IN DEFENSE OF THE SECOND DECIMAL

FELIX CHAYES, *Geophysical Laboratory, Washington, D. C.*

ABSTRACT

Synthetic and practical examples are used to show that excessive rounding must be avoided whenever quantitative results are to be subjected to extended numerical analysis. In work reporting original measurements it is usually preferable to record the complete observation. Where this is not desirable, it should always be possible for an interested reader to reclaim the original observations. This can not be done if the hundredths place is rounded out of conventional silicate analyses or the tenths place is rounded out of modal analyses.

The problem of the extent to which measurements are to be carried and recorded plagues every natural science and has no general solution. In fact, it is not a general problem at all, but an infinite family of specific problems, each of which must be faced in its own context. Basically these problems are psychological, and the mathematical or physical conventions we adopt in treating them reflect the demands of the situations which give rise to them.

For a long time petrologists and mineralogists seem to have been unduly impressed by the precision of chemical analysis. As late as the first decade of the 20th century, silicate analyses were not infrequently stated to three places, and sometimes a fourth was added for good measure. The pendulum has now swung the other way; and in a summary of the most extensive study of the reproducibility of silicate analyses so far made, the standard practice of quoting through the second place has been strongly questioned (Fairbairn, 1951). Recently a similar attack has been made on the fairly common practice of retaining the first decimal in statements of modal analyses (Hamilton, 1952). These objections reflect a rather widespread skepticism but are based on reasoning which I feel is not applicable. Both Fairbairn and Hamilton base their argument on the *interpretation* of measurements, and while this is no doubt a major consideration for every reader, the immediate problem facing the the author of the measurements is their *presentation*.

In the standard rock analysis, weighings are recorded to the "nearest tenth" of a milligram and the principal fusion, out of which come the values for silica, alumina, lime, magnesia, titania, and total iron, is carried through on a one gram sample. A tenth of a milligram is one hundredth of a per cent of a gram, and when the chemist tells us that he finds 73.24 per cent of silica in a rock, he is telling us what he has observed. He does not say that he or any other chemist would get exactly 73.24 per cent on a second sample, but *only that he got 73.24 per cent on this one.* If he reported 73.241, he would be telling us more than he could

possibly observe in a single analysis. If he reported 73.2 he would not be telling us all that he observed, and that, after all, is the first thing we want to know.

Before asking him to run the analysis we probably guessed that the sample contained between 70 and 75 per cent of silica, and he probably was informed of this guess. We assume that composition is a continuous variable; and in asserting that the sample carries between 70 and 75 per cent of silica, we imply that the true silica *may* have any value whatever —to any number of decimals—between 70.000. . . and 75.000. . . . The analyst's position is a little different. It is his job to make observations, not assumptions, and he has no way of making direct observations past the second decimal place unless he works on a 10 gram sample. Even when he is favored by a large gravimetric factor, as when he weighs magnesium phosphate and reports MgO, the last entry which may have any value at all between 0 and 9 is the hundredths place. To him this is the observational limit of continuity, and quite sensibly he records his result to that limit. This is what we want him to do, or what we should want him to do.

For reasons concerned more with the interpretation than the presentation of data, Dr. Fairbairn (1951, p. 70) feels that we " . . . should ruthlessly discard unnecessary detail beyond the decimal point," and the sense of my own contribution to Bulletin 980 (Fairbairn 1951, p. 64) is the same, though the wording is less forthright. On reconsideration, however, I am convinced that from the point of view of representation—a point of view too seldom considered by geologists—the premature mutilation of the chemist's report is most unwise. And whether for interpretive purposes the detail beyond the decimal point is unnecessary, whether, in fact, it is to be regarded as detail at all, will depend entirely on what information we wish to obtain from the data and how we propose to obtain it.

Dr. Hamilton's argument that modal analyses should not be quoted beyond the nearest per cent is based on the same reasoning which led to Dr. Fairbairn's explicit and my implicit dictum concerning the hundredths place in chemical analyses. It is open to somewhat the same objection, and for one of my experience it is much simpler to examine the problem in terms of modal analysis. The discussion which follows, therefore, centers on the advisability of retaining the first place after the decimal in modal analysis; but most of the argument is directly applicable to the question of whether or not the second place should be retained in conventional silicate analyses.

For count lengths between 1500 and 2000 points the theoretical precision for major constituents, as a standard deviation, is about 1 per cent

of the whole. There is thus some justification for discarding the tenths, and this is particularly so if we merely wish to scan the table and declare solemnly that the rock either is or is not "variable." But the validity of the table *as a record of the observations* is impaired by this process, and if any serious evaluation of the data is to be attempted the results may vary widely and erratically depending on whether the calculations have been performed before or after rounding.

Suppose, for instance, that an analysis of count-length 1502 yields the frequencies

Quartz	449
Microcline	502
Plagioclase	530
Biotite	21

A single count is 1/1502 or .03 per cent of the whole. In the hundredths place the analysis thus moves in units of .03, and the first entry which may have any value regardless of its neighbor to the left is in the tenths column. When we see a value of 10.03 per cent, for instance, we immediately assume that it *might* have been 10.02 or 10.04, and much of the meaning we attach to the result is based on this assumption. But for our chosen count length the two nearest neighbors of 10.03 are 10.00 and 10.06. The entry in the tenths place controls the entry in the hundredths. But the entry in the tenths place is not at all influenced by that in the per cent column; what we may call the "continuous least count" of the procedure is thus 0.1 per cent, just as it is 0.01 per cent in the standard silicate analysis. Reducing the observed frequencies to the nearest relative least count we have:

Quartz	29.9
Microcline	33.4
Plagioclase	35.3
Biotite	1.4

With the tenths rounded out this becomes:

Quartz	30
Microcline	33
Plagioclase	35
Biotite	1

which gives a total of 99. Now if the table is to be an adequate record of the observations, we must be able to reclaim the original frequencies from it. This may be attempted in any of three ways. First, we may simply ignore the missing 1 per cent. Alternatively, we may "recalculate" the recorded analysis to 100 before reconstructing the observed frequencies, and this may be done in either of two ways. If we are still adamant about the tenths place, we merely assign the missing per cent to the largest

value, plagioclase. We may feel, however, that our calculation is to be preferred to the analyst's observation, and in this event we shall regain the lost per cent by adding 4 parts of it to plagiocase and 3 each to quartz and microcline. (In fact every one of these procedures assumes that guessing is better than observing, but the assumption is more obvious in the last case.)

Frequencies reconstructed from the original statement and each of the rounded versions are shown in Table 1. If the tenths place is left

TABLE 1. OBSERVED AND RECONSTRUCTED FREQUENCIES

Mineral	*A*	*B*	*C*	*D*	*E*
Quartz	449	449	451	451	456
Microcline	502	502	496	496	500
Plagioclase	530	530	526	541	532
Biotite	21	21	15	15	15

A—Observed.
B—Recalculated from total and percentages carried to tenths.
C—Recalculated from total and rounded percentages, ignoring closure error.
D—Recalculated from total and rounded percentages, assigning lost per cent to most abundant constituent.
E—Recalculated from total and rounded percentages, distributing lost per cent over three major constituents.

standing in the analysis, the combination of total count-length and percentage composition permits reconstruction of the observed frequencies, so that the frequencies themselves need not be recorded. If the tenths place is rounded out,[1] the observations can not be reclaimed correctly from the table. If we then want to know what the analyst found—something the table is supposed to tell us—we must write him a letter.

It may be argued, and no doubt for specific purposes it is sometimes true, that whether the analyst recorded 526 or 541 counts of plagioclase, or 15 or 21 counts of biotite, is a rather trivial affair. Correct as it may be in particular cases, however, this attitude betrays a fundamental lack of sympathy for the relation between tables which record data and arguments which interpret them. Rounding which makes recovery of the original observations impossible is a form of interpretation and should be done by text discussion, by separate tables, or not at all. But the writer presenting original data will have fallen short of his responsibility if an interested reader can neither find the measurements nor reclaim them from information recorded in the paper. If a table of percentages is to provide an adequate record of observations, most of the conventional

[1] 5's have been rounded to the nearest number.

rounding, e.g. rounding dictated by the results rather than the experimental procedure, must be left for the reader to do. The comforting sense of superiority which creeps over us when we examine unrounded results published by others is suitable recompense for our performance of this service.

Regarding a table simply as a description or record, the damage done by excessive rounding may be summarized as a blurring or actual deletion of the observations. If we plan to do more than publish or inspect the table, however, the effects of over-rounding may be far more complex.

Suppose we are given a string of ten results differing by 0.1, e.g.,

$$60.1,\ 60.2,\ 60.3,\ \cdots,\ 61.0.$$

The standard deviation for such a run is 0.303 and does not vary with the numbers to the left of the decimal point, or with the first number of the string on the right. Thus the series $hk\cdot 0\ \cdots\ hk\cdot 9$ has exactly the same standard deviation as the series $hk\cdot 5\ \cdots\ h(k+1)\cdot 4$. This holds for any index of dispersion and is the basis for the coding procedures which so often facilitate extended numerical analysis. If the tenths place is rounded out, each entry in the series becomes hk or $h(k+1)$, and if we next subtract hk from each member, any such series is resolved into a run of 0's followed by a run of 1's which, we may represent by

$$x(0)y(1),$$

where $0 \leqq x \leqq 10$ and $y=10-x$. The series we began with would now be written

$$0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 1,\ \text{or}\ 5(0)5(1).$$

The standard deviation of the rounded run obviously varies with the value of x, and this in turn is controlled by the first term, in the fashion shown in Table 2.

TABLE 2. EFFECT OF FIRST TERM OF SERIES ON VARIANCE (V) AND STANDARD DEVIATION (s)

1st term	x	y	V	s
$hk.0$	5	5	.278	.527
$hk.1$	4	6	.267	.516
$hk.2$	3	7	.233	.483
$hk.3$	2	8	.178	.422
$hk.4$	1	9	.100	.316
$hk.5$	0	10	0	0

If in a group of such series the initial terms are equally likely, a condition which could be met only if the continuous least count of the pro-

cedure yielding them were not larger than 0.1, the average variance obtained after rounding would be .183 as compared with .092 computed directly from any one of the unrounded series *regardless of its initial term*. This would amount to an average upward bias of more than 30 per cent in the standard deviation, produced entirely by rounding. If this bias were constant we might even accept it as providing an extra margin of safety in certain situations. But this is very clearly not the case; rounding will sometimes lead to an underestimate of dispersion and will quite frequently overstate it by considerably more than the long run average of 30 per cent. In fact, the probability of obtaining the exact average bias in a single run is zero.

TABLE 3. BIOTITE VALUES, WESTERLY TEST STRIP*

2.7	2.2	3.0	3.4
3.1	3.8	4.2	3.8
3.7	3.5	3.2	3.3
2.3	3.2	2.7	2.8

* From Table 25, *U. S. Geol. Survey, Bull.*, **980.**

This example is of course highly synthetic. The result could be generalized by repeated enumeration and summation, but the procedure would no doubt alienate most of the audience for which this note is intended, without making any original contribution to analysis of the purely statistical problem involved. On the whole it will be best to leave mathematics to the mathematicians and pass at once to a practical illustration. Modal analyses of the test strip of Westerly granite, used as one of the materials in the analytical testing program described in Bulletin 980 and cited by Dr. Hamilton as a particularly egregious instance of "pseudo-precision," will serve as an example. The biotite values recorded in Table 25 (p. 61) of the Bulletin are shown below in Table 3. The mean biotite content for the strip is 3.2 per cent, the standard deviation calculated from the unrounded data—e.g. the data recorded to the continuous least count—is 0.5552 per cent, and the variance is 0.3083.

As with the hypothetical example, we round the data as recorded, add 0.1 to each of the original values and round again, add 0.2 and round once more, and so forth. From each rounded series we next calculate the variance and standard deviation. The results are shown in Table 4.

The measurements were made in an attempt to detect inhomogeneities in the test material prior to crushing. Do the data of Table 3 indicate significant sample variation with regard to biotite content? The theoretical analytical error of the counting procedure is binomial, and for an

TABLE 4. EFFECT OF ROUNDING ON VARIANCE (V) AND STANDARD DEVIATION (s) OF DATA IN TABLE 3

	V	s
Original data, unrounded	.3083	.555
Original data, rounded	.4292	.655**
Original data plus 0.1 and rounded	.4667	.684**
Original data plus 0.2 and rounded	.4958	.704**
Original data plus 0.3 and rounded	.4000	.632*
Original data plus 0.4 and rounded	.3625	.602*
Original data plus 0.5 and rounded	.3333	.577
Original data plus 0.6 and rounded	.3333	.577
Original data plus 0.7 and rounded	.2958	.544
Original data plus 0.8 and rounded	.3292	.574
Original data plus 0.9 and rounded	.3833	.619*

average content of 3.2 per cent and an average count length of 1500, the expected variance attributable to analytical procedure is 0.206. This is readily tested against the observed variances, before and after rounding, and the results are shown by the asterisks forming the last column of Table 4. A double asterisk indicates that the excess of observed over expected variance is highly significant, a single asterisk that the excess is suggestive but not conclusive, and a blank that an analytical error alone might account for the observed variation. To the original question the double asterisk thus answers with a firm "yes," the single asterisk says "perhaps," the blank says "no," or, at any rate, "not found." The results after rounding are highly variable; there are 3 double asterisks, 3 single asterisks, and 4 blanks. The occurrence of a blank in the first line and a double asterisk in the second is particularly striking; no evidence of sample variation is found in the original unrounded data, but the operation of rounding makes it appear that observed variation is considerably larger than can be accounted for by analytical error alone. Now the addition of a constant has no effect on dispersion, and it is dispersion with which we are concerned. If we had refrained from rounding, the entries in each column would have been identical, from top to bottom of the table, *as they should be.* Variation in variance, and hence in standard deviation, is generated entirely by rounding.

In this case the apparently innocuous practice of rounding leads to a conclusion exactly opposite to that given by the unrounded data; the rounded values actually mislead us. In other cases the effect of rounding will be merely to throw the data out without seeming to do so. In the tenuous atmosphere in which "broad" or "fundamental" geological problems are discussed, there may often be no practical difference be-

tween rounded and unrounded data—but my own feeling is that this happy circumstance usually arises when the data, whether rounded or not, are no more than window dressing.

The experimental situation which led to the modal analyses we have been discussing is forthright, and the hypothesis being tested could hardly be simpler. With reasonably complicated experimental designs the uncertainty generated by rounding propagates at an alarming rate and in a fashion both devious and complex. Data which have not been butchered by improper rounding are an absolute prerequisite for numerical analysis of even moderate maturity.

Now the effect of all this is bound to be rather puzzling. For most geologists, statistics is still an extracurricular activity, and one of its major charms is that it reinforces our skepticism. In mineralogical and petrographic subjects, and to some extent in structural geology as well, the earliest focus of skepticism is likely to be some field of mensuration. And our first urge is to " . . . ruthlessly discard unnecessary detail beyond the decimal point." At this stage of development we are concerned chiefly with the interpretation of data and regard the information yielded by tests of precision or reproducibility as offering an objective index of how we are to think about data obtained under less exacting circumstances. We tolerate the "unnecessary" detail only because by using it we can show that in certain situations it may in fact be unnecessary or even misleading. The experimental designs into which we cast our data are still likely to be very simple, so simple that statistics is something of a sheep in wolf's clothing. For the most part we are still making important discriminations by inspection and merely reinforcing them by calculation. In work of this type it is certainly desirable to eliminate unnecessary detail.

As we make the numerical analysis do more and more of our work, the situation changes. The information we formerly preferred to carry in our heads is now removed from the data, where it was once either unnecessary or potentially misleading, only at the peril of rendering the whole analysis invalid.

The point at which we stop recording and start interpreting our measurements is not as clearly defined as we like to suppose, and its location will surely vary with both the data and the observer. As a practical guide it may help to remember that rounding can be done at any time by anyone, that only one person can *un*round, that even he can do so only for so long as he preserves his notebooks, and that unrounded or inadequately rounded data mislead only those willing to be misled. The unwary should not be wasting their time reading technical literature.

The primary purpose of a table presenting new data is to record ob-

servations; if the original observations can not be read directly or reconstructed correctly from the table, they are lost. This is just what happens when the hundredths place is rounded out of conventional silicate analyses, or the tenths place out of modal analyses based on 1000 or more counts. If the loss of the original measurements is a matter of no consequence, it is fair to suppose that there was no real need to make them in the first place.

Finally, I believe there is a strong relation between imagined and realized precision. If a measuring procedure is intrinsically bad, no amount of enthusiasm and labor will help much. But many types of measurement we use in geology are neither intrinsically good nor intrinsically bad; they are sound if we use them properly and unsound if we do not. The prior conviction that such a method is "rough" or "approximate" practically assures that the results will be much worse than that. Whatever may be the case with football players, the good professional naturalist is always an amateur willing to make the old college try. We have seen an example of the damage done by rounding prior to calculation. The man who makes "rough" measurements by an "approximate" method to "determine an order of magnitude" is guilty of rounding prior to observation. The attitude and procedure are sometimes unavoidable, but that is their only recommendation.

I recall dimly that there is always an error of closure in mapping. From brief but disastrous experience I also recall that this error may be very large unless one is willing to fight all the time to keep it small. An index of refraction determination is not likely to be much good in the third place unless the mineralogist who made it did considerable fussing over the fourth. A petrographer overimpressed by a precision error of one per cent is likely to produce modal analyses which do not even report the order of dominance of the major constituents correctly. For this last type of error, which I have encountered frequently and described at length elsewhere (Chayes, 1952, pp. 226–232) there seems to be no completely rational explanation. No doubt it is sometimes due to faulty material or inadequate technique. But that is by no means the whole explanation, for in some cases the equipment was suitable and the procedure—on paper—proper. What seems to have been lacking was the firm conviction that good measurements are worth making. The absence of this conviction—what I have called above rounding prior to observation—is even more wasteful than rounding prior to calculation. A too easy tolerance of the second practice creates a mental set in favor of the first, with the result that we play with numbers instead of working with them or making them work for us. This is not a statistical problem and the statistician usually evades or ignores it, flattering us and clearing

his own skirts by supposing that we think our measurements worth making.

Obviously the psychological fault of rounding before observation should be avoided wherever possible. Refraining entirely from the use of numbers is often preferable to the all too common practice of dressing up our guesses in them. The measurement we are about to make should always be the best of its kind ever made.

Rounding which makes the class interval broader than the unit of measurement—the continuous least count where that is deducible—should never be performed before the data are recorded and the calculations run.

What this comes to is that tables recording original analytical data will often contain considerable "unnecessary detail" to the right of the decimal point. It is much better to be charged with "pseudo-precision" than to be guilty of pseudo-skepticism.

References

CHAYES, F. (1952), The finer-grained calcalkaline granites of New England: *Jour. Geol.*, **60**, 207–254.

FAIRBAIRN, H. W., et al. (1951), A cooperative investigation of precision and accuracy in chemical, spectrochemical and modal analysis of silicate rocks: *U. S. Geol. Survey, Bull.*, **980**, 71 pp.

HAMILTON, W. B. (1952), Precision of geological data: *Bull. Geol. Soc. Am.*, **63**, 323–324.

Manuscript received Nov. 15, 1952.