OFFICIAL INSIGNIA OF KAPPA DELTA PI

Orders on official blanks must be approved by a chapter officer and the Recorder - Treasurer of the Society.

PRICE LIST

Badges

<table>
<thead>
<tr>
<th>Size</th>
<th>No. 0</th>
<th>No. 1</th>
<th>No. 2</th>
<th>No. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. 0</td>
<td>$2.75</td>
<td>$4.00</td>
<td>$5.25</td>
<td>$6.00</td>
</tr>
</tbody>
</table>

Guard Pins

SMALL OR MEDIUM SIZE

<table>
<thead>
<tr>
<th>Single Letter</th>
<th>Double Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain</td>
<td>$2.25</td>
</tr>
<tr>
<td>Close Set Pearl</td>
<td>4.50</td>
</tr>
<tr>
<td>Crown Set Pearl</td>
<td>6.00</td>
</tr>
</tbody>
</table>

LARGE SIZE

<table>
<thead>
<tr>
<th>Single Letter</th>
<th>Double Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain</td>
<td>$2.75</td>
</tr>
<tr>
<td>Close Set Pearl</td>
<td>5.50</td>
</tr>
<tr>
<td>Crown Set Pearl</td>
<td>7.50</td>
</tr>
</tbody>
</table>

ADDITIONAL FOR WHITE GOLD GUARDS

Plaint | $1.00 | $2.00 |
Close or Crown Set Jeweled | 1.50 | 2.50 |

COAT OF ARMS GUARDS

Miniature, Yellow Gold | $2.75 |
Scarlet, Yellow Gold | 3.25 |

TAXES

To prices quoted must be added a Federal tax on jewelry of twenty per cent. In addition a use or occupation tax is charged in some states as indicated: Alabama, 2%; Colorado, 2%; Illinois, 2%; Iowa, 2%; Kansas, 2%; Michigan, 3%; North Dakota, 2%; Ohio, 3%; South Dakota, 2%; Utah, 2%; Wyoming, 2%. Since state taxes vary from time to time, officers should make a check on taxes in their own states to determine the amount of payment.

GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN
The Monopoly of Objective Tests

FRANK N. FREEMAN

The thesis of this paper is that the use of objective examinations is greatly overdone in the present-day school and college in this country, and that this excessive use is bound to have a harmful effect on study and learning in America. I believe that our students are acquiring a wrong concept of what study means and that the real aims of instruction are being distorted. This distortion of aims, produced by the instruments of examination that we use, comes at a time when we are adopting in theory quite different aims and are trying to reorganize our curriculums so as to achieve them. We are coming to realize that the mere acquisition of isolated facts is not an education, and that unless the student goes forth from our schools and colleges with well established concepts or ideas his education has been a failure. At the moment when the importance of seeing things in relation to each other and to the major concerns of human life is recognized, our schools and colleges are dominated by a technique of examination that leaves this type of thinking out of account. It is time we realize that our aims in teaching and our methods of evaluating teaching are at cross purposes and that we sought to bring them into harmony.

I

The use of so-called objective examinations has become very widespread. They were first employed by teachers in departments of education and by teachers in high schools. They have now been adopted by many teachers in other departments in colleges and universities. Along with this extensive use has grown up the dogma that objective tests provide the only satisfactory examining instrument, and that the essay examination is entirely outmoded. I shall contend, on the contrary, that the objective test has definite limitations, is useful only for certain purposes, and is much inferior to the essay examination for other purposes. This view is not new, but it needs renewed emphasis.
Behind the By-Lines

The Monopoly of Objective Tests is the leading educational psychologist, Frank N. Freeman, now Dean of the School of Education of the University of California. For thirty years he was a member of the staff of the University of Chicago. He accepted the deanship at the University of California in 1939. He is the author of a dozen books on education and science, and co-author of several others. He also has contributed frequently to leading educational and psychological journals. He has occupied many positions on various commissions and has been prominent in research.

Written especially for The Educational Forum is the article, The Educator Helps America, by John Edgar Hoover, since 1924 Director of the Federal Bureau of Investigation of the United States Department of Justice. Known throughout the country for his efficient management of the Bureau he has been awarded honorary doctorates by thirteen colleges and universities. He has been active in Boys Clubs and the Boy Scouts of America, and has received many medals for his contributions to good citizenship. He has written many articles for magazines, especially for law reviews and police journals.

Dean J. B. Edmonson, of the School of Education, the University of Michigan, evaluates the work of one of our international organizations in his Implications of UNESCO for Education. For eight years Dr. Edmonson was a member of the Educational Policies Commission of the National Education Association. He is the author of two books, and joint author of twelve others. He has written often for educational magazines.

What's Ahead in the Colleges? is the vital question posed and answered by Orde- way Tead, President of the Board of Higher Education of New York City, and editor of the economic books of Harper Brothers. He is author or co-author of ten volumes and has also been a frequent contributor to magazines.

Looking backwards from a century ahead (2046), the article The Role of Education in the Twentieth Century Revolution analyzes the present social scene. The author is Harold Benjamin, Director of the Division of International Educational relations in the United States Office of Education. He has been editor of a series of educational books and has performed editorial duties on several magazines.

Recent Changes in Soviet Education is a lucid and compelling article written by George S. Counts, of Teachers College, Columbia University, a member of the Laureate chapter of Kappa Delta Pi. A year ago Dr. Counts wrote Education and the Promise of America, a volume which appeared in the Kappa Delta Pi Lecture Series. Dr. Counts has been president of the American Federation of Teachers and has written much in the field of education.

Their Own Good Earth, the short story for the issue, is by Nettie Wyso who contributed to our columns previously. She has been engaged in editorial work and writing.

Edmund deS. Brunner is the author of Agricultural Extension Service in the United States. He is Professor of Rural Education in Teachers College, Columbia University. He was a member of President Roosevelt's Advisory Committee on Education. He has written much in the field of country life in its religious and economic aspects and in the

(Continued on page 506)
It may be helpful to sketch briefly the history of the objective test. There are three forms most commonly used: the true-false, multiple choice and completion forms. Others, such as the matching tests, are less often used. These forms of tests are so familiar, even to readers of popular magazines, that they do not need to be illustrated.

All three of these forms of test originated as means of testing ability rather than as means of examining knowledge. The completion test was first hit upon by H. Ebbinghaus, in 1897. He was seeking a method of measuring mental fatigue, but found that he had discovered a means of measuring mental capacity. The Ebbinghaus test was introduced widely in this country by Trabue's Language Test Completion Scales. This title was somewhat of a misnomer since the scales measure mental ability more than language ability or the effect of instruction in language. When used as part of an examination rather than as a test of ability, the device is differently set up. In the first case the part of the sentence which is given should make it evident what conditions must be fulfilled by the word or words which are omitted, whereas in the second case, the words which are given do not make these conditions obvious. In the latter case it requires ingenuity to put together the cues supplied by the words which are given and to figure out what they mean. Hence, when the completion form is used in an examination it measures a different kind of ability than when it is used in a mental test. It determines whether the student can supply a given fact or item of knowledge when the nature of the information required is evident rather than whether he can combine cues so as to infer what the nature of the missing part is. When the completion device was taken over from the mental test to be used in examinations and achievement tests, then, its essential character was altered and it demanded a different kind of thinking.

The true-false and alternative or multiple choice devices have a similar history, though their historical development from their use in mental tests to their use in examinations is foreshortened. They were first used extensively in the Army Alpha Scale in the first World War. Within a few years they were coming to be widely used in examinations.

There is an essential difference between examinations and mental tests which is not always recognized. In mental tests the measurement is largely relative—that is, it consists in a comparison between individuals. Granted that we know something about the kind of ability that is measured, though this is usually little enough, the amount of this ability that a person making a given score possesses is defined only as being greater or less than that possessed by somebody else, or by a group of other persons. In an examination the returns should signify more than this. They should reveal in a more substantive way what knowledge the individual possesses or what ability he commands. They must be scrutinized carefully in order to determine whether they require the kind of knowledge and the kind of
ability which it is the aim of instruction to produce, and whether the amount of knowledge and ability that is produced satisfies the aim.

These questions regarding the substantive character of tests and examinations have not been sufficiently considered in recent discussions of the theory and technique of testing achievement, and in the construction of tests and examinations. In technical terms, validity is not given enough attention in contrast to reliability. In fact, the true meaning of validity disappears altogether in the quite respectable technique of test construction in which the validity of an item is defined as its consistency with all other items of the test. The criterion of validity then becomes purely an internal one, and the question whether the test as a whole measures what it is supposed to measure is a matter of purely gratuitous assumption. It seems likely that this neglect of the question whether the test or examination measures what it is supposed to measure grows in part out of the predominant use of forms of tests which do not bear too close scrutiny in respect to what they measure. I shall return to this point.

II

The great argument for the use of objective forms of examination was in fact their superior reliability. The early studies by Starch and Elliott in 1912-13, have become classics in the sense that everybody quotes them but nobody reads them. In each of their three studies these investigators sent an examination paper in a high school subject to a large number of high school teachers and asked them to grade it on a scale of one hundred per cent. The grades covered a wide range. Of 116 teachers of mathematics, for example, two marked the paper over 90 and one below 30. It is obvious that a large part of this variation may be due, not to the unreliability of scoring but to a difference in the standard required for passing. This variation in standards also affects the scores on objective tests when these scores are transformed into marks. On the other hand, marks may be made more uniform by agreement concerning the standards to be employed. In fact, they may be made approximately as uniform, by appropriate preparation of the readers, as are scores on objective tests.

The notion that marks based on the older types of examination, including the essay examination, are almost entirely arbitrary, has persisted for over thirty years, in spite of the fact that it has been abundantly proved that these examinations can be made substantially as uniform as are objective examinations. As long ago as 1929 a comparison between the reliability of objective tests and essay tests was made by Ben C. Wood and the writer in their study of motion pictures.\(^1\) Eight reliability coefficients from the objective tests, corrected by the Spearman-Brown formula, were .708, .718, .723, .752, .770, .789, .804 and .941. From the essay tests the three coefficients calculated were .74, .76 and .81. Obviously these coefficients are

---

of the same order. In the thirties close agreement in scoring essay-type tests was achieved in experiments conducted by Traxler and Anderson, Frutchey, Sims and Stalmaker. There is no longer any ground for the belief that it is necessary to resort to the objective type examination to secure uniformity, or reliability. It is merely a question of taking sufficient pains.

Pains, by the way, are as necessary in constructing objective examinations as essay examinations. As grievous sins against true education have been committed by objective tests—even of the "standardized" variety—as by essay tests. This is particularly true of tests that deal with principles or general ideas. They indicate the difficulty, if not the impossibility, of dealing adequately with general questions in tests of this kind. Some of the published tests, which are doubtless widely used, are enough to make courses in Education the laughing stock of the academic world. The following examples, chosen from one of the most respectable tests

---


on An Introduction to Education, which shall here be nameless, are by no means the worst that could be found. The student is directed to indicate whether the following statements are true or false:

Teaching should make use of the familiar in the child's experiences.

The real test of method depends upon its utilization of the laws of learning.

Every successful teacher will avoid making cross connections between the subject-matter of the lesson and the more direct experiences of life.

The school can well be irresponsible concerning the health of children in its care.

A good question must be closely related to the experiences of the learner.

A memory question develops initiative and creative thinking.

Educational tests are valuable means of comparison and diagnosis.

Inexperienced teachers should not attempt to formulate thought provoking questions.

A wise teacher will attempt to kill off by punishment the instinct of mastery.

A child always learns best by imitating other pupils' responses.

Responses to situations are entirely unrelated to the general mental state of the person confronted by the situation.

Individual differences should be given consideration in the junior high school.

These questions are taken from one of the four parts of the test. Perhaps they are intended to eliminate the Mortimer Snerds from the group. If anybody got as far as a course in Introduction to Education and was unable to answer these questions before he started the course, he should look into some of Mortimer's explanations of why he is so stupid. Either a course on which such an examination is based is full of
platitudes or the examination is a very poor representation of the course. One hopes that the latter is the case. It would give some indication, for example, of a prospective teacher's competence to hear him tell how "teaching should make use of the familiar in the child's experiences"; but his assent to this proposition would reveal just nothing at all concerning his competence. In fact, it would be worse than nothing, because it would seem to give knowledge when none exists.

These may be poor examples of true-false questions, but if such questions are prepared by prominent educators and put out by a leading publisher, one shudders to think of the kind of questions that must be formed by the average teacher. Perhaps a method should be judged only by its best examples. At any rate these will reveal its possibilities and limitations. Let us then look at an example selected from one of the best objective tests I know, one in educational psychology:

In acquiring a manual skill most learners benefit greatly from having an instructor or mechanical device to guide their hands through the movements involved.

This is a good question to find out whether the student knows one of the facts which has been discovered through scientific experiment. However, so far as the question and the answer go the fact is an isolated one, unrelated either to the supporting evidence or to its implications for education. Furthermore, there are no qualifications expressed or implied, and there is no opportunity for the student to suggest that qualifications might be made. By such questions, furthermore, the student is not encouraged to consider the reasons for the facts that are in him but only to accumulate them. He is encouraged, rather, to try to anticipate what facts the instructor is likely to call for, and to file them away in neatly arranged pigeon holes in his mind.

This is not mere speculation. It is proved to the hilt in every day teaching experience as well as by scientific experiment. Why do students insist on knowing whether they will have an objective or an essay examination? There can be only one answer. They will prepare differently for the two types. They frankly say so. Their reports are backed up by observation of the methods they actually use. Terry and Meyer both found that students report that their methods in studying for the two examinations differ widely. A quotation from one of Meyer's students is typical of all:

"For an essay type test, I usually try to fix the general outline, the major drift of the subject, in my mind, and then add as many details to the general absorption as my time and energy permit. I usually outline the material on paper and try to think it through several times. When false and true, completion or multiple-choice tests are expected, I concentrate my attention on learning details, definitions, words, figures. I stuff my memory with as many facts as

---

I think it likely to retain for the required time, until and including the test, and then quickly forget every thing except the few points that appealed to me as most important."

Meyer also confirmed the students' reports of the methods they used by examining records of their study. Meyer, in another experiment also found that students could recall material in organized fashion and know facts when cues were not given better when they prepared for essay examinations than when they prepared for objective examinations.

III

What is the purpose of teaching and of examination, and what is the relation of this purpose to the student's later life, personal or professional? I think we may say that one of the chief purposes is to produce understanding or comprehension. This is something much broader and deeper than the accumulation of facts, though it has an essential relation to facts. Facts are the means, or part of the means to understanding, but they are not understanding itself. The possession of facts does not guarantee the attainment of understanding and the fading out of facts from the memory does not prove that understanding has been correspondingly dimmed. In truth, the function of facts is largely to serve as a sort of scaffolding to understanding. After the main edifice has been built

the scaffolding may be taken down.

What is it that will carry over into everyday life, or into teaching, the recollection that somebody found that pupils were helped in learning to write by tracing the letters in grooves or the development of a conception of learning which is the outgrowth of various forms of evidence and which, in turn, is applicable to a wide variety of situations? To draw an example from another field, what approach to the history of education will illuminate our current issues the best, one which impresses the student with the requirement that he learn that Sturm influenced Jesuit Education, or one that undertakes to build up in the student's mind a picture of the dominant ideas, beliefs and social practices of an epoch and the connections between these and the educational system. Granted that there is not a complete dichotomy, there is at least a vital difference in emphasis. The stress we place in our evaluation will be reflected in the stress the students place in their study.

There is a certain precision and neatness in objective measurement and a susceptibility to statistical manipulation that is very attractive,—and for certain purposes, useful. As an exclusive, and perhaps even a chief, method of evaluation of learning, however, it is fundamentally inadequate. Only by rigidly channeling the student's thoughts can they be represented by a single number. By doing the infinite variety of the individuals' thinking is ignored, and the possibility of checking and guiding the development of the students' ideas, con-

---

ceptions and points of view is abandoned. This, to be sure, is a delicate process, but it lies at the heart of teaching.

Consider what is revealed regarding the comprehension of the problem by two following actual answers to an examination question:

Question: Why is it difficult to express the range of differences in abilities by a ratio?

Answer, Student A: Ratio is merely a means of establishing a relationship between two factors. The difficulty in trying to express the range of one's abilities by a ratio, would come about due to the fact that one's abilities are overlapping and so varied that it would be impossible to establish a ratio.

Answer, Student B: A ratio of difference in ability can be expressed only when the test is calibrated from 0. This rarely can be done unless one goes back to the abilities of the very young child. As a rule, it is inexpedient to go farther back than the probable range of the examinees. Hence the ability to perform one arithmetic example as compared with the ability to do six of the same difficulty does not imply a ratio of 1:6 in ability, the ability present in the first case being considerable, and the improvement in ability for six right perhaps negligible.

No mere number would do justice to the difference between these two statements, nor would it give a clue as to where to begin in the education of the student who wrote the first statement—perhaps at some relative, if not at absolute zero.

I would not cast out the objective tests on which so much ingenuity has been expended. They have their uses. I would, however, protest vigorously against the casting out of what is contemptuously called the essay examination. I maintain that the most delicate and direct means of exploring the student's mind, as of instructing him, is still the method of exposition and discussion. Nothing comparable to it has been devised as a means of revealing the students' thinking, or as a means of cultivating the ability to think. I suggest that we recover our balance, confining objective tests to those uses to which they are fitted, and restoring the free expression of thought through language to the position of dignity which it deserves.

"A man may write at any time if he will set himself doggedly to it."

—Samuel Johnson