THE THEORY OF THE CONSTANCY OF INTELLIGENCE*

By William Stern

Translated by Heinrich Klüver,
University of Minnesota

The concept of the intelligence quotient (IQ) has been introduced[1] with the view to find an expression of the degree of intelligence characteristic of a certain individual: this means a kind of intellectual *constant* which is independent of the variation in the actual intelligence performance as it takes place with increasing age. This concept, therefore, supplemented Binet's measure of the "mental age" (MA). The MA expresses the *existing* status of the intelligence in terms of the age level whose normal status the individual has reached. The IQ, however, relates the MA of an individual to his chronological age (CA) :IQ = MA/CA. On each age level, this value must be = 1 with reference to normal intelligence. For intelligence below normal we have < 1 and above normal > 1.

The question whether, and in what sense, it is possible to speak of such an intellectual quantitative constant is of no slight theoretical importance; the further question whether the above mentioned value proves to be empirically constant to such an extent as to warrant not only diagnostic but also prognostic statements may become of great significance in practical and psychological respect. Both questions already have been discussed in my book concerning intelligence.[2]

Psychological work in Germany in recent years has not been concerned with the above questions, due to the fact that the scale of Binet and his successors no longer continued to be the chief object for investigation. The problems presented by the selection of gifted individuals made it necessary for one to try, again and again in each case, to devise new tests. While these tests were

* Compare W. Stern: *Neue Beiträge zur Theorie and Praxis der Intelligenzprüfung.* Pp. 146-155. *Hamburger Arbeiten zur Begabungsforschung.* No. VI. Leipzig: Joh. A. Barth, 1925. Pp. XII + 193. The translator.

[1] W. Stern: *Die psychologischen Methoden der Intelligenzprüfung. Bericht über den 5. Kongress f. exp. Psychologie,* 1912.

[2] W. Stern: *Die Intelligenz der Kinder und Jugendlichen* (third edition). Leipzig, Barth, 1920. Pp. 143, 171, etc.

standardized with respect to different localities and compared with regard to their results, no effort was made to relate them to a standardized scale comprising different MA levels. Now the basic deficiency of all scales is to be sought in the fact that they have to use a rigid set of tests which makes it impossible to exclude the influence of previous knowledge and coaching. It has never been quite clear to me how the American psychologists have been able to overcome this difficulty.

On the other hand, the advances made by these psychologists by means of standardized scales in the field of intelligence examinations are of such significance that we have every reason to consider them most seriously. Terman submitted the Binet scale to a revision and extension which has been known as the "Stanford Revision."[3] Recently its importance and usefulness has been proved by results based on a large amount of data. It seems as if in America the epoch has passed in which "Binet testers" by means of dilettantish and mechanical examinations discredited the test method: examinations performed in conformity with Terman's instructions seem—so far as it is possible to pass judgments from afar—to meet the requirements of scientific method. At the same time, they have a bearing on every day problems of school life. No special emphasis is laid upon the testing of subnormal children; the great mass of school children, and especially the gifted ones, are also successfully examined.

In the succeeding presentation the results of the Stanford examinations are used to support a *theoretical* consideration dealing with the constancy of mental development.

W. Peters, in a critical discussion of my book about intelligence, was concerned with the problem of the constancy of development and its measurement.[4] He is inclined to consider the problem of the constancy of the IQ essentially as a mere "imaginary problem": for this constancy is for him *presupposed* in every properly constructed scale.

His argument runs somewhat as follows: a scale is properly constructed only when one places at each specific age those tests which correspond to the normal ability of this age. The individuals

[3] Compare e.g. *Die Intelligenz* (3rd edition) p. 136. Further: Stern-Wiegmann: *Methodensammlung.* p. 215.

[4] W. Peters: *Das Intelligenzproblem und die Intelligenzforschung. Zsch. f. Psychol.*, 1922, *89*, 1.

of this age level who are above or below average surpass or fail to reach this performance in approximately equal degrees and frequencies. Having a sufficiently large number of unselected individuals at a certain age level, the average of all values must be therefore MA = CA; if this does not hold, it is necessary to shift the tests from one level to the other until the equation is fulfilled. Since this is true for each age level, the average IQ will be = 1/1 everywhere; therefore it is constant—independent of age. The fact that the application of the scale leads to the average constancy of the IQ merely proves, therefore, the efficiency of the standardizer in placing appropriate tests at specific ages. It does not, however, prove anything concerning the children examined, or throw light upon the mental development of the non-average child.

Peters' statements are indeed entirely adequate with reference to averages obtained from group investigations, and those investigations carried on in America or elsewhere, which consider this statistical constancy of the IQ something more than a criterion for the standardization of scales, are on the wrong track.

But it was not *this* constancy, however, which led to the introduction of the IQ. It is *back of* this methodical constancy that the problem of psychological (or, better, psychogenetic) constancy begins. This is an *intra*-individual problem. The question is: in what sense and within what limits may the IQ *of the same person,* at different ages, be considered constant? This question certainly cannot be answered by group investigations which are carried on at the same time with children of different ages; it is when identical individuals are examined again and again in the course of longer periods that a solution of the problem is possible.

A priori it would be quite possible that there is a methodical (inter-individual) constancy. Suppose the same unselected 1,000 children were examined in four successive years. If the tests are appropriately chosen and the children represent a naturally distributed group, the average of all MA's for each year will be equal to the chronological age. In spite of this fact it might be possible that the IQ of the individual child would vary considerably from year to year, in some children irregularly, in others always in one direction. It may be possible also that the IQ would gradually decrease, e. g., in the case of subnormal children, and increase in the case of supernormal children. Furthermore, there could be the possibility that there exists a far-reaching individual constancy,

considerable deviations from which could be found only in some children. That would mean that the characterization of a child gained by means of the IQ has great prognostic probability. Indeed, these problems as to constancy are not imaginary problems. They represent extremely important research problems which can only be solved on the basis of extensive *empirical* material obtained from children who have been *frequently re-tested.*

Concerning this psychogenetic problem, Peters also holds a very definite view to which I may refer very briefly. The constancy of the IQ has been sometimes regarded as if the intellectual development of the child were a *uniformly* progressing development. Indeed, in the case of the constancy of the IQ the graphic representation is a *straight line* (abscissa axis = CA, ordinate axis = MA) which, for instance, for IQ = 1 ascends in an angle of 45°. Peters again is right in pointing out that this straightness cannot be considered as proof of a uniformly progressing development since it represents an essential characteristic of the system. However, he is not satisfied with this negative statement, but adds the positive assertion that according to exact findings now available the mental development shows a gradually decreasing rate.

Such a thesis seems to me no more justified than that of the uniformity of the rate of development, for I do not see *any possibility at all* of reaching objectivity as to the absolute form of the curve of development. It is, of course, possible to get intuitively the "absolute" impression that during certain periods of youth the general mental development is accelerated (e. g., in puberty compared with childhood), but an exact measurement is altogether impossible. Where are we to find this measure?

Peters, in his very brief statements, seems to point to the fact that the examination of *one* circumscribed and isolated achievement (e. g. completion of missing elements) does not reveal a uniform, but a decreasing rate of development. Dauber[5] had based his thesis of the decrease in rate of general mental development on this argument. But this special decrease has, after all, nothing to do with the general tempo of development but merely with the well-known fact of differential psychology that the more the psychic functions vary the more they involve difficulties. So far as a *definite* per-

[5] J. Dauber: *Zur Entwicklung der psychischen Leistungsfähigkeit-Fortsch. d Psychol.,* 1918, *5.*

formance is concerned, the younger children for whom the solution is difficult vary in their achievements to a greater extent than the older children; and still older children show hardly any progress at all since the performance has now become too easy. Is it, however, justifiable to speak of an arrest of mental development at this stage?

Progress in development, considered as the *total process* of growing individualities can never be understood by the juxtaposition of part developments. This progress is something qualitative: each age level is characterized not only by a "plus-performance" in fields formerly dealt with in a less successful way, but also by the appearance of novel achievements which did not exist up to that time (e.g., understanding of abstract concepts) and by qualitatively new aspects of the former kinds of achievements. This explains the fact that the different age levels in the scales of Binet and Terman do not contain only identical problems whose difficulties increase, but also qualitatively different tasks. This fact is not to be considered as a deficiency of the method preventing exact comparison of the different levels, but it is, on the contrary, a necessary and spontaneously found adjustment to the above mentioned qualitative change in mental development, an adjustment which has not yet been made sufficiently. Primarily, the fundamental intellectual change occurring at puberty has not found adequate recognition in the scales now in existence. The fact that the tests which were adequate *before* puberty and which now reveal less progress in development during puberty should not lead to the conclusion "that the absolute rate of the development of intelligence during puberty decreases in comparison with the rate in the pre-pubertal period."

Therefore, the absolute measure mentioned by Peters (p. 28) does not exist at all as an exact expression of the rate of mental development. The different levels in the scales of Binet and Terman are "equally distant" only with reference to the specific age to which they correspond (e.g., with reference to a non-psychological moment); it is, however, meaningless to state that from the tenth to the eleventh year the achievements differ *psychologically* less (just the same: or more) than from the sixth to the seventh year. It is merely a methodical device if one considers these steps as "equally large" and thus represents the average achievement by means of a straight line ascending from one age level to the other.

This methodical construction, however, furnishes us with a measure *relative to which* the actual course of development of different individuals or of specific groups (e.g., the feeble-minded) may be exactly determined. These *relative measures* form the real problem for the empirical investigation of the development of intelligence. The question whether the curve of an individual develment shows a *constant relation* to this measure or a bent form (e. g., whether it shows that the rate of development decreases or increases with increasing age) now becomes meaningful. In the first case, the IQ remains constant, in the second it decreases, and in the third it increases.

Peters also develops a definite thesis with reference to these relative curves of development. According to him, they show the same form as the curves representing absolute development, namely decreasing steepness. "As long as experience does not prove the opposite, we shall assume that the mental development of the backward child first increases rapidly; then by and by more slowly until finally it almost ceases."

The data to which Peters refers can apparently be interpreted in different ways; since they led me to the conclusion that the IQ remains practically constant[6] while Peters believes there is a tendency to decrease. Today it is no longer necessary to discuss those very limited data since the investigations of American psychologists which have become known recently places at our disposal a material of far greater range and exactness. *This material*—which the following paper will report in detail[7]—*confirms our assumptions and disproves the hypotheses of Peters.*

The most important results so far as they have a bearing on the foregoing discussion may be summarized:

I. Re-tests of the same children show that the IQ varies within certain limits which are, however, in most cases so narrow that practically one can speak of constancy. The probable range of variation of the IQ of a child is 10 points (5 points up and 5 points down). The chance that the IQ will either increase or decrease as much as 15 points is only 1/20.

[6] Of course, this holds only for a limited period. It goes without saying that during the time preceding the complete cessation of intellectual development a considerable retardation of development takes place which expresses itself in a decrease of the IQ. Compare: *Die Intelligenz* (3rd edition, p. 144.)

[7] Compare: *Neue Beiträge* etc. p. 156. The translator.

(The American psychologists do not state the normal IQ as =1 as we do, but as = 100, according to the formula IQ $= \frac{MA}{CA} 100$. There is a chance that the IQ of a child who tests at 85 (according to our computation as .85) will fall within the range between 80 and 90 in re-tests. In the totality of cases investigated, the range of five points has been exceeded as often as not reached.)

II. Considerable intra-individual variations occur now and then especially in the case of psychopaths. But neither superior nor feeble-minded children show a constant tendency to change the IQ in certain directions. There is especially no verification of Peters' assumption that the IQ of feeble-minded individuals is liable to decrease with increasing age.

III. Changes to such an extent that pronounced subnormality becomes normality or normality becomes supernormality, or vice versa, hardly ever occur.

But what do these findings *mean?*

While we had to emphasize their positive importance against attempts to deny them it is no less essential to show the *limits* of their importance.

The American psychologists tend to a certain extent to an over-evaluation, assuming that the IQ furnishes a kind of general formula for the total mental status of man. This view cannot be held. The IQ merely expresses the degree of the reactive, general intelligence, that means the capacity to handle definite objectively determined tasks for thinking. The constancy of the IQ makes it probable that the degree of this capacity belongs to the native and permanent characteristics of the individual; we are then justified in assuming in the case of the individual a level of mental alertness and adaptability which is characteristic for him. But: *level is not personality!* Original impulses, spontaneous interests, special talents, qualitatively determined tendencies, and volitional traits contribute to the formation and development of the intellectual side of the individual in a way which does not necessarily affect directly the achievements in an intelligence examination. Some of these traits may also be submitted to experimentation, but not to the kind of experiment that is represented by an intelligence examination (in this connection one may think of examinations for vocational fitness); other traits cannot be subjected to experimental

methods at all, they can be appreciated only by general observation. Thus the individual and irrational structure and the unpredictable development of every personality remains a fact even if we succeed in finding a rational and quantitative index for some of its aspects, especially for its *level* of mental reactivity.

From this there result certain limitations of the *prognostic* value of the IQ. In the first place, it makes possible *negative* predictions of high probability, namely the determination of the schools and professions for which the individual in question is not fitted, since they presuppose a higher mental level than that which is characteristic of the person. The IQ, however, seems to be of less importance for the *positive* prognosis as to school and profession, for the mental capacity of a man is strongly determined by the above mentioned personality traits in such a way that the mere determination of the level on which his achievements take place does not allow us to picture his future in a clear-cut way. The same motive that is operative in Hamburg in supplementing the intelligence examination by other criteria in the case of selection of pupils should also be effective to exclude the use of the IQ as the only prognostic index. Especially in the case of the selection of gifted individuals, the IQ involves some difficulties which, as it seems, have not always been mastered in America.[8] The IQ designates those individuals in the mass which have the strongest *reactive* intelligence; but it is not certain at all, it is not even probable, that these versatile and "clever" individuals are always the ones who possess the greatest creative and productive qualities.

The *positive value of the IQ* seems to me a fourfold one.

I. It gives us a quantitative measure for the mental level of adjustment of an individual which can be readily used and interpreted.

II. It is—on account of its clearness and independence from the chronological age—an expression easily handled in comparative statistical studies, e.g. in comparing children of different origin, of different social levels, of different school achievements, etc.

[8] Questions of this kind arise when we hear that in California the intellectual leadership of the next generation is to be identified by means of the IQ. In a group of 250,000 children in the larger cities, the 600 most intelligent ones (only children with an IQ of 140 or more) have been located; the goal is to observe them psychologically in the future, and to give them all possible educational opportunities. This is thought to be the best way of getting a properly selected intellectual leadership.

III. As to negative prognoses, the IQ renders valuable service supporting (but not establishing) our judgments concerning the degree of feeblemindedness, the unfittedness for certain school systems and school classes, for training in certain professions and for the professions themselves. It also is of value in the determination of responsibility with respect to criminal proceedings, etc.

IV. As to positive prognoses and conferences, the IQ furnishes a preliminary orientation which, however, must be controlled, supplemented and corrected by other psychological methods.

But even these more limited tasks are important enough to justify our endeavor to turn the attention of German investigators to these long neglected problems. In the first place, it would be necessary to arrive at a desirable German scale. Though Bobertag's revision of the Binet scale was full of merit, it does not meet modern claims. The tests must be selected and standardized in taking into consideration the newly obtained results in the field of intelligence investigation. A more far-reaching emancipation from Binet's sketchy and poorly arranged list of tests must be achieved.