Columbia University

Contributions to Education

Teachers College Series

THE MEASUREMENT OF FAIR-MINDEDNESS

BY
GOODWIN B. WATSON, PH.D.

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
CONTRIBUTIONS TO EDUCATION, NO. 176

Published by
Teachers College, Columbia University
New York City
1925

ACKNOWLEDGMENTS

The author wishes to express his deep and genuine gratitude to Professor George A. Coe for the stimulus of his keen thought, to Professor William H. Kilpatrick for the richness of the experience and counsel which he offered, to Professor William A. McCall for most interesting and quite indispensable guidance through the technical problems of the study, and to Professor Harrison S. Elliott of Union Theological Seminary for encouragement of the entire enterprise.

To Miss Mildred Barthelmess of the State Normal School at Trenton, New Jersey, to Professor Jerome Davis of the Yale Divinity School, and again, particularly to Professor Elliott, the author is indebted for opportunities to carry forward this study in groups which could not otherwise have been reached.

Unfortunately, because of the confidential nature of their assistance, direct personal acknowledgment cannot here be made of the service of many persons who contributed significantly to the enterprise. However, the author does wish to recognize the valuable aid so graciously rendered by his brethren in the ministry of the Methodist Episcopal Church, and by his teachers and fellow students in Teachers College and in Union Theological Seminary.

G. B. W.

New York City
April, 1925

CONTENTS

CHAP.		PAGE
I.	The Need for Tests of Fair-Mindedness	1
II.	The Construction of the Tests	8
III.	The Scoring of the Tests	13
IV.	What do the Tests Really Measure?	18
V.	Suggested Uses for the Tests	36

Appendix

1.	Individual Profiles	46
2.	Group Profiles	68
3.	Statistical Tables	88

CHAPTER I

THE NEED FOR TESTS OF FAIR-MINDEDNESS

Swift changes have come in educational practice since the days of not so long ago, when the efficiency of the instruction which had been meted out to the youth of the countryside was judged by the formal visit of school committee-men who were more possessed of a sense of responsibility and dignity, than of techniques for measuring achievement. Now, one who would ask regarding the efficiency of a teacher or curriculum must be prepared for a volley of I. Q.'s, E.Q.'s and A.Q.'s, with a running accompaniment of norms, standard deviations, correlations, and probable errors. If, perchance, some of the tests have been too hastily thrown together, and if some of the results have been much too hastily interpreted, that would be but natural. The measurement movement in education is not the first advance of science which some have "loved, not wisely, but too well."

Advance, it surely has been. Whatever mistakes shall have to be discovered and corrected, there will remain a large residue of gain. Future generations will be able to select curricula and methods with far greater effectiveness, because of those indefatigable workers who have discovered how to answer some questions beginning, "How much," and "How many."

Much of the dissatisfaction that has arisen in connection with this movement has expressed itself in a demand for tests of something more important than the abilities thus far measured. Some who believe that "out of the heart are the issues of life," have been searching for more accurate methods of evaluating motives, feelings, and purposes. Indeed, they may have had considerable reason for their fear that because certain qualities were easily measured, these qualities might be taken as criteria of desirable learning. It is possible that the contrary would be nearer the truth. Too often has schoolroom practice been controlled by the fact that supervisors could measure changes in ability to manipulate figures but could not measure the purposes which the child had in manipulating the figures. Hence, it has sometimes been said that the school was more concerned in teaching pupils how to write than in helping them to discriminate between forgery and an emancipation proclamation as an end for which the writing was to be

used. Most skills and techniques of the kind which has been measured are two-edged swords,—they can be used for the welfare of society or for its detriment.

A number of attempts have been made to measure something other than intelligence or practical educational skills. Among a large number of attempts, only four or five have at the present date been standardized. One is Miss Downey's Will-Temperament Test. With its many useful features, it is limited by the fact that temperament is measured in so small a realm, namely, that of expression through handwriting. A second test is the Pressey Cross-Out Test. This affords considerable aid in the discovery of persons who are emotionaly unstable. Third, we have the Woodworth Questionnaire, which sheds some light on the causal factors in emotional conditioning. A fourth test is Koh's Test of Ethical Discrimination, on which not many norms are now available. The Hart Test of Social Attitudes and Interests contains some very excellent material, but has not yet been fully standardized. The Kent-Rosanoff Association Test is also of considerable worth in discovering emotional complexes.

Among the attitudes which are considered most desirable by a considerable group of religious educators is one which is called variously, Open-Mindedness, Freedom from Prejudice, Scientific-Mindedness, and Fair-Mindedness.

Something of the sort seems to be suggested by Professor Coe's criticism[1] of religious instruction which leads not to knowledge but to partisanship, and his emphasis of fair-mindedness and teachableness as objectives in the education of modern youths.[2]

Even more clearly such a viewpoint is set forth by Hocking in *Human Nature and its Remaking*[3] when he sets up an ideal of a

[1]This kind of instruction in childhood produces not only in Catholicism but also in Protestantism an easily recognized adult type, the man who settles historical or scientific questions without historical or scientific study, and by the result judges whether his neighbors are sheep or goats. . . .

There is no security for worthy ends short of the habit of considering others' points of view.

Coe, Geo. A., *A Social Theory of Religious Education*, pp. 65, 40.

[2]An educated man . . . must have command of the method of the mind, and he must be—to some extent—a thinker, not a mere imitator. He may or may not have more opinions than other persons, but he has more opinions to which he has a right . . . .

To the extent that society is organized in the interest of the common good, and chooses its means by scientific types of analysis, the attitude of teachableness supplants that of dogmatic assertion.

Coe, Geo. A., *What Ails our Youth?* pp. 39, 77.

[3]Why do we not display with complete equableness all views of the best way of life, and say 'Now choose; think out your course for yourself'? Instead of teaching our children our morality, why not teach them ethical science? Instead of religion, metaphysical criticism? Instead of political faith, political philosophy? Instead of manners, the principles of aesthetics? In short, why not make thinkers of them rather than partisans?

Hocking, W. E., *Human Nature and Its Remaking*, pp. 230–23.

rational scientifically-minded generation, free to choose its own best course.

Likewise Miss Follett in *The New State*, a well-known exposition of a new social and political philosophy, finds open-mindedness one of the central prerequisites.[4] In every disagreement she believes men should find a new challenge to investigation and rethinking, not an occasion for intellectual strife.

Frequently, in his writings, Professor John Dewey has set forth very clearly an educational objective of open-mindedness.[5] He

[4]The lesson of the new psychology is, then: Never settle down within the theory you have chosen, the cause you have embraced; know that another theory, another cause exists and seek that. The enhancement of life is not for the comfort-lover. As soon as you succeed,—real success means something arising to overthrow your security . . .

And throughout our participation in the group process we must be ever on our guard that we do not confuse differences and antagonisms, that diversity does not arouse hostility. Suppose a friend says something with which I do not agree. It may be that instantly I feel antagonistic, feel as if we were on opposite sides, and my emotions are at once tinged with some of the enmity which being on opposite sides usually brings. Our relations become slightly strained, we change the subject as soon as possible, etc. But suppose we were really civilized beings, then we should think, 'How interesting this is; this idea has evidently a larger content than I realized; if my friend and I can unify this material we shall separate with a larger idea than either of us had before!"

Follett, M. P., *The New State*, pp. 38, 40.

[5]Some attitudes may be named which are central in effective intellectual ways of dealing with subject matter. Among the most important are directness, open-mindedness, single-mindedness (or whole-heartedness) and responsibility . . .

Open-mindedness. Partiality is, as we have seen, an accompaniment of the existence of interest, since this means sharing, partaking, taking sides in some movement. All the more reason, therefore, for an attitude of mind which actively welcomes suggestions and relevant information from all sides . . . Openness of mind means accessibility of mind to any and every consideration that will throw light upon the situation that needs to be cleared up, and that will help determine the consequences of acting this way or that . . . The worst thing about stubbornness of mind, about prejudices, is that they arrest development; they shut the mind off from new stimuli. Open-mindedness means the retention of the childlike attitude; closed-mindedness means premature intellectual old age . . .

Open-mindedness is not the same as empty-mindedness. To hang out a sign saying, 'Come right in; there is no one at home' is not the equivalent of hospitality. But there is a kind of passivity, willing to let experiences accumulate and sink in and ripen, which is an essential of development, Results may be hurried; processes may not be forced. They take their own time to mature. Were all instructors to realize that the quality of mental process, not the production of correct answers, is the measure of educative growth something hardly less than a revolution in teaching would be worked . . .

Dewey, John, *Democracy and Education*, pp. 204–206.

While it is not the business of education to prove every statement made any more than to teach every possible item of information, it is its business to cultivate deep-seated and effective habits of discriminating tested beliefs from mere assertions, guesses and opinions; to develop a lively, sincere and open-minded preference for conclusions that are properly grounded, and to ingrain into the individual's working habits methods of inquiry and reasoning appropriate to the various problems that present themselves. No matter how much an individual knows as a matter of hearsay and information, if he has not habits and attitudes of this sort, he is not intellectually educated. He lacks the rudiments of mental discipline. And since these habits are not a gift of nature (no matter how strong the aptitude for acquiring them); since, moreover, the casual circumstances of the natural and social environment are not enough to compel their acquisition, the main office of education is to supply conditions that make for their cultivation. The formation of these habits is the Training of the Mind.

Dewey, John, *How We Think*, pp. 27, 28.

has pointed out the confining consequences of prejudice. Moreover he has made it clear that he is interested in a sort of open-mindedness which is not empty-mindedness, but is akin to the meaning of "fair-mindedness" as used in these tests.

One of the most popular presentations of this same idea is found in *The Mind in the Making* by James Harvey Robinson.[6] The whole thesis of this luminous book is set forth in a clear-cut demand for a "critical, open-minded attitude" as the solution for most of the world's ills.

Speaking particularly from the viewpoint of the function of higher education, President E. M. Hopkins of Dartmouth College[7] said: "It is one of the functions of the college to promote open-mindedness—a quality sorely needed today in a period of intolerance that is worse than any other in history."

It seems clear, then that any attempt to study open-mindedness will not be labor in a field which is deemed trivial and insignificant. Whatever may have been true of some aspects of educational measurement, certainly the construction of tests in this field will not be a business of testing simply for the fun of testing,—of sticking in prods to see what happens, quite regardless of the social and philosophical significance of the results, if any are attained.

There remains another important question with regard to the worth of the enterprise. Some who would grant that the field is one in which social consequences are very significant might well ask: "But why tests? Why not experimentation to see what kinds of teaching, what subject matter, and mental training tend toward open-mindedness? Why not a treatise setting forth even more clearly the dangers of prejudice in modern society and an attempt to convince men, yes, to convert them?"

In this field, as in every other, selection of desirable method waits

[6]If some magical transformation could be produced in men's ways of looking at themselves and their fellows, no inconsiderable part of the evils which now afflict society would vanish away or remedy themselves automatically. If the majority of influential persons held the opinions and occupied the point of view that a few rather uninfluential people now do, there would, for instance, be no likelihood of another great war; the whole problem of 'labor and capital' would be transformed and attenuated; national arrogance, race animosity, political corruption, and inefficiency would all be reduced below the danger point. As an old Stoic proverb has it, men are tormented by the opinions they have of things, rather than by the things themselves. This is eminently true of many of our worst problems today . . .

I am not advocating any particular method of treating human affairs, but rather such a general frame of mind, such a critical, open-minded attitude, as has hitherto been but sparsely developed among those who aspire to be men's guides, whether religious, political, economic, or academic.

Robinson, J. H., *The Mind in the Making*, pp. 3, 12.

[7]Reported by the *New York Sun*, Dec. 4, 1924.

upon the creation of measuring instruments. Schools and theorists set forth their schemes; and one may appear effective, another quite hopeless, but there is no proof of the pudding without some standard by which the eater can make his judgment.

> Richards, in his presidential address before the American Association for the Advancement of Science said: "Plato recognized, long ago, in an often-quoted epigram, that when weights and measures are left out, little remains of any art. Modern science echoes this dictum in its insistence on quantitative data; science becomes more scientific as it becomes more exactly quantitative".[8]

Before we can decide whether this course or that, whether lecturing or discussion, whether accusation or persuasion will be more effective in the creation of fair-mindedness, we must have first some method of telling in how great a degree this quality of mind is present before and after the application of these devices.

"After all," some may say, "while fair-mindedness is certainly an essential objective in the educative process, and while it is true that we must first have some method of judging as to how great a degree it is present, can we not tell by unrefined observation? Do not people know prejudiced persons from fair-minded persons? Can not people tell you themselves the points at which they are prejudiced and those at which they are open-minded?"

All of the evidence seems to indicate that such commonsense judgments are highly unreliable. Probably the best study of the reliability of ratings on traits of character has been made by Rugg.[9] After an exhaustive study under conditions unusually favorable he finds that character ratings are valid only when (1) the ratings are made by three or more independent judges, (2) who have been trained under a leader skilled in scaling so that they are in perfectly clear agreement as to the scale being used, and (3) who are thoroughly acquainted with the subjects. These conditions can very rarely be fulfilled. "A single rating by a typical school officer will only rarely locate a person within his proper 'fifth' of the entire scale." "Hence the apparently dogmatic answer to the question, 'Can human character be rated on point scales accurately enough for practical uses in education?' No! It would be far better to give our energies to the attempt to measure it objectively than to make subjective judgments on point scales. The point cannot be made too emphatic that we should discard these loose methods of rating once and for

[8]Quoted by W. A. McCall in *How to Measure in Education* p. 8.
[9]Rugg, H. O., *Journal of Educational Psychology*, Feb.-Nov. 1921.

all." It does not seem probable that we may hope for any greater reliability in ratings of fair-mindedness than he has there discovered with reference to intelligence, leadership, industry, team-work, etc. If we include self-ratings, we may well expect less validity. Dewey, in a recent lecture, said, "It is of the very nature of prejudice that it is largely unconscious. Everyone is pretty sure that the matters which seem to us to be our prejudices, aren't."

Further light in the same direction came from an investigation carried on for other purposes, described here on pages 25 and 26. One individual, who was rated by a number of the students who were in classes with him as one of the most fair-minded persons in the school, was described by an intimate friend as "perfectly hopeless" in his extremism and prejudice upon a number of religious and economic issues. "Conclusions which he has drawn by 'intuition' he is willing to die for, however incongruous they may be with the results of his rational thinking." Another person described this same individual as "one of the most lovable personalities I have ever dealt with, but quite likely to go off half-cocked on some new idea." In this case at least, there seems to be clear evidence of what Thorndike has called the "halo effect."

Again, evidence comes from a class of eighteen normal school students studying educational measurements, who were asked to rank themselves and other members of the class on open-mindedness. They did it under protest, saying they did not trust their ratings, but they did try to do it conscientiously. The ratings which a pupil gave himself yield a correlation of .07 with the average ratings which he was given by the other members of the class. Fifteen of the eighteen pupils were given ranks which ranged from 3 to 17, or further.

A little more evidence is found in the vote of students and faculty members of two institutions with regard to the two most open-minded members of their group. While there was considerable agreement upon a few individuals, the remainder of the votes scattered to include 73 per cent. as many individuals as there were persons voting and an average of 51 per cent. of the possible choices in each group.

While all of this evidence may not be conclusive and there is need for further study of the conditions under which ratings in this field will be reliable, this evidence does make it clear that in general such ratings are not dependable or trustworthy. If they are to

be proven useful at all, the proof itself will require a test, the validity and reliability of which have already been established.

In general, there seems to be clear evidence that there is a need and demand for a type of education which shall bring about progress toward fair-mindedness. Such progress is at best, uncertain, until some instrument can be constructed which will measure the degree to which fair-mindedness, or freedom from bias, has been achieved.

The construction of such a test, however, is not wholly in the interests of those who wish to promote fair-mindedness as an educational objective. While the author finds himself in very great sympathy with this point of view, the test itself is an equally good instrument for those who believe fair-mindedness undesirable upon certain questions. There is no attempt here to insist that fair-mindedness rather than prejudice is desirable. Many competent persons feel that upon certain questions the reverse is true. From either point of view the test should prove useful. It merely indicates what the situation is within the individual at the time at which the test is administered. In the light of this evidence, the educator can more discriminatingly select methods which will lead in the direction in which he deems it desirable to work.

In order to make the problem compassable, the issues with which this test deals have been limited to religious and economic issues, including in the former certain "moral," or "amusement" questions which have become tied up with religious sanctions. The choice of these two types of issue as the field in which to work was determined in part by the author's interest, but in large degree by a realization of the tremendous consequences of prejudice and intolerance, as contrasted with fair-open-scientific-mindedness in these realms.

CHAPTER II

THE CONSTRUCTION OF THE TESTS

In the endeavor to construct tests which shall objectively measure fair-mindedness, there are two obvious dangers. One is that the test shall be a measure of agreement with the opinions of the person who constructs the test. It is all too easy to feel that those who differ radically from our own point of view are thereby showing strong prejudice. A second danger is that the test shall prove to be merely a measure of convictions, a test of opinions reasonably arrived at and held with a fair degree of confidence.

To avoid these dangers, these tests were constructed to measure prejudice defined as: Any tendencies, however produced, (*a*) to cross out, as distasteful, terms which represent one side or another of religious or economic controversies; (*b*) to call sincere and competent persons who hold different opinions on religious and economic issues incompetent or insincere; (*c*) to draw from given evidence conclusions which support one's bias but which are not justified by that evidence; (*d*) to condemn in a group which is disliked, activities which would be condoned or approved in some other group; (*e*) to regard arguments, some of which are really strong and others of which are really weak, as all strong if they be in accord with the subject's bias, or all weak if they run counter to that bias; and (*f*) to attribute to all the people or objects in a group, characteristics which belong to only a portion of that group. Most of the tests were so constructed that opinions could be expressed on either side of the issue without going to the extreme of registering one of the tendencies which have been defined as prejudices. All of the tests were so constructed that prejudice or lack of fair-mindedness could be registered upon several aspects of each of the religious and economic issues considered. One might be prejudiced, for example, in the direction of fundamentalism, or of modernism, or of radicalism upon religious questions.

Form A, the Word Cross Out Test, is based upon the principle employed by Pressey, that a tendency to cross out an unusually large number of words, feeling that they are annoying or distasteful, is an indication of some sort of emotional set, or conditioning. The following are typical of the fifty-one words included in this test:

Bolshevist	Dancing
Mystic	Holy Communion
Sunday Blue Laws	Unitarian

Form B, the Degree of Truth Test, consists of a collection of statements about religious and economic matters, upon each of which it would be possible to find sincere and competent authorities in disagreement. The following are typical of the fifty-three statements included in this test:

The churches are more sympathetic with capital than with labor.
Jesus was more interested in social, than in individual salvation.
Poor men cannot get justice in the courts today.
The ordinary Catholic priest is well above the average in his community, so far as learning and good judgment are concerned.
Prohibition, in the experience of the United States, has been a failure.

In each case, the subject is given the opportunity, by checking on a scale, to say that each statement is:

(2)—so true that no one with a fair understanding of the subject could sincerely and honestly believe it false.
(1)—probably true, or true in large degree.
(0)—uncertain, or doubtful.
(−1)—probably false, or false in large degree.
(−2)—so false that no one with a fair understanding of the subject could sincerely and honestly believe it true.

The emphatic portions of the first and last positions are underlined in the directions, so that people will be more certain to notice the extreme character of the statement. Both this test and the Form A test force many thoughtful individuals into a dilemma, neither horn of which is wholly satisfactory. For some reasons they wish to cross a given word out, for other reasons they would prefer not to. They cannot compromise here. Likewise in the Form B test, they dislike to admit that the statement is only probably true or false but they hesitate to go the full length of the extreme position. The theory of the test is that in such an uncomfortable situation, the choice is more than ordinarily significant.

Form C, the Inference Test, presents a statement of fact, followed by several conclusions which some persons might draw from that fact. Thus, for example, it is stated:

1. Statistics show that in the United States, of one hundred men starting out at an age of 25, at the end of forty years, one will be wealthy, while fifty-four will be dependent upon relatives or charity for support.

☐ The present social order cheats the many for the benefit of the few.

☐ The average young man, under present conditions, cannot count on being wealthy at the age of 65.

☐ Most men are shiftless, lazy, or extravagant, otherwise they would not need to be dependent.

☐ The one man is living upon luxuries ground out of the bones of the masses of common people.

☐ Some day the workers will rise in revolt.

☐ No such conclusion can fairly be drawn.

The subjects are asked to check only such conclusions as are established by the facts given in the statement above, drawing upon no other evidence. They are warned to check only those inferences which are certain,—none that are merely probable. It is permissible to check either the final alternative under each case, stating that none of the conclusions can fairly be drawn, or else that one in each case (the second, in this example) which is merely an innocuous repetition of the evidence given in the preliminary statement.

Form D, the Moral Judgments Test, consists of fifteen instances, with an opportunity beneath each, to approve the act, to declare it a matter of indifference, or to disapprove of it. Each instance in the test is parallel to one or two other instances in the type of situation it presents. Thus, "faith cures" are presented, once through the medium of a Japanese idol, once at a Roman Catholic shrine, and once under the leadership of a Protestant evangelist. In another instance unwarranted search is made of a suspected "radical" headquarters, while in another, the same procedure is carried out with a big business corporation, suspected of dishonesty. These instances are, of course, scattered through the test so as to conceal as far as possible the parallelism. It matters not, for the purposes of the test, whether the subject shall approve, be indifferent to, or disapprove, the action in any situation, if only he will be consistent in the parallel act under slightly different circumstances.

Form E, the Arguments Test, is based upon the tendency of an individual to feel that all of the arguments upon his side of the case are strong, while those on the opposite side are weak, irrelevant, or very easily refuted. Twelve issues, such as the desirability of the Roman Catholic church, the probability of immortality, the usefulness of the Ku Klux Klan, etc., were selected. Upon each were gathered the principal arguments used by persons supporting

each side of the question. Some were presumed to be weak, while others seemed to the author to be strong. After selection by a few seminar groups, the arguments were submitted to a group of twelve judges, chosen because of their supposed ability to rate religious and economic arguments fairly. They included a professor of philosophy, a professor of education, a professor of economics, a professor of sociology, two professors of public speaking and argumentation, two teachers of religious education in different institutions, one leader in inter-denominational activities in the industrial field, a member of the International Committee of the Y. M. C. A. and a psychologist. Extraordinary difficulty was found in getting any unanimity. Arguments had to be revised, and resubmitted. Apparently, wide deviations in judgment occurred, less because of bias on the question as a whole, than because of unusual reactions to some feature of the premise in the argument. For example, certain judges were not greatly concerned by the fact that the particular argument favored the child labor law, but did seem to be influenced by the fact that the premise contained the word *profit*, or *employer* or *farmer*.

No argument was finally retained in the test upon which there were less than 75 per cent. of the judges in agreement. As further evidence, twelve of the most fair-minded persons, as recorded by the other forms of the test, were studied to see how well their opinions agreed with the judges. The results are tabulated in Table III of the Appendix. It is not at all improbable that many capable students of each issue will look over the ratings given to arguments by other sane and capable individuals, and find themselves dismayed. Few, if any, individuals have been discovered who do not differ emphatically from the great mass of their fellow students, on some of these arguments.

Form F, the Generalization Test contains a number of generalizations about Jews, ministers, I. W. W.'s, business men, miracles, missionaries, etc. Each of these generalizations is true of some of the members of the group, but not of all of the members. The subject is given a chance to say that the statements are true of "All, Most, Many, Few, or No" members of the group. Any answer except "All" or "No" is accepted, and ignored in the scoring.

Each of these test-forms has been revised six, seven, or, in some instances, ten times. Preposterous as some of them may seem, no items were retained which were not reacted to in the extreme prej-

udice form, by several individuals among the first two hundred to take the test.

At the close of the six test-forms is appended a personal data sheet, asking for sex, age, occupation, education, religious training, occupation of father, and years of residence in communities of less than 5,000 people, and more than 100,000 people. The tests may be anonymous, or the name of the subject may be signed. Usually a *key*, in which the subject can recognize and recover his own paper is desirable.

A brief Manual of Directions is published separately, giving the purpose of the tests, the results of standardization, directions for giving the tests, and for obtaining the gross score. This Manual of Directions, together with complete copies of the test and full directions for finding both the gross score and the analytical score, may be obtained from the Bureau of Publications, Teachers College, Columbia University, New York.