A COMPARISON OF SAMPLE SPLITTING METHODS

By Francis J. Flanagan and Richard C. Kellagher

Trace Elements Investigations Report 505

UNITED STATES DEPARTMENT OF THE INTERIOR

GEOLOGICAL SURVEY

IN REPLY REFER TO:

UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
WASHINGTON 25, D. C.

AEC-699/5

April 13, 1955

Mr. Robert D. Nininger, Acting Assistant Director
Division of Raw Materials
U. S. Atomic Energy Commission
Washington 25, D. C.

Dear Bob:

Transmitted herewith are three copies of TEI-505, "A comparison of sample splitting methods," by Francis J. Flanagan and Richard C. Kellagher, March 1955.

We are asking Mr. Hosted to approve our plan to combine this report with TEI-371 and submit it for publication in the Journal of Sedimentary Petrology. The Commission has already approved the publication of TEI-371.

Sincerely yours,

John H. Eric
for W. H. Bradley
Chief Geologist

Geology and Mineralogy

This document consists of 16 pages.
Series A.

UNITED STATES DEPARTMENT OF THE INTERIOR

GEOLOGICAL SURVEY

A COMPARISON OF SAMPLE SPLITTING METHODS*

By

Francis J. Flanagan and Richard C. Kellagher

March 1955

Trace Elements Investigations Report 505

This preliminary report is distributed without editorial and technical review for conformity with official standards and nomenclature. It is not for public inspection or quotation.

*This report concerns work done on behalf of the Division of Raw Materials of the U. S. Atomic Energy Commission.

USGS - TEI-505

GEOLOGY AND MINERALOGY

Distribution (Series A)	No. of copies
Argonne National Laboratory	1
Atomic Energy Commission, Washington	2
Battelle Memorial Institute, Columbus	1
Carbide and Carbon Chemicals Company, Y-12 Area	1
Division of Raw Materials, Albuquerque	1
Division of Raw Materials, Butte	1
Division of Raw Materials, Casper	1
Division of Raw Materials, Denver	1
Division of Raw Materials, Hot Springs	1
Division of Raw Materials, Ishpeming	1
Division of Raw Materials, Phoenix	1
Division of Raw Materials, Richfield	1
Division of Raw Materials, Salt Lake City	1
Division of Raw Materials, Washington	3
Dow Chemical Company, Pittsburg	1
Exploration Division, Grand Junction Operations Office	1
Grand Junction Operations Office	1
National Lead Company, Winchester	1
Technical Information Service, Oak Ridge	6
Tennessee Valley Authority, Wilson Dam	1
U. S. Geological Survey:	
Fuels Branch, Washington	1
Geochemistry and Petrology Branch, Washington	25
Geophysics Branch, Washington	1
Minerals Deposits Branch, Washington	1
E. H. Bailey, Menlo Park	1
A. L. Brokaw, Grand Junction	1
N. M. Denson, Denver	1
C. E. Dutton, Madison	1
W. L. Emerick, Plant City	1
A. H. Koschmann, Denver	1
R. A. Laurence, Knoxville	1
J. D. Love, Laramie	1
L. R. Page, Washington	1
P. C. Patton, Denver	1
J. F. Powers, Salt Lake City	1
C. B. Read, Albuquerque	1
Q. D. Singewald, Beltsville	1
A. E. Weissenborn, Spokane	1
TEPCO, Denver	2
TEPCO, RPS, Washington, (including master)	2
	74

CONTENTS

TABLES

A COMPARISON OF SAMPLE SPLITTING METHODS

By Francis J. Flanagan and Richard C. Kellagher

ABSTRACT

A sample design has been set up by which three methods of sample splitting for grain counting--the microsplit, the cone splitter and hand-quartering--may be compared. The methods of sample splitting are used for one classification in the design and the sample weights of 5, 10 and 20 grams of known grain composition for the other classification. Only one subset of data--that of the 5-gram sample split by the cone splitter--showed a value of χ^2 for precision that was equal to or less than that expected from chance alone. Similar calculations for estimates of accuracy show that the 5-g sample by the cone splitter and the 10-g sample by hand-quartering do not exceed the χ^2 values due to chance alone. If the subtotals of the 3 x 3 experimental design are used to calculate χ^2, it may be seen that: (1) the splitting of the 5-g sample by all methods is both more accurate and more precise than the splitting of the other two sample weights; and using the three weights for each method (2) the microsplit is the least accurate and the least precise of the three methods, (3) the hand-quartering is slightly more precise than the cone splitter and (4) the cone splitter is slightly more accurate than hand-quartering.

INTRODUCTION

In connection with studies of the monazite sands of the southeastern United States, made by the U. S. Geological Survey on behalf of the Division of Raw Materials of the U. S. Atomic Energy Commission, a large number of samples was submitted for grain-count analysis. To reduce the disproportionate amount of time being spent in splitting the samples, a cone splitter was designed (Kellagher, 1953), and preliminary tests of the operation of the splitter were made. The results of these tests proved not amenable to statistical analysis and further tests were set up in which the operation of two other splitting methods--the microsplit and hand-quartering--could be compared with the performance of the cone splitter. The microsplit used in the tests measures 1 in. by 1 in. at the top and has 14 alternating chutes.

Otto (1953), reporting test data obtained on three methods of sampling used eight ternary mixtures so selected that influencing factors could be studied nearly independently of each other, concluded that the microsplit was more accurate and more rapid than the improved methods of Krumbein and Pettijohn (Otto, 1933). Unfortunately Otto does not give the original grain composition of his mixtures. Wentworth and others (1934) devised a rotary type of sample splitter and after testing concluded that their splitter achieves a large improvement in accuracy over the Jones splitter. After comparing their deviations and those of Otto, they conclude that all methods are approaching the theoretical values.

We wish to acknowledge our discussions of this problem with R. M. Garrels and G. J. Jansen of the U. S. Geological Survey and to thank

W. J. Youden and W. H. Clatworthy of the National Bureau of Standards for their advice on the statistical solution of the problem.

TIME OF OPERATION

The time required for splitting a sample in half using a micro-split may be estimated at half a minute. A 500 grain sample of -80+100 mesh quartz weighs about 5 mg. To reduce a 5-g sample of this material to 500 grains would require about ten splits and consequently about six minutes. Hand-quartering requires a longer time to make ten splits. This time is sufficiently small to be of little consequence for a small number of samples, however, the sampling operation alone would require two days for 100 samples and more than two weeks for 1000 samples.

Using a 10-g sample, the cone splitter can reduce this sample to about 500 grains by taking two 2 1/2 percent splits of the sample. These two operations of the cone splitter take about one minute and result in a saving of about 75 percent in sampling time.

PRECISION AND ACCURACY

The cone splitter has been shown to have an advantage in that the time required for its operation is much less than that required for the other methods. This comparison of the rapidity of operation is a straightforward procedure and estimates of the time necessary for splitting may be made by mathematical considerations; the accuracy and precision of the method are not so simple to estimate. Shape, size, density, and numbers of grains may affect the splitting operation as may also the sample weight and the number of passes required for the final split.

Although most investigators of splitting operations use the weight percent or the volume percent obtained by splitting samples of unknown composition and term their results accuracy instead of precision, it is axiomatic that to estimate the accuracy of the splitting operation one must know the composition of the original sample to be split. The best estimate of this composition is the actual grain frequency percent of the constituents of the sample used. To obtain this composition six samples of approximately 600 grains of each of the pure minerals were counted and weighed. From these data the average weight per grain, its reciprocal the number of grains per gram, and the grain frequency composition were calculated (table 1).

Table 1.--Grain frequency composition of the test samples.

Mineral	Average wt/grain (g)	Number grains/gram	Grams/5 g of sample	Grain frequency (percent)
Quartz	8.46×10^{-6}	1.186×10^{5}	1	38.57
Ilmenite	17.36×10^{-6}	5.760×10^{4}	2	37.46
Monazite	27.13×10^{-6}	3.686×10^{4}	2	23.97

DESCRIPTION OF TESTS

A simple 3 x 3 two-way classification experimental design with triplicate replications was set up in which one classification was the methods used, that is, the microsplit, cone splitter, and hand-quartering, and the other classification was increasing weights of sample to be split. The three samples were artificial mixtures of quartz, monazite, and ilmenite, all -80+100 mesh in a weight ratio of 1:2:2, respectively, and weighed 5, 10 and 20 grams. Each sample was split to final or counting size (approximately 600 grains) by each method. All splitting operations were performed by one man. The final split was counted by operator A, checked for gross counting errors by operator B, and the split returned to the original sample before resampling. The grain counts of these final splits are shown in table 2.

CALCULATIONS

The observations made in this kind of test are total grain counts of each mineral constituent counted on a slide under a binocular microscope. As seen from table 2, the problem resolves itself into one of enumeration statistics, calculations for which are shown in any introductory statistics book, for example, Dixon and Massey (1951). The problem is further resolved into two distinct classifications: (1) the estimation of the precision of the splitting operation, and (2) the estimation of the accuracy.

Calculations in enumeration statistics may be carried out using the statistic χ^2 which is readily calculated by the formula

$$\chi^2 = \sum_{i=1}^{k} \frac{(G_O - G_E)^2}{G_E}$$

Table 2.--Grain counts of splits of three samples.

Method	Split	Operator	5 g Q	5 g I	5 g M	5 g Total	10 g Q	10 g I	10 g M	10 g Total	20 g Q	20 g I	20 g M	20 g Total
Microsplit	1	A	219	145	243	607	157	326	190	673	311	233	299	843
		B	223	144	243	610	174	328	190	692	310	232	292	834
	2	A	321	315	224	860	216	377	272	865	186	195	160	541
		B	340	330	230	900	198	374	268	840	180	209	149	538
	3	A	162	180	107	449	139	185	135	459	322	328	205	855
		B	166	183	108	457	133	185	132	450	310	328	209	847
Cone splitter	1	A	357	339	219	915	401	436	239	1076	249	217	138	604
		B	356	336	222	914	390	430	242	1062	240	216	135	591
	2	A	226	206	144	576	338	280	204	822	254	155	84	493
		B	219	208	141	568	324	269	210	803	243	158	83	484
	3	A	231	215	151	597	248	268	163	679	434	295	219	948
		B	236	214	151	601	242	247	169	658	420	288	227	935
Hand-quartering	1	A	201	217	178	596	238	240	150	628	219	257	149	625
		B	181	218	194	593	223	232	149	604	204	238	144	686
	2	A	418	308	248	974	237	250	169	656	270	180	121	571
		B	396	308	247	951	211	240	152	603	251	178	128	557
	3	A	262	248	187	697	233	197	110	540	347	189	132	668
		B	260	255	187	702	230	203	112	545	352	201	132	685

Average grain count = 695

Q - Quartz
I - Ilmenite
M - Monazite

where G_O is the observed grain counts of a mineral and G_E the expected counts. The expected counts are obtained from the experimental grain counts for precision and from the original sample composition for accuracy. For example, using the counting data of operator A for the 5-g sample split by the cone splitter and inserting the marginal totals as shown in the following tabular data,

Sample	Quartz	Ilmenite	Monazite	Total
1	357	339	219	915
2	226	206	144	576
3	231	215	151	597
Total	814	760	514	2088

one may obtain the expected number of quartz grains for sample 1 by multiplying the grain total of sample 1 by the ratio of the total quartz grains in the three samples, to the total number of grains. Numerically this would be

$$915 \times 814/2088 = 357.$$

This expected number of grains is used for estimates of precision. For the estimation of accuracy the ratio above is replaced by the proportion of the mineral in the original sample. The expected number of quartz grains then becomes

$$915 \times 0.3857 = 353.$$

A typical calculation of χ^2 for precision for the 5-g sample split by the cone splitter (operator A) is shown in table 3.

χ^2 thus calculated may be used in either a qualitative or quantitative manner. In the qualitative sense we may use the calculated chi squares, especially where they exceed table values, to rank the variables, in this case by the method of splitting or the sample weights.

Table 3.--Typical calculations of χ^2.

Sample no.	Mineral	Grains observed G_O	Grains expected G_E	$\frac{(G_O - G_E)^2}{G_E}$
1	Quartz	357	357	0
	Ilmenite	339	333	0.108
	Monazite	219	225	0.160
2	Quartz	226	224	0.018
	Ilmenite	206	210	0.076
	Monazite	144	142	0.028
3	Quartz	231	233	0.017
	Ilmenite	215	217	0.018
	Monazite	151	147	0.109

$$\chi^2 = \sum \frac{(G_O - G_E)^2}{G_E} = 0.534$$

To draw quantitative conclusions from the χ^2 values one must refer to a table of the χ^2 distribution to determine if the calculated values exceed the theoretical values. These tables have two variables (1) the probability interval that the investigator chooses, in this case, 95 percent, and (2) the degrees of freedom (d.f.) which are dependent on the experimental design. In normal enumeration problems where an estimate of precision is the goal, the degrees of freedom may be calculated from the expression d.f. = (n - 1)(k - 1) where the design has n rows and k columns of data.

Logically, the degrees of freedom may be developed as follows: Referring to the tabular data on page 10, there is originally one degree of freedom for each mineral in each sample, or in this case a total of 9. If we total the minerals across the rows, we use up one degree of freedom for each total as, knowing the total, only two of the three minerals can be independent. For three row totals we lose three degrees of freedom. The grand total may be then determined from the row totals without loss of a degree of freedom. As we now know the grand total, one column total will be determined once we know the grand total and the other two column totals. From the original 9 degrees of freedom we must subtract 3 lost in calculating row totals and 2 lost for columns total, leaving 4 degrees of freedom available for estimate of precision.

For estimates of accuracy, we know the grain frequency proportions, which, for precision, are calculated from the grain column totals and the grand total. Hence, we do not lose the two degrees of freedom associated with the column totals and we have therefore six degrees of freedom available for estimates of accuracy. The calculated χ^2 values for precision for each of the subsets of data in table 2 are shown in table 4 and those for accuracy in table 5.

If one compares the calculated χ^2 values for precision in table 4 with the theoretical values of 0.484 and 11.14 for the 95 percent confidence interval for 4 d.f., it is seen that the calculated value for only one of the subsets lies within the theoretical limits. It may be concluded then that the observed grain counts for this subset of data do not differ from those expected from chance alone.

Comparison of the values for accuracy, for which the limits for the 95 percent confidence interval for 6 d.f. are 1.24 and 14.45, shows that only two subsets of data show grain counts which differ from those expected from chance alone.

Table 4.--χ^2 for precision using subsets of counting data.

Weight (g)	Microsplit	Cone splitter	Hand-quartering
5	114.60	0.67	34.90
10	19.32	19.17	15.68
20	88.18	31.55	77.17

Table 5.--χ^2 for accuracy using subsets of counting data.

Weight (g)	Microsplit	Cone splitter	Hand-quartering
5	100.47	2.13	30.17
10	93.53	22.87	8.09
20	79.15	46.04	73.39

One may also draw some qualitative conclusions concerning the methods and the weights. The subtotals of each of the nine subsets of the data, the values for which are shown in table 6 may also be used to calculate χ^2 in the same manner as shown previously. As the data presented in this table represent all the data in the experiment we shall classify our calculations according to (1) methods of splitting and (2) sample weights, and each of these classifications may be treated with respect to (a) precision and (b) accuracy. The results of these two types of calculations for the nine subtotals are shown in table 7.

CONCLUSIONS

In order to test the cone splitter a series of samples with weights of 5, 10, 20 g of a mixture of three minerals has been split three times to grain counting size to compare the performance of three methods of sample splitting--the microsplit, the cone splitter and hand-quartering. All splitting was performed by the same person. Calculations of the statistic χ^2 for precision show that only the value for the 5-g sample split by the cone splitter does not exceed the value that might be due to chance alone. Similar calculations as estimates of accuracy show that the 5-g sample split by the cone splitter and the 10-g sample by hand-quartering do not exceed the χ^2 values due to chance alone. If χ^2 is calculated using the subtotals of these nine subsets with respect to both precision and accuracy--the individual methods by the three sample weights or the individual sample weights for all three methods--it may be concluded that (1) the splitting of the 5-g sample by all methods

Table 6.--Subtotals of data in table 2 tabulated by method and by sample weight.

Method	Microsplit				Cone splitter				Hand-quartering			
	Quartz	Ilmenite	Monazite	Total	Quartz	Ilmenite	Monazite	Total	Quartz	Ilmenite	Monazite	Total
5 g	702	640	574	1916	814	760	514	2088	881	773	613	2267
10 g	512	888	597	1997	987	984	606	2577	708	687	429	1824
20 g	819	756	664	2239	937	667	441	2045	836	626	402	1864
Total	2033	2284	1835	6152	2738	2411	1561	6710	2525	2086	1444	5955
Weight	5 grams				10 grams				20 grams			
Micro-split	702	640	574	1916	512	888	597	1997	819	756	664	2239
Cone splitter	814	760	514	2088	987	984	606	2577	937	667	441	2045
Hand-quartering	881	773	613	2267	708	687	429	1824	836	626	402	1864
Total	2397	2173	1701	6271	2207	2559	1632	6398	2592	2049	1507	6148

Table 7.--χ^2 for precision and for accuracy, calculated by methods and by sample weights.

	Methods			Weights		
	Microsplit	Cone splitter	Hand-quartering	5 g	10 g	20 g
Precision	183.30	57.88	55.05	27.06	195.94	177.44
Accuracy	435.72	82.78	90.33	110.79	287.99	210.06

is both more accurate and more precise than the splitting of the other two sample weights; and using the three weights for each method (2) the microsplit is the least accurate and the least precise of the three methods, (3) the hand-quartering is slightly more precise than the cone splitter and (4) the cone splitter is slightly more accurate than hand-quartering.

REFERENCES

Dixon, W. J., and Massey, F. J., Jr., 1951, Introduction to statistical analysis: New York, McGraw-Hill Book Company, Inc.

Kellagher, R. C., 1953, A multiple cone splitter for rapid sampling in the laboratory: U. S. Geol. Survey Trace Elements Inv. Rept. 371.

Otto, G. H., 1933, Comparative tests of several methods of sampling heavy mineral concentrates: Jour. Sedimentary Petrology, v. 3, p. 30-39.

Wentworth, C. K., Wilgus, W. L., and Koch, H. L., 1934, A rotary type of sample splitter: Jour. Sedimentary Petrology, v. 4, p. 127-138.