NBSIR 78-1574

A Suggested Experimental Plan for the Determination of the Accuracy and Reliability of Coal Mine Dust Measurements

Harry H. Ku
Joan R. Rosenblatt

Center for Applied Mathematics
National Bureau of Standards
Washington, D.C. 20234

NBSIR 78-1574

A SUGGESTED EXPERIMENTAL PLAN FOR THE DETERMINATION OF THE ACCURACY AND RELIABILITY OF COAL MINE DUST MEASUREMENTS

Harry H. Ku
Joan R. Rosenblatt

Center for Applied Mathematics
National Bureau of Standards
Washington, D.C. 20234

June 30, 1979

Final

Issued October 1979

Prepared for
United States Department of the Interior
Bureau of Mines

TABLE OF CONTENTS

i

0. Introduction

In December 1975, a report entitled "An Evaluation of the Accuracy of the Coal Mine Dust Sampling Program Administered by the Department of Interior" [1] was prepared by the Analytical Chemistry Division of the National Bureau of Standards (NBS), in response to a request by Senator Harrison A. Williams, Chairman, Committee on Labor and Public Welfare. The report attempted to evaluate the adequacy of (1) the personal sampler unit used to measure the coal dust concentration in the mines, (2) the procedures used by the Department of Interior (DOI) to analyze the dust samples obtained from mine operators, and (3) any recent studies made by the Department or private industry concerning the coal dust sampling equipment. Recommendations in the report of NBS included continued research efforts by the appropriate organization to correct the personal sampler deficiencies and to consider the use of area or stationary samplers as a supplement to the personal samplers.

Concurrent with the NBS report cited above, the General Accounting Office (GAO) also issued a report in December, 1975 [2], suggesting that improvements are still needed in coal mine dust sampling programs and procedures relating to penalty assessments and collections. Among recommendations made by GAO, the following paragraph specifically called for actions by the agencies involved:

> "GAO further recommends that the Secretary of the Interior
> instruct the Mining Enforcement and Safety Administration
> [MESA, now MSHA] and the Bureau of Mines and the Secretary
> of HEW instruct the National Institute for Occupational Safety
> and Health to conduct a joint study to determine quantitatively
> the accuracy and reliability of dust measurements when taken
> with the current equipment by coal miners in underground
> mines. GAO suggests that the assistance of the equipment
> manufacturers and the National Bureau of Standards be
> solicited."

To implement the above recommendations, an interagency agreement [3] was made between the Bureau of Mines (BOM) and the NBS on February 1, 1977. The agreement has as its technical objective "the quantitative determination of the accuracy and reliability of dust measurements when taken with the current equipment by coal miners in underground mines," and the work is to be conducted in two phases. In Phase I, NBS is to develop an experimental plan for the determination of the accuracy and reliability of coal mine dust measurements at three levels of funding, and submit a report to BOM in accordance with terms of the agreement. This report was reviewed by BOM, with comments and suggestions from the Mining Safety and Health Administration (MSHA) and the National Institute for Occupational Safety and Health (NIOSH), to serve as a basis for decisions in carrying out the experimental program in the field and in laboratories. The actual experimental program will be conducted in Phase II. This document is the final report called for at the conclusion of the Phase I part of the agreement.

In Section 1 of this report, the terms "reliability" and "accuracy" are clarified and defined for the purpose of this investigation. Certain pertinent information, collected as results of personal contacts and study of existing literature, is presented and discussed. In Section 2 an experimental design is suggested to measure the various sources of variability as relating to the use of approved personal samplers "by coal miners in underground mines." In Section 3, a dry run is suggested to test the efficacy of the design. The results of this dry run will be used to modify the protocol to reflect the effect of increased or diminished monitoring (supervision), and for the design of experimental programs at different levels of funding. Pilot experiments on 10 to 12 mines are also recommended to obtain preliminary estimates of variabilities due to different sources. In Section 4 an overall view is presented and factors that may contribute to the resulting uncertainty of coal mine dust measurements are discussed in that context. A statistical model and corresponding analysis of variance for the design is given in Appendix A.

1.   Accuracy and Reliability of the Coal Mine Dust Measurements

The terms "reliability" and "accuracy" have been used in many fields of measurement so much so that clarification of their meanings for the purpose of this report becomes mandatory. In general, inaccuracy, or uncertainty, refers to the departure of a measured value from an agreed reference base; reliability refers to the frequency that a system will perform satisfactorily its intended function when used under stated conditions.

Following these concepts, which we believe are reasonable for the purpose of this report, we will discuss in more detail the various factors (sources of systematic errors, and extent of random variability) that contribute to the uncertainty and unreliability of the results of dust measurements.

1.1 Accuracy/Uncertainty of the Coal Mine Dust Measurements

By the term "uncertainty" of a reported value we mean the credible limits to its likely inaccuracy [4] as measured from some agreed basis of reference. Two components of uncertainty are usually present: the uncertainty due to systematic error and the uncertainty due to imprecision. The distinction between the two types, however, is by no means rigid and fixed, and is only meaningful in relation to the frame of reference.

1.1.1 Reference Base for the Coal Mine Dust Measurements

In the "Federal Coal Mine Health and Safety Act of 1969", the reference base for the concentration of respirable dust is the average concentration of respirable dust if measured with the instrument

developed by the Mining Research Establishment (MRE). Measurements made with any other instrument must be converted to an MRE equivalent for the purpose of the Act.

The currently used conversion factor for all personal samplers is a multiplicative factor of 1.38 based on comparison data of personal sampler and MRE collected in 200 coal mines in 1973, and recommended in MHSA Report of Investigations RI 7772 [5]. The uncertainty of this value, however, was not given in the report.

The same data were reanalyzed by BOM and essentially the results supported the value of 1.38 recommended by MESA [6]. Using different criteria for the selection of "valid" data, and different assumptions on the error structure, the conversion factors calculated for several levels of concentrations range from 1.19 to 1.84. Laboratory measurements and measurements in return airways gave much higher results. The reasons for these high values were not known. [6]

Among the many difficulties encountered in the determination of a conversion factor, we note the following major ones:

a. Data editing. In the analysis by MHSA or BOM, nearly 50 percent of the data collected were rejected for reasons that these sets of data were impossible, inconsistent, or unreliable.

b. Treatment of data at lower range of concentration. At the lower range of concentration, the relationship between the two results in this range was difficult to characterize. From an inspection of the results of BOM's alysis, it is not at all certain that one value of conversion factor could serve adequately throughout the dust concentration range from 0.0 to 5.5, mg/m^3. BOM, however, indicated to us that precise measurements at such low concentrations are not of particular interest.

When the Federal Coal Mine Health and Safety Act was enacted in 1969, it was based on data from Great Britain, using MRE as the standard instrument for measurements of coal mine dust. The inherent imprecision of the MRE measurements, however, is about twice that of the personal samplers, as reported by Eastern Associated Coal Corp. (EACC) in 1976 [7]. Other investigators, however, claimed a "consistency" of about ±8% [18]. The use of the MRE as a reference basis, and hence the conversion factor contributes a substantial uncertainty to results obtained in the coal mine dust sampling program.

1.1.2 Possible systematic errors in the Coal Mine Dust Measurement Program

For the purpose of this report, a systematic error is defined as an error in the measurement process that stays constant in magnitude and direction for all, or a large part of the resulting data. The conversion

factor used to calculate MRE equivalent from results of personal samplers certainly would contribute a systematic error, due to the uncertainty of the number 1.38 used for all results of dust concentrations. However, results of all mines will be treated the same way, and its presence would have the same effect on all mines in the administration of the program.

Other sources of systematic errors are less evident, and may in many cases be uncovered only with improvement of precision of measurements, or with experiments designed specifically to check on particular sources of systematic errors. For example, the conversion factor of 1.38 was derived using data ranging from 0.5 to 5.5 mg/m^3, and a straight line relationship is forced through zero by the addition of a large number of zero pairs. It is conceivable that this factor may work well within this range, but could be improved upon in the range of 0.0-0.5 mg/m^3.

A number of sources of systematic errors belong to this category, where there are no experimental data, or insufficient data, to ascertain their existence or magnitude. Consequently, we can only discuss those systematic errors that are conceptually possible.

(i) Possible differences in results as measured by personal samplers manufactured by different companies.
The existence of a significant difference between the MSA Bendix personal samplers was alluded to in the 1975 NBS Report [1], and the difference was attributed to a loss of tare weight of the Bendix filter capsule. This problem of weight loss was corrected, but subsequently a weight gain problem caused Bendix units to be taken off the approved list again in early 1977. A difference of 20% between average values of MSA and UNICO was also reported in a 1974 BOM study [6]. Obviously possible systematic error associated with the design, or material and process in manufacture, must be closely guarded against.

Recent reports [7,11] that we are aware of did not address this problem. The difficulty in comparing the results of two or more makes of personal samplers seems to stem from the fact that variations due to location or positioning of samplers are usually confounded with variations due to differences among individual personal sampler units. A difference among parallel samplers set up in the mines could be due to a difference in unit performance, or due to differing dust concentrations seen by each sampler in the package. The Pittsburgh Technical Support Center (PTSC) of MESA routinely checked only the preweight of cassettes of different manufacturers, but not the total performance of the units. We are not aware of any routine monitoring program of unit performance by NIOSH, the agency that approves the use of particular sampling units, or by the manufacturers of the personal sampler units.

4

Upon the request of the Administrator of MHSA in June 1977, an NBS team visited the processing laboratory of MHSA Pittsburgh Technical Support Center, and reviewed the quality control program used in the weighing of respirable dust samples. One of the present recommendations by NBS deals with the monitoring of the manufacturing process of each type of filter on a continuing basis to detect the extent of change of the weight of filters with time. A report of NBS findings and recommendations is under preparation [16].

(ii) Treatment of oversized particles in samplers

By the current practice, filters with weight gain of 6 mg or more are examined for oversized particles (greater than 10 micrometers, nonrespirable) either visually or micro-scopically. The procedure for the determination of oversized particles is outlined in PTSC Standard Method No. A4 and Addendum [8]. The justification for using 6 mg as the cut-off point is demonstrated in a memorandum dated October 28, 1976 from the Chief of Dust Branch, PTSC, to the Assistant Admin-istrator. Essentially, an experiment was performed on 500 samples, with the result that 14 out of 15 samples weighing more than 6 mg contained an excessive number of oversize particles, whereas only 7 out of 485 samples weighing less than 6 mg contained excessive oversized particles.

Assuming the above figures are representative, then among those filters with weight gain less than 6 mg, about 1.5 percent would contain excessive oversized particles. Depending on the number of oversized particles present on these filters, the concentration of the respirable dust level as defined would be biased upward. Although the magnitude of this bias may be negligible, its presence should continue to be recog-nized.

(iii) The presence of a disproportionate number of low concentration of 0.1 mg/m^3

Records of "Notice of Compliance" of all occupations of a section in each of two mines, provided to us by MESA, were examined. The Notice listed dust concentrations measured over ten shifts on ten more or less consecutive days. Dust concen-trations of the "Intake Air" for a particular day were also given, but not always. A sample of these records is shown in Figure 1.

There were 230 measured values for the mine in Pennsyl-vania from February 1972 to January 1977 with a range of values from 0.1 to 7.6 mg/m^3. For the mine in Illinois, there were 344 measured values from November 1970 to December 1976,

with a range of values from 0.1 to 9.9 mg/m^3. Frequencies of dust concentrations from 0.1 to 2.0 mg/m^3 are listed in Table 1. We note that among the 20 categories, there are 40 and 41 values of 0.1 mg/m^3 out of totals of 175 and 193 values, or about 23 percent and 21 percent respectively. These dis-proportionately large numbers of 0.1 mg/m^3 values raise questions as to their validity, especially when the dust concentrations of the intake air were somewhat higher, in the 0.2 and 0.3 range.

At the weighing laboratory in PTSC, we observed the practice followed in the evaluation of filters which showed a final weight equal to or less than the initial weight that may account for a portion of large frequency of 0.1 dust concen-trations. These filters are set aside and examined visually or with a microscope. If dust particles are observed the sample is considered valid and the value 0.1 mg/m^3 is assigned to the sample. If no dust particles are observed, the sample is declared invalid and the serial number of the cassette is forwarded to MHSA Headquarters for investigation of causes leading to the invalid measurement.

Whatever the cause of the large number of 0.1 mg/m^3 measurement values, malfunctioning of the unit or tampering, the procedure adopted has the effect of biasing the concen-trations downward. It is reasonable to assume that dust concentration in the coal mine should not be less than that of the intake air, and to establish a procedure based on this assumption.

United States Department of the Interior

BUREAU OF MINES

****NOTICE OF COMPLIANCE****

NOVEMBER 23, 1973   210--0003
MINE 36-00803-0   SECTION 0060
ALLEGHENY   (AC) PENNSYLVANIA

IN ACCORDANCE WITH SECTION 70.260 OF THE MANDATORY HEALTH STANDARDS, UNDERGROUND COAL MINES
THE FOLLOWING IS A CUMULATIVE CONCENTRATION OF RESPIRATORY DUST, RELATIVE TO THE
RESPECTIVE SECTION OF YOUR MINE.
IN ACCORDANCE WITH SECTION 70.222 AN ALTERNATE SAMPLING CYCLE IS ESTABLISHED.

		INTAKE AIR	
CASSETTE NO.	DATE	CONCENTRATION	
41106234	11-16-73	.3	

CASSETTE NO.	DATE	CONCENTRATION	PROD.
41110635	08-15-73	.6	0330
41106500	08-16-73	.3	0290
41106212	08-17-73	1.0	0400
41110278	09-05-73	.5	0310
41110282	09-06-73	.8	0270
41108166	09-12-73	.1	0390
41106489	11-13-73	.1	0330
41111124	11-14-73	.1	0350
41106493	11-15-73	.1	0270
41108151	11-16-73	.8	0420
CUMULATIVE:		4.4 AVE: .336	

HARMAR COAL CO
OAKMONT MINE
ATTN: PAUL MOHR
P O BOX 500
LIBRARY   PA 15129

INTERIM COMPLIANCE PANEL PERMIT
PERMIT NO.   DATE-ISSUED   DATE-EXPIRES   LIMIT
000000   00-00-00   00-00-00   20 MILLIGRAMS

R. E. BARRETT (SIGNED)   11-12-73   01-12-74
U.S. BUREAU OF MINES REPRESENTATIVE   NEXT SAMPLING CYCLE

RECEIVED
NOV 27 1973

Figure 1.

7

TABLE 1

FREQUENCIES OF DUST CONCENTRATION UNDER 2 mg/m^3

Dust Concentration mg/m^3	Frequency of occurrence	
	Pennsylvania Mine	Illinois Mine
0.1	40	41
0.2	9	12
0.3	11	20
0.4	11	12
0.5	16	21
0.6	15	11
0.7	3	8
0.8	15	10
0.9	0	1
1.0	14	7
1.1	8	9
1.2	5	4
1.3	3	6
1.4	1	3
1.5	9	7
1.6	1	9
1.7	5	1
1.8	3	5
1.9	0	2
2.0	6	4
Total	175	193

1.1.3 Imprecision of the Coal Mine Dust Measurements

In a measurement process, precision refers to the typical closeness together of successive independent measurements of a single magnitude generated by repeated applications of the process under specified conditions [4]. In coal mine dust measurements, the dust level changes continuously and the measured concentration represents the average level over the time period sampled and the space traveled by the sampler during this time period. Since the measured magnitude is defined as an average, the imprecision of the personal sampler is usually determined from results of 2 to 4 samplers held in a sampler package with inlets 6 to 8 cm (2 1/2 to 3 inches) apart. The standard deviation of a single measurement is then calculated from the scatter of these values, assuming that the dust level within this short range of distance is the same for all samplers in a package. Some typical values of standard deviations using this procedure and reported by Eastern Associated Coal Corporation are reproduced in Table 2. The precisions can be also characterized by the coefficients of variation in percent, i.e., the standard deviation divided by the average and multiplied by 100. Bituminous Coal Research, Inc. (BCR), in a letter to BOM in February 1977 [9], reported coefficients of variation of 7.6, 5.4, and 7.8 percent in three different mines under monitored conditions. These values are perhaps representatives of coefficients of variations for monitored field experiments conducted nowadays. Even then, there were 36 packages out a total 145 packages for which the concentration range was greater than 1 mg/m^3 for each set of four samplers.

Many sources of variability contribute to the imprecision of the samplers using a multiple sampler or a sampler package procedure. In laboratory experiments, we may list the following possible sources:

      (1) weighing
      (2) correctness in assembling the cyclone unit to pump
      (3) ability of a sampler to repeat a measurement
      (4) differences between samplers of the same make
      (5) differences between samplers of different makes
      (6) possible gradient in dust chamber

For monitored field experiments, additional sources include:

      (7) operation under hostile conditions (insufficient light, difficulty in adjustment, movement of samplers, etc.)
      (8) presence of over-sized particles
      (9) gradients of dust concentration as a result of ventilation arrangement (to replace 6 above)

For unmonitored field experiments, additional possible sources could be due to:

      (10) lack of calibration and checking of air flow rate

TABLE 2

Precision of Paired Personal Samplers in the Laboratory,
and in the Field

EACC REPORT [7]

	Laboratory	Field, monitored[d]	Field, not monitored
number of pairs	116[a]	693[b]	101
S	0.14	0.26	0.89[c]

S = standard deviation (in units of mg/m^3)

a. Bendix, MSA, and Wilson Casella personal samplers, used with a modified Lippman Sampler.

b. Dust level ranged from 0 to 5 mg/m^3; about 10 percent of paired data was eliminated because of "oversized material, pump failure during testing or damaged filters, and differences between pairs considered improbable."

c. 28 pairs showed a difference of 0.7 mg/m^3 or more.

d. We note that the difference between the "monitored" and "not monitored" field experiments in Table 2 was the presence or absence of an experienced researcher throughout the work period underground. The sampler packages (two) were located on the continuous mining machine or loader. The operators were not involved with the sampler packages in the course of their work.

(11) incorrect placement of samplers
(12) inaccurate reported time period of sampling
(13) vibration and bumping during the use or transit

For the imprecision of personal samplers as used "by coal miners in underground mines", all the above sources are present with the exception of one, i.e., the gradient factor (listed as 9) which may inflate the imprecision of the sampler package results.[1]   However, the introduction of an operator will add a new component of variability due to the interaction of the operator with the personal sampler in the course of his work. To the best of our knowledge, no attempt has been made to determine the magnitude of this factor.

1.2 Reliability of the Coal Mine Dust Measurements

For the personal sampler to perform its intended function satisfactorily, the physical design, construction, and manufacture of the sampler must be such that a valid dust concentration measurement results at the conclusion of its use. Pump failure, or other mechanical failure where the sampler failed in a physical sense, may be classified as the first type of unreliability.

A second type of unreliability is inherent in the procedures used when the identity or the validity of the filter is in question. Cassette mismatched with the mine data card, incorrect data, missing essential information, etc., cast doubts on the result and it is declared to be invalid.

The third type of unreliability pertains to the measured values themselves, i.e., whether the measured values represent the respirable dust concentration at the particular time and place. Presence of excessive over-sized particles, weight loss or gain of filters, and incorrect flow rates will yield measured values but do not provide a reliable estimate of the correct respirable dust concentrations.

The physical condition of the personal sampler units, including mechanical breakdowns, improper maintenance, etc., is the responsibility of the mine operator; procedural problems can be detected by inspection and dealt with by additional training and educational program under MSHA.

---

[1] We note that the difference between the "monitored" and "not monitored" field experiments in table 2 was the presence or absence of an experienced researcher throughout the work period underground. The sampler packages (two) were located on the continuous mining machine or loader. The operators were not involved with the sampler packages in the course of their work.

Hence, we will be concerned only with the last mentioned component of unreliability: the presence of excessive oversized particles on filters, filters showing no weight gain or even showing weight loss after use, and the flow rates during use.

It is our understanding that the design of the personal samplers does not preclude the deposit of oversized particles (10 micrometers or larger) on the filters. The decision as to whether an overweight (6 mg or more) filter sample contains an "excessive" number of oversized particles is made at the weighing laboratory at PTSC by visual inspection with a hand magnifier or examination under a microscope. In the latter case, 10 fields of 0.25 mm^2 are examined. If the total number of oversized particles is 30 or more in 10 fields, then the sample is voided. If the total number is 29 or less, the sample is deemed valid. About 80 percent of the overweight samples can be screened for oversized particles by visual inspection, and 20 percent by the microscope. In the experience of PTSC over a six month period, about 2.7 percent of samples were voided for containing oversized particles (7,380 out of 273,000).

Caplan [15], in studying performance characteristics of the sampler, noted difficulty encountered in assembling the vortex finder into the top of the cyclone body. Cracks between the vortex finder and shoulder due to misalignment in assembly, or occurring after assembly, introduced high variability between replicate tests. It is possible that the presence of an excessive number of oversized particles is also due to leakage between the two components of the sampler. We understand some corrective action has been taken on the MSA sampler since Caplan reported his findings. The problem is believed to be less severe for the Bendix samplers.

The evaluation of filters with low dust concentration values (0.1 mg/m^3 or less) was discussed under 1.1.2 (iii), and will not be repeated here. Possible causes for low weight gain or weight loss are:

     a.   incorrect initial or final weight
     b.   filter losing weight over time
     c.   sampling time overestimated
     d.   incorrect rate of sampling (air flow rate)
          used in computation

Of the four causes, it appears that the estimated sampling time is under complete control of the miner who reports the number of minutes on the mine data card. Some elapsed time device on the unit itself, which we understand is being developed, would significantly reduce this source of error and cause dust concentration measurements to be more reliable.

The present protocol calls for a flow of 2.0 liters per minute and calls for the miner to check the air flow rate after the first hour of operation. Again there is no assurance that the flow rate is being checked as required. Recognizing this problem, BOM has initiated contract to design and fabricate an improved personal sampler pump to control the air sampling flowrate automatically.

2. Some Unit Designs Appropriate for the Estimation of the Error Variance of Interest

In arriving at an estimate of the uncertainty of measurements made with the personal sampler "as used by coal miners in underground mines," the most difficult task is, perhaps, the estimation of the imprecision of the sampler under actual working conditions. A direct estimation would require burdening the miner with sampling packages, which is objectionable because of bulk and weight. In any case, the miner would be aware of the fact that he is performing an experiment and the results obtained could be biased. Section 2 addresses this problem in particular and suggests an experimental design where the imprecision of the sampler under working conditions can be estimated indirectly.

The indirect way of obtaining an estimate of experimental error is the realistic way in the present context. This estimate would include subtle changes in the use of the personal sampler. If the causes of these changes cannot be identified and eliminated, then their effects must be included as part of the component of imprecision. In addition, the experimental setup is similar to those checking on the "area" or "fixed point" sampling concept, and the miner need not know that he is part of the experiment.

The renewed interest toward the study of "Area Sampling" or "Fixed Point Sampling" to supplement, or even to replace, the personal sampler [7, 10, 11] is another consideration that motivates the choice of the unit design described below. Report L-792 of the Bituminous Coal Research, Inc. [11] suggests that there is a dust level gradient as established from measurements made with personal samplers mounted on the front cab, the machine operator, and the rear cab. Comparison of dust concentrations plotted by the Eastern Associated Coal Corporation [7] for three test series also showed reasonable correlation between machine operator and fixed points on the machine. It is expected that data generated from the unit design can be used to investigate further the advantages of fixed point sampling methodology as well.

2.1 The 5x5 Latin Square Design

This unit design measures the dust concentrations of a section of a mine for five shifts (five days) using five personal samplers. One of the samplers will be used by the operator of a continuous miner in the usual manner; the other four samplers will be made into two sampler packages, to be hung from the front and rear cab posts of the canopy of the mining machine. Figure 1 on page 7 of BCR report [11] gives a photographic representation of the proposed arrangement.

13

The measurement schedule is given in the diagram below, using the following notation.

Days   :   1, 2, 3, 4, 5

Samplers :   A, B, C, D, E

Locations: RL – rear cab left
               RR – rear cab right
               M – miner operator
               FL – front cab left
               FR – front cab right

Days \ Location	RL	RR	M	FL	FR	
1	A	B	C	D	E	d_1
2	B	D	E	C	A	d_2
3	C	E	B	A	D	d_3
4	D	C	A	E	B	d_4
5	E	A	D	B	C	d_5
	L_1	L_2	L_3	L_4	L_5	DC

Here the mean dust concentration for a day, d_i, represents the average over five samplers at five locations about the operator. The mean dust concentration for a location, L_i, represents the average over five days with five samplers, and the mean dust concentration as measured by a sampler, A for example, is averaged over five days and five locations. DC, of course, represents the grand average of dust concentration over five days as measured at five locations with five samplers. These main effects of day, location, and sampler may be estimated.

The rotation of the samplers used at a given location assures that there is no confounding of location effect and sampler effect – a precautionary measure necessary in deciding if adjacent samplers, six to eight centimeters apart, do or do not measure the same dust concentration.

The main interest of the experiment, however, is not the estimation and comparison of these main effects. Here we supplement the classical method in the analysis of a latin square design experiment. After estimating the main effects from the data by the method of least squares, the "unexplained" variability is represented by the "residuals" remaining

14

in each of the 25 cells after the differences due to main effects are removed. Standard deviations calculated from these residuals are measures of imprecision of the experiment. The imprecision of the personal sampler as used by the operator in the course of his work would be represented principally by the five residuals under the column M. By accumulating these residuals over a number of unit experiments, the imprecision of the personal sampler for the miner can be estimated and compared with that for the fixed point samplers.

The mathematical model and analysis of this design is given in Appendix A. The assumptions made in the analysis are spelled out and discussed in the Appendix.

2.2 Other Unit Designs that may be Considered

Designs other than the class of Latin Squares may be used to avoid the restriction that the number of days, the number of locations and the number of personal samplers are equal to one another. The general principle -- estimating the imprecision of the personal sampler from the residuals after the main effects are removed -- however, remains the same.

Using the same notations, a Youden Square with three locations, seven days, and seven personal samplers is shown below:

Location \ Days	1	2	3	4	5	6	7	
Cab Front	A	B	C	D	E	F	G	L_1
Operator	B	C	D	E	F	G	A	L_2
Cab Back	D	E	F	G	A	B	C	L_3
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	DC

A least squares fit to the data would yield seven residuals for the mining machine operator. The efficiency of this design is not much different from the 5x5 Latin Square, except a unit experiment is stretched to seven days.

Chapter 11 of Experimental Statistics [12] lists a number of Youden Square arrangements that may be considered. All of them, however, involve more personal samplers and longer periods of time.

15

2.3 Advantages and Limitations of the Designs

The main advantage of the unit designs given above is that the imprecision of the personal sampler as used by miners during the course of their work can be realistically estimated. This estimate, adjusted for day-to-day dust level variations, would include variability due to the interaction between the miner and the personal sampler under hostile conditions -- a factor that has not been considered in other experiments performed to date. Comparison can also be made between supervised and unsupervised experiments suggested under pilot experiments discussed in the next section.

On the other hand, the present design is restricted to investigations for a continuous miner with a cab and canopy. Perhaps it can be adapted to other "high risk" miners who work in fairly constant position with respect to a machine; e.g., a roof bolter. During the pilot experiment stage, we feel that the experiment should apply to the continuous miner only. Another limitation is that the information obtained for the main purpose of this experiment is rather low per unit experiment. Even ten experiments will only give us a barely adequate number of degrees of freedom for estimating the standard deviation. Other information collected, however, may be also of value; e.g., to check if two samplers mounted side by side will yield the same measurements.

2.4 Checking the Performance of the Sampler in the Laboratory

The precision of the personal samplers has been studied in the laboratory and reported in a number of publications [6,7,13]. The resulting data, however, usually represented the behavior of these samplers under "ideal" conditions. How well these samplers work under "hostile" conditions similar to those encountered in the mines has not been studied in the laboratory. Do these samplers possess sufficient "ruggedness" to be used in an environment which they were designed for?

Youden suggested that a procedure used in a collaborative test [14] must be checked out first by the deliberate introduction of changes of levels of factors likely to be encountered in actual use. His design, now called "Youden's Ruggedness Test", appears to be appropriate in locating possible sources of trouble in the use of personal samplers in the mines. A design for checking on seven factors in eight combinations is shown in Youden's paper. The factors could be, for example, the following:

		Condition 1	Condition 2
1.	Shock or vibration	no	yes
2.	Flow rate	2.0 ℓ/min	1.5 ℓ/min
3.	Rotation in position	no	yes
4.	Assembly of sampler	researcher	other
5.	Oversized particles input	none	present
6.	Dust input	continuous	intermittent
7.	Position of sampler	center	side

The larger effects of these factors would show up in the analysis of the resulting data, and a realistic measure of imprecision can be computed from the seven differences of the averages, each for condition 1 and condition 2.

3. Dry Run and the Pilot Experiments

Prior to the conduct of a full scale experiment, it would be prudent to perform a dry run at two or three mines, to check on:

1. practical aspects of the design presented in section 2, and
2. modifications in procedure and protocol that may be necessary.

For the dry run, we suggest the use of the unit design (5x5 Latin Square) on the continuous miner with cab and canopy arrangement. The first five-day run would be under close supervision of BOM, MSHA, or their contractors. A second five-day run using the same machine would be repeated by the mine operator, without supervision, within a specified time period while the mining conditions remain essentially unchanged, [2]/ say, four weeks. (The experiments performed by the mine operators are included to check if the unit experiment is workable without supervision. Presumably the section foreman, in addition to training as to protocol to be followed, would also learn from the actual performance of the monitored experiment.)

The results of the dry run would be analyzed, and the procedure for further experimentation revised based on the experience gained from the dry runs. A pilot experiment, involving a minimum number of mines that would provide a reasonable amount of information on the imprecision of the personal sampler as used by the miners in the mines, is recommended following the dry run.

To accumulate 25 to 30 degrees of freedom for the estimate of imprecision under each of the two conditions, supervised and unsupervised, we need to perform the experiment in 10(12) mines. For convenience, these mines may be selected in two clusters of 5(6) each; e.g., in West Virginia and in Pennsylvania. If the information obtained is useful to NIOSH in their Third Round of National Coal Study, these 10(12) mines could be selected from the list of mines prepared by NIOSH for their health study. Since few mines report a dust concentration of $2mg/m^3$ or more, 2(3) mines in each cluster should represent mines that have a high day-to-day variability, and the remaining a low variability. The variabilities may be determined from compliance notices given in recent past. At this stage of the experiment, we feel that all the mine sections selected should have seams high enough to give best working conditions for the supervision staff.

2/ In the sense that the same miner and same type of cutters are used, the ventilation system stays about the same condition, etc.

Since two types of personal samplers are in general use, the Bendix and the MSA, two and three of each type could be used for the five samplers called for in the design for half of the 10(12) experiments, and reversed for the balance. Differences of samplers of the same type and those of different types in a sampler package could then be studied.

3.1 Treatment of Void Samples

In the EACC Report [7], Harris rejected 11 percent of on-man samples and 0.6 percent of the fixed point samples "mainly because the dust filter cassettes were contaminated with oversized material...". These figures, however, included sampled values in shifts monitored and not monitored. BCR monitored all shifts in their experiment [11], and reported a lesser percentage (about 2%) of void samples, most of which were due to mechanical failure.

Assuming these percentages of void samples are typical in an experiment of this type, we can expect a few missing data in the monitored experiments due to mechanical failures. If the sampler on the operator fails, the day's experiment would have to be repeated on another day after repair of the sampler. If a sampler in a package fails, a missing value technique may be applied or an auxiliary sampler may be added to the package in anticipation of possible failures. In either case, the interpretation of the results will not be straightforward. We would prefer to repeat the day's measurements, provided the period of a unit experiment is not lengthened enough to be impractical.

Since the purpose of the experiment is to determine the imprecision of field measurements, the cassettes (with special marking) should be sent to PTSC for processing in the routine manner. A sample voided because of contamination of excessive oversized particles may, nevertheless, have to be included in the analysis because the time elapsed between sampling and weighing would not allow a repeat measurement. If the weighing can be performed in the field offices of MSHA, it may be possible to check back and trace the reasons for the malfunction of the unit.

3.2 Protocol for Monitored Experiments

In the following, we assume that six unit experiments are to be performed at each of two general locations and the samplers are to be supplied by BOM, MSHA, or NIOSH.

a. For each mining section selected for the experiment, assign five regular personal samplers to be painted or tagged as A, B, C, D, E. The sixth one is reserved as a spare. For example, A and B could be Bendix samplers, D and E MSA samplers. C will stand for Bendix or MSA alternatively in successive experiments.

b. Fixtures on the front and rear post of the cab of the machines are to be fixed ahead of time for locating the sampler packages. Distances

between the operator sampler and the two sampler packages are to be measured.

  c. The monitoring personnel, or monitor, will check these samplers for air flow, pump operation, leakage, cleanliness of cyclone, and other routine points. He will insert the four samplers in the two packages, using the design. The days 1 to 5 are to be permuted randomly and the arrangement of the samplers in the row corresponding to that day are to be used.

  d. The personal sampler for the operator will be carried in and out of the mine by the operator; the two packages by the monitor. All samplers are to be turned on and off from portal to portal. Possible inadvertent abuse of the sampler by the operator in carrying it in and out of the mine may be noted.

  e. The monitor will check the samplers in the package during the shift, leaving the checking of the on-man personal sampler to the operator. The monitor may collect other data of interest, e.g., the shift production, cutting frequency, air flow and brattice location, etc.

  f. Upon completion of the shift, the monitor checks conditions of all samplers, notes leakage of any other irregularities, etc. He will rearrange the five samplers for the next day.

3.2.1 Unmonitored Experiments

  After the completion of a monitored experiment in a mine section, the foreman or group leader for the section will replace the monitoring personnel and carry out the latter's duties. Personal samplers routinely used in the mine will replace those samplers supplied by government agencies. Some additional explanation or instruction to the foreman may be necessary to clear up possible misunderstandings.

3.3 Estimate of time required for each unit experiment

  In view of a number of possible interferences, such as strikes, work stoppage, insufficient coal production, etc., BOM estimated from their experience that two calendar weeks should be allowed for each unit experiment employing one monitor. Therefore, for the 12 mine sections contemplated in the pilot experiment, about 6 man-months will be required. We suggest the use of four to six monitors so that the result of the pilot experiment may be available in about 2 months time.

3.4 Experimental Plans at Different Levels of Funding

  Several outcomes are possible at the conclusion of the pilot experiments. Decisions regarding further experimentation will have to be conditioned on the results and their interpretation.

First, the results of the pilot experiment may be sufficiently precise for the purpose at hand. In that case no further experimentation is necessary.

Second, the results may indicate that the information obtained is useful, but not sufficiently precise or general in nature to allow the necessary inferences. Then a large-scale experiment will be necessary to firm up the various estimates of components of variability. The extent and scope of this experiment will necessarily depend on the results of the pilot experiments.

Third, if for some reason the pilot experiments do not yield the needed information, then the design and its execution would have to be reexamined as to the causes, and further experimentation would be based on a revised scheme.

4.0 General Remarks on the Coal Mine Dust Sampling Program

In evaluating accuracy and reliability of measurements for mandatory standards and regulatory requirements, it is essential to keep in mind several criteria:

1. Purpose and intent of the regulation,
2. Measurement capabilities, and
3. Operational procedures and definition of "compliance".

Obviously no one set of criteria will work satisfactorily for all types of regulatory requirements.

The coal mine dust sampling program is typical of a class of mandatory regulations which came into being only recently -- say in the past ten years. These regulations generally require measurements of low concentration health hazards over a long period of time. Air and water pollution, radiation are other examples. Because the concentrations of the toxic substances are generally extremely low, measurements made on them are usually state-of-the-art type, and variability of individual measurement is in general large enough to obscure the signal. Averaging over time and space helps, the extent of averaging being dependent on the variability of the measurements. The loss and injury, if any, are over a long term rather than immediate.

The purpose of the coal mine dust sampling program is to determine that the respirable dust in the mines is at a level sufficiently low to prevent black lung disease (pneumoconiosis) among the miners. Personal samplers, carried by the "high risk" miner in a mining section, are used to collect a quantity of respirable dust. The respirable dust for each shift collected in a cassette is weighed before and after use. The regulation specifies that the average of ten shifts should not exceed 2.0 milligrams per cubic meter of air being sampled.

The various sources of variabilities that are inherent in the dust level determinations include:

1. Spatial variability in a section of the mine. At different locations in a section, the dust levels vary considerably. The personal sampler is worn by the "high risk" miner, usually the mining machine operator. But, the "high risk" miner also moves around in the course of his work.

2. Day to day variability, or shift to shift variability. This source of variability is estimated by BOM to be about 50-70 percent, mainly because of mining and ventilation conditions.

3. Variability among samplers of the same make, or different makes. Based on experiments conducted by BOM using samplers with collectors in parallel and close together, the relative standard deviation of a measured value is about 5% when conducted in the laboratory, 5-10% when conducted in mines by BOM research workers, and as much as 50% when obtained by miners [17]. Hostile environment no doubt contributes to the increase, but the magnitudes of components due to various possible causes are not known. MSHA believes that the "gradient" in the air flow around the mining equipment may contribute a large portion of the variability.

Given that the above brief description of the various sources of variabilities is a fair representation of dust concentration measurements when taken with current equipment by miners in the mine, a model may be postulated to relate these sources of variabilities and to indicate appropriate estimates of their magnitudes. The current procedure, specifying a ten-day average, suggests implicitly that the dust concentration of a "high risk" occupation follows some probability distribution, and the shift-to-shift changes are relatively independent. It follows then that the problem is to estimate the relevant parameters of this distribution -- the mean value, the variance, and possibly the skewness. These parameters are assumed to remain fairly constant over time, as long as the conditions of mining remain unchanged, to allow estimation of their numerical values.

A determination of the concentration of respirable dust collected in one shift then is a sample taken from this distribution, plus collection and measurement errors. In symbol, we may express the measured weight, y, as a function of these parameters:

$$y = \mu + \Delta + \varepsilon + e \qquad\qquad (4.0\text{-}1)$$

where:
 μ : the mean value
 Δ : bias that may be inherent in the sampler or in the use of the sampler,
 ε : collection and measurement error with mean zero and variance

$$\sigma_\epsilon^2 ,$$

e: variability due to distribution of dust levels in a mine over space and time, with mean zero and variance σ_e^2.

In the above formulation, Δ can be reduced by calibration, by quality control of the manufactured samplers, and by protocol in the use of samplers through training programs. The "ϵ" term can be reduced by improving the sampler's design to reduce chance of tampering, to assure constant flow rate during use, and to improve weighing procedures. The "e" term, however, will remain with variance σ_e^2 unless the condition in the mine is changed.

It seems reasonable to assume that the purpose of the coal mine dust sampling program is to assure that the value of "μ" in a mine is less than 2 mg per cubic meter. The difficulty is that even if "Δ" and "ϵ" can be reduced to negligible magnitudes, the uncertainty in the average of ten shifts as an estimate of "μ" still depends on the variability of e -- a component that is large in comparison with the other two sources of variability.

A reasonable approach to supplement the ten-day average decision rule appears to be that of the control chart for each mine section or each mining machine. Decisions based on control charts are useful in the following sense:

1. Enforcement is based on the long term behavior of measurement results -- based on the distribution of dust level over space and time.
2. Enforcement can be applied on a graduated series of actions rather than a violation-compliance decision.
3. Assignable causes for increase and decrease of dust levels can be investigated, to detect tampering, misuse, and/or to assure the improvements are real.
4. Almost all data are available, no additional work is involved except to set up procedures for control charts for averages and ranges for each mine (or mining machine), and to study the results.
5. A control chart on ranges of, say, measurement results of five shifts, can be used to compare shift-to-shift variability between different mines. Study of mines which have extremely large shift-to-shift variability may reveal causes of such erratic behavior, thereby leading to possible improvement in ventilation and other dust suppression systems.

4.1 A Generalized Model

If the assumptions are made that the dust concentrations are time dependent and that shift-to-shift concentrations are correlated, then the problem becomes much more complicated than that associated with the model given in subsection 4.0. Some preliminary discussions and suggestions

22

for a respirable dust sampling program under these assumptions can be found in a thesis by Seppanen [17].

5. Acknowledgement

REFERENCES

[1] LaFleur, P. D. An Evaluation of the Accuracy of the Coal Mine
     Dust Sampling Program Administered by the Department of the
     Interior. A final report to the Senate Committee on Labor and
     Public Welfare, prepared by staff of the Analytical Chemistry
     Division, National Bureau of Standards, December, 1975.

[2] General Accounting Office. Improvements Still Needed in Coal
     Mine Dust-Sampling Program and Penalty Assessments and Collections.
     Report to the Congress by the Comptroller General of the
     United States, December 31, 1975.

[3] Contract between U. S. Bureau of Mines and the National Bureau of
     Standards, Number J0177013 dated February 1, 1977. Determination
     of the Accuracy and Reliability of Coal Mine Dust Measurements.

[4] Eisenhart, C. Expression of the Uncertainties of Final Results.
     Science, 160, June 14, 1968. pp 1201-1204.

[5] Tomb, T. F., Treaftis, H. N., Mundell, R. L., and Parobeck, P. S.
     Comparison of Respirable Dust Concentrations Measured with MRE
     and Modified Personal Gravimetric Sampling Equipment. Report
     of Investigations 7772, Bureau of Mines, U. S. Department of
     the Interior, 1973.

[6] Breslin, J. A. Precision of Gravimetric Sampling of Respirable
     Dust in Coal Mines. Unpublished report, U. S. Bureau of
     Mines, 1974.

[7] Harris, H. E. Development of Area Sampling Methodology. Final
     Report by the Eastern Associated Coal Corporation under
     Contract H0133108 with the U. S. Bureau of Mines. October,
     1976, Revised, 1978.

[8] Goldberg, S. A., and Ambrosia, D. Standard Method No. A4,
     Dust Group, Technical Support Center, Pittsburgh, Mining
     Enforcement and Safety Administration, U. S. Department of
     Interior. June, 1972.

[9] Letter from J. A. Kost of Bituminous Coal Research, Inc.,
     Pittsburgh, Pa. to W. Courtney, U. S. Bureau of Mines, dated
     February 7, 1977.

[10] Tomb, T. F., and Ondrey, R. S. Determining the Feasibility of
     Area Sampling to Enforce the Respirable Dust Standard in
     Underground Coal Mines. MESA Informational Report 1037, 1976.

[11] Kost, J. A. and Saltsman, R. D. Evaluation of the Respirable
      Dust Area Sampling Concept as Related to the Continuous Miner
      Operator. Report L-792 of the Bituminous Coal Research, Inc.,
      Monroeville, Pennsylvania, January 1, 1977.

[12] Natrella, M. G. Experimental Statistics, NBS Handbook 91,
      Revised Edition, 1966.

[13] Kaplan, K. J., Doemeny, L. J., and Sorensen, S. Evaluation
      of Coal Mine Dust Personal Performance. Final Report to NIOSH
      under Contract No. PH CPE-R-70-0036, November 1973.

[14] Youden, W. J. The Collaborative Test, Journal of the Assoc.
      Official Agricultural Chemists 46, 5562, 1963. (Reprinted as
      paper 3.4 in NBS Spec. Pub. 300, Vol. 1. Precision Measurement
      and Calibration: Statistical Concepts and Procedures,
      Feb. 1969.)

[15] Caplan, K. J., Doemeny, L. J., and Sorenson, S. D. Performance
      Characteristics of the 10 mm Respirable Mass Sampler: Part
      II - Coal Dust Studies. American Industrial Hygiene Assoc.
      Journal 38, 162-172, April, 1977.

[16] Parobeck, P., Tomb, T., Ku, H., and Cameron, J. Measurement Assurance
      Program for Weighings of Respirable Coal Mine Dust Samples,
      in preparation.

[17] Seppanen, M. S. The Design of an Airborne Respirable Dust
      Monitoring System for Underground Coal Mines. A Thesis submitted
      to the Faculty of the Graduate School of the University of
      Minnesota. December, 1975.

[18] Dunmore, Joan H., Hamilton, R. J., and Smith, D. S. G. An
      Instrument for the sampling of Responsible Dust for Subsequent
      Gravimetric Assessment. J. Scientific Instruments, Vol. 41, 1964.

APPENDIX A

The purpose of this Appendix is to give a statistical model and corresponding analysis of variance for the unit design (experiment) proposed in Section 2.1, and to indicate how data could be combined from a group of unit experiments

CONTENTS

A.1 The Unit Design

First, we define notation for the observed values (measured weights) obtained in carrying out one unit design. In the diagram below, y_i ($i=1,2,\ldots,25$) denotes the weight of respirable dust collected at the ith day-location-sampler point. The letters A to E identify the five samplers.

Location	(1)	(2)	(3)	(4)	(5)	
Day	RL	RR	M	FL	FR	Sum
1	A/y_1	B/y_2	C/y_3	D/y_4	E/y_5	D_1
2	B/y_6	D/y_7	E/y_8	C/y_9	A/y_{10}	D_2
3	C/y_{11}	E/y_{12}	B/y_{13}	A/y_{14}	D/y_{15}	D_3
4	D/y_{16}	C/y_{17}	A/y_{18}	E/y_{19}	B/y_{20}	D_4
5	E/y_{21}	A/y_{22}	D/y_{23}	B/y_{24}	C/y_{25}	D_5
Sum	L_1	L_2	L_3	L_4	L_5	G

For analysis of the Main effects of days, samplers, and locations, one uses the totals D_d (for the dth day), S_s (for the sth sampler), and L_ℓ (for the ℓth location). The day and location totals appear as row and column sums when measured values are entered in the tabular scheme above. Sampler totals are calculated as indicated in the diagram, e.g., for Sampler B, $S_B = y_2 + y_6 + y_{13} + y_{20} + y_{24}$. Finally, G denotes the grand total, $G = y_1 + y_2 + \ldots + y_{25}$.

A.2 Statistical Model for Unit Design

At the first stage of analysis, for an experiment conducted during five consecutive operating shifts for one mine section, it is convenient to assume provisionally that there are fixed (constant) values for

- average dust level,
- day effects,
- location effects,
- sampler effects,

so that the expected value of the observation y_i may be represented by $E(y_i) = \mu + \theta_d + \theta_\ell + \theta_s$ where

27

μ = average dust level,

θ_d = average effect of day d,

θ_ℓ = average effect of location ℓ,

θ_s = average effect of sampler s,

and the appropriate day, location, and sampler are given in the diagram. (In the initial analysis, the day, location, and sampler effects are thought of a adjustments to the average dust level and for each set of effects, we assume $\theta_1 + \theta_2 + \ldots + \theta_5 = 0$).

In the statistical model for the unit design, the measured values differ from their expected values because of "error" due to imperfections of the measuring instruments (samplers) and fluctuations of the dust level. The model hypothesizes further that there may be additional variability in the observed dust levels for the "machine operator" position. The assumptions about errors are expressed as follows: for observations made at locations where the sampler is fixed on the mining machine ($\ell=1,2,4,5$), $y_i = E(y_i) + e_i$, and for observations made where the sampler is carried by the machine operator ($\ell=3$), $y_i = E(y_i) + e_i + m_i$. We assume that the measurement errors ($e_1, e_2, \ldots, e_{25}$) and additional operator-related variations ($m_3, m_8, \ldots, m_{23}$) are statistically independent. They have zero means by definition, and we define

σ_e^2 = variance of observation error,

σ_m^2 = variance due to machine operator.

A.2.1 Variants of the Model

In practice, it may be necessary to make certain technical modifications of this model. For example, the magnitudes of errors may tend to be proportional to the level of dust concentration, so that the two variances σ_e^2 and σ_m^2 are not constant from day to day.

One of the purposes of the Dry Run recommended in Section 3 of the report is to examine whether the additive model appears to be appropriate for the direct observations or for their logarithms, and whether it is the variance or the relative variance (coefficient of variation) that appears to be constant. Depending on the outcome of exploratory analyses, one may rewrite the model as a representation of $\log(y_i)$, or specify the use of weights to equalize variances.

If the model should need to be modified, the corresponding modifications in the analyses described below would be straightforward.

A.2.2 Discussion of the Assumptions

The question has been raised whether the location effects will be "fixed" during the five days of the unit experiment. More precisely, will the differences between front and rear locations be approximately constant from day to day?

If the unit experiment could be analyzed as a factorial experiment (i.e., if the five samplers were identical), a departure from the fixed-effect assumption would show up as a Days-by-Locations interaction effect. An important task in the Dry Run is to investigate the validity of this assumption. To facilitate this, it would be desirable in the Dry Run to have samplers as nearly identical as possible, or calibrated so as to have independent estimates of the differences among them. Furthermore, the two "left-right" differences between samplers packaged together could be expected to be relatively negligible.

If the Location effects are not approximately constant, then it may not be feasible to obtain an indirect estimate of any special variability (with variance σ_m^2) due to the mine operator. There are two kinds of departures from the assumption to look for: (1) front-to-rear difference constant but front-to-operator difference variable or (2) all location differences variable. In the first case, the estimate of σ_m^2 will be "inflated", including day-to-day variations of the average concentration at the mine operator's location in addition to any other variability introduced by the mine operator's handling of the personal sampler. Evidence of this kind of variability may be of interest. In the second case, measurement errors are completely confounded with location-to-location variability within days.

It is generally agreed that there are large day-to-day changes in the average dust level, so that the fixed-effect assumption for Days has not been questioned. The assumption that differences among samplers are constant is plausible unless some daily calibration activity is introducing rather than preventing errors.

In this section, no assumption has been made about the shape (normal, lognormal, or other) of the error distributions. No such assumption is required for estimating the variances σ_e^2 and σ_m^2, although it will be needed for performing tests of significance such as F-tests. The Dry Run will provide initial information about the shape of the error distributions.

A.3 Analysis of Variance, Unit Design

The analysis of variance has two goals:

(1) to extract information about the magnitudes of the pro-visionally fixed average dust level and effects of days, locations, samplers (for use in later analyses where these sources of variability may be modeled and interpreted in a different way) and

(2) to extract information about the magnitudes of σ_e^2 and σ_m^2.

For the first goal, quantities are calculated according to the established textbook methods for analysis of Latin Square designs. For the second goal, we add a special orthogonal decomposition of the "sum of squares for error". The computations for the analysis of variance are displayed in Table A-1.

In Table A-1, the total of 25 measurements (25 degrees of freedom) is allocated to five principal sources of variability. For locations and for "error", it is indicated that a further orthogonal decomposition of the sum of squares may be of interest. These (indented) portions of Table A-1 will be discussed later in Sections A.3.2 and A.4. (It is not customary to show the over-all average as a "source of variability", and the quantity $G^2/25$ listed as its sum of squares is usually called the "correction factor". The first row of Table A-1 may be ignored during the discussion of analysis of the unit design.)

A.3.1 Main Effects

In the textbook analysis of a Latin Square design, the "extra" variance component σ_m^2 is not present. The tests for the main effects of days, samplers and locations are F-tests. For example, the mean square for Days is the sum of squares divided by the associated degrees of freedom (d.f.), MSD=SSD/4, the mean square for Error is MSE=SSE/12, and F=MSD/MSE with 4 and 12 d.f. is the test statistic calculated to assess the significance of day-to-day differences.

Two points must be considered in the interpretation of these F-ratios under the assumptions of the statistical model stated in Section A.2.

(1) It must be assumed additionally that the errors e_i are normally distributed--or approximately so.

(2) It must be recognized that the F-tests are approximate unless $\sigma_m^2=0$--or is small in comparison to σ_e^2.

Although the statistical tests are approximate, the interpretation of the F-ratios is guided by the "expected mean squares" which are calculated from the assumptions stated in Section A.2.

A.3.2 Comparing Locations

An orthogonal decomposition of the sum of squares for Locations is indicated in Table A-1. This analysis is of potential interest for several reasons:

(1) If the Location effects are signficant, it may be possible to identify meaningful structure such as front-to-rear gradient.

(2) Unlike the Day and Sampler effects, the Location effects are "unsymmetrical" with respect to the possible presence of the variance component σ_m^2 associated with the mining machine operator.

TABLE A-1

ANALYSIS OF VARIANCE FOR THE UNIT DESIGN

Source of Variability	Degrees of Freedom	Sum of Squares	Expected Mean Squares (a)
Average dust level	1	$G^2/25$	$25\mu^2 + \frac{1}{5}\sigma_m^2 + \sigma_e^2$
Days	4	$\sum D_d^2/5 - G^2/25 = SSD$	$\frac{5}{4}\Sigma\theta_d^2 + \frac{1}{5}\sigma_m^2 + \sigma_m^2$
Samplers	4	$\sum S_s^2/5 - G^2/25 = SSS$	$\frac{5}{4}\Sigma\theta_s^2 + \frac{1}{5}\sigma_m^2 + \sigma_e^2$
Locations	4	$\sum L_\ell^2/5 - G^2/25 = SSL$	$\frac{5}{4}\Sigma\theta_\ell^2 + \frac{1}{5}\sigma_m^2 + \sigma_e^2$
Front (Left vs. Right)	1	$(L_4 - L_5)^2/10$	
Rear (Left vs. Right)	1	$(L_1 - L_2)^2/10$	(See text, Sec. A.3.2)
Front vs. Rear	1	$(L_1 + L_2 - L_4 - L_5)^2/20$	
Machine Op. vs. Other	1	$(L_1 + L_2 + L_4 + L_5 - 4L_3)^2/100$	
Error	12	$\sum y_i^2 - SSD - SSS - SSL - G^2/25 = SSE$	$\frac{1}{5}\sigma_m^2 + \sigma_e^2$
Machine Operator	4	SSM (See text, Sec. A.4)	$\frac{3}{5}\sigma_m^2 + \sigma_e^2$
Residual	8	SSE - SSM	σ_e^2

(a) Mean Square = (Sum of Squares)/d.f. and the expected values have been calculated under the assumptions stated in Section A.2

Following are the expected mean squares for the four orthogonal comparisons among locations that were listed in Table A-1.

Source	Expected Mean Square
Front (Left vs. Right)	$\frac{5}{2}(\theta_{FL}-\theta_{FR})^2+\sigma_e^2$
Rear (Left vs. Right)	$\frac{5}{2}(\theta_{RL}-\theta_{RR})^2+\sigma_e^2$
Front vs. Rear	$5(\theta_F-\theta_R)^2+\sigma_e^2$
Machine Op. vs. Other	$4\left(\theta_M-\frac{\theta_F+\theta_R}{2}\right)^2+\frac{4}{5}\sigma_m^2+\sigma_e^2$

The subscripts on the θ's correspond to column headings in the diagram in section A.1, except for the abbreviations $\theta_F = (\theta_{FL}+\theta_{FR})/2$, $\theta_R = (\theta_{RL}+\theta_{RR})/2$. If the location effects do appear to be significant, notice that some location comparisons should be based on error estimates that exclude σ_m^2. On the other hand, if the location effects appear not to be significant (i.e., $\theta_\ell=0$ for all subscripts, at least approximately), then the decomposition of the Location Sum of Squares gives additional estimates of the relative magnitudes of σ_e^2 and σ_m^2. In any case, the two "right-left" differences should be small since the two samplers at the front (rear) of the mining machine are packaged together. The sums of squares for these two comparisons may be compared (possibly pooled) with SSE.

A.4 Variance Components, Unit Design

The textbook sum of squares for error (SSE in Table A-1) is an estimate of $\sigma_e^2+\sigma_m^2/5$, combining both of the variances that are hypothesised in the statistical model considered here. SSE can, however, be expressed as the sum of two terms of which one is an estimate of σ_e^2 alone. The general idea is to use the five observations made at location 3 (Machine operator) to obtain an estimate involving σ_e^2, and to estimate σ_m^2 from the other 20 observations. The exact details are a little more complicated because there are only 12 degrees of freedom for estimation of error.

Table A-1 shows a convenient (desk-calculator) formula for calculating SSE. The result is identical to the result that would

be obtained by calculating the sum of 25 squared "residuals". The residual r_i associated with the observation y_i is the difference between y_i and the value $\hat{y}_i$ calculated using the estimates $\hat{\theta}$ of the main effects. E.g., $r_3 = y_3 - \hat{y}_3$, where $\hat{y}_3$ is the sum of

$$\hat{\mu} = G/25,$$

$$\hat{\theta}_d = D_1/5 - \hat{\mu} \quad \text{(day one)},$$

$$\hat{\theta}_\ell = L_3/5 - \hat{\mu} \quad \text{(location three)},$$

$$\hat{\theta}_s = S_C/5 - \hat{\mu} \quad \text{(sampler C)}.$$

Although there are 25 such residuals, they are not statistically independent-- the average μ and the three sets of main effects have been estimated and subtracted from the observed values..

The linear algebra of least squares estimation can produce for each of the 25 residuals a formula expressing it as a linear combination of the 25 original observations y_i. For example, again using the third observation, $r_3 = \frac{1}{25}[-3y_1 - 3y_2 + 12y_3 \ldots]$; to complete the formula, a condensed notation is introduced. The 5x5 array below displays the coefficients of $y_1, \ldots, y_{25}$ (in the order specified in Section A.1) for calculation of r_3:

$$\frac{1}{25} \begin{bmatrix} -3 & -3 & 12 & -3 & -3 \\ 2 & 2 & -3 & -3 & 2 \\ -3 & 2 & -3 & 2 & 2 \\ 2 & -3 & -3 & 2 & 2 \\ 2 & 2 & -3 & 2 & -3 \end{bmatrix}$$

The set of 25 residuals can be replaced by a set of twelve statistically independent residuals $r'_1, r'_2, \ldots, r'_{12}$ in such a way that $\sum_{i=1}^{25} r_i^2 = \sum_{j=1}^{12} r'^2_j$; the method is indicated in a paper by Cameron [1]. This has been done in such a way that only the first four orthogonal residuals contain information about the value of σ^2_m. The first four

are displayed below in two ways. First, for convenient calculation (if the residuals are already available from a least squares analysis by computer):

$$r'_1 = \frac{5\sqrt{3}}{6} r_3$$

$$r'_2 = \frac{\sqrt{5}}{6} (r_3 + 4r_8)$$

33

$$r'_3 = \frac{\sqrt{10}}{6} (r_3 + r_8 + 3r_{13})$$

$$r'_4 = \frac{\sqrt{30}}{6} (r_3 + r_8 + r_{13} + 2r_{18})$$

The quantity SSM shown in Table A-1 may be calculated from $SSM = \Sigma_{j=1}^{4} r_j'^2$, but it is readily verified that SSM is actually proportional to the sum of squares of the five residuals in the middle (machine operator) column: $SSM = \frac{5}{3} (r_3^2 + r_8^2 + \ldots + r_{23}^2)$. There is no need to show the remaining 8 r_j', since the last sum of squares, SSR, is obtained by subtraction.

The second display of $r'_1, \ldots, r'_4$ shows them as linear combinations of original y_i, in the condensed notation introduced above, omitting the display for r'_1 since it is proportional to r_3, whose coefficient array was shown above:

$$r'_2 = \frac{\sqrt{5}}{30}
\begin{bmatrix}
1 & 1 & 0 & 1 & -3 \\
-2 & -2 & 9 & -3 & -2 \\
1 & -2 & -3 & -2 & 2 \\
2 & 1 & -3 & -2 & 2 \\
-2 & 2 & -3 & 2 & 1
\end{bmatrix}$$

$$r'_3 = \frac{\sqrt{10}}{30}
\begin{bmatrix}
1 & -2 & 0 & 1 & 0 \\
-2 & 1 & 0 & 0 & 1 \\
-2 & -2 & 6 & -1 & -1 \\
2 & 1 & -3 & 1 & -1 \\
1 & 2 & -3 & -1 & 1
\end{bmatrix}$$

$$r'_4 = \frac{1}{\sqrt{30}}
\begin{bmatrix}
-1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & -1 & 1 \\
0 & -1 & 3 & -1 & -1 \\
1 & 0 & -3 & 1 & 1
\end{bmatrix}$$

Notice first the middle column of each array: there are four orthogonal comparisons among the observations made at the machine-operator location. The smaller coefficients in other parts of the array are due to correlations among the original residuals, but the row-sums, column-sums, and sampler-sums of these coefficients are all zero, so it is clear they are not estimating any of the main effects (under the assumptions of our statistical model).

The ratio of the two mean squares obtained from this decomposition of SSE provides an F-test which is not approximate, $F=\dfrac{SSM/4}{SSR/8}$ with 4,8 d.f. One such test will not provide a definitive result, but an accumulation from several runs of the unit design can increase confidence in the evaluation.

As a final remark, note that the orthogonal decomposition is not unique: for example, it could have started with

$$r_1'' = 5\sqrt{3}\ r_{23}/6$$

and used the middle-column residuals from the bottom up. But the sum of squares SSM would always have the same value.

A.5 Combining Results from Several Designs

The proposals for analysis given in this section are less specific than those for the unit design, since the approach to analysis depends on (1) relative magnitudes of main effects and variance components defined for the unit design and (2) the purpose of the analysis.

The two chief purposes that can be served by analysis of a group of unit designs (say, the group proposed for the pilot experiment) are (a) to assess the uncertainty of the 10-day (or 5-day) average calculated for checking compliance with the dust level standard under current regulations and (b) to assess the uncertainty of the dust level determinations made under current regulations when they are considered as measurements to assess the effectiveness of dust control. In the latter case, the large day-to-day variations of dust level are important and must be treated in a new way in a statistical model for analysis of the data.

Analyses for these two purposes will be discussed separately. Section A.5.1 is on the uncertainty of a 10-day average and covers the analyses that would be considered for the pilot experiment, including those needed to determine whether a large survey would be desirable. Section A.5.2 introduces the need for a new model for analysis of the uncertainty of dust level measurements in the broader context.

A.5.1 Uncertainty of Ten-Day Average

In the basic statistical model for evaluation of the uncertainty

35

of a ten-day average, it is assumed that there is a fixed dust level μ_d for each of the $d=1,2,\ldots,10$ days:

$$E(y_d) = \mu_d \quad ,$$

where y_d is the observation on the dth day. The quantity of interest for regulatory action is

$$\mu = (\mu_1 + \ldots + \mu_{10})/10 \quad ,$$

and its value is estimated by

$$\bar{y} = \Sigma_d y_d/10 \quad .$$

The observed values y_d differ from their expected values because of "error" and, since the machine operator's personal sampler is used for these measurements, the statistical model includes both types of error defined in the statistical model for the unit design. Observation y_d is represented by

$$y_d = E(y_d) + e_d + m_d \tag{M-1}$$

$$= \mu_d + e_d + m_d$$

The errors are assumed to be statistically independent and to have variances σ_e^2, σ_m^2, respectively.

Data obtained in the pilot experiment would be used to estimate σ_e^2 and σ_m^2 and the variance of $\bar{y}$,

$$Var(\bar{y}) = (\sigma_e^2 + \sigma_m^2)/10 \quad ,$$

and to assess the precision of the estimate of $Var(\bar{y})$. Data obtained in the pilot experiment would also be used to check the adequacy of the simple model (M-1) stated above and the assumptions about the errors.

The model (M-1) is closely related to the model adopted for analysis of the unit design. Here μ_d includes the over-all average dust level plus the effect of the dth day. An equivalent expression would be

$$\mu_d = \mu + \theta_d \quad (\theta_1 + \ldots + \theta_{10} = 0) \quad .$$

The effect of location is not included, since all measurements are made at the machine-operator location. Omission of sampler effects from the model is discussed below.

The chief analyses of data from the pilot experiment would be directored to assessing the relative magnitudes of variance components. (1) Each unit experiment provides a test of the hypothesis $\sigma_m^2 = 0$ (the F-ratio derived at the end of Section A.3.2). (2) Each of the unit experiments provides an estimate of σ_e^2 based on 8 degrees of freedom.

36

These would be tested for equality (e.g., by Bartlett's test, see
Dixon and Massey [2], pages 179-180). If they should appear not to
be equal, possible reasons such as (2a) monitored vs. unmonitored
experiments, and (2b) differences between machine operators or mines,
would be tested. Depending on the outcome of analyses for equality of
variance, there could be (3) one or more sets of pooled estimates of
σ_e^2 and σ_m^2, the latter being estimated only if it appears to be signif-
icantly different from zero.

The effect of the sampler (or samplers) used might have to be
included in the statistical model, and does have to be included in the
complete uncertainty analysis. Two cases can be described briefly,
though the actual analysis would probably require additional information:
(1) the machine operator selects "at random" a sampler from the collection
maintained at the mine; (2) the machine operator uses the same sampler on
each of the ten days. In the first case, the statistical model might
need to be amplified by adding an additional "error" due to sampling of
samplers. The magnitude of this error could be estimated by laboratory
tests of company samplers; it is possible that it would be small enough
to be neglected (in effect, incorporated into the other errors). In the
second case, there would need to be an estimated limit for any systematic
error that might be present due to the particular sampler used, to be
added to the statistical uncertainty limit for the estimated average
dust level.

Differences among samplers are estimated in the pilot experiment
but the pilot experiment would not provide information about variability
of company samplers if Bureau of Mines samplers were used. Unless there
is strong reason to believe that the accuracy of samplers as maintained
by the company has an effect on the way they are used by the machine
operators, it appears advantageous to estimate the variability due to
differences among samplers in a separate investigation. (A larger
survey following the pilot experiment might include the use of company
samplers if evaluation of this source of variability seemed to require
further investigation).

The summary of analyses of the pilot experiment would indicate
strategies for stratified sampling in the larger population of mines,
and give variance estimates for use in determining sample sizes needed
in a larger survey.

A.5.2 Over-all Uncertainty

When the ten-day average dust level is thought of as a sample
estimate of the average dust level in a mine, the daily average dust
level (even if measured without error) is a value obtained on a sample
day. A statistical model appropriate for analysis in this broader
context postulates a long-run average dust level μ for a mine and
represents a combination of explanatory and random variables that may
be involved in departures from the long-run average.

Examples of explanatory variables that might be considered are

type of coal seam or any other operating conditions that may be associated with dust level.

The model might be expressed in the following form:

$$y_d = \mu + \sum_{k=1}^{P} \beta_k x_{kd} + M_d + e_{wd} \quad ,$$

where

y_d = observed level on day d,

x_{kd} = value on day d of the kth explanatory variable (e.g., 1 or 0 for presence or absence of some condition),

β_k = regression coefficient for kth explanatory variable,

M_d = random effect of day d,

e_{wd} = within-day error on day d (includes e_d and m_d).

Statistical assumptions about the within-day errors e_{wd} are derived from those made above for e_d and m_d. The random day-to-day differences M_d are assumed to be statistically independent of the other errors, to have zero expected values, and variance σ_d^2.

Analysis of the pilot experiment provides estimates of σ_d^2, since each unit experiment provides a mean square for days. With the revised model, the expected mean square for days is given by the expression in Table A-1 with $(\Sigma\theta_d{}^2/4)$ replaced by σ_d^2. The F-ratio comparing SSD/4 to SSE/12 gives an approximate F-test for the hypothesis $\sigma_d^2=0$. Available evidence suggests that this statistical test will not be needed--that the day-to-day variability is vastly larger than within-day variability due to differences among samplers and machine operators.

The variance of a ten-day average, in terms of the expanded model, is

$$(\sigma_d^2 + \sigma_m^2 + \sigma_e^2)/10 \quad .$$

A principal outcome of the pilot experiment would be to assess the need for precise information about the values of σ_m^2 and σ_e^2, considering the magnitude of their contribution to the over-all uncertainty.

A caution should be expressed.

In all the suggested analyses of the pilot experiment, it must be expected that "problems" would arise. If samplers malfunction or if the rotation scheme for a unit experiment is not followed exactly the statistical analysis becomes more difficult. The formulas given in Sections A.3 and A.4, and the analyses suggested in this section cannot be applied mechanically, but should be preceded or accompanied by graphical analyses to check the general assumption that the statistical distributions of errors are reasonably well-behaved. Should such checks fail, the interpretation of estimated variances would need to be made with great care.

A.6 A Numerical Example

The following is a numerical example carrying out the computations required for the analysis of variance exhibited in Table A-1. The experiment in which these data were obtained did not exactly follow the unit design, in that the samplers were assigned to fixed locations, but the numerical computations have been done as if the samplers had been rotated. Interpretation follows the computations.

Data From Test Shift 1,2,3,6,7, EEAC

Test Series F, A-Shift

(Attached to Letter From Welby Courtney to H. H. Ku, Nov. 7, 1977)

Location

Day	RL	RR	M	FL	FR	
1	1.10	1.12	1.42	2.01	2.78	8.43
2	2.44	2.41	2.54	4.57	6.44	18.40
3	0.52	0.64	2.23	1.75	1.68	6.82
4	3.57	3.46	4.19	5.04	6.36	22.62
5	4.43	4.51	9.18	7.89	7.91	33.92
	12.06	12.14	19.56	21.26	25.17	90.19

Sampler Totals

Sampler	A	B	C	D	E
	1.10	2.44	0.52	3.57	4.43
	4.51	1.12	3.46	2.41	0.64
	4.19	2.23	1.42	9.18	2.54
	1.75	7.89	4.57	2.01	5.04
	6.44	6.36	7.91	1.68	2.78
	17.99	20.04	17.88	18.85	15.43

39

Estimates of Main Effects

for the Model: $E(y_i) = \mu + \theta_d + \theta_\ell + \theta_s$

$\hat{\mu} = \dfrac{90.19}{25} = \qquad\quad +3.6076$

$\hat{\theta}_d$; d=1	8.43/5	$-$	3.6076	$=$	-1.9216	
2	(18.40/5)	$-$	"	$=$	$+\ .0724$	
3	(6.82/5)	$-$	"	$=$	-2.2436	days from average
4	(22.62/5)	$-$	"	$=$	$+\ .9164$	
5	(33.92/5)	$-$	"	$=$	$+3.1764$	

$\hat{\theta}_\ell$; ℓ=1	(12.06/5)	$-$	"	$=$	-1.1956	
2	(12.14/5)	$-$	"	$=$	-1.1796	
3	(19.56/5)	$-$	"	$=$	$+\ .3044$	locations from average
4	(21.26/5)	$-$	"	$=$	$+\ .6444$	
5	(25.17/5)	$-$	"	$=$	1.4264	

$\hat{\theta}_s$; s=A	(17.99/5)	$-$	"	$=$	$-\ .0096$	
B	(20.04/5)	$-$	"	$=$	$+\ .4004$	
C	(17.88/5)	$-$	"	$=$	$-\ .0316$	samplers from average
D	(18.85/5)	$-$	"	$=$	$+\ .1624$	
E	(15.43/5)	$-$	"	$=$	$-\ .5216$	

See pp. 13–32, <u>Experimental Statistics</u> (reference [12] of report).

Computation of r'

	3	8	13	18	23
$\hat{\mu}$	3.6076	3.6076	3.6076	3.6076	3.6076
$\hat{\theta}_d$	-1.9216	.0724	-2.2436	.9164	3.1764
$\hat{\theta}_\ell$	.3044	.3044	.3044	.3044	.3044
$\hat{\theta}_s$	- .0316	- .5216	.4004	- .0096	.1624
y	1.9588	3.4628	2.0688	4.8188	7.2508
y	1.42	2.54	2.23	4.19	9.18
r	- .5388	- .9288	.1612	- .6288	1.9292
r'	- .7777	-1.5764	- .5154	-2.3351	(undefined)

$SSM = \frac{5}{3}\Sigma r'^2 = 8.8084$

Residual $= \underline{\ 3.9419}$

SSE $\quad$ 12.7503

Check: $SSM = \frac{5}{3}\Sigma r^2 = 8.8084$

41

Analysis of Variance of Unit Design

Sources of Variability	Degrees of freedom	Sum of Squares	Mean Square
Correction factor	1	325.3694	
Days (SSD)	4	98.3042	24.5761***
Samplers (SSS)	4	2.2993	.5748
Locations (SSL)	4	26.8173	6.7043***
Front (left vs. right)	1	1.5288	1.5288
Rear (left vs. right)	1	.0006	.0006
Front vs. Rear	1	24.7086	24.7086***
Machine op. vs. other	1	.5791	.5791
Error (SSE)	12	12.7501	1.0625
Machine operator (SSM)	4	8.8082	2.2020*
Residual (SSR)	8	3.9419	0.4927

* Significant at .05 level

*** Significant at .01 level

The main purpose of the numerical example is to illustrate the computations required for the analysis of variance of a unit experiment. In addition, the interpretation of the results may be discussed--with caution.

In the analysis of variance table, notice that the Samplers effect has a small mean square. This is consistent with the fact that the samplers were not rotated, and the Samplers mean square (MSS) is actually estimating the same thing as the Error mean square (MSE). The "Locations" effect, on the other hand, includes all of any sampler differences that may have been present.

When the Locations effect is broken down into four parts, it appears that the left-right differences (samplers packaged together) are small or non-existent, unless there were compensating differences between the samplers.

The estimate of σ_e^2, from MSR, is

$$\hat{\sigma}_e^2 = 0.4927 \quad ,$$

and the estimate of σ_m^2, using the formulas for the expected mean squares shown in Table A-1, is

$$\hat{\sigma}_m^2 = \frac{5}{3} \; (MSM-MSR)$$

$$= 2.8488 \quad .$$

These numerical results illustrate those that might be expected from a unit experiment, but the verification of assumptions using a Dry Run is nevertheless still necessary.

The computations for this numerical example were repeated using the logarithms of the data, but the results were the same qualitatively (including the tests of significance) and are not shown.

A.7 References for Appendix A

[1] J. M. Cameron, An Algorithm for obtaining an orthogonal set of individual degrees of freedom for error, J. Research NBS - B. Math. & Math. Physics, Volume 67B (1963), pp. 19-22.

[2] Dixon, W. J., and F. J. Massey, Jr., Introduction to Statistical Analysis, Second Edition, McGraw-Hill, 1957.

NBS-114A (REV. 7-73)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO.	2. Gov't Accession No.	3. Recipient's Accession No.
4. TITLE AND SUBTITLE A Suggested Experimental Plan for the Determination of the Accuracy and Reliability of Coal Mine Dust Measurements			5. Publication Date
			6. Performing Organization Code
7. AUTHOR(S) Harry H. Ku and Joan R. Rosenblatt			8. Performing Organ. Report No.
9. PERFORMING ORGANIZATION NAME AND ADDRESS NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234			10. Project/Task/Work Unit No.
			11. Contract/Grant No. J0177013
12. Sponsoring Organization Name and Complete Address *(Street, City, State, ZIP)* U. S. Department of the Interior Bureau of Mines 4800 Forbes Avenue, Pittsburgh, PA 15213			13. Type of Report & Period Covered
			14. Sponsoring Agency Code

15. SUPPLEMENTARY NOTES

16. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

In an agreement between the Bureau of Mines and the National Bureau of Standards dated February 1, 1977, on the "Determination of the Accuracy and Reliability of Coal Mine Dust Measurements," NBS is to submit a report to BOM at the conclusion of the Phase I part of the work. This document is the report called for in the agreement.

In Section I of this report, the terms "reliability" and "accuracy" are clarified and defined for the purpose of this investigation. Certain pertinent information, collected as results of personal contacts and study of existing literature, is presented and discussed. In Section 2 an experimental design is suggested to measure the various sources of variability as relating to the use of approved personal samplers "by coal miners in underground mines." In Section 3, a dry run is suggested to test the efficacy of the design. Pilot experiments on 10 to 12 mines are also recommended to obtain preliminary estimates of variabilities due to different sources. In Section 4 an overall view is presented as to factors that may contribute to the resulting uncertainty of coal mine dust measurements.

17. KEY WORDS *(six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons)*

Coal mine dust; accuracy; reliability; design of experiment; personal sampler; pneumoconiosis.

18. AVAILABILITY ☐ Unlimited	19. SECURITY CLASS (THIS REPORT)	21. NO. OF PAGES
☒ For Official Distribution. Do Not Release to NTIS	UNCLASSIFIED	
☐ Order From Sup. of Doc., U.S. Government Printing Office Washington, D.C. 20402, <u>SD Cat. No. C13</u>	20. SECURITY CLASS (THIS PAGE)	22. Price
☐ Order From National Technical Information Service (NTIS) Springfield, Virginia 22151	UNCLASSIFIED	

USCOMM-DC 29042-P74