THE CONVENTIONAL EXAMINATION IN CHEMISTRY AND PHYSICS VERSUS THE NEW TYPES OF TESTS: PART 1

EARL R. GLENN

The Lincoln School of Teachers College, Columbia University.

In the attempt to improve upon the traditional examination in physical science it is important that we keep in mind certain facts with respect to the intelligence of pupils and the present practice of teachers in giving school marks:

I. According to the psychologists*, pupils may be distributed over three different scales; one represents *inherited capacity*, one represents *ability*, and one represents *specific performance*.

1. Inherited capacity has to do with the "start in life." We do not measure *inherited capacity* directly in school work.
2. If the educational conditions are favorable in the home, school, etc., the rank of the individual on the *ability* scale will approximate his true rank on the *inherited capacity* scale.
3. If a student's ability is his *ability-to-do* then what he does represents his *specific performance*; that is, the problem that he solves, the experiment that he performs, etc.
4. These *specific performances* may be measured in school practice. *It is these that school marks ought to measure.*
5. We do not measure ability, however; we estimate it by the specific performances of the student.
6. There are two important characteristics associated with the performances of a student:
 - (a). Specific performance is *unstable* and *one test* does not give an accurate estimate of the ability of a given student.
 - (b). The average of several (3 or 4) performances is relatively stable. The tables below show these two facts.

TABLE I: CHEMISTRY

A table showing the results obtained on a series of tests given near the end of the course, 1920-1921:

No. 1

No.	Name and rank on composite of 12 tests.	Rank on Stanford revision of Binet-Simon test. ¹	Rank by (1) quality of experimental work.	Rank by (2) accuracy in solution of chemical problems	Rank by (3) quality of a paper written upon a chemical industry.
1	A	1 (137)	1	1	1
2	B	4 (115)	4	4	4
3	C	2.5 (119)	2	2	2
4	D	2.5 (119)	3	3	3
5	E	5 (113)	5	7	7
6	F	7.5 (104)	8	9	9
7	G	6 (107)	7	8	6
8	H	7.5 (104)	6	6	8
9	I	9 (95)	9	5	5

¹Tests given January to March, 1920.

No. 2

No.	Name and rank on composite of 12 tests. Student.	Rank by (4) ability to assemble chemical information in graphical form.	Rank on (5) technical information test. (100 items.)	Rank on (6) Stenquist mechanical ability test (Picture test)	Rank on (7) silent reading test (Burgess PS-1)
1	A	1	3	4	1
2	B	2	4	3	3
3	C	7	8	7	8
4	D	5	6	6	4
5	E	8	1	1	6
6	F	4	2	2	5
7	G	6	7	8	7
8	H	3	9	9	2
9	I	9	5	5	9

II. A school mark or an examination grade should have the same meaning to student, teacher, parent, and administrator.

No. 3

No.	Name and rank on composite of 12 tests. Student.	Rank (8) chemistry test M C (Biography)	Rank (9) chemistry test No. 1 (Formulas symbols, etc.)	Rank (10) on chemistry test No. 2 (Range of information)	Rank (11) on chemistry test No. 3 (Problem-solving)	Rank (12) on chemistry test C-1
1	A	1	1	2	1	1
2	B	4	5	3	5	3
3	C	2	2	1	4	2
4	D	3	3	5	2	6
5	E	5	4	4	3	5
6	F	7	8	6	8	8
7	G	6	6	8	6	4
8	H	8	7	7	7	9
9	I	9	9	9	9	7

TABLE II: PHYSICS

A table showing the results obtained on a series of tests given near the end of the course, 1920-1921:

No. 1

No.	Name and rank on composite of 14 tests. Student.	Rank on Stanford revision of Binet-Simon test. ²	Rank (1) on silent reading-test (Burgess PS-1)	Rank (2) on Stenquist mechanical-ability test	Rank (3) on problem solving test V	Rank (4) on physics test MC (Biography)
1	A	6 (109)	4	3	1	1
2	B	9 (93)	9	1	4	2.5
3	C	3 (114)	2	4	2	6.5
4	D	4 (111)	6	2	3	8
5	E	7 (108)	3	8	5.5	5
6	F	1 (129)	7	6	8	2.5
7	G	2 (117)	1	5	5.5	4
8	H	5 (110)	8	7	9	6.5
9	I	8 (98)	5	9	7	9

²Tests given January to March, 1920.

No. 2

No.	Name and rank on composite of 14 tests. Student	Rank (5) on physics test X (Completion test on laws)	Rank (6) on physics test A (True-false on mechanics heat)	Rank (7) on physics test LP (Completion, whole course)	Rank (8) on physics test S (Problem solving)	Rank (9) on physics test T F (True-false, whole course)
1	A	1	1	1	1	1
2	B	2	4	2	7	3
3	C	3	3	4.5	2	4
4	D	7	6	3	4	2
5	E	4	2	6	8	8
6	F	5.5	5	7	5	6
7	G	5.5	7	4.5	3	5
8	H	9	9	8	9	7
9	I	8	8	9	6	9

No. 3

No.	Name and rank on composite of 14 tests. Student.	Rank (10) on technical information test (100 items)	Rank (11) by quality of experimental work.	Rank (12) by quality of home-written work.	Rank (13) by quality on paper on special topic.	Rank (14) on accuracy of numerical problems
1	A	2	1	1	1	1
2	B	1	2	2	2	3
3	C	3	5	6	3	4
4	D	5	9	8	4	6
5	E	7	3	3	9	2
6	F	4	4	4	6	5
7	G	6	8	9	7	8
8	H	8	6	5	8	7
9	I	9	7	7	5	9

There are some fundamental objections to the marks as ordinarily given upon the traditional examination papers:

1. There are large variations in teachers' marks of the same student, in the same subject, on the same examination. The variation is practically never less than 5% and is often as much as 15%. Few teachers can use a marking interval smaller than 8%.
2. The relative difficulty of the questions is unknown. One investigation shows a range of variation of 10 to 298 points on a single examination.
3. Due to "passing" grades and other causes, the usual examination grades are not consistent and are not distributed in accord with the known range of specific performance of students.
4. The mark given on a conventional examination does not measure *any single thing*. It measures a variety of specific performances involving four factors; i. e., quality of product, difficulty of task, time consumed, and quantity of product or work done.
5. In physical measurement the thing to be measured must be consistent, so that one unit is like another. Such consistency will be secured in educational measurements by arranging the test to measure a *single variable*. The other factors should be kept relatively constant.
6. Some additional objections to the traditional examination are:
 - (a). The strain produced upon students and teachers.
 - (b). The delay of promotion until the examination period.
 - (c). Only a few points of subject matter are covered in the test.
 - (d). It requires little care in preparation on the part of a teacher.
 - (e). It requires *too much* time to score.
 - (f). Students spend too much time in writing and too little in thinking.
 - (g). It has *no legitimate use* as a basis for determining how many pupils should fail in a subject.

III. Some advantages for the new types of test are:

1. The definite testing of some selected performance without leaving much to the personal equation of the examiner.
2. Freedom from strain of the two or three-hour examination.

3. Absence of the great labor of grading conventional "essay" examination papers.
4. Compels a decision as to what is being taught before undertaking to measure it.
5. The order of difficulty approximates the varying abilities of the class.
6. The wide range covered makes it possible to discover what topics have not been effectively taught.
7. The labor of scoring is reduced to a minimum.
8. Permits pupil to spend time thinking rather than in the mechanical work of writing.
9. Eliminates handwriting, style, etc., as disturbing factors in testing a specific performance.
10. Permits pupils to be tested on a specific performance, and excludes consideration of general ability in English composition.
11. Makes possible construction of another test next year of entirely different questions but of the same difficulty for testing a given performance.
12. The pupils *enjoy* these tests and the personal judgment of the teacher is not a factor in deciding the grade.

The chief disadvantages of the new tests are:

1. The units of thought may be too short.
2. The questions must be selected with great care, otherwise the poor student will make a high score.
3. Each pupil must have a copy of the test. This raises the item of expense, which is a factor in most schools.
4. Most teachers do not understand the methods to be used in preparing these tests, and many of them do not care to learn.

* References:

1. Burgess, May Ayres. *The Measurement of Silent Reading* (1921), Russell Sage Foundation, New York.
2. Rugg, H. O. Teachers' Marks and the Reconstruction of the Marking System. *Elementary School Journal*, 18, 701-719 (1918).
3. Trabue, M. R. Greater Objectivity in School Examinations. *Bulletin of New York Soc. for Exp. Study of Ed.* 2, No. 18. (1920).

PLATINUM PRODUCTION IN 1920.

Estimates of the production of crude placer platinum in the United States in 1920 give Alaska, 27 ounces; California, 656 ounces; Oregon, 23 ounces; and Washington, 8 ounces.

Refiners reported a production of 41,544 troy ounces of new platinum metals in 1920, of which 36,015 ounces was platinum, 418 ounces iridium, 409 ounces osmiridium, 4,309 ounces palladium, and 393 ounces of the minor metals, including rhodium and osmium. This represents a decrease of 3,565 ounces, as compared with the production in 1919.

There were also produced 57,710 ounces of secondary platinum metals in 1920, of which 51,255 ounces was platinum, 3,355 ounces iridium, and 3,100 ounces palladium. This represents a decrease of 3,806 ounces, as compared with the production in 1919.

The platinum metals imported for consumption in 1920 included 80,955 ounces of crude platinum and unmanufactured ingots, bars, sheets, etc., of which approximately 58,009 ounces is believed to be refined metal; 4,718 ounces of iridium; 4,473 ounces of osmiridium; 6,944 ounces of palladium; 2,675 ounces of the minor metals; and 781 ounces of manufactured platinum ware. The imports of platinum for consumption in 1920 were 32,492 ounces greater than in 1919.

The consumption of platinum metals in the United States in 1920 was 141,041 ounces, of which the jewelers took 57 per cent, the electrical industry 19 per cent, the dental industry 11 per cent, the chemical industry 10 per cent, and miscellaneous uses 3 per cent.