counting,* and they express nothing except that result. In counting the objects in a collection, one treats each object as if it were interchangeable with every other. Thus one disregards their individualities, which nominal numbers denote, and also their several places in the collection, which ordinal numbers denote. The cardinal number of any collection denotes *how many* members it contains; if its members are *units* of some *distributable* property, then its cardinal number denotes *how much* of the property is distributed among the objects that are considered.

Consider the property called *weight* or *heaviness.*[2] Given a suitable scale-balance, we place in one of its pans a pile of sand and in the other pan two bodies B_1 and B_2 successively. If we should find that B_1 causes its pan to sink while B_2 does not, we assert that B_1 is 'heavier than' B_2. If we should find that the same pile of sand exactly balances another body B_3 and also still another body B_4 when B_4 replaces B_3, we assert that B_4 and B_3 are 'equally heavy.' If we should put both B_4 and B_3 into one pan and add sand to the pile in the other pan until it counterbalances the first, we say that the second pile is 'twice as heavy as' the original pile. Thus, we can define a set of standard weights, and by means of this procedure we can say that

[2] Cf. Cohen and Nagel, *op. cit.,* 289-301.

any member of the set bears some specified relation to any other member. Of course, we are limited by the capacity of the instrument and also by its construction. We feel sure that if these limitations could be removed, or if we could replace this instrument by other instruments in certain ranges of weight, or if we could correct for friction, distortion, etc., then these principles would hold perfectly both within and without our range of experience. But this is pure assumption.

These operations suggest that heaviness differs from hardness in being an *additive* property. Nevertheless, we must not be too hasty in assigning numbers to two or more bodies according to their operationally defined weights, and in assuming that the result of adding these numbers expresses the result of placing these bodies in the same pan of the balance. In general, if a property is truly additive, and measurable, it must satisfy all the following criteria.

(1) The relation 'x-er than' must be unequivocally defined by the behavior of the detector of the property x, and in such wise that this operational definition shows the relation to be *asymmetrical*. Given a collection of objects which have the property x, the detector must show that between any two objects in the collection, such as B_i and B_j, one and only one of these relations holds: (a) $B_i > B_j$; (b) $B_i = B_j$; (c) $B_i < B_j$; with respect to x-ness. Thus, if x is weight, we assert (a), (b), or (c) according as B_i counterbalances more sand than B_j, the same sand as B_j, or less sand than B_j counterbalances, when B_i and B_j are interchanged in the detector. Outside a determinable range of uncertainty these judgments are not confused; if this range is greater than is permitted by the degree of precision which we demand, then the operation of counterbalancing on this instrument does not define the relation 'heavier than.' But if the detector permits the unequivocal judgment $B_i > B_j$, then it always yields the judgment $B_i \neq B_j$ and $B_j \not> B_i$, and thence the judgment $B_j < B_i$. Thus the relation 'heavier than' is *asymmetrical* if it is *determinate*.

(2) The defining operations must show that the relation 'x-er than' is *transitive: e.g.* if $B_i > B_j$ and $B_j > B_k$, then $B_i > B_k$.

(3) The collection of objects which have the property x must in its turn have the so-called *group-property*.[3] For example, if two objects B_i and B_j severally excite the detector of the property x, then the collection must contain another member B_k, such that $B_k > B_i$ and $B_k > B_j$ and $B_k = B_i + B_j$ in the sense that B_k excites the detector of the property x in exactly the same manner as B_i and B_j together excite it. In other words, if the effects of B_i and B_j on the detector are additive, then the detector must show an effect which is equivalent to their sum. This implies that the number of objects which have the property is *infinite*.[4]

(4) The addition in the detector of bodies that have the property x must satisfy the *commutative law* of the addition of numbers. Thus, if $B_i + B_j = B_k$, then $B_j + B_i = B_k$.

[3] Cf. C. J. Keyser, *Mathematical Philosophy,* 1922, 202.

[4] What about adding *negative quantities* of the same property? Strictly speaking, no such quantities exist. A surface, for example, cannot carry fewer units of positive charge than none. When we speak as if it does, we mean that we must add some positive charges to the surface in order to bring the detector into a certain equilibrium state. We interpret this as implying that the surface carries more negative charges than positive charges; all that we detect is the excess of the one kind of charge over the other. We cannot detect the absolute quantities of either kind of charge. The same principle applies to any other instance that we may select.

(5) The operation of empirical addition must satisfy the *associative law* of addition of numbers. Thus, $(B_i + B_j) + B_k = B_i + (B_j + B_k)$, etc.

(6) The operation of empirical addition must satisfy the *axiom of equals* which holds for the addition of numbers. Thus, if $B_i = B_i'$ and $B_j = B_j'$, then $B_i + B_j = B_i' + B_j'$ and $B_i + B_j' = B_i' + B_j$.

In particular, if the property x has an antagonist $\bar{x}$, which we agree to call the negative of x, then the operation of empirical addition must satisfy two additional criteria; namely:

(7) The collection of objects, ordered with respect to the relation 'x-er than,' must contain a member B_o which does not excite the detector of the property x, and which, if placed in the detector along with another object B_i will always yield the result $B_o + B_i = B_i + B_o = B_i$. Hence B_o is called the *neutral member* of the collection.

(8) Finally, if the collection contains an object B_i which excites the detector of the property x in one manner, it must contain a corresponding object B_{-i}, which excites the detector in the opposite manner, and such that if B_i and B_{-i} are placed together in the detector, they do not excite it. Since the sum of their several effects on the detector is imperceptible, as is also the effect of B_o upon it, we say that $B_i + B_{-i} = B_o$.

If a property is measurable it satisfies all these criteria; if it fails any of them it is non-additive and non-measurable. In other words, the addition of 'quantities' of the property x has to satisfy all the axioms of addition of cardinal numbers: otherwise it cannot be expressed by the result of numerical addition.

It is at least doubtful whether any observable property exists which *satisfies* all these axioms. For example, the class of all weights that are detectible by means of the scale-balance does not have the group-property *(3)*; and its members do not, in general, obey the commutative law of addition *(4)*; both criteria remain unsatisfied if the balance is overloaded. Nevertheless, there is a range of weights, which we can determine empirically, and within which we can *pretend* that scale-balance weight is additive, and prove that we have not so introduced an uncertainty which exceeds some specified standard of tolerance. Now some mental and social data are of this kind: they *approximate* the criteria of measurability within a given standard, although they do not *satisfy* the criteria. But other kinds of mental and social data fail them utterly. We shall mention examples of both kinds.

Are perceptible brightnesses measurable? Perceptible brightnesses are not identical with what the physicist calls the brightnesses or luminosities of surfaces. To determine the physical brightness of a surface, one first ascertains the wavelengths of the radiation which the surface emits, transmits, or reflects toward the detector. Next, with respect to each minute range of wavelengths, one determines (a) its surface-rate of power, and (b) its so-called 'visibility factor,' taken from a standard wavelength-luminosity curve. Taking the product of (a) and (b) for each minute range of wavelengths, one then summates the products. These operations define the physical brightness of the surface. It is imperfectly correlated with perceptible brightness within certain limits, but it is not identical with the latter.

Consider a surface illuminated by two sources S_1 and S_2 in succession. When S_1 is used alone, the observer perceives a brightness of the surface which we may call B_1; when S_2 is used alone, he perceives a brightness B_2 on the same surface. Using the flicker-method of photometry, or else the method of direct comparison, let us balance B_1 against a comparison-field B_1', and also balance B_2 against another comparison-field B_2'. Now, expose the surface to both sources S_1 and S_2 at once. If we agree that in so doing, we add B_1 and B_2, then by the axiom of equals *(6)*, $B_1 + B_2 = B_1' + B_2'$, and $B_1 + B_2' = B_1' + B_2$. But this is not generally true. Suppose, for example, that the sources which produced B_1 and B_2 respectively, emitted only lithium light ($\lambda = 671$ mμ), while the sources that produced B_1' and B_2', respectively, emitted only thallium light ($\lambda = 535$ mμ). Suppose moreover that $B_1 = B_1'$ is high, while $B_2 = B_2'$ is low. Then $B_1 + B_2'$ is the sum of a bright red and a dim olive-green, while $B_1' + B_2$ is the sum of a bright olive green and a dim red. Although observation yields the separate equations $B_1 = B_1'$, $B_2 = B_2'$, it is very likely to yield $B_1 + B_2' \neq B_1' + B_2$. It may also yield $B_1 + B_2 \neq B_1' + B_2'$. The operations may not satisfy or even approximate the axiom of equals. Some have suggested that heterochromatic brightnesses *would be* additive *if* it were not for Purkinje's effect, and they may be right: anything probably would be different from what it is if only it did not remain the same. But since Purkinje's effect characterizes the brightnesses that we perceive, anything that lacked it would not be a perceptible brightness.

But are perceptible brightnesses measurable if they are *homochromatic?* No, their addition is not *commutative (4)*. Suppose that direct comparison yields $B_1 > B_0$, $B_2 > B_1$. Suppose also that B_2 is *much* greater than B_1. Then, direct comparison may yield $(B_1 + B_2) > (B_0 + B_2)$; $(B_2 + B_1) = (B_2 + B_0)$; whence, if $(B_0 + B_2) = (B_2 + B_0)$, then $(B_1 + B_2) \neq (B_2 + B_1)$. Some have surmised that the axiom of equals would be satisfied, together with the commutative law, if it were not for the fact that brightness-perception goes by jumps; nevertheless, differential thresholds characterize brightness-perception, and anything that is not infected with them is not perceptible brightness.

Again: has the class of all perceptible brightnesses the so-called *group-property?* If so, then the addition of any perceptible brightness B_i to any other perceptible brightness B_j must be equivalent to another perceptible brightness B_k, such that *(3)* $B_k > B_i$ and $B_k > B_j$. By definition of perceptibility, if B_i and B_j are perceptible while B_o is not, then $(B_i + B_j) > (B_i + B_o)$, and $(B_i + B_j) > (B_o + B_j)$. But if either B_i or B_j is near the 'terminal threshold' of perceptible brightness, direct comparison is very likely to yield that $(B_i + B_j) = (B^i + B_o)$ or that $(B_j + B_i) = (B_j + B_o)$; whence $B_i = B_o$ or $B_j = B_o$, which contradicts the hypothesis.

Hence, perceived brightnesses are not, in general, additive or measurable. Like scale-balance weights they fail some of the criteria of measurability, although, within certain limits, they approximate these criteria within some determinable standards. If we are content with these standards, then we can set limits to a range of brightnesses within which we can pretend that brightnesses are measurable and outside which we cannot pretend that they are. But we now have to consider some classes of mental and social data of which not even this is true.

Are perceptible hues measurable? First, can we arrange all perceptible hues in the order of their resemblance to some standard hue, such as Helmholtz's primary chlorine green? Yes; if we exclude neutral gray from the collection, we can so arrange them, but not in a rectilinear series, as axiom *(1)* requires. The arrangement must correspond to a closed curve, on which the purple that complements the standard chlorine green will be at the pole opposite the latter, since it resembles the chlorine-green less than any other hue resembles it. But the relation 'resembles' is non-transitive, and so violates axiom *(2)*: for example, one may find a red and a blue green which resemble the chlorine green, but which (being complementaries) do not resemble each other. We have noted that neutral gray is not in this ordered series. But unless it belongs in the collection, axiom *(3)* is not satisfied, for the rule that the addition of one hue to another gives a third hue requires that the neutral hue be in the collection, otherwise the rule does not provide for the addition of complementary hues in certain proportions. Moreover, unlike the series of natural numbers, the number of distinguishable hues is finite. For example, there is a chlorine green which is "chlorine-greener" than any other hue, and there is a purple which is less chlorine-green than any other hue. The same assertion holds in principle for any primary hue that we may select. Axioms *(1)*, *(2)* and *(3)* are violated: *perceptible hues are non-additive and non-measurable.*

Consider next a class of *social facts,* which some authors falsely have treated as measurables: *e.g.* attitudes, interests, intelligences, aptitudes, skills, drives, and the like. Thurstone,[5] for example, has asserted that attitudes can be measured. Let us look into this question.

We say that a person holds one attitude or another toward some specified change (now occurring or in prospect), according as he is predisposed to make one kind of response or another kind of response to the change. Thus an attitude is a preparation for action. We suppose that it is correlated with a corresponding pattern of tensions and conductivities in the nervous system, and thus determines the *equilibrium state* of the latter. Just as solid carbon obeys one set of thermodynamic laws when it is in the form of soot, another set when it is in the form of graphite, and still another set when it is in the form of diamond: so a person, while he holds the 'attitude' of a pacificist toward a war, will make certain responses to banners, posters, pageants, poems, and to appeals to enlist, to buy bonds, to encourage the government, etc., which he would not make if he were holding the attitude of a patriot toward the war. Of course, we cannot describe his attitude physically, probably we would not so describe it even if we could. Rather, we *name it after* the course of action which we think it predisposes the person to execute, but the properties of the attitude do not depend on our way of describing it.

These assertions together imply that there are as many attitudes as there are distinguishable patterns of action, or distinguishable patterns of brain-tension. Each attitude predisposes the subject to behave according to its own corresponding set of laws. It is meaningless to ask *by how much* one attitude differs from another as it is meaningless to ask *by how much* the molecular arrangement of carbon in the form of graphite differs from its arrangement in soot or in diamond.

Thurstone does not make the mistake of supposing that these questions are meaningful, but he does assert that it is possible to measure some of the properties of attitudes, such as their *predisposing tendency* toward a given course of action, such as actively participating in a war. He assumes, moreover, that this tendency is additive, in the sense that by taking the sum, or else the mean, of the tendencies of two or more attitudes toward a given course of action, he gets the equivalent of

[5] L. L. Thurstone, A law of comparative judgment, *Psychol. Rev.,* 34, 1927, 273-286; Attitudes can be measured, *Amer. J. Sociol.,* 33, 1928, 529-554; Theory of attitude measurement, *Psychol. Rev.,* 36, 1929, 222-241.

the predisposing tendency of another attitude toward the same course of action. This assumption I invite you to examine.

Before we can proceed to *measure* any property of any attitude, we have to *detect* the attitude itself. But why is it proposed to detect a person's attitude toward a given change, in the first place? Most attitude-testers are practical men: they wish to know, for example, what one needs to do to a person to cause him to execute some course of action, such as joining the church, voting against a bond-issue, buying accident-insurance, enlisting in the army, or the like. Or, if they believe that some tendencies toward action are not amenable to propaganda, then they wish to know what means of persuasion they must avoid. Hence they wish to know what the person is already set or tending to do about the issue; or, in other words, to detect his attitudes. But, unless they intend to start working on him as soon as they have finished the test, if they use the results of the test, they presuppose that he will be holding on another occasion the attitudes which he held when they tested him. Unless the second occasion is dated, they presuppose that the test indicates what attitudes the person most probably will hold on *any* other occasion within a reasonably short time. In other words, they presuppose that the result of a single test will indicate what the person's attitudes are, *habitually.*

This presupposition is not, in general, plausible. The sinner today may become a missionary before tomorrow night; a man who is ready to give all for the love of his lady today may be glad within a fortnight that she quit him. But if the practical man denies the presupposition, he thereby denies that the detection of attitudes is useful to the propagandist.

How do the attitude-testers try to detect what a person is tending to do about a given change at the moment? *They ask him questions!*[6] Does what he *says* indicate what he is tending to *do?* Thurstone says merely that *if* it does, *then* the tester, by taking the average of the numbers which many judges set down to denote 'how favorable' toward the question they guess his statement of opinion to be, can determine 'how strong' is his predisposing tendency with respect to a corresponding *action.* The 'predisposing tendency' is not operationally defined; those who adopt this hypothesis thereby assume that the average of the guesses of the individual judges indicates the result of a set of operations that are either inconceivable, infeasible, or else merely not yet performed. Thurstone explicitly refuses to make this assumption, but his procedure is pointless unless what a person *says,* at one instant, about an hypothetical situation reliably indicates what he is most likely to *do* if the situation should *ever* become actual; this he *implicitly* presupposes.

Thurstone assumes that all possible opinions toward a given question (and with them, the corresponding tendencies to action) can be arranged in a series that is ordered according to the relation *more favorable than;* that 'favorableness' is additive according to the laws of algebraic addition; and moreover, that if B_i and B_k

[6] I once knew a certain devout Methodist who used to testify that his local church was more precious to him than all his worldly possessions, and who contributed less money toward its annual support than he spent in half a year for chewing tobacco. Also, I know certain individuals who talked pro-German until 1917, pacificistically through 1917, but who entered the army in 1918 and fought, although they were entitled to exemption. And I know certain capitalists who praised the New Deal ardently in 1933, talked communism in 1934, and raised money for the Liberty League in 1936.

are *any* two opinions in the ordered collection, the latter contains another opinion B_j, such that B_j is more favorable toward this course of action than is B_i and less favorable than B_k toward it. This assumption implies that between any two ordered opinions there are an infinite number of similarly ordered opinions, so that the numbers that express their 'favorableness' constitute a 'dense set.'

Consider these: If war should be declared, I would fight,
(A) under all conditions;
(E) under no conditions;
(I) under some conditions; namely, (a), (b), (c), ...;
(O) not under some conditions; namely, (a′), (b′), (c′),............

Obviously the classes (*A*) and (*E*) of opinions contain only one member each; classes (*I*) and (*O*) may each contain many opinions. Note especially that (*A*) and (*O*) contradict each other, as do (*E*) and (*I*). The axiom of excluded middle implies that between the sole opinion in class (*A*) and the least unfavorable of the opinions in class (*O*), no opinions can be inserted in the ordered series; and that between the sole opinion in class (*E*) and the least favorable opinion in class (*I*) no opinions can be put into the ordered series. *Opinions do not stand in a continuum.* If opinions are the transforms of attitudes, then *attitudes do not stand in a continuum.*

We cannot now review the elaborate procedures which Thurstone employs[7] for ordering statements of opinion respecting any given issue, and for assigning scale-numbers to the statements. The careful reader will discover clearly that *each procedure merely defines a rule for assigning* ordinal *numbers* to the statements. Nothing in the procedure provides a means of determining 'how much' favorableness any statement indicates, or 'how much more' favorable to a given decision one statement of opinion is than another.

From another direction let us examine the assertion that a person who endorses two or more statements expresses the sum of the 'favorablenesses' of each. Is any statement more favorable to the prospective war than the sole statement in (*A*)? No, for it includes all the statements in (*I*). Hence the property of 'favorableness to war' is non-additive, because the objects which have it do not form a collection that has the *group property.* If a person endorses (*A*) and also, for example, (*I. c*), he has not indicated more 'favorableness' than if he endorsed only (*A*). But if both are indicated by finite scale-numbers, their sum would be greater than the scale-number of (*A*). If one averages the scale-numbers, as Thurstone recommends because of the possibility of ambiguity in both question and answer, then the individual who endorsed (*A*) and also every member of (*I*) separately would earn a lower patriotic score than a person who endorsed only (*A*). Perhaps the procedure should be worked over.

In brief, no property of any attitude can be measured unless the attitude can be first detected. It is not evident that it can be detected from the person's own assertions of 'opinion.' Nor does the procedure which we mentioned provide a means of measuring the 'favorableness' of opinions themselves. It yields a counterfeit measure of the strength of a person's tendency toward any course of action. Similar defects inhere in the so-called measurement of interests, drives, skills,[8] X-O emotional stability, intelligences, and the like.

---

[7] *Op. cit.,* cf. footnote 5.

[8] H. M. Johnson, Some neglected principles in aptitude testing, this JOURNAL, 47, 1935, 159-165.

But why bother about principles? If these procedures are unsound, surely the 'Movements' will die of themselves! Why not let them alone?

The logician may answer, "Because if an invalid method leads one to factual truth, it is only by accident." The practical man may add, that the time and effort that are wasted in futile endeavor are subtracted from the amount that is available for effective work. There are questions concerning human traits which are interesting; they may be important; they may be answerable; but they probably will remain unanswered as long as they are attacked by attempts to measure what is intrinsically non-measurable.

For example, consider the attempts made during the past 30 years to determine certain psychological effects of drugs, partial asphyxiation, fatigue, insomnia, and sleep. The plan of nearly every experiment presupposed that the effect on some arbitrarily selected function would be *graded* according to the magnitude or the duration of the agent of impairment. In nearly every instance no such gradation was found; whereupon the investigator drew the conclusion that the agent was ineffective. The conclusion was false, because it rested on the false presupposition, that if the agent had an effect, the effect would be of the kind that he was seeking. But if the experimenter, like Dunlap,[9] and Bagby[10] and others in their study of asphyxiation, had asked, not 'how much' the function was affected, but *in what manner* it was performed under the compared conditions, he probably would have had to draw a different and valid conclusion. Elsewhere,[11] I shall present this field of investigation in detail. For the present, it is enough to say that it illustrates, clearly and almost tragically, the practical waste that results from reliance on false presuppositions and on pseudo-mathematics. Those data should be measured which can be measured; those which cannot be measured should be treated otherwise. Much remains to be discovered in scientific methodology about valid treatment and adequate and economic description of non-measurable facts. Their detection as such, however, is logically simple.

American University H. M. JOHNSON

ERRATA

Professor Edmund Jacobson wishes to correct the following errors in his article on "The course of relaxation in muscles of athletes that appeared in the last number of the JOURNAL. On p. 99, line 37, the numbers should be "1/3 or 1/4" in place of "3 or 4." The sentence should then read, "String tension in the galvanometer is so adjusted as to yield an excursion of 1/3 or 1/4 cm. per millivolt applied to the string terminals."

Professor Karl M. Dallenbach also wishes to correct two errors that were made in his apparatus note, "Two new A.C. chronoscopes." On p. 147, the caption to the ninth column of Table II should be "10th 10" instead of "4th 10," and the legend of Fig. 1 on p. 151 should read "Dial of Model SWC-1" instead of "SC-1."

[9] Knight Dunlap, Psychological research in aviation, *Science,* 49, 1919, 94-97.
[10] English Bagby, The psychological effects of oxygen deprivation, *J. Comp. Psychol.,* 1, 1921, 97-113.
[11] H. M. Johnson, *Human Sleep,* (in press).