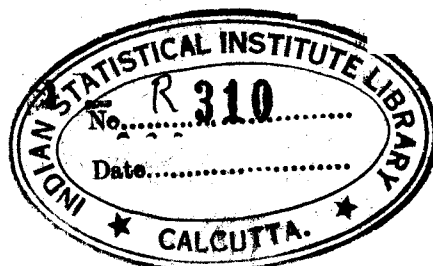# AGENDA FOR

SECOND CONFERENCE ON ACCURACY AND PRECISION

OF CHEMICAL ANALYSES

by

Walter A. Shewhart

Conference with Messrs. L. A. Wooten of B.T.L., W. B. Tall of W. E. Company and members of their staffs to be held at Murray Hill early in November, 1946.

INTRODUCTION

Object of Investigation

The original object of the investigation is well stated in a letter of November 18, 1944 from L. I. Shaw to B. L. Clarke as follows:

> "Our main concern is to find an acceptable measure of the overall precision of the complete analytical process, as performed in practice in our laboratories, rather than the best precision obtainable by a certain method, under ideal conditions. In discussions with our inspection organizations or other interested persons, we would then be able to state that if a particular result, for example, which was near the specification limit, were to be checked by additional analyses on the same sample, these could be expected to agree with the first result to within certain limits. Also in comparing a number of analyses made over a period of time on different lots of the same material, we would be able to state that a certain variation in the results did or did not indicate an actual variation in the material."

Conference on February 22, 1945

The first conference between representatives of the W. E. Company and B.T.L. was held at Murray Hill on February 22, 1945. The first three pages of the agenda for that meeting outlined briefly the problem of measuring overall precision, and suggested some steps that might be taken in carrying out a cooperative program of measuring such precision and of using the results in routine chemical analyses of materials. The suggestions therein set forth were considered in some detail and it was informally agreed to use them as a guide in making a study of the methods of chemical

analysis of some material to be agreed upon later by B.L. Clarke and L. I. Shaw. Shortly thereafter, 45% permalloy was chosen as the material. In what follows, the results of the statistical analysis of the chemical analyses of 45% permalloy for manganese, iron, nickel, carbon, and silicon are set forth in accord with the outline given in the previous agenda.

Sample of Material

As a matter of record, it should be stated that the sample of material was (as I understand) collected at Hawthorne in the form of drillings and turnings. This material was placed on a large sheet of paper and an attempt was made to mix it thoroughly. Then approximately one-half of the material was sent to the Laboratories and the other half was kept by Hawthorne. Subsamples for chemical analyses were taken from these two portions of the original sample. It is my understanding that the sizes of the particles present in the original sample varied over a wide range.

Results of Chemical Analyses

Most of the results of the chemical analyses are shown in Table 1. Each analysis was repeated eight times and the sequence of values thus obtained is shown in the table, reading from top to bottom for each set of eight. The BTL and the fourth set of Hawthorne for iron shown in this table were made with a 1 ml. indicator, whereas the other three

Table 1

PER CENT MANGANESE, IRON, NICKEL, CARBON, AND SILICON IN 45% PERMALLOY

	Laboratory B Analyst								Laboratory H Analyst						
	ALB	CLL	TPJ	WEB	GSM	WCJ	MEC	—	BJS	MIN	JAS	WEB	CWD	HJM	JD
% Manganese	.804	.787	.798						.764	.762	.760				
	.799	.774	.789						.786	.739	.744				
	.776	.765	.804						.768	.742	.770				
	.770	.775	.785						.762	.770	.785				
	.786	.771	.785						.758	.798	.768				
	.802	.767	.803						.770	.799	.758				
	.791	.785	.797						.775	.766	.735				
	.780	.826	.801						.772	.773	.745				
% Iron			53.95	53.80	53.95				53.46		53.58	53.85			53.60
			53.85	53.87	53.74				53.43		53.51	53.68			53.67
			53.85	53.94	53.71				53.51		53.40	53.71			53.55
			53.79	53.71	53.76				53.54		53.40	53.79			54.08
			53.76	53.68	53.76				53.55		53.50	53.81			54.42
			53.74	53.85	53.63				53.45		53.60	53.60			54.33
			53.76	54.06	53.74				53.57		53.78	53.90			54.38
			53.70	53.76	53.80				53.47		53.56	53.81			54.32
% Iron - K3-4574			53.58	53.65											
			53.58	53.65											
			53.50	53.70											
			53.50	53.57											
			53.54	53.58											
			53.60	53.66											
			53.55	53.66											
			53.56	53.44											
% Nickel			45.35	44.96	45.13				45.21				44.28	44.96	
			45.28	45.14	45.14				44.95				44.88	45.05	
			45.31	45.01	45.24				45.21				45.19	45.21	
			45.19	45.18	45.38				44.98				45.07	45.04	
			45.24	44.90	45.15				45.17				45.22	45.09	
			45.32	44.97	45.10				45.25				45.16	45.04	
			45.38	44.96	45.38				45.22				45.14	45.08	
			45.31	45.05	45.07				45.42				44.60	45.15	
% Carbon H - gravimetric Mfg. Std. 17000				.044		.047	.047	.048	.040	.059					.050
				.041		.058	.046	.041	.038	.055					.049
				.045		.043	.048	.046	.041	.050					.059
				.040		.042	.040	.047	.037	.050					.057
				.058		.058	.042	.041	.043	.045					.017
				.040		.040	.046	.059	.044	.043					.045
				.056		.041	.050	.041	.058	.050					.054
				.058		.056	.047	.059	.051	.046					.058
% Carbon H - Volumetric B - A.S.T.M. E39-42T									.034						
									.032						
									.037						
									.037						
									.044						
									.044						
									.045						
									.038						
% Silicon H - Residue B - Mfg. Std 17000				.151	.150		.161		.151					.160	.152
				.138	.144		.150		.155					.162	.140
				.134	.146		.154		.158					.156	.144
				.146	.146		.120		.155					.158	.154
				.146	.174		.155		.147					.159	.159
				.141	.158		.156		.159					.157	.159
				.147	.146		.153		.140					.164	.147
				.146	.119				.154					.145	.156
% Silicon H - HF treatment B - K3-4574				.151				.160	.158					.160	.148
				.138				.157	.157					.160	.140
				.134				.154	.150					.153	.138
				.146				.150	.151					.154	.159
				.146				.119	.149					.164	.137
				.141				.121	.150					.163	.167
				.147				.153	.141					.162	.156
				.146				.130	.154					.159	

Hawthorne sets for iron were made with a 10 ml. indicator. The BTL results obtained with a 10 ml. indicator are not shown in the table but are given in Fig. 6.

ANALYSIS OF DATA

Problem from Viewpoint of Analysis of Data

In order to answer the questions raised in the paragraph quoted above from Mr. Shaw's letter of 11/18/44, it is necessary to analyze the total variance of each of the sets of 48 measurements on each of the chemical constituents Mn, Fe, Ni, C, and Si. This involves the following two steps:

Step 1. Breakdown of total variance into components attributable to <u>known</u> factors or sources of possible variation such as:

1.1 Differences between repetitive observations made by an observer.

1.2 Differences between observers.

1.3 Differences between laboratories.

Step 2. Detection of the presence of <u>unknown</u> factors or sources of variation within each set of repetitive observations.

Stated in another way it is apriori reasonable to assume that the estimated total variance

$$V = \frac{\sum_{i=1}^{48} (x_i - \bar{x})^2}{48-1} \qquad (1)$$

for a given determination may be contributed to by the three
kinds of factors listed above. In this sense these three
factors are "known" as possible sources of variation. For
example, it is quite likely that if two observers were to
take an indefinitely large number of observations, the averages
of their results would differ by more than might be explained
as a random fluctuation upon the assumption that the expected
values for the two were identical. This is particularly true
in experiments where volumetric and/or "color-end-points" are
involved. Likewise differences may be introduced by different
laboratories through differences in chemical reagents, and
equipment as well as differences in minor details of conduct-
ing a chemical analysis.

It is but natural therefore to start with an analysis
of the overall or total variance to segregate the components
of variance attributable to such "known" factors and to deter-
mine if the observed differences between these components differ
by more than may reasonably be attributed to chance. The
statistical methods of carrying out such an analysis are those
customarily described in statistical literature under the head
of "analysis of variance".

We must not stop, however, with such an analysis to
check whether such known potential sources of variation do in
fact contribute measurably to the overall variation (or pre-
cision to use the term introduced in Shaw's letter) because

there may and usually does exist one or at most a few unknown factors. For example, let us consider any one of the sets of eight repetitive measurements by a given observer. It is customary in the analysis of variance to assume that such a set constitutes a random sample. Work within the Bell System has shown, however, that this assumption is seldom justified except as a first approximation. In other words, there usually exist unknown <u>assignable</u> causes of variation.

For example, attention has already been directed to the fact that the sample of material weighed out consisted of some fairly large particles of material. Unless, therefore, permalloy is a <u>strictly homogenous</u> material it is possible that the percentage of a given component such, for example, as iron may vary from one weighed out portion to another by an amount greater than can be reasonably considered as a random fluctuation. Perhaps also it is possible that some of these small particles may have picked still smaller particles of iron that did not come from the permalloy. That is to say, if the total sample of permalloy collected at Hawthorne could have been put into solution and then samples could have been weighed out and distributed to the six observers at both laboratories, the variance between samples might well have been less than exists when test samples of comparatively large pieces of solid material

are weighed out and distributed. Other possible sources of variance may readily suggest themselves to the reader.

Then too it is possible, in some kinds of chemical analyses at least, that the reagents used and even the material under test may change during the time that the series of repetitive measurements are being made. Likewise, the successive results obtained by an observer may be correlated because the observer allows the measurement made at any point in the series to influence him in making the next measurement. This may happen particularly in volumetric methods and perhaps in certain others.

In general, two types of statistical methods are required in analyzing a sequence of repetitive measurements schematically represented by

$$x_1, \; x_2, \; x_3, \; \ldots, \; x_n \qquad (2)$$

where the order in the sequence is that in which the observations were made. These two types are:

A. Examination of certain characteristics of the distribution of such numbers irrespective of the order that they occur in sequence (2).

B. Examination of certain characteristics of the sequence (2) taking account of the order in which the observations were made.

We shall see in what follows that the analysis of variance reveals that some of the known potential factors do contribute to the variance and that the two types of statistical analysis considered in the previous paragraph do reveal the existence of at least two kinds of unknown factors.

Had we, for example, found no evidence for believing that known and unknown factors contributed to the total variance, the answer to Mr. Shaw's questions would have been quite simple. For example, the best estimate of "overall precision" would simply be the total variance given by (1) multiplied by a constant. How this answer is modified by our findings will appear later in this memorandum.

ANALYSIS OF VARIANCE

Variance between Observers and Variance within Sets of Observations

Let us first consider the breakdown of total variance into two parts, namely, that between observers and that within sets of eight repetitions.

Let N = total number of observations.

n = number of observations per observer.

m = number of observers.

v_1 = variance within sets of observations.

v_2 = variance between observers.

v = total variance.

x_{ij} = jth observation for ith observer.

$\bar{x}$ = average for all of observations.

$\bar{x}_{i\cdot}$ = average for n observations for the ith observer.

s_0 = estimate of standard deviation of differences between observers.

s_e = estimate of standard deviation within sets of observations.

Then $(N-1)v = m(n-1)v_1 + (m-1)v_2$ \hfill (3)

or

$$\sum_{1=1}^{m} \sum_{j=1}^{n} (x_{ij} - \bar{x})^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} - \bar{x}_{i\cdot})^2 + n \sum_{i=1}^{m} (x_{i\cdot} - \bar{x})^2 \quad (4)$$

In terms of this notation we then have the breakdown shown in Table 2.

Source of Variation	Sum of Squares	Degrees of Freedom	Variance
Within sets of n observations	$\sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij}-x_{i\cdot})^2$	$m(n-1) = N-m$	$v_1 = s_e^2$
Between observers	$n \sum_{i=1}^{m} (\bar{x}_{i\cdot} - \bar{x})^2$	$m - 1$	$v_2 = ns_0^2 + s_e^2$
	$\sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij}-\bar{x})^2$	$mn-1 = N-1$	$v = s^2$

Table 2

ANALYSIS OF VARIANCE - ALL OBSERVERS

		% Mn	% Fe	% Ni	% C	% Si
Within sets of measurements	42 degrees of freedom	Sum of squares .01014325	.36966250	1.2129000	.00140150	.00360115
		v_1 .00024151	.00880149	.02887857	.00003337	.00008574
Between observers	5 degrees of freedom	Sum of squares .00759475	.95503542	.62286667	.00108517	.00137685
		v_2 .00151895	.19100708	.12457333	.00021703	.00027537
Total	47 degrees of freedom	Sum of squares .01773800	1.32469792	1.83576667	.00248667	.00497798
		v .00037740	.02818506	.03905887	.00005291	.00010591
$F = \dfrac{v_2}{v_1}$ or $\dfrac{v_1}{v_2}$		6.29**	21.70**	4.31**	6.50**	3.21*
$s_e = \sqrt{v_1}$		.0155	.0938	.1699	.0058	.0093
$s_o = \sqrt{\dfrac{v_2 - v_1}{8}}$		.0126	.1509	.1094	.0048	.0049
$s_{oe} = \sqrt{s_e^2 + s_o^2}$		.0200	.1777	.2021	.0075	.0102
$s = \sqrt{v}$		.0194	.1679	.1976	.0073	.0103

Table 3

Table 3 gives the components v_1 and v_2 for all six observers considered together.

Interpretation of Results

The question of primary interest at this point is whether the variance between observers is significantly greater than the variance within sets of observations. To answer this question we may make use of the so-called F test where F is defined by

$$F = \frac{v_2}{v_1} \qquad (5)$$

assuming that $v_2 > v_1$ and that all of the sets of repetitive observations have the same variance. In Table 3, values of F for all observers are statistically significant at either the .01 or .05 levels (represented by the double asterisk for .01 and the single asterisk for .05).

In terms of this analysis it appears that, at least for Mn, Fe, Ni, and C, the observer variance $s_o{}^2$ is of the same order of magnitude as $s_e{}^2$ or that contributed by differences between repetitive observations, and hence should be taken into account in estimating the overall precision called for in the paragraph quoted above from Mr. Shaw's memorandum.

It is my understanding that a single observation is made on a sample at Hawthorne and if we were to base our

estimate s_{oe} of overall precision on the two components s_o and s_e, including both Laboratory and Hawthorne observers, we would get

$$s_{oe} = \sqrt{s_o^2 + s_e^2} \qquad (6)$$

We shall return later to a consideration of the use of this overall estimate in setting tolerance limits.

Differences between Laboratories

In the preceding analysis any effect of existing significant differences between the two laboratories appeared in the two components v_1 and v_2. In other words, the estimate of overall variance given by (6) implies that the differences between BTL and Hawthorne are negligible. Likewise the analysis of the preceding paragraphs rests on the assumption that, for each component, the variances for different analysts do not differ significantly. We must therefore investigate the validity of these two assumptions.

First let us apply the "Student" test to the differences between averages. This is done in Table 4 and we see that for Mn and Fe the differences are significant on the .001 level. That is to say the difference in level shown in Fig. 1 for these two components is greater than can reasonably be attributed to chance.

Comparison of Averages for Two Laboratories

		% Mn	% Fe	% Ni	% C	% Si
Average	Laboratory B ($\overline{X}_B$)	.788000	53.8025	45.1592	.042083	.149625
	Laboratory H ($\overline{X}_H$)	.765000	53.5904	45.0592	.041585	.152667
Variance	Laboratory B (v_B)	.000216	.009202	.016843	.000016	.000113
	Laboratory H (v_H)	.000279	.024926	.057756	.000092	.000099
$s_{diff.} = \sqrt{\dfrac{v_B}{n_B} + \dfrac{v_H}{n_H}}$		.004544	.037709	.055752	.002121	.003020
$\overline{X}_B - \overline{X}_H$		.0230	.2121	.1000	.0005	.0030
$t = \dfrac{\overline{X}_B - \overline{X}_H}{s_{diff.}}$		5.062	5.624	1.794	.236	- 1.007
$P(t)$		<.001***	<.001***	.1	.8	.3
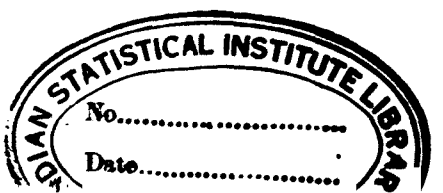
Table 4

CHEMICAL ANALYSIS OF 45% PERMALLOY
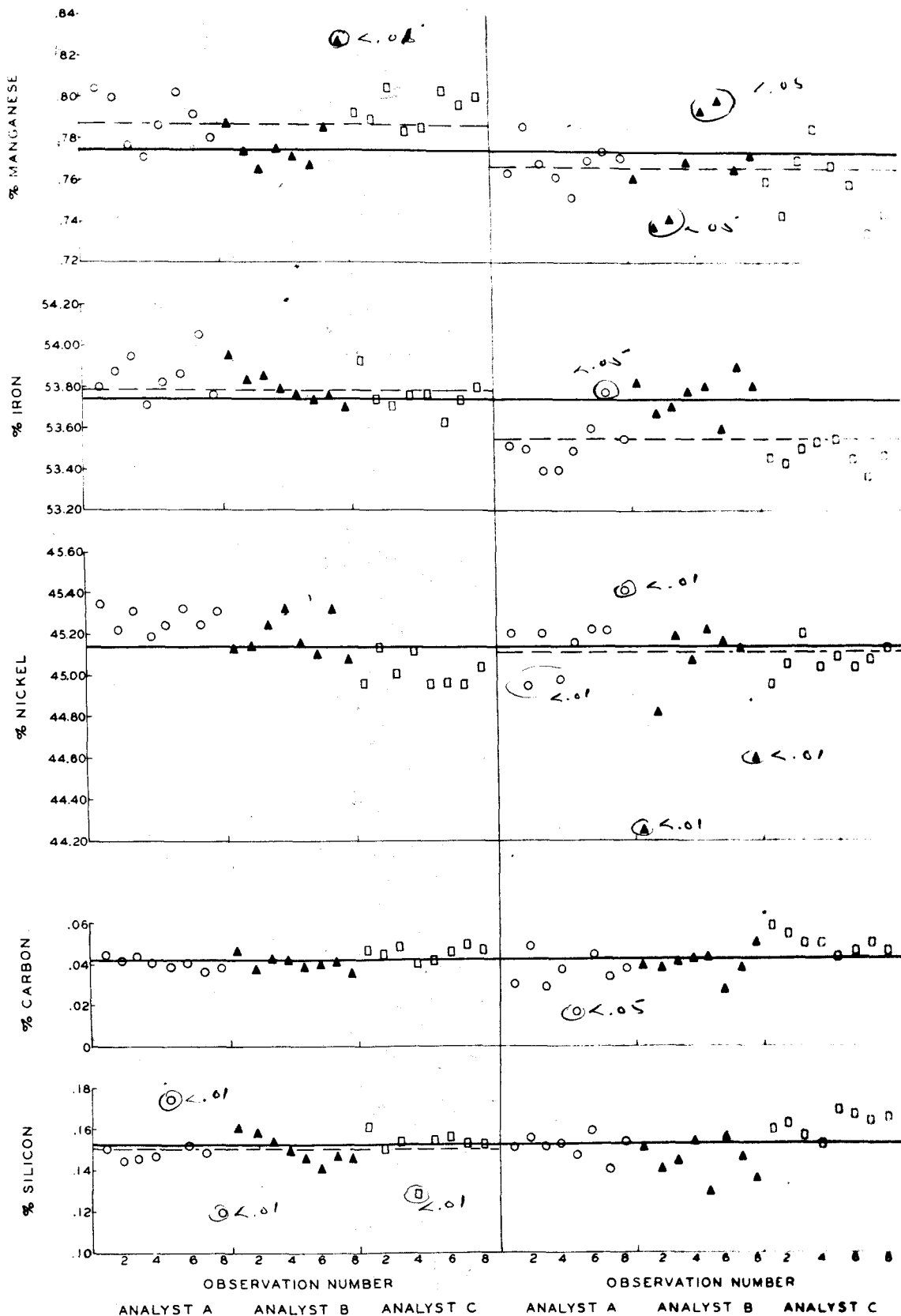
Fig. 1

Hence let us consider the breakdown of the total variance for each laboratory into two components v_1 and v_2 as shown in Tables 5 and 6. We see that for both Mn and Fe, the BTL data do not reveal any evidence of significant observer variance. Likewise for Hawthorne, there is no indication of significance for Mn; however, for Fe a significant difference is indicated. This means that we are not strictly justified in using equation (6) as an estimate of overall variance in the case of Fe and Mn if we are to allow for the indicated significant differences between laboratories. However, if the values of s_o^2 and s_e^2 for Mn and Fe are taken from Table 6, and used in equation (6), the values of s_{oe} thus obtained for Mn and Fe might be used as estimates of overall variance for Hawthorne provided we do not find below any evidence to indicate that the variances for the individual observers are significantly different among themselves.

For Ni, we see that v_2 is significantly larger than v_1 for BTL but not for Hawthorne. However, in this case we shall see in a later section that the lack of indication of significant difference between Hawthorne observers is probably caused by the effect of other factors or assignable causes. Much the same type of story applies in the case of C and Si and hence we shall return to a consideration of these in a later section.
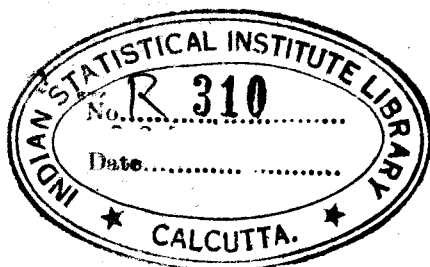
ANALYSIS OF VARIANCE - LABORATORY B

		% Mn	% Fe	% Ni	% C	% Si
Within sets of measurements	Sum of squares	.00429900	.17712500	.12762500	.00020775	.00252688
21 degrees of freedom	v_1	.00020471	.00843452	.00607738	.00000989	.00012033
Between observers	Sum of squares	.00067900	.03452500	.25975833	.00015208	.00007275
2 degrees of freedom	v_2	.00033950	.01726250	.12987916	.00007604	.00003637
Total	Sum of squares	.00497800	.21165000	.38738333	.00035983	.00259963
23 degrees of freedom	v	.00021643	.00920817	.01684275	.00001564	.00011303
$F = \dfrac{v_2}{v_1} \text{ or } \dfrac{v_1}{v_2}$		1.66	2.05	21.37**	7.69**	3.31
$s_e = \sqrt{v_1}$		.0143	.0918	.0779	.0031	.0110
$s_o = \sqrt{\dfrac{v_2 - v_1}{8}}$		.0041	.0332	.1244	.0029	
$s_{oe} = \sqrt{s_e^2 + s_o^2}$		.0149	.0977	.1468	.0043	
$s = \sqrt{v}$		.0147	.0959	.1297	.0040	.0106

Table 5

ANALYSIS OF VARIANCE - LABORATORY H

		% Mn	% Fe	% Ni	% C	% Si
Within sets of measurements	21 degrees of freedom	Sum of squares				
		.00584425	.19253750	1.08527500	.00119375	.00107425
		v_1 .00027830	.00916845	.05167976	.00005684	.00005115
Between observers	2 degrees of freedom	Sum of squares				
		.00056775	.38075833	.24310833	.00093008	.00119308
		v_2 .00028387	.19037916	.12155416	.00046504	.00059654
Total	23 degrees of freedom	Sum of squares				
		.00641200	.57329583	1.32838333	.00212383	.00226753
		v .00027878	.02492591	.05775580	.00009234	.00009858
$F = \dfrac{v_2}{v_1}$ or $\dfrac{v_1}{v_2}$		1.02	80.76**	2.35	8.18**	11.66**
$s_e = \sqrt{v_1}$		.0167	.0958	.2273	.0075	.0072
$s_o = \sqrt{\dfrac{v_2 - v_1}{8}}$		.0001	.1505	.0935	.0071	.0083
$s_{oe} = \sqrt{s_e^2 + s_o^2}$		.0167	.1784	.2458	.0104	.0109
$s = \sqrt{v}$		.0167	.1579	.2403	.0096	.0099

Table 6

Differences Between Variances for Individual Analysts

Table 7 gives the estimates of variance for each set of 8 repeated measurements. Let us consider, as an example, the six sets of 8 observations for iron. Let us assume that each set is a random sample from a normal distribution of error of measurement and that the true standard deviations are σ_{e1}, σ_{e2}, ... σ_{e6}. The hypothesis to be tested is that

$$\sigma_{e1} = \sigma_{e2} = \cdots = \sigma_{e6} \qquad (7)$$

There are two reasons why the results of such a test are important. In the first place, the method used above in breaking down the total variance into two components v_1 and v_2, involves the assumption that (7) holds. In the second place it is desirable to know if the errors of measurement differ significantly from observer to observer.

We shall make use of the Bartlett test.[1] Table 8 gives the detailed computations for iron. If the ratio B/C is greater than X^2 at the chosen level of significance, it is unlikely that the six estimates of variance would have given the observed value of B/C if the hypothesis was correct that the error of measurement was the same for all six observers.

- -

[1] See for example, Properties of Sufficiency and Statistical Tests by M. S. Bartlett, Proc. Roy. Soc. Lon. vol 160, 1937, pp 268-282.

PER CENT MANGANESE, IRON, NICKEL, CARBON AND SILICON IN 45% PERMALLOY

AVERAGES FOR DIFFERENT ANALYSTS

| | Laboratory B Analysts | | | | | | | Laboratory H Analysts | | | | | | |
	AIB	CLL	JFJ	HEJ	GSM	WCJ	MEC	BJS	MIN	JAS	MEB	CMD	HDM	JD
% Manganese	.789	.781	.794					.769	.768	.758				
% Iron			53.80	53.85	53.76			53.47		53.53	53.77**			53.54
% Iron - KS-4574			53.53	53.61										
% Nickel			45.27**	45.02	45.18**			45.17				44.93	45.08	
% Carbon - gravimetric				.040		.041	.046**	.040	.050**					.035
% Carbon - volumetric							.042	.059						
% Silicon - residue				.150	.147		.151	.151					.162*	.145
% Silicon - HF							.151	.152*					.159**	.144

Variances for Different Analysts

	AIB	CLL	JFJ	HEJ	GSM	WCJ	MEC	BJS	MIN	JAS	MEB	CMD	HDM	JD
% Manganese	.000159	.000388*	.0000673					.0000997	.000468*	.000267				.0548
% Iron			.00611	.0120	.00784			.00356		.0148*	.00917			
% Iron - KS-4574			.00162	.00645*										
% Nickel			.00325	.00553	.00946			.0825*				.127**	.00540	
% Carbon - H - gravimetric B - Mfg. Std. 17000				.00000714		.0000120	.0000106	.0000446	.0000867				.0000993*	
% Carbon - H - volumetric B - A.S.T.M. E29-42T							.0000686	.0000251						
% Silicon - H - residue B - Mfg. Std. 17000				.0000462	.000225*			.0000528					.0000324	.0000882
% Silicon - H - HF treatment B - KS-4574							.0001994	.00000771					.0000160	.0000324*

Table 7

Comparison of Variances for Fe
by Bartlett's Test

Analysts				Observations					Σx	Σx^2	$\frac{(\Sigma x)^2}{8}$	$\Sigma(x-\bar{x})^2$	S_i^2	$\log_e S_i^2$
	1	2	3	4	5	6	7	8	(1)	(2)	(3)	(4)=(2)−(3)	(4)/n−1	
JFJ	.95	.83	.85	.79	.76	.74	.76	.70	6.38	5.1308	5.0880500	.04275000	.00610714	−5.098296
HGJ	.80	.87	.94	.71	.82	.85	1.06	.76	6.81	5.8807	5.7970125	.08368750	.01195536	−4.426576
GSM	.93	.74	.71	.76	.76	.63	.74	.80	6.07	4.6563	4.6056125	.05068750	.00724107	−4.927986
BJS	.46	.43	.51	.54	.55	.45	.37	.47	3.78	1.8110	1.7860500	.02495000	.00356429	−5.636791
JAS	.52	.51	.40	.40	.50	.60	.78	.56	4.27	2.3825	2.2791125	.10338750	.01476964	−4.215189
MEB	.83	.68	.71	.78	.81	.60	.90	.81	6.12	4.7460	4.6818000	.06420000	.00917143	−4.691662
												Sum .05280893		−28.996500

Mean of Variances $\overline{S^2}$ = .00880148

$\log_e \overline{S^2}$ = −4.732833

m = number of analysts = 6

n = observations per analyst = 8

$$B = (n-1)m \log_e \overline{S^2} \; - \; (n-1) \sum_{i=1}^{6} \log_e S_i^2$$

$$= 42\,(-4.732833) \; - \; 7\,(-28.99650)$$

$$= -198.7790 \; + \; 202.9755 \; = 4.1965$$

$$C = 1 + \frac{m+1}{3m(n-1)} \;=\; 1 + \frac{7}{3(42)} \;=\; 1.0556$$

Bartlett's Ratio $\dfrac{B}{C} = \dfrac{4.1965}{1.0556} = 3.9756$ $P\!\left(\dfrac{B}{C}\right)$ for $m-1$ degrees of freedom = .557

NOTE: The tabulated observations are the original observations minus 53.00

Table 8

Column 3 of Table 9 gives the values of P(B/C) **for**
all five elements for six observers. We see that the **results**

Bartlett's Ratio and Associated Probability P(B/C)

	All Observers			BTL Observers			Hawthorne Observers		
	$\frac{B}{C}$	Degrees of Freedom	P(B/C)	$\frac{B}{C}$	Degrees of Freedom	P(B/C)	$\frac{B}{C}$	Degrees of Freedom	P(B/C)
Mn	8.920	5	.116	4.799	2	.093	3.603	2	.172
Fe	3.975		.557	0.828		.666	3.078		.217
Ni	35.846***		.001	1.836		.412	14.460***		.001
C	16.979**		.007	0.457		.800	2.906		.239
S	9.940		.081	4.024		.142	2.335		.314

Table 9

for Ni and C are significant at the .001 and .01 levels. This
table shows similar computations for BTL and Hawthorne **analysts,**
and only one statistically significant value of B/C **appears,**
namely that for Ni at Hawthorne.

Stated in engineering language (leaving out all **if's**
and but's as symbols of caution) my interpretation of **these**
results is as follows. There is no evidence that the **six**
analysts differ in respect to precision of their **repetitive**
measurements in the case of Mn, Fe, and Si. For C, **the**
significant difference for the case of all six observers **is**
apparently attributable to the fact that all observers **at**
BTL appear to have a higher precision than those at **Hawthorne.**
For Ni, we shall see later that the large dispension at

Hawthorne is likely produced by assignable causes. A glance at Fig. 1 will reveal that these conclusions appear quite reasonable.

ASSIGNABLE CAUSES OF VARIATION

Thus far we have considered the breakdown of variance to see if known factors such as differences between laboratories and differences between observers do in fact contribute in a significantly measurable way to the overall variance. In this section we are to see if we can find any evidence of the influence of unknown factors or assignable causes upon the variance for each of the sets of eight observations.

$P(\lambda)$ Test

Let us assume that a sample of n has been drawn from a normal universe with standard deviation σ. Some types of assignable cause may produce one or at least a few large deviations. For example, if a few small particles of iron had been picked up with the sample of permalloy one of these might have gotten into one of the subsamples and produced an abnormally high observed percentage of Fe. It is therefore often worth while to test the observed difference d_1 between the largest (or smallest) observation and the one nearest to it or the difference d_2 between the 2nd from the largest (or smallest) and the 3rd. Let us call

$$\lambda_1 = \frac{d_1}{\sigma} \quad \text{and} \quad \lambda_2 = \frac{d_2}{\sigma}$$

Irwin[*] has given tables of the probabilities $P(\lambda_1)$ and $P(\lambda_2)$ of getting values of λ_1 and λ_2 as large or larger than that observed.

In the present case, we do not know σ and hence we take estimates s_e given later in Table 11. In Fig. 1 certain points are circled in red and the values of $P(\lambda)$ are indicated.

This test gives evidence of the presence of an assignable cause of variation in the results of i.

    a. The second BTL observer for Mn.

    b. The second Hawthorne observer for Ni.

    c. The first and third BTL observer for Si.

For several reasons I am inclined to believe that in each instance the assignable cause lies at the door of the analyst and is not attributable to lack of homogeneity in the subsamples.

η Test

We shall now look in on additional evidence of the presence of an assignable cause that may be laid at the door of the analyst — evidence to suggest that an analyst is influenced by the results of his previous measurements. The tool for doing this is the η test.

Let x_1, x_2, ..., x_i, ..., x_n be a sequence of n analyses on a component such as Fe, by one analyst, and let $\bar{x}$ be the average of these n observations. Then

- -

[*]Biometrika, Vol. XVII, pp 238-250.

$$\eta = \frac{\dfrac{\sum\limits_{i=1}^{n} (x_i - x_{i+1})^2}{n - 1}}{\dfrac{\sum\limits_{i=1}^{n} (x_i - \overline{x})^2}{n}} \qquad (8)$$

The distribution of η for samples of n from a normal universe was recently derived by Von Neumann[1] and has been tabled by Miss B. I. Hart.[2]

Values of η for each of the samples of eight are given in Table 10 and are plotted on the η charts of Figs. 2 and 3. The expected value $\overline{\eta}$ and three probability levels are also given on the η charts as a function of n.

We note that in general too many values of η fall below the expected value - in fact seven fell below the 95% limit. This effect may be produced by either of the following two kinds of causes: 1) a trend in the successive values such as that shown in Fig. 1 by the second BTL observer for Fe, 2) abrupt shifts in the averages of subgroups of the n observations. For example, the eight analyses of each observer were for the most part carried through in four groups of two each, and if the averages of these pairs differ by more than is

- -

[1] "Distribution of the Ratio of the Mean Square Successive Difference to the Variance", Annals of Mathematical Statistics, Vol. 12, (1941), pp. 367-395.

[2] "Significance Levels for the Ratio of the Mean Square Successive Difference to the Variance", Annals of Mathematical Statistics, Vol. 13, No. 4, December, 1942.

PER CENT MANGANESE, IRON, NICKEL, CARBON AND SILICON IN 45% PERMALLOY

T'S FOR DIFFERENT ANALYSTS

			Laboratory B Analysts							Laboratory H Analysts				
	ALB	CLL	TFI	HET	GSM	WCJ	MEC	BTS	MIN	JAS	HEB	CWD	HTM	JD
% Manganese	1.576	1.005*	2.568						1.083*	1.329				1.832
% Iron			0.634**/1.783	2.865/1.910	1.626			1.590		1.249	3.051	1.070*	2.168	
% Nickel			2.994	2.815	2.742			1.960						
% Carbon				1.143		2.085	1.887	3.055	0.801*					3.530
H - gravimetric														
B - mfg. Std. 17000														
% Carbon							1.383	1.040*						
F - volumetric														
B - ASTM E30-42T														
% Silicon				0.487***/1.083			2.073	3.793*					1.808***/3.351	
H - residue														
B - mfg. Std. 17000														
% Silicon							1.909	2.243					1.645/3.396	
H - HF treatment														
B - E3-1074														

OR η

ate Sets of

9/23/46

PROBABILITY LIMITS FOR η

$$\eta = \frac{1}{n-1} \sum_{i} (x_i - x_{i+1})^2 / \frac{1}{n} \sum_{i} (x_i - \bar{x})$$

SAMPLE SIZE n

SUBJECT __Laboratory B - Values of η for Separate Sets of__
__Eight Determinations.__

ENGINEER _____   DATE __9/23/46__

Fig. 2

$$\eta = \frac{1}{n-1} \sum_{1} (x_i - x_{i+1})^2 / \frac{1}{n} \sum_{1} (x_i - \bar{x})$$

SAMPLE SIZE n
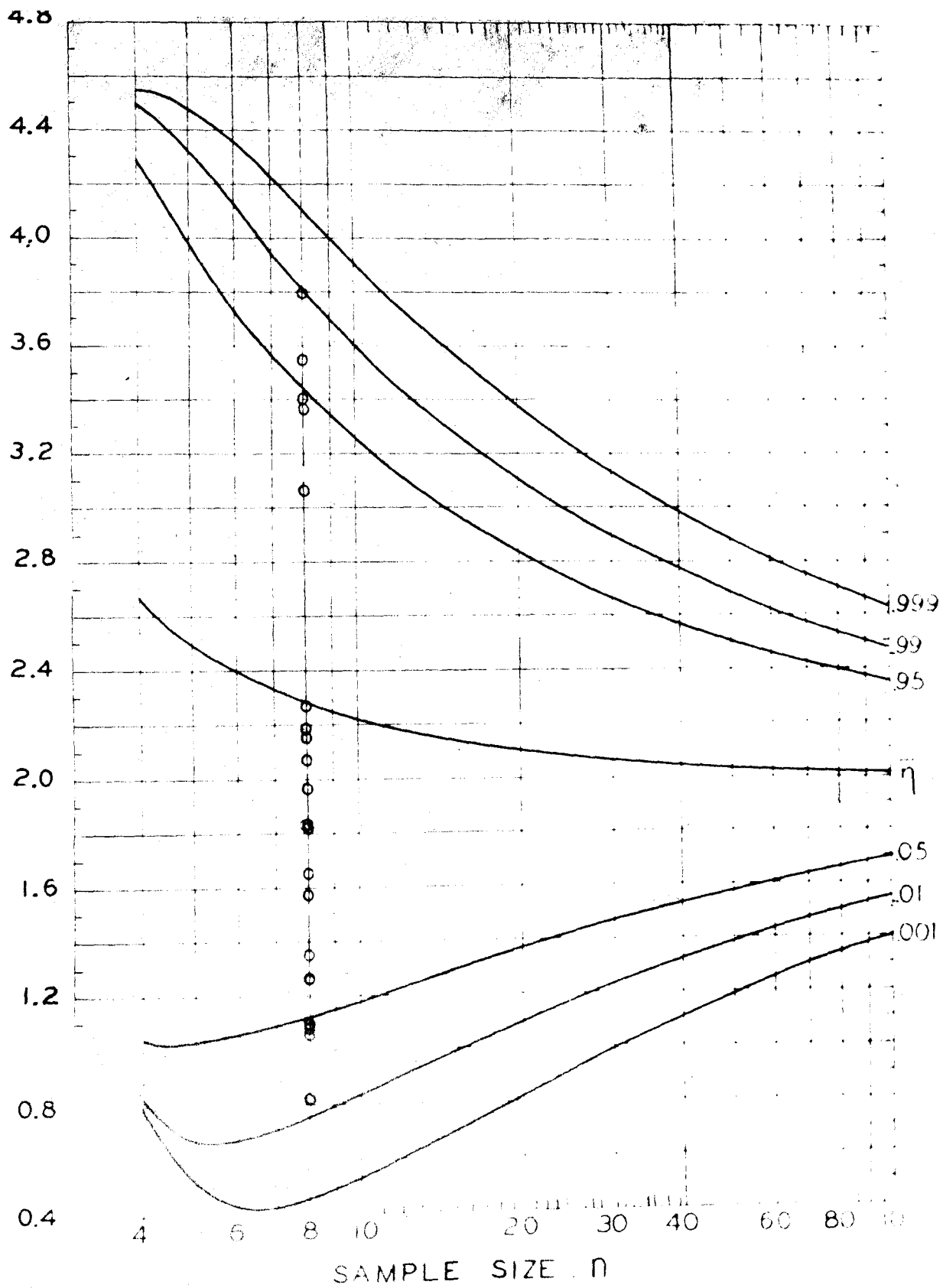
PROBABILITY LIMITS FOR η

SUBJECT Laboratory H - Values of η for Separate Sets of
Eight Determinations.

DATE 9/23/46

ENGINEER

Fig. 3

attributable to chance fluctuations of a constant chance cause system, we would get values of η that are on the average smaller than they would otherwise be. I suspect that this is the assignable cause.

If the averages for pairs of values do not differ significantly but there is a tendency for an analyst to be influenced by his first value when taking the second, then the value of η will tend to go outside the higher limits. From this viewpoint, it is interesting to note that three of the four highest values of η belong to Hawthorne observer JD. It is interesting also to note that the tendency toward small values of η exists for both BTL and Hawthorne.

It is possible, of course, that a tendency toward correlation between pairs of measurements may be partly attributable to the fact that when two subsamples are carried simultaneously through the steps of the analytical procedure, involving heating, cooling and the like, these steps may introduce correlation. It is to be hoped that time will be available at the conference to explain in some detail the practical significance of the assignable cause revealed by the η test.

DISCUSSION OF RESULTS

Let us first consider the bearing of the results upon the questions raised in Mr. Shaw's memorandum.

. <u>Overall Precision</u>

Mr. Shaw says:

"Our main concern is to find an acceptable measure of the overall precision of the complete analytical process, as performed in practice in our laboratories, rather than the best precision obtainable by a certain method, under ideal conditions."

An answer to this question, based upon the results of chemical analyses shown in Table 1, is given in Table 11, and will now be explained.

<u>Estimates of Overall Precision</u>

	s_e	d.f.	s_o	d.f.	s_{oe} single observations	s_{oe} paired observations
ˌanese	$.0155_{BH}$	42	$.0001_H$	2	.0156	.0110
n	$.0938_{BH}$	42	$.1505^{**}_H$	2	.1773	.1644
kel	$.0779_B$	21	$.0935_H$	2	.1217	.1085
ˌbon	$.0075_H$	21	$.0071^{**}_H$	2	.0104	.0091
icon	$.0093_{BH}$	42	$.0083^{**}_H$	2	.0124	.0109

Table 11

We should note first that Mr. Shaw was interested in the overall precision for his own laboratory, and not in the overall precision allowing for differences between laboratories. We must, however, allow for error of measurement and the differences between observers. Estimates s_e and s_o of the standard

Typical 1% Tolerance Limits on Single Observations

	% Mn	% Fe	% Ni	% C	% Si
s_e	.0155	.0938	.1699	.0058	.0093
s_{oe}	.0200	.1777	.2021	.0075	.0102
s	.0194	.1679	.1976	.0073	.0103
L_1	.30	53.30	44.00	0	0
L_2	.80	55.30	46.00	.05	.20
$L_{11} = L_1 + 2.326\ s_e$	.3361	53.5182	44.3952	.0135	.0216
$L_{12} = L_1 + 2.326\ s_{oe}$	.3466	53.7134	44.4701	.0174	.0237
$L_{13} = L_1 + 2.326\ s$	.3452	53.6906	44.4597	.0170	.0240
$L_{21} = L_2 - 2.326\ s_e$	.7639	55.0818	45.6048	.0365	.1784
$L_{22} = L_2 - 2.326\ s_{oe}$	.7535	54.8866	45.5299	.0326	.1763
$L_{23} = L_2 - 2.326\ s$	.7549	54.9094	45.5403	.0330	.1760

Table 12

deviation of repetitive measurements and the standard deviation
of differences in expected values for different observers are
given in columns 2 and 4 of Table 11. Columns 6 and 7 give
estimates of overall standard deviation (or measure of pre-
cision) for single measurements and for the averages of pairs
of measurements. These are respectively

$$\sqrt{s_o^2 + s_e^2} \qquad \text{and} \qquad \sqrt{s_o^2 + \frac{s_e^2}{2}}$$

First let us consider the estimates of s_e^2 given in
Table 11. Since the Bartlett test does not reveal any signifi-
cant differences between the individual observer variances
for Mn, Fe, and Si, it seems reasonable to pool the results
for the six observers in obtaining an estimate of s_e in these
instances. One way of doing that is to take the estimates
based upon 42 degrees of freedom as given in Table 3. We
have seen, however, that for both Ni and C, the Bartlett test
indicates significant differences between the observer vari-
ances when all six observers are taken together but not between
the three observer variances within a given laboratory. In
the case of Ni, however, there is, as we have seen, an indica-
tion that assignable causes are present. In fact, it is
quite likely that the Hawthorne value of $s_e = .2273$ for Ni
in Table 6 is affected to a great extent by such causes all

of which should be removed from the Hawthorne technique. For
example, Table 5 shows that the corresponding BTL value is
only .0779 or about one third of the Hawthorne value. When
the assignable causes are removed at Hawthorne, I should ex-
pect their value to approach that of BTL and therefore I
suggest using the BTL value. The s_e for C is taken from the
Hawthorne Table 6.

For one reason or another it appears desirable to
take the estimate of s_o for each of the five elements from
Table 6 showing the analysis of variance for Hawthorne. It
will be noted in Table 6 that the ratio of $\frac{v_2}{v_1}$ = F for Fe, C,
and Si is large enough to indicate that differences between
expected values for observers are significant. Although F
for Ni in Table 6 is not significant, this fact is largely
explained on the basis that the assignable causes present
at Hawthorne helped to produce an abnormally high value
s_e = .2273. Had it not been for the presence of these as-
signable causes, it is reasonable to believe that F for Ni
would also have been significantly large.

If we look at Table 11, we see that the values of
s_o except for Mn are either of the same order of magnitude
or larger than the corresponding values of s_e. Hence in all
cases except for Mn it is important that we take the factor
s_o into account in estimating the overall variance or pre-
cision.

2. Tolerance Limits on Single Observations

Mr. Shaw's second sentence reads:

"In discussions with our inspection organizations or other interested persons, we would then be able to state that if a particular result, for example, which was near the specification limit, were to be checked by additional analyses on the same sample, these could be expected to agree with the first result to within certain limits."

Let us next see what can be said on this comment by Mr. Shaw.

First let us look at Fig. 4 . Here the observed range in each of the five sets of 48 observations is shown alongside the specification or tolerance range. (These tolerance ranges are taken from Mr. Kott's memorandum of 4/22/46). For Mn and C some of the observations fall outside the tolerance limits and for Fe, Ni and Si some fall near a specification limit in the sense implied by Mr. Shaw.

Now let us consider the situation that arises in practice. As I understand, it is customary practice in routine analyses conducted at Hawthorne to make only one analysis on a sample of material and from day to day such analysis may be made by several different observers. Assume that on a certain day some analyst is instructed to take a sample from the previous day's production of permalloy and to analyze subsamples for each of the five components. For convenience let us consider only one of these analyses, let us say for Fe. Suppose that the single observation for iron lies just above the lower
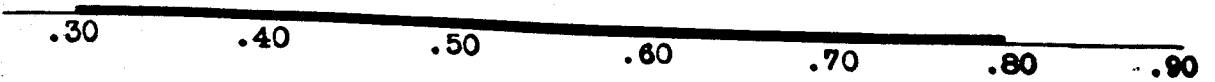
Percent Manganese

.30   .40   .50   .60   .70   .80   .90

Percent Iron

53.0   54.0   55.0   56.0

Percent Nickel

44.0   45.0   46.0   47.0

Percent Carbon

0   .05   1.0

Percent Silicon

0   .10   .20
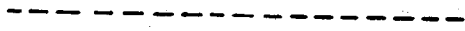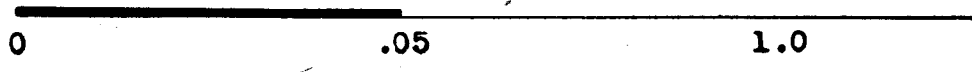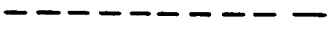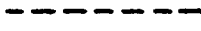
━━━ Tolerance Range

---- Observed Range

tolerance limit L_1. Mr. Shaw raises a very simple question - one to which an answer will be suggested shortly - what would we expect to get if additional analyses on <u>this same sample</u> were to be made? Since he does not specify how many additional analyses, let us assume that this is a very large number just in order to increase the precision of this average so that for the sake of the present discussion, this average may be assumed to be the "true" value. Would you expect this average to fall within the tolerance range?

It should be noted, of course, that Mr. Shaw does not indicate by what method these additional analyses are to be made nor does he state that they shall be made by an observer from the same laboratory. If the check analyses may be made by different methods, accuracy as well as precision must be taken into account, and if the check analyses may be made by another laboratory we must also consider the component of variance attributable to differences between laboratories.

The practical problem, however, is much more involved, because we are not commercially interested in where the true iron content of the sample lies but instead we are interested in where the true iron content of the entire lot lies. To get at the basis for attacking this practical problem, we would need to consider the following factors or sources of variance in addition to those already considered:

1. Differences between iron content of all possible samples from the given lot.

2. Differences between methods of measurement. This brings us face to face with the problem of determining how the original sample should be taken from the lot. This is a big and somewhat involved subject and should be given careful consideration at our conference in November. The importance of solving this practical problem is obvious if it is desirable that the quality of product as represented by chemical properties be kept within tolerance limits.

Now let us return to the simpler problem posed by Mr. Shaw in which it is supposed that the additional analyses are to be made from the same sample. Let us also assume that such analyses are to be made by the same method and by observers at Hawthorne. Assuming for the present that s_o^2 and s_o^2 given in Table 11 are the true components of variance, let us see within what limits a single routine measurement would have to be held. The probability of accepting a sample of tolerance quality in which the % iron is less than the lower tolerance limit L_1 and the probability of accepting a sample of tolerance quality in which the % iron is greater than the higher tolerance limit L_2 are each not greater than .01. For this purpose, let us take the following values of L_1 and L_2 suggested by A. E. Kott in his memorandum of 4/22/46.

$$L_1 = 53.3 \% \text{ iron}$$

$$L_2 = 55.3 \% \text{ iron}$$

Fig. 5 shows two sets of limits, namely L_{11} and L_{12} and L_{21} and L_{22} in addition to the tolerance limits L_1 and L_2 for iron. If it were not for the fact that observers differ significantly among themselves, we could get along with the limits L_{11} and L_{12} but if we are to allow for both components of variance present it would be necessary to reject whenever single observations by a given observer fell outside the limits L_{12} and L_{22}. In other words, single observations would have to lie within the limits L_{12} and L_{22} in order that the chance that the material under test lies below L_1 and the chance that the iron content of the material under test does not lie above L_2 are each not greater than .01.

We may, of course, look at this situation in another way. So long as we allow single observations to fall any place within the limits L_1 and L_2 instead of L_{12} and L_{22}, we are giving at least 99 percent assurance that the true iron content of the permalloy does not fall below $L_1 - .4133 = 52.89$ and at least 99% assurance that the true iron content does not fall above $L_2 + .4133 = 55.71$.

Table 12 gives similar values of s, s_e, and s_{oe} for Mn, Fe, Ni, C, and Si. In the last four rows of this table

Fig. 5

MSH-10/14/46
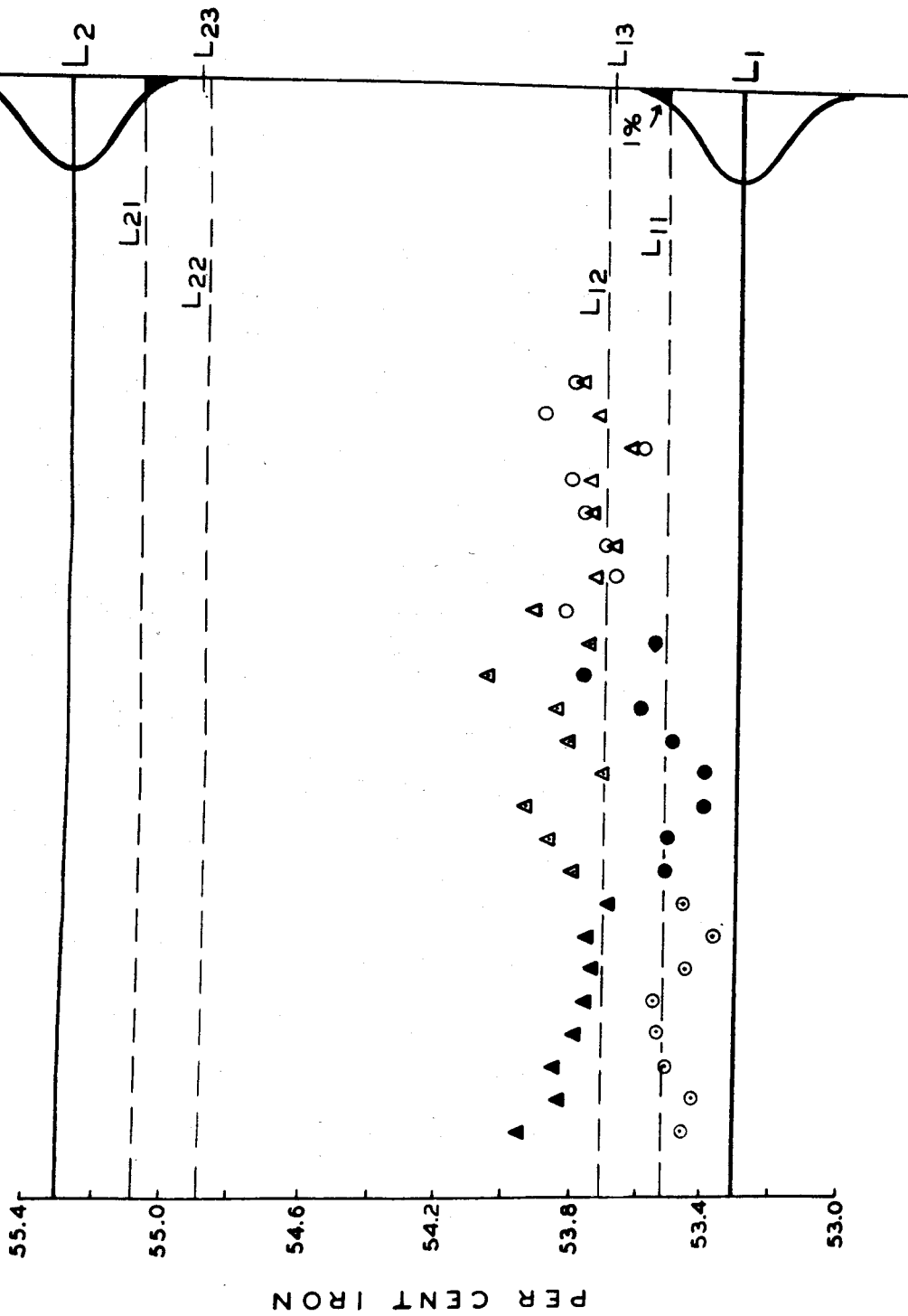
values are also given for the limits L_1 and L_2 together with L_{12} and L_{22} from which figures similar to Fig. 5 may be constructed.

It should, however, be kept in mind that, if we were to allow for differences between samples as we should do in the practical case but cannot do without further data, we should get even a narrower range within which single observations must lie in order to give the assurance herein assumed.

3. Application to the Control of Quality

Let us now see what can be said on the point made by Mr. Shaw in the third sentence of the paragraph quoted above, viz.

"Also in comparing a number of analyses made over a period of time on different lots of the same material, we would be able to state that a certain variation in results did or did not indicate an actual variation in the material."

Fundamentally the problem here involved is one of control.

A. Accuracy

First let us consider the following question: How can we be reasonably sure even when no variation in results is observed that there is no "actual variation in the material", to use Mr. Shaw's words? This is a particularly pertinent question in the case of Fe.

First of all, we should note that we have good evidence to indicate that there is a significant difference

between observers at Hawthorne. This means that even if an observer could and would take an infinite number of observations, we might expect that the average of these would differ from the actual Fe content by what is customarily termed a constant error. To complicate matters, we should recall also that the BTL and Hawthorne averages were found significantly different. If we look now at the Fe data in Fig. 1, we see that the second Hawthorne observer, to speak in engineering terms, seems to be quite like the BTL observers. In fact, the observed significant difference between BTL and Hawthorne is largely attributable to the fact that Hawthorne observers one and three are much lower than the other four observers.

Another troublesome fact is that the three analysts at BTL had great difficulty in carrying out the analysis with the 10 ml. indicator used by Hawthorne. As evidence, I give in Fig. 6 the final set of BTL observations with the 10 ml. indicator alongside the final BTL set with the 1 ml. indicator. For the 10 ml. indicator, the observers differ significantly and the average of the 10 ml. data is significantly higher than for the 1 ml. data. Even for the 1 ml. indicator the second observer shows a significant trend. In fact the η for this set is .634 and falls below the .01 level on the η chart.

Finally let us look at the difference between the results obtained by the two BTL observers using the KS-4574

Twenty-four determinations of % Iron in 45% Permalloy by three different analysts

FS   Fig 6

method (see Table 1) of analysis and the other results shown
in Table 1. In the first place, we find that on the Fisher t
test, the chance is less than one in a thousand that the
average of the 48 BTL observations (using the Hawthorne method)
shown in Table 1 would differ from the average of the 24
observations made by method KS-4574 if there were no "true"
difference in the methods. It is also interesting to note
that the first and third Hawthorne observers (see Fig. 1) are
down on about the same level as the two sets of BTL data taken
by the KS-4574 method.

In the light of all of these significant discrepancies,
what is the "actual" Fe content of the sample? If that question
is not a sufficient puzzler let us consider this one which is
still nearer to the practical problem. In the course of routine
inspection, what assurance does one observed value give you
about the "actual" Fe content of the lot of permalloy from which
the sample was taken?

If we are concerned about the "actual" Fe content
of 45% permalloy, it would appear that further work is required
in order to discover the assignable cause that produces the
significant differences indicated above.

B. Problem of Detecting "Actual Variation in the Material"

Of course the solution of the problem of detecting
"actual" variation in the quality of material as posed by

Mr. Shaw does not demand that the accuracy be known, but it does demand <u>that the accuracy though unknown be kept constant</u>. It also imposes the condition that the precision of repetitive measurements on a given sample be kept constant. In the light of these requirements let us see what could be done to meet them.

a. In order to insure that the accuracy of <u>an observer</u> be kept constant it seems desirable to have a <u>standard sample</u> of material constantly on tap. It would be sufficient to check from time to time to see that the average of a set of n observations on this sample by a given observer did not vary more than it is reasonable (at some prescribed probability level) to attribute to a sampling fluctuation. This could be done at the start by standard control chart methods or perhaps more easily by a run chart method that I shall discuss at the conference.

b. In order to insure that the precision for repetitive measurements be kept constant it would be possible to make use of the repetitive measurements made from time to time (see preceding paragraph) on the standard sample by some statistical procedure, the simplest though not necessarily the most efficient being a control chart on variance. It is, however, of great importance to check the precision of the actual observations made on samples taken in the course of routine inspections. If a pair of observations be taken on each such sample, then the circle control chart discussed at the previous conference could be used.

c. Let us assume that repetitive observations on the standard sample by a given observer fluctuate in a random manner and are distributed normally about an expected value $\bar{x}'$ with a standard deviation σ'. Note that this expected value $\bar{x}'$ need not be the true Fe content of the standard sample.

It would then be possible to draw a figure like Fig. 7. If the given observer were then to take pairs of observations on the standard sample and plot each pair as indicated in Fig. 7, 99% of such points would be expected to fall within the circle, so long as his accuracy and precision did not change significantly. If a point representing a pair of observations on the standard sample falls outside the circle but within the dotted limits this could be taken as indication that his accuracy had changed. If such a point falls outside both the circle and the dotted lines, this fact may be taken as indicating that the precision has changed and, under certain conditions to be discussed at the conference, it may be taken that the accuracy has also changed.

If pairs of observations on samples taken in the course of inspection are plotted on this chart, the probability would be .99 that all points would fall within the dotted limits if the precision remains unchanged. Hence if a point falls outside the dotted limits, we might take this as an indication that the precision had changed.
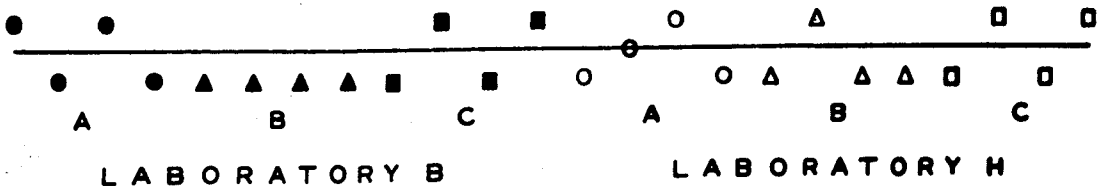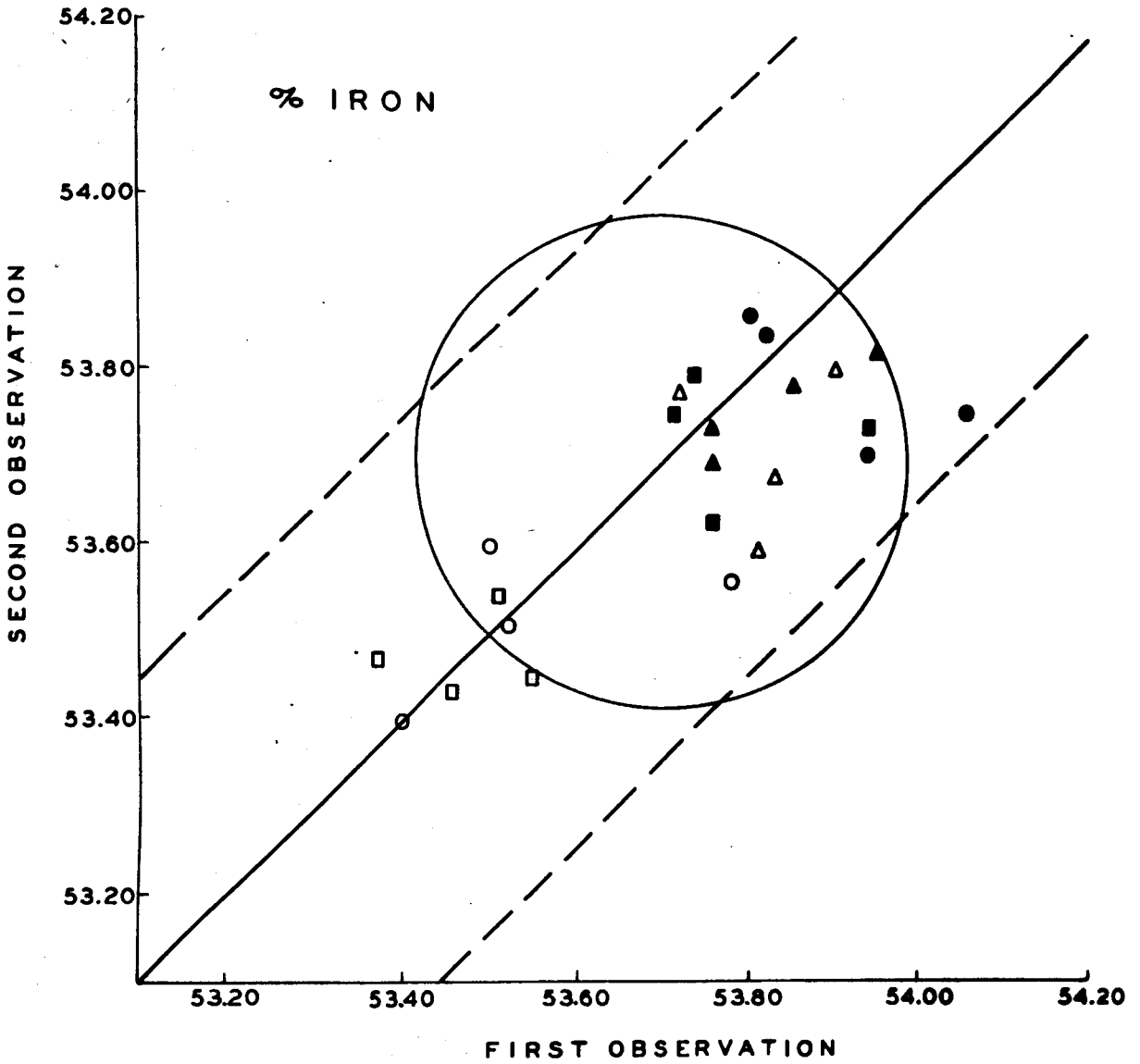
Fig. 7

Of course we do not know either the true Fe content
of the sample of permalloy or the true variance σ_e^2 of repeti-
tive measurements. We have seen, however, that s_e^2 may be
taken as a reasonable estimate of σ_e^2. The circle in Fig. 7
makes use of this estimate s_e^2 and the center of the circle
is arbitrarily placed at the average Fe content as determined
by the 48 observations shown in Fig. 1. Just for illustrative
purposes, the 24 pairs of points determined by these 48 ob-
servations are plotted in Fig. 7. We note that all points are
within precision limits but five points fall outside the circle
but within the dotted limits thus indicating that it is un-
likely that these pairs of observations could have arisen as
random fluctuations from a normal population of measurements
having an expected value of 53.70% and a standard deviation
s_e. This conclusion is, of course, consistent with our pre-
vious findings that there exist significant differences between
observers and between laboratories. One interesting feature
of this circle control chart is the fact that when once set
up, all that an analyst has to do is to plot each point cor-
responding to a pair of observations on a given sample.

This circle chart is interesting in still another
way. It will be noted that 15 points fall below the diagonal
through the center of the circle and only 8 points fall above.

Here we have a slight indication at least that there is a
tendency for the second observation in a pair to be less than
the first. Also it should be noted that the points tend to
hug this diagonal and this is consistent with the information
previously revealed by the η test that pairs of observations
appear to be closer together than would be expected if the
observation in each set of 8 were not correlated serially.
In other words, this simple chart is very helpful in suggesting
changes in average and precision as well as suggesting the
presence of correlation, that the more efficient but more
involved methods show up in a quantitative way.

If in routine control of permalloy each observer makes
use of such a circle control chart as well as the customary
control chart, the indications given by the customary type
of control chart may be taken as an indication of an actual
change in the quality of permalloy provided pairs of observa-
tions on routine samples fall within the dotted limits on the
circle control chart and provided pairs of observations taken
from time to time on the standard sample fall within the circle.

What has been said in the last few paragraphs is
offered as one way of checking on the actual change in quality
without knowing what that actual quality is. It is essential,
however, that such charts be kept for each observer, because
as we have seen, observers at Hawthorne are significantly dif-
ferent. Much more needs to be said at the conference on the
different topics considered in this section.

SOME QUESTIONS AND COMMENTS FOR CONSIDERATION

1. Assume that the sample of material taken at Hawthorne had been put into solution and one half of the solution had been sent to BTL, would you expect the precision of repetitive measurements as measured by s_e to have been materially reduced? If it is not feasible in such an instance to put all of the material into solution, would it be possible to grind or pulverize the original sample, before subsamples were weighed out?

Some of the observed variance in repetitive measurements as conducted is attributable to the differences in the composition of the subsamples and therefore the observed variance is presumably larger than that contributed by the method of analysis. In any work that may be done to compare accuracies as well as precisions of different analytical methods it would be well to eliminate the effect of differences in the chemical composition of the samples used.

2. Are the tolerance limits now implied in the specification of permalloy economically sound?

3. How should a sample of material from a lot be taken?

The answer to this depends upon the meaning of the specified requirements. For example, do the specified limits $L_1 = 53.3$ and $L_2 = 55.3$ on iron apply to the content of the whole lot? If so, some form of representative sample may be desirable. If, however, the specified limits imply that the

iron content should not vary outside these limits from point to point, the method of sampling should be different.

4. In order to set rational and economic tolerance limits it is desirable for a specification engineer to have data on the variance of each of the chemical components of the material. Would it be desirable to get such data on permalloy for the use of specification engineers? If such data were made available might it not lead to changes in the specified limits?