Irish Journal of Education, 1971 v, 2, pp 70-85

THE EFFECTS OF STANDARDIZED TESTING*

GEORGE MADAUS, PETER AIRASIAN
Boston College
and
THOMAS KELLAGHAN
Educational Research Centre
St Patrick's College, Dublin

Many arguments for and against the use of standardized tests have been advanced The present paper reviews studies of the effects of testing on examinees, teachers, organizations sponsoring tests, and parents Almost without exception, these studies have failed to distinguish between evaluation derived from test results and evaluation based on other information Thus, the effects of information based on standardized tests must still be considered an open question

THE ISSUES

Controversy over the effects of examination practices on individuals and social institutions is certainly not a recent phenomenon O'Meara (34), in his excellent historical survey of examination practices from ancient times to the present, points out that, over the centuries, examinations have been accused of being responsible for a wide range of ills including

> carelessness, hatred, favouritism, labour unrest, unprogressiveness, defective art, dishonesty, discontent, poverty, fraudulency, laziness, a generator of mental defectiveness and physical degeneration, serfdom, radicalism, suffering, death, strikes and war (34, p 10)

On the other hand, he concludes that examinations are so valuable educationally and socially that they are 'a virtual necessity, which will continue until some new device is discovered that will fulfil in a better way the function of the examination both from the educational and social view points (34, p 11)'

Since about 1960, the controversy over testing has been particularly active both in the United States and Great Britain The controversy, for the most part, has been concerned with the use of objective, rather than more conventional, kinds of tests The advantages and disadvantages of such tests and their effects on children, teachers, parents, schools and

*The preparation of this paper was supported by a grant from the Russell Sage Foundation

society in general have been discussed An extensive literature which in general is sympathetic to testing (e g , 7, 10, 11, 16, 19) is countered by one which is unsympathetic and often hostile (e g , 25, 26, 46, 48) Holt's views illustrate the latter position

> The threat of a test makes students do their assignments, the outcome of the test enables us to reward those who seem to do it best The economy of the school, like that of most societies, operates on greed and fear Tests arouse the fear and satisfy the greed (26, p 2)

A favourable view is expressed by Gardner

> Anyone attacking the usefulness of the tests must suggest workable alternatives It has been proven over and over again that the alternative methods of evaluating ability are subject to gross errors and capable of producing grave injustices (19, p 56)

These two positions illustrate the somewhat diffuse nature of the debate in the United States In much of the literature there is a failure to define clearly the type of test which is being attacked or defended For example, Holt seems to rail primarily against informal teacher-made tests (although one could infer from the tone of the entire essay that he opposes standardized tests as well) Gardner, on the other hand, is primarily defending standardized tests of intelligence, achievement or aptitude Characteristics of test instruments, such as bias, item ambiguity, triviality and irrelevancy, are the focus of the attack of other critics Banish Hoffman's (25) book is typical of this line of criticism He focuses more on weaknesses in the instruments themselves, including the wording of individual questions, than on the function to which these tests are put One might ask whether critics that argue along these lines would allow testing if a 'perfect' test could be designed

In much of the writing about testing there is a failure to distinguish between the test itself and the function of the test While Holt appears to attack tests directly, he really seems more upset by their functions—grading, evaluation, and judging class work—than by the tests themselves His argument might have more cogence had he said he was against teachers evaluating student performance at all, since he claims that such evaluation casts a backward shadow, with deleterious effects, on learning and teaching

In Britain the contemporary debate over testing is more sharply defined in that it centres primarily around the assets and liabilities associated with the function of selection rather than characteristics of tests *per se*

Recently, the debate has had wide coverage in the popular press following the publication of the three *Black Papers* (10, 11, 12) and the replies in the *Red Paper* (13). The *Black Papers* constitute a strong defence for the continued use of standardized tests to select pupils as well as a sharp attack on 'progressive education' in general. Writing in *Black Paper Two*, Eysenck claims that:

> Without the help of IQ tests, advancement into the higher paid, better educated groups of society will be barred to many able working-class children, thus bringing to the top a large number of people of mediocre ability, while keeping submerged many people of superior ability. Thus the use of a new mediocracy is socially unjust, naturally disastrous, and ethically unacceptable (14, p. 40).

A quotation from *The Red Paper* illustrates the fact that in Britain the issue does not concern the test or testing *per se* as much as the function of *early* selection:

> A minor theme in this symphony of lamentation [*Black Papers One* and *Two*] is that intelligence tests, if properly constructed and used, *are* of value in revealing unexpected potentialities in children. This *cri de coeur* can be got out of the way quite quickly. Of course intelligence and other tests are useful indicators of unsuspected powers in a child, but their role is not, *and should never have been*, to give a scientific veneer to the slaughter of the innocents which *any* selection procedure at any early age involves when it results in a *minority* being passed on for *privileged* schooling.
>
> Man is a status-motivated animal and to suggest that the selection of a privileged minority of young children from the mass can result in anything but intense parental anxiety and, therefore, in devastating discouragement to the children who 'fail', is so obvious that there is no need to press the point. Selection at a later age, after individual differences and interests have been established, and largely accepted, on the basis of the experience of both parents and children, is quite another matter (24, p. 2).

Two other points about the English controversy should be noted. First, the argument over age of selection and educational goals has strong political overtones. There is a definite socialist-conservative polarization along *Black-Red Paper* lines. A second factor is the sharp division along the hereditary-environment continuum. Simon (39), among others, has noted that there is a hereditarian assumption underlying many of the

Black Papers essays. This assumption is manifest in Eysenck's position and in Richard Lynn's more stringent charge that the 'progressives'

> raise false hopes that much more can be done for slum children than is actually possible. The same is true of comprehensive and fashionable new methods of teaching. False premises lead to false remedies and ultimately to disappointment. If it is thought desirable to improve the intelligence of the population, money would be much better spent on helping less intelligent people to limit the size of their families (31, p. 30).

The hereditary argument of the *Black Papers* has spilled over to include the Jensen controversy raging in the United States prompting Simon (39) to quote from the *New Scientist* that in the resultant controversy over the genetic basis of intelligence, 'science, politics and prejudice have become inextricably mixed.'* Although it is not always recognized, the debate over whether or not to test is part of a larger philosophical, ideological and political controversy concerning both the goals of education and the methods to be employed in arriving at these goals. Unfortunately, the debate on the benefits and liabilities of testing is confined for the most part to statements deduced from philosophical and/or political positions; rarely is there empirical evidence to support the conclusions reached (cf. 30); when there is, persons with different philosophical orientations will weight the data differently.

Complicating the matter still further is the fact that tests are designed to do unpopular jobs and hence it becomes difficult to remain calm and reasonable when discussing them (19). Emotional involvement is liable to cloud or distort not only the 'expert's' vision but even laymens' views of tests, since they, or at any rate their progeny, have all either passed or failed some examination or other during their life times (45).

Speculatively and even emotionally based statements, however, cannot lightly be dismissed, though it must be realized that the validity of such assertions has never been carefully investigated. Unquestionably there are legitimate issues that can be raised on the effects of testing, but to wage the debate without sufficient data in this age of technology is hardly satisfactory. Interestingly Burt, almost sixty years after he helped introduce 'new-type' tests to Great Britain, writing in the second *Black Paper*, levels the following criticism against 'progressive education':

> Most of the practical changes . . . in educational organization and in educational methods, have been suggested not so much by systematic

*For further treatment of the Jensen controversy in Britain, see 6 and 15.

> observation or the analysis of experimental studies as by the prevailing ideology—by what some writers like to style the *Zeitgeist* Before they are adopted on any large scale it is imperative that they should be systematically tested, and where necessary modified, by carefully controlled research (5, p 24)

With hind-sight, it is perhaps too easy to point out that this statement by one of the foremost pioneers in the testing movement applies not only to 'progressive education', but also to many testing practices Given the enthusiasm of the moment, it is not surprising that the pioneers in testing on both sides of the Atlantic, ignored or simply failed to take into consideration what the side effects of proliferated testing might be over a long period of time In 1963, Mahler and Smallenburg (32) observed that considering the vital importance of questions concerning the effects of testing, 'relatively little attention has been given to them (p 103) ' The situation changed very little over the following six years In the 1969 edition of *The Encyclopedia of Educational Research*, Womer (47) noted that 'in many areas research evidence is not available to substantiate either claims made for or criticisms made of test use (p 1462) ' Womer went on to reiterate a fact pointed out by Goslin (21) that most research in testing is directed towards improving test accuracy rather than with problems concerning the consequences of these tests or the functions for which they have been designed Kirkland's (30) most recent review does not change earlier conclusions In the remainder of this paper we propose to re-examine studies of the effects of standardized testing As should be clear from our discussion so far, the number of studies that throw any real light on this topic is very small indeed

Our review will be confined to a consideration of the use of objective tests by schools or school systems for their own internal needs, such as guidance, diagnosis, placement or programme evaluation We shall not consider studies of external testing programmes, i e those administered by agencies external to the school and frequently used to certify to the successful completion of some level of education or to select students for further education

EFFECTS ON EXAMINEES

It seems reasonable to assume that standardized test score information most directly affects the examinee Goslin (21) suggests two levels of effect which the widespread use of ability and achievement tests are likely to

have on the individual The first is the *direct* impact of additional information about his own abilities For example, the result of a test or examination (pass-fail) might affect an individual's self-concept, level of aspiration or plans for further education At the second level of effects, information comes to an individual in a more indirect way Effects at this level follow when the sponsoring agency uses test results as a basis for decisions about examinees In the case of external examinations, second level effects may be very direct and very powerful, as when an individual fails an examination, the passing of which is necessary for entry to a profession or to a course of further education In internal examinations, effects at this level are more likely to be indirect and subtle

First level effects

Bloom sees effects at this level as being most important 'If these test effects are understood and utilized properly,' he writes, 'they can do much to enhance the student's learning as well as his regard for himself (2, p 41) ' How the tests are used seems important Tests are frequently used to categorize or sort individuals In this case, the effects on the student may not always be beneficial Gardner has pointed out that if a society 'sorts people out efficiently and fairly according to their gifts, the loser knows that the true reason for his own lowly status is that he is not capable of better This is a bitter pill for any man (19, p 42) ' The sorting function probably more obviously relates to the second level of effects than to the first However, the mere reporting of a test score can easily carry classificatory overtones (a percentile rank or simply 'better' or 'worse' than another student's score)

In assessing the impact of ability and achievement test information on pupils, it is important to bear in mind that testing programmes form only a part of the evaluative procedures of the school and classroom The school is basically an evaluative institution where the student must become 'used to living under the constant condition of having his words and deeds evaluated by others (27, p 10)' In this context, a key role within the classroom is that of the teacher-evaluator, and any consideration of the effects of formal testing must take into account the network of constant evaluations made by teachers in their everyday work With this in mind, our search of the research literature dealing with the effects on examinees of information from formal testing programmes, led us to the conclusion that, with one exception, such research is inadequate in its conceptualization and interpretation The inadequacy of these and other studies lies in the failure to place the information from formal testing

programmes within the context of the informal ongoing and more pervasive evaluations made within the school and the classroom

One study in which different forms of evaluation are taken into account is that of Torshen (43) Torshen investigated the theoretical proposition that the form of evaluation which has the greatest impact on a person's immediate environmental context will also have the strongest effects on his personality, self concept and mental health She hypothesized that in most classrooms, teachers' grades have a strong direct impact on students, while standardized achievement test scores have a more indirect influence Underlying her hypothesis was the belief that achievement test scores do not play an important role in the daily routine of American classrooms since they are given at most once a year and the results are generally available only to the teacher or other school personnel Further, she assumed that the teacher's grade incorporated non-academic assessments not included in the standardized test information Such assessments may themselves indirectly influence performance on standardized tests Using a multiple partial regression technique, Torshen found that when the effects of teachers' grades and various control variables, such as sex and IQ, were removed, the remaining relationship between achievement test scores and measures of self-concept and mental health were not statistically significant Further, she showed that there is a significant relationship between student grades and measures of self-concept and mental health which is independent of any influence of the control variables, or of achievement test scores As a result of her research, Torshen concluded that grades contain an element which affects students' self-concept and mental health but is not related to purely academic achievement This non-academic element may be composed, at least in part, of teachers' evaluations of students on what Parsons (35) has labelled the 'moral' dimension of achievement Torshen also concluded that results of standardized tests may not be without impact, since such results may have affected teachers' grades, which in turn were related to mental health The importance of Torshen's research lies in the implications that when more informal classroom evaluations are taken into consideration, relationships between standardized test scores and various personality measures tend to disappear

In some reviews there is an assumption that tests have marked effects upon many pupils (32, 40) These effects, it is claimed, may be positive when test information proves to be reinforcing and provides motivation for further learning, or they may be negative as when such information creates anxiety and feelings of inadequacy and lack of self-confidence It

is further claimed that negative effects tend to increase as the child goes through school (40) None of the studies on which these reviews was based however, had carefully examined the numerous variables relevant in a study of the effects of testing, particularly the relationships of formal testing programmes to the more informal assessment of teachers For example, in a number of studies it is suggested that classroom evaluation is related to pupils' self-concepts (3, 9, 17, 33, 38), Torshen observes, however, that

> The researchers seem to select the measure of evaluation which was easiest to obtain without regard for the impact that type of evaluation had on students This writer was unable to find any research which investigated the possibility that one form of evaluation has a greater impact on students' personalities than another form of evaluation has (43, p 27)

Thus, because of the consistent failure to distinguish between the effects of test results and those of other evaluative information, any firm generalization on the basis of the studies cited is not possible

A variable that logically should affect the impact of formal test information on student personality or cognitive variables is the amount and quality of the information he receives about his performance on tests of ability or achievement Goslin points out that

> Very little of a systematic nature is known about the effects on children of providing them with specific information about their abilities
>
> Obviously, the effects depend upon the information given, previously held conceptions of ability, the way in which the information is presented, the strength of competing estimates, and various other factors (22, p 136)

None of the studies reviewed attempted to control for this information variable Torshen (43) we saw, assumed that standardized test information was not made available to students but instead was placed on the permanent record file available only to the teacher or other school personnel The findings of Goslin (22) and of Brim *et al* (4) substantiate the validity of Torshen's assumption Goslin (22) reports that very few teachers, even those who felt students should receive intelligence and achievement test information, ever gave them such information Brim *et al* (4) found that while a majority of the student respondents would be interested in knowing their IQ test results, most had not received such

information Brim further concluded that such feedback of information, when it does occur, is unevenly distributed in the American population In families where a student's intelligence is known in a fairly specific way, the family tends to be better educated, and the student tends to score well on tests and to have high educational aspirations

The fact that teachers are loathe to pass on test information to students suggests either that they think the information is of little value or that they fear the effects the information might have on students There are important issues involved here—ones which deserve attempts at empirical resolution

Second level effects

Second level effects, according to Goslin (21), and as we have already seen, happen as a result of sponsoring organizations, in this case the school, using test results as a basis of decisions about the examinee 'In this way', Goslin points out, 'tests influence the individual's opportunities for receiving the best education, getting a good job, and in general improving his social position (21, p 184)' These effects of testing raise some complex philosophical, moral and legal issues and relate to the reasons minority groups in the United States are often strongly opposed to standardized tests, they are also a principal issue in the bitter debate now being waged in Britain over whether or not to abandon their elitist system of education Goslin (22) reports the percentage of time that elementary school principals reported using various kinds of test as a basis for grouping The following is a list of the type of test followed by the percentage of time it was used for grouping reading readiness 61 per cent, individual intelligence 30 per cent, group intelligence 39 per cent, reading achievement 42 per cent, arithmetic achievement 29 per cent, achievement battery 29 per cent, other tests 52 per cent, non-standardized reading tests 46 per cent The total percentage of time tests were reported as being used for grouping was 41 Tests were more frequently reported as being used for diagnosing learning difficulties Such a use was reported 71 per cent of the time (In the questionnaire, principals could list up to four main uses for each type of test)

Despite these statistics on the use of formal test information on grouping, Rist (36) and Jackson (27) point out that many first and second grade teachers form intra-class groupings without the benefit of formal test data Rist contends that these evaluations of student potential, made as early as the kindergarten, are relatively firm and are later passed on to

teachers in higher grades The degree to which formal standardized test data modify these early assignments is open to argument

Whether it is the result of teacher assessments based on social, affective and cognitive cues, standardized test performance or some combination of these two factors, it seems reasonable to hypothesize that the formation of ability groups may indirectly have adverse consequences on the child's self-image Smith observed that

> the child is highly aware of his standing in the entire ability group structure, and that his expectations of success are influenced more by the total structure than by the other students in his particular class (40, p 243)

If the child perceives his standing in the ability group structure as low this could lead to a vicious circle As Coleman *et al* (8) put it 'if a child's self-concept is low, if he feels he cannot succeed, then this will affect the effort he puts into the task, and thus, his chance of success (p 281)'

Ability grouping is something of which people are aware, and concerning which some people have very strong feelings Brim and his colleagues (4) in their survey of American beliefs and attitudes about intelligence found that respondents generally were aware of ability grouping during their school years and for the most part they believed that grouping is based on standardized test results Further, they found that upper-class respondents (with higher intelligence and aspirations) held more favourable opinions about homogeneous ability grouping in schools than did respondents from poorer educational backgrounds In a British study, where the majority of parents favoured streaming, opinion was not found to be related to social class (23)

Despite the indication that ability grouping is widely practised, that people are aware of it, and that membership of a particular group is likely to affect an individual's self-concept, motivation, attitudes and achievement, we are a long way from having any solid evidence on the effects of grouping, to date, empirical studies in this area have produced conflicting findings (29)

EFFECTS ON TEACHERS

Goslin (21) suggests that the teacher can influence a person's score This may result from something as gross as being poorly trained in test administration or from something much more subtle as when a teacher's expectations from the child's performance act as a self-fulfilling prophesy.

Mahler and Smallenburg (32), in their review of research on the effects of testing, uncovered little research regarding effects on teachers except in relation to external examinations Attitudes and practices that arise as a result of adapting to external examinations may not, of course, be the same as those that result from using standardized tests within the school The only systematic research in this important area is Goslin's (22) study of American teachers' practices and attitudes This research, given the American situation, was necessarily of the survey type However, it offers valuable insight into how testing practices influence American teachers First of all, Goslin found that while less than 40 per cent of American teachers have had formal training in tests and measurements, virtually all have access to their pupils' test scores In fact, in most schools IQ results are routinely distributed to teachers Goslin points out the difficulty of making inferences from survey data about the extent that possession of such information has on teacher behaviour towards pupils He notes that teacher sophistication in testing and measurement might be an important variable in this context, and he suggests that 'explicit considerations need to be given to the problem of teacher training in the field of measurement (p 129)' Other interesting findings from Goslin's survey are that teachers tend to view standardized tests as relatively accurate measures of a student's intelligence and achievement They also feel that the abilities measured by these tests are important determinators of subsequent academic success and should be used, along with grades, for grouping and admission purposes Further, teachers view the skills measured by these tests as more influenced by learning than by innate capabilities Finally, all teachers feel that *some* children should be given some information on their test performances Despite these beliefs and attitudes, teachers reported a rather low degree of use of test scores

The well known study of Rosenthal and Jacobson (37), *Pygmalion in the classroom*, was an attempt to examine experimentally the influence of test results on teachers and, through teachers, on pupils At the beginning of the school year a non-verbal IQ test was administered to a group of children Teachers were then given the names of children who, on the basis of the test findings, could be expected to show dramatic intellectual growth Actually, the names of the children were chosen randomly and consequently the only difference between 'high expectancy' children and the ordinary children was in the mind of the teacher and of the investigators The investigators claim that group IQ tests administered by the teachers on three occasions during the year showed a significantly larger gain in IQ in the 'high expectancy' children than in their classmates The

results have been strongly criticized on statistical and design grounds (e g , 20, 28, 41, 42) The study, however, is open to criticism not just on methodological grounds but also on substantive ones, particularly those relating to the types of cues and information that go into the formation and communication of expectancies If we are to be realistic, formal test information has to be fitted into the network of constant ongoing evaluations made by teachers, peers and the individual himself during the daily grind of classroom life The formation of expectancies then is embedded in complex personal, social and perceptual facets of the classroom and is too subtle and complex to be accounted for simply by giving teachers intelligence test information in September and measuring pupils' IQs at some later point If it were that simple, then formal internal testing programmes would indeed constitute a direct and extremely powerful treatment or intervention It can also be readily argued, however, that at the level of teacher expectancies, formal test information is an indirect and relatively weak treatment, and that the really important cues on which such expectancies are based probably can be manipulated

EFFECTS ON ORGANIZATIONS SPONSORING TESTS

Tests are primarily administered so that the organization sponsoring the test can learn something about the examinee's abilities This information is then used to screen, place or select individuals Goslin (21) lists two potential secondary effects of testing on the sponsoring organization firstly are changes in the personnel of the organization resulting from the use of tests for the selection of personnel, and secondly are changes in the social structure of the organization or its method of operation due to increased knowledge about the aptitude of existing personnel

An example of the former effect can be seen in changes in enrolment patterns in American higher education The social class composition of college applicants in the United States of America has become much more heterogeneous than in the past, due in no little measure to the use of the College Board Examination in admissions (21, p 176) Similarly in Britain (here the examination concerned was really an external one, we cite the example for lack of evidence about internal examinations), the composition of grammar schools changed following the 1944 Education Act The availability of scholarships and the extended use of intelligence and attainment tests to select pupils for grammar schools resulted in an increased participation by pupils from lower social class homes (18, 44) The schools changed from being socially selective towards being socially

comprehensive and from being intellectually comprehensive towards being intellectually selective

An example of a change in the social structure of a sponsoring organization (in this case, the school) following the introduction of tests would be the formation of groups on the basis of test results We have already considered the question of ability grouping in dealing with the effects of testing on students Here we are concerned with effects on the institution —its social structure and method of organization—rather than on the individual By definition, grouping should reduce the amount of variance in ability within a class or group By implication, it should also reduce the amount of variance in the social class composition of the group Thus if test results are used to reduce the heterogeneity of class groups, a change in the social structure of the school results For one thing, children of different levels of ability will have less contact with each other Alternatively, the use of test results could lead to a type of organization that would facilitate individualized instruction This could result in a system of non-grading or in the reduction of the amount of group instruction, which incidentally might require physical arrangements different from those normally found in schools While consequences of testing such as these can be hypothesized for a school, whether or not testing in itself would lead to such changes is an open question As Goslin (21) has pointed out, systematic research within organizations is required before we will be able to say precisely what impact testing has on organizations

EFFECTS ON PARENTS

While a little is known about how adults perceive the effects of tests they had taken themselves (4), we know next to nothing about how a parent's knowledge of a child's test performance or teacher evaluations affect family interaction, perceptions and aspirations Brim and his colleagues feel that keeping parents ignorant of their child's test results is based on the school's assumption that such information may be injurious to the child's self-esteem or motivation, or mental health They conclude that 'it is shocking and astonishing to find so little solid social research testing this fundamental assumption (4, p 12)' Then they go on to point out quite correctly that

> If a school were to develop a systematic policy of dissemination of test scores, it could replace the irrational, unevaluated, and probably damaging procedures of the schools at the present time in handling differences in intelligence (4, p 13)

CONCLUSION

Despite controversies concerning the consequences of testing, empirical evidence on the subject is almost completely lacking There has been a number of surveys concerning such things as teachers' attitudes to testing and their use of test results, while these contain much valuable information, they were necessarily post facto in nature, given the entrenched status of testing practice in American schools There also has been a number of experimental investigations, but these, almost without exception, have failed to distinguish between evaluation derived from test results and evaluation which may have been based on a multiplicity of other cues the child's learning record, his home background, his co-operativeness, attentiveness and general decorum, how he interacted with other children Thus, there is a need for experimental studies of the effects of classroom evaluation practices (including standardized testing) Such studies are necessary before we will be in a position 'to plan programmes of evaluation for the most productive consequences to students (11)'

There are several areas of effect where the future researcher might look The controversialists have already indicated many of them At the personal level, the possible effects of test information on student learning, self-concept and level of aspiration might be examined The role of test information on the formation of teacher attitudes and expectancies might also be looked at At the non-personal level, an examination of the possible effects on patterns of educational participation by social class and on such things as school curricula and the organization of classes within the school is indicated Only when a great deal more empirical evidence on these matters is available,* will it be possible to answer many of the claims and charges about standardized testing being made by controversialists today

*It may not be possible to examine experimentally all these issues in school systems where objective testing is already widely practised

REFERENCES

1 BALCH, J The influence of the evaluating instruments on students' learning *American Educational Research Journal*, 1964, *1*, 169-183

2 BLOOM, B S Theoretical issues in evaluation In Tyler, R W (Ed), *Educational evaluation New roles, new means* The Sixty-eighth Yearbook of the National Society for the Study of Education, Part II Chicago National Society for the Study of Education, 1969

3 BORESLOW, B Self evaluation and academic achievement *Journal of Counseling Psychology*, 1962, *9*, 246-254

4 BRIM, O G, GLASS, D C, NEULINGER, J, FIRESTONE, I, and LERNER, S C *American beliefs and attitudes about intelligence* New York Russell Sage Foundation, 1969

5 BURT, C The mental differences between children In Cox, C B, and Dyson, A E (Eds), *Black papet two Crisis in education* London The Critical Quarterly Society, 1969

6 CAMBRIDGE SOCIETY FOR RESPONSIBILITY IN SCIENCE *A public debate on education, race, IQ, intelligence* Cambridge Cambridge Society for Responsibility in Science, 1971

7 CHAUNCY, H, and DOBBIN, S E *Testing Its place in education today* New York Harper & Row, 1963

8 COLEMAN, J, CAMPBELL, E O, HOBSON, C J, MCPARTLAND, J, MOOD, A M, WEINFRIED, F D, and YORK, R L *Equality of educational opportunity* Washington, DC US Government Printing Office, 1966

9 COMBS, C F Perceptions of self and scholastic underachievement in the academically capable *Personnel and Guidance Journal*, 1964 *43*, 47-52

10 COX, C B, and DYSON, A E (Eds) *A black paper Fight for education* London The Critical Quarterly Society, 1969

11 COX, C B, and DYSON, A E (Eds) *Black paper two The crisis in education* London The Critical Quarterly Society, 1969

12 COX, C B, and DYSON, A E (Eds) *Black paper three Goodbye Mr Short* London The Critical Quarterly Society, 1971

13 CUDDAHY, R, GOWAN, D, and LINDSAY, C *The red paper* Edinburgh Edinburgh Islander Publications, 1970

14 EYSENCK, H J The rise of the mediocracy In Cox, C B, and Dyson, A. E (Eds), *Black paper two The crisis in education* London The Critical Quarterly Society, 1969

15 EYSENCK, H J *Race, intelligence and education* London Temple Smith, 1971

16 FINDLEY, W G (Ed) *The impact and improvement of school testing programs* The Sixty-second Yearbook of the National Society for the Study of Education, Part II Chicago National Society for the Study of Education, 1963

17 FINK, M B Self concept as it relates to academic under-achievement *California Journal of Educational Research*, 1962, *13*, 57-62

18 FLOUD, J Social class factors in educational achievement In Halsey, A H (Ed), *Ability and educational opportunity* Paris OECD, 1961

19 GARDNER, J W *Excellence* New York Harper & Row, 1961

20 GEPHART, W J Will the real Pygmalion please stand up *American Educational Research Journal*, 1970, *7*, 473-475

21 GOSLIN, D A *The search for ability Standardized testing in social perspective* New York Russell Sage Foundation, 1963

22 GOSLIN, D A *Teachers and testing* New York Russell Sage Foundation, 1967

23 GREAT BRITAIN DEPARTMENT OF EDUCATION AND SCIENCE *Children and their primary schools* A report of the Central Advisory Council for Education (England) London HMSO, 1967

24 HEMMING, J Introduction In Cuddahy, R, Gowan, D, and Lindsay, C (Eds), *The red paper* Edinburgh Edinburgh Islander Publications, 1970 Pp 2-4

25 HOFFMAN, B *The tyranny of testing* New York The Crowell-Collier Press, 1962

26 HOLT, J W *On testing* Cambridge, Mass Pinck Leodas Association, Inc, 1968

27 JACKSON, P *Life in classrooms* New York Holt, Rinehart, and Winston, 1968

28 JENSEN, A R How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 1969, *39*, 1-123

29 KELLAGHAN, T The organization of classes in the primary school *The Irish Journal of Education*, 1967, *1*, 15-36

30 KIRKLAND, M C The effects of tests on students and schools *Review of Educational Research*, 1971, *41*, 303-350

31 LYNN, R Streaming Standards or equality? In Cox, C B, and Dyson, A E (Eds), *Black paper three Goodbye Mr Short* London The Critical Quarterly Society, 1971 Pp 26-29

32 MAHLER, C, and SMALLENBURG, H Effects of testing programs on the attitudes of students, teachers, parents, and community In Findley, W G (Ed), *The impact and improvement of school testing programs* The Sixty-second Yearbook of the National Society for the Study of Education, Part II Chicago National Society for the Study of Education, 1963 Pp 103-125

33 MITCHELL, J V Goal setting behavior as a function of self acceptance, over and underachievement and related personality variables *Journal of Educational Psychology*, 1959, *50*, 93-104

34 O'MEARA, J F A critical study of external examinations and of their influence on secondary education Unpublished MA thesis, University College, Cork, 1944

35 PARSONS, T The school class as a social system Some of its functions in American society In Halsey, A H, Floud, J, and Anderson, C A (Eds), *Education, economy and society A reader in the sociology of education* New York The Free Press, 1961 Pp 434-455

36 RIST, R C Student social class and teacher expectations The self-fulfilling prophecy in ghetto education *Harvard Educational Review*, 1970, *40*, 411-451

37 ROSENTHAL, R, and JACOBSON, L *Pygmalion in the classroom Teacher expectations and pupils' intellectual development* New York Holt, Rinehart & Winston, 1968

38 SHAW, M C, EDSON, K, and BELL, H H The self concept of bright underachieving high school students as revealed by an adjective check list *Personnel and Guidance Journal*, 1960, *39*, 193-196

39 SIMON, B *Intelligence, psychology and education A Marxist critique* London Lawrence & Wishart, 1971

40 SMITH, C P (Ed) *Achievement-related motives in children* New York Russell Sage Foundation, 1969

41 SNOW, R E Unfinished Pygmalion *Contemporary Psychology*, 1969, *14*, 197-199

42 THORNDIKE, R L A review of *Pygmalion in the classroom American Educational Research Journal*, 1968, *5*, 708-711

43 TORSHEN, K P The relation of classroom evaluation to students' self concept and mental health Unpublished PhD dissertation, University of Chicago, 1969

44 WESTERGAARD, J, and LITTLE A Educational opportunity and social selection in England and Wales Trends and policy implications In OECD Study Group in the Economics of Education, *Social objectives in educational planning* Paris OECD, 1967 Pp 215-232

45 WHITEHEAD, F External examinations examined I A conflict of functions *Journal of Education*, 1955, *87*, 485-487

46 WOMER, F B Testing programs—misconceptions, misuse, overuse *Michigan Journal of Secondary Education*, 1961, *2*, 153-161

47 WOMER, F B Test use In Ebel, R (Ed), *Encyclopedia of Educational Research* (4th ed) New York Macmillan, 1969 Pp 1461-1469

48 YOUNG, M *The rise of the meritocracy* Harmondsworth, Middlesex Penguin Books, 1961