RJR INTEROFFICE MEMORANDUM

AUTHORS: W. T. Morgan
R. E. Sells

DATE: July 5, 1988

Subject: Tobacco Blended Subgroup Variability

In February of 1988, a study was begun at Tobaccoville to quantify the variability observed in blended subgroup tobaccos (nicotine and sugar). This included the flue cured subgroups CG1, CG2, and CG3, burley subgroups KG1, KG2, and KG3 (nicotine only), and Salem KS FBS. Observed cigarette tar variability had raised questions about blend uniformity, leading to the formation of a group to study subgroup variability. Also of interest in the study was comparing subgroup values to Final Blended Strip (FBS) values. The test was planned by representatives of Brands R&D and Quality Assurance, and run with the complete cooperation of Manufacturing. All laboratory testing was done by Quality Assurance. This memorandum will discuss the details of test design, sample collection, data analysis and results.

Test Design

As all tobaccos are handled by bulkers at Tobaccoville, the natural sources of variability in the subgroup blends are

1) Analytical variability
2) Variability within samples
3) Variability within bulkers
4) Variability between bulkers.

Analytical variability is variability of the analytical method, that is, variability of the measurement of identical samples caused by the measurement technique. Analytical variability was not measured in this study because it has been measured in other studies. However, it will be discussed here relative to the other sources of variability.

Variability within samples was observed by taking side by side samples from the belt at bulker discharge, rather than by splitting samples. This was done to separate it from analytical variability, and also to provide an estimate of variability from the 'rational subgroup' perspective of traditional quality control methodology. Thus, this is really more of a 'short term' variability measurement than a true 'within sample' variability measurement.

Variability within bulkers was observed by taking four samples from each bulker, spaced throughout the bulker, and comparing them to the bulker average, for each bulker. Variability between bulkers was observed by comparing bulker averages to each other, to see if they differed by more than would be expected as a result of the other components of variability.

Variance Components Model

These components of variability are used to describe the variance of a single measurement. The variance components model used hypothesizes that the process (CG1 nicotine, for example) has an overall average level, and that deviations from that overall level are caused by the sources of variability discussed above. Thus, the observed value for a single measurement is the overall average, plus a random deviation of that particular bulker from the overall average, plus a random deviation of that particular sample from the bulker average, plus a random deviation of the particular part of the sample from the sample average, plus a random deviation of the particular measurement from the value of the part of the sample used.

These random deviations will differ from bulker to bulker and sample to sample, etc., and thus describe populations for which variances can be estimated. The statistical analysis described here performed this estimation. If these distributions are assumed to be normal, then tests of statistical significance can be performed to see if any of these variance components are significantly different from zero. If not, then it can be argued that the corresponding source of variation does not actually contribute anything to the variance of the response.

Sampling

For this test, 30 bulkers of each type of material were sampled. Samples were taken four times from each bulker, two of which were duplicates (side by side) for a total of six samples. Duplicate samples were not taken for all cases because of the increased lab work, while those samples would not contribute to knowledge about within or between bulker variability. Thus the study provided 60 independent observations of within sample variability (degrees of freedom), 90 independent observations of within bulker variability, and 29 independent observations of between bulker variability. These are adequate numbers of samples to use to estimate the variance components of interest with good precision.

Samples were collected by Manufacturing and sent to Quality Assurance for analysis. All sampling was done on day shift so that the bulkers sampled were not necessarily consecutive. If a bulker began discharge during the other shift or too late in the day shift to get four samples, the bulker was not sampled. As the sample collector was not available after May 1, the full sampling was not completed for CG2 (19 bulkers sampled), KG1 (29 bulkers sampled), and FBS (22 bulkers sampled). These numbers of samples were judged adequate for the test rather than introducing additional variability that might be caused by a different sampler.

Results and Discussion

Sampling for the test began the week of February 15 and ended the week of April 25, 1988. The statistical procedure used to estimate variance components as discussed above was Analysis of Variance. Results of overall summary statistics are shown in Table I. It should be noted that the standard deviations and coefficients of variation are calculated on a raw basis, that is, without recognition of the sampling structure, and thus should not be compared to values from other tests.

Raw data are plotted against sample number in time sequence with bulkers separated by vertical lines in Figures 1-11 for all variables.

Variance Component Analysis

Table II shows estimated variance components and indicates components that were not significantly different from zero with footnotes. Only some of the within bulker components were not significant, with between bulker components significant in every case. Total variances for nicotine were approximately the same for the flue cured subgroups and FBS, and for the burley subgroups, but not for sugar. Table II also shows that expected analytical variances were .0004 and .0144 for nicotine and sugar, respectively. This component is included in the within sample variance component.

Table III expresses the results of the variance components analysis in terms of the percentage of variability attributable to each source. In general, the analytical variance and the within bulker variance components account for small proportions of the total variability, while the between bulker component is the largest component except in the case of FBS. The fact that the within bulker variance is small indicates that the bulkers are mixing the tobacco well, so that two samples taken from different parts of the bulker will be nearly as alike as two samples taken side by side.

The large between bulker component indicates that the bulker averages are different, that is, the materials going into each bulker are different on a grand average basis. Changes would have to be made to the way materials were selected for inclusion in a given bulker to decrease this component of variance. The relatively small component for FBS indicates that these averages were more consistent than those of the subgroups. Table IV displays basic statistics for the bulker averages for each material.

The large within sample standard deviation for FBS for both nicotine and sugar show the heterogeneity of the materials in the FBS.

The magnitude of the variance components are shown graphically in Figures 12 and 13 for nicotine and sugar, respectively. They are plotted as accumulated variance components, and expressed in standard deviation units (the square root of the variance component) to make them more meaningful. Looking at Figure 12, it is not clear that the between bulker component for CG1 nicotine represents 61% of the variability, as Table III indicates. That is because the height of the bar is the square root of the total variability, and the portion of the bar due to between bulker variability is the excess of between bulker variability above the square root of the sum of the other components. If the components were plotted in units of variance, the proportions plotted would match the values given in Table III.

Confidence Intervals

The variance components can be used to calculate confidence intervals for bulker and process averages. The process average is the overall average level discussed above, about which samples deviate. Thus, a confidence interval for the process average must include variability between bulkers, while a confidence interval for a bulker average includes only the variability within a bulker.

Table V shows 95% confidence intervals for nicotine and sugar for these averages, assuming that a single sample is taken. Thus, if a sample of CG1 is taken and measured to be 2.10% nicotine, the true bulker average will be between 1.97% and 2.23% with 95% confidence (2.10+.13%), and the overall process average (which includes an estimate of the deviation of this particular bulker from the overall average) will be between 1.90% and 2.30% with 95% confidence (2.10+.20%). The width of this confidence interval can be reduced with additional sampling, if desired. The process average confidence interval is not much larger than the bulker average confidence interval for FBS because the between bulker variance component was relatively small for FBS.

Prediction of FBS Nicotine

As discussed above, one of the goals of the test was to see how well FBS nicotine values could be predicted from subgroup values. Since bulkers empty with different cycles, daily averages were calculated for each material by averaging all the bulker averages for all bulkers sampled in that day. Generally this was simply a single bulker average, but in some cases two bulkers were sampled the same day for a single material. These daily averages were used with assumed nicotine values of .90% in G-7 and Turkish to predict FBS nicotine daily averages by using the component values weighted according to the blend formula for Salem KS.

Table VI shows these predicted daily values and the FBS bulker averages observed. Because materials were sampled at different rates, only the first 13 FBS bulkers were sampled before all 30 bulkers of CG3 were sampled, thus only 13 comparisons can be made. The data are plotted in Figure 14 (observed and predicted) and indicate that the observed values were generally less than the predicted values. Figure 15 shows the predicted values plotted against the observed values. They have a correlation coefficient of .40, which is significantly different from zero at a 90% confidence level, but not at 95%. Thus, there seems to be a weak correlation and a bias between these values. The reason for this bias is unclear.

To explore these relationships further, finished product and G-13 values were obtained from QA. The G-13 data are summarized in Table VII and plotted in Figure 16, and show good consistency. Unfortunately, only two finished product samples were taken during the test period. Calculations were made to predict the finished product nicotine using the appropriate values for FBS, G-13, and an assumed value of 1.2% for shorts. Again, there seemed to be a bias with predicted values greater than those observed. These results are given in Table VIII.

Conclusions

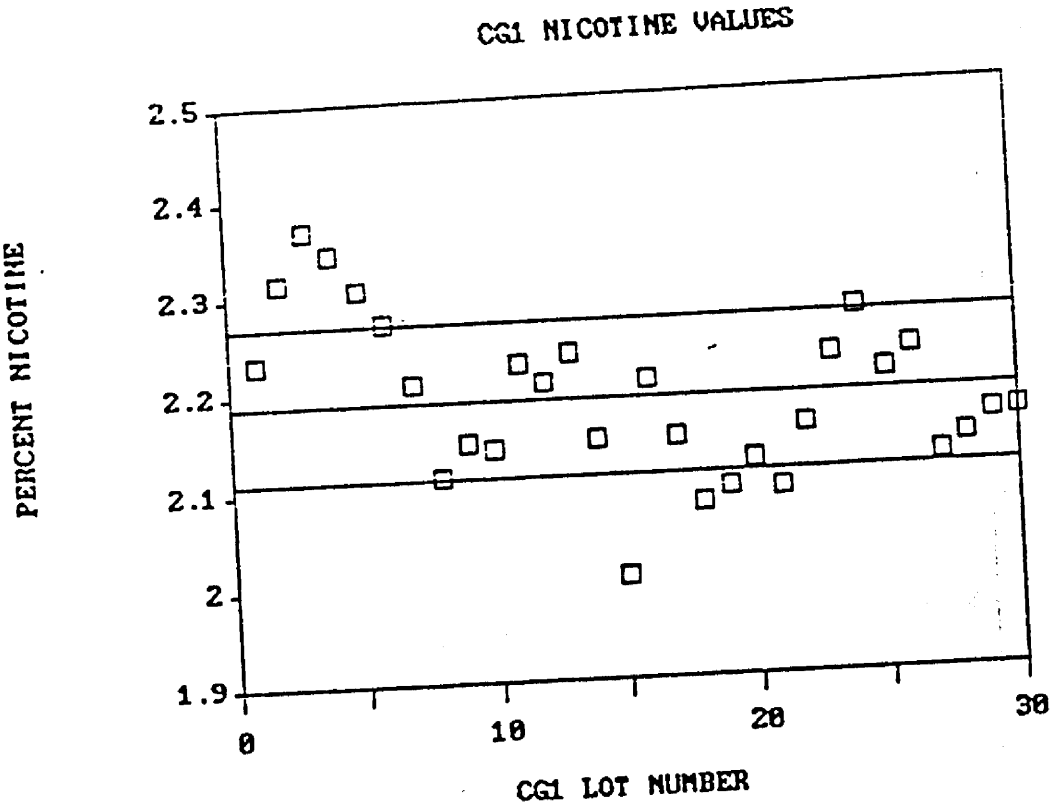
This study was conducted from February to April of 1988, to look at subgroup and FBS nicotine variability, and the use of subgroup values to predict FBS nicotine. The analysis includes estimates of analytical variability, within sample variability, within bulker variability, and between bulker variability for each of CG1, CG2, CG3, KG1, KG2, KG3, and Salem KS FBS. Results showed that

1) Within bulker variability was generally small. This shows that samples from different parts of the bulker are as similar as samples close together.

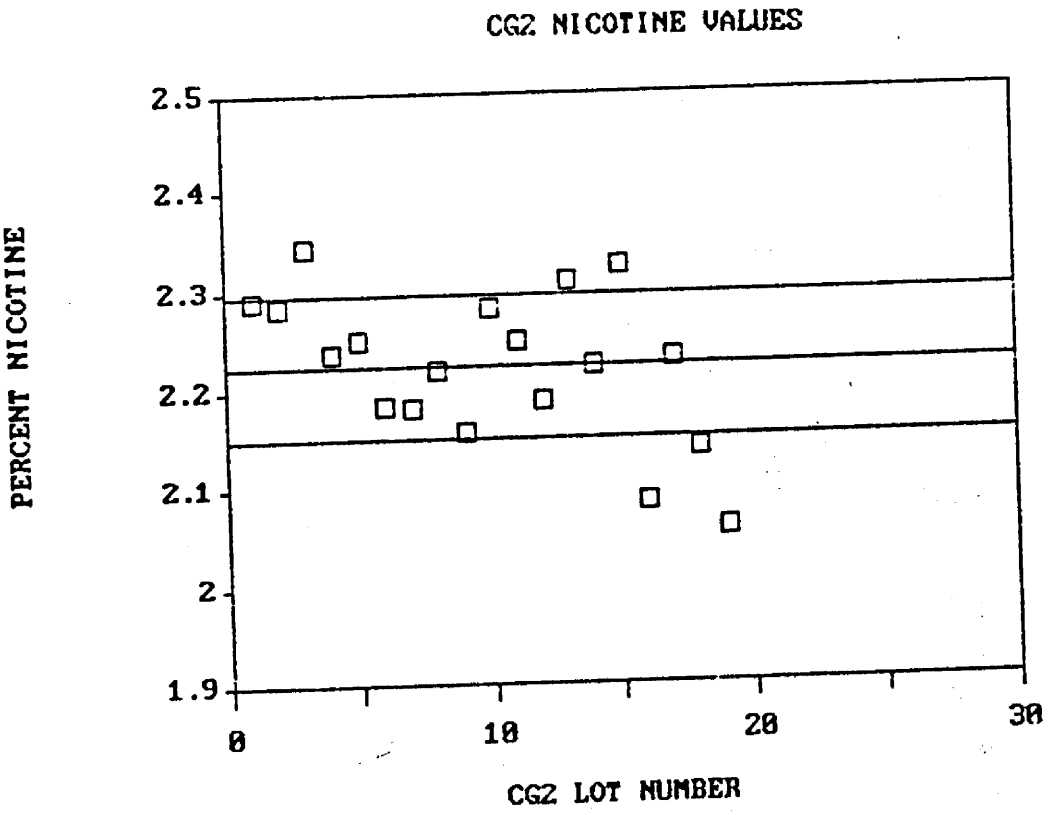
2) Between bulker variability represented a large portion of total variability for the subgroup blends, but not for FBS. Bulker averages for FBS were more consistent than for the subgroup blends.

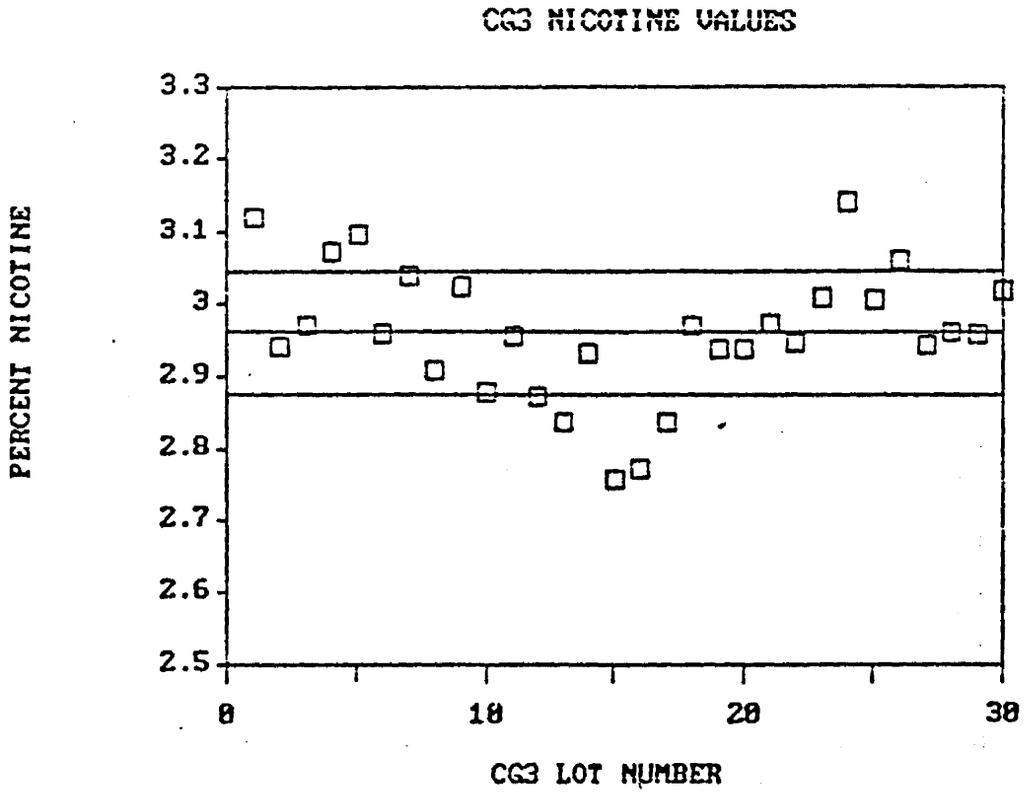
3) Within sample variability was larger for FBS than the subgroup blends. This is most likely due to the heterogeneity of materials in FBS.

4) FBS Nicotine values predicted from subgroup values showed a weak correlation with observed values. However, predicted values were generally about .3% higher than the observed values.

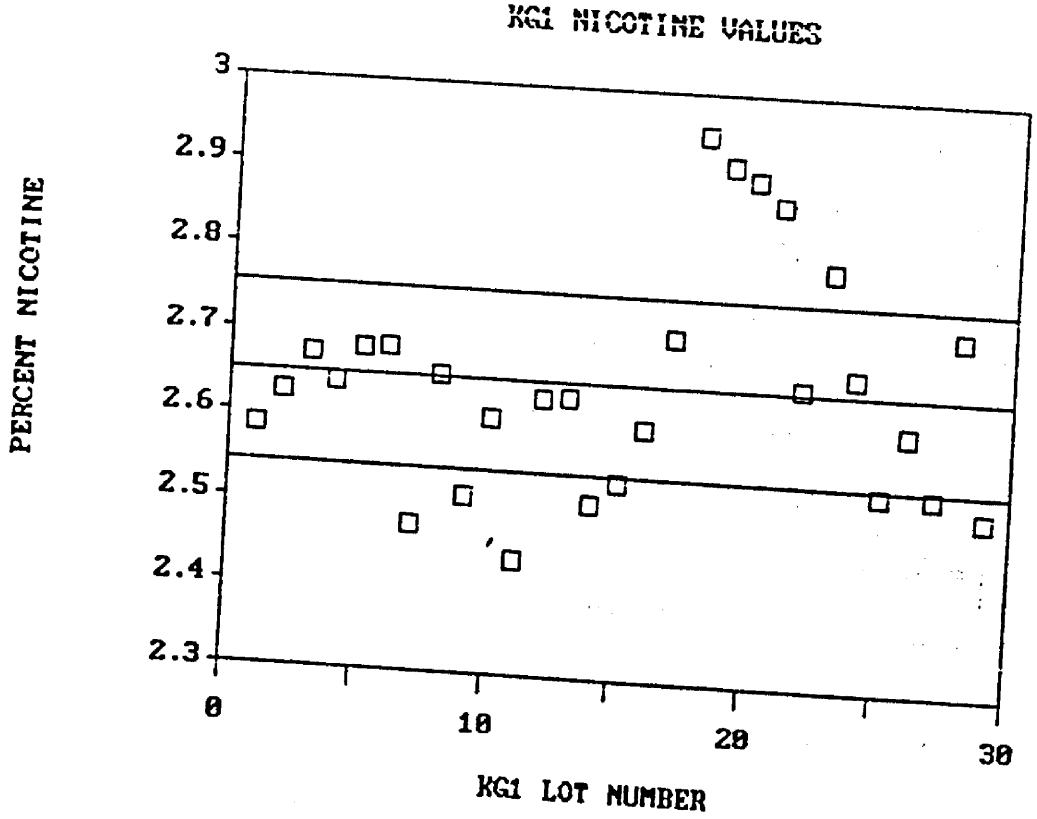


51895 8392

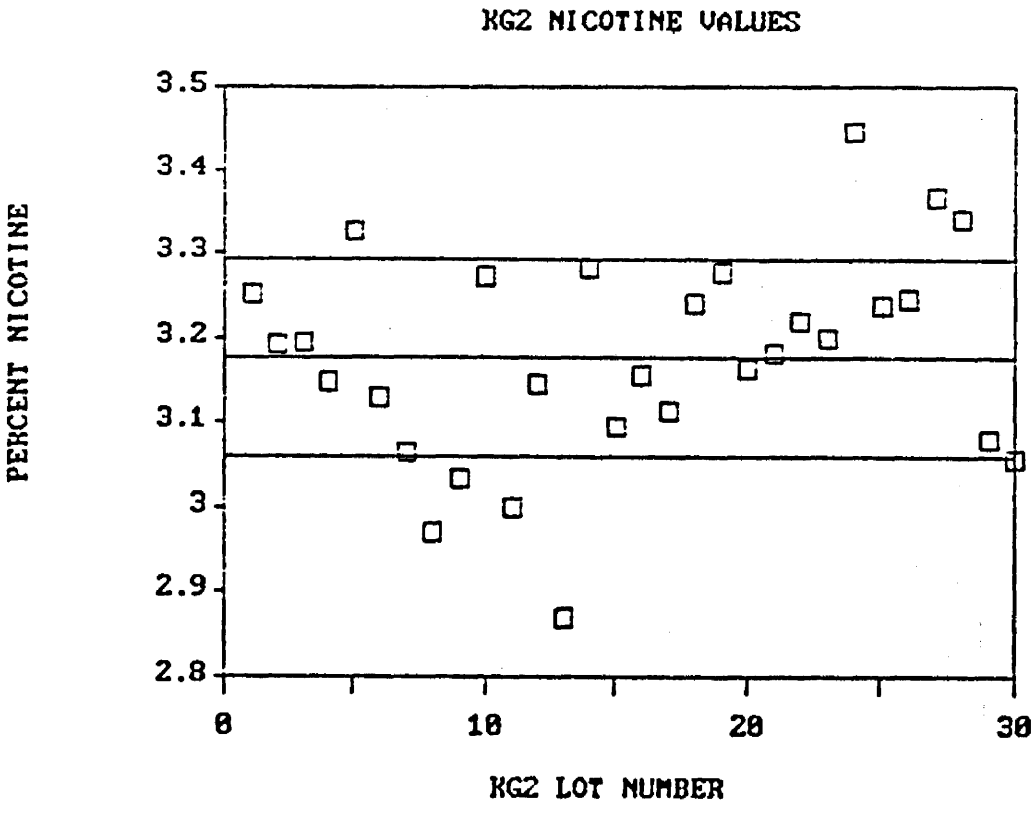


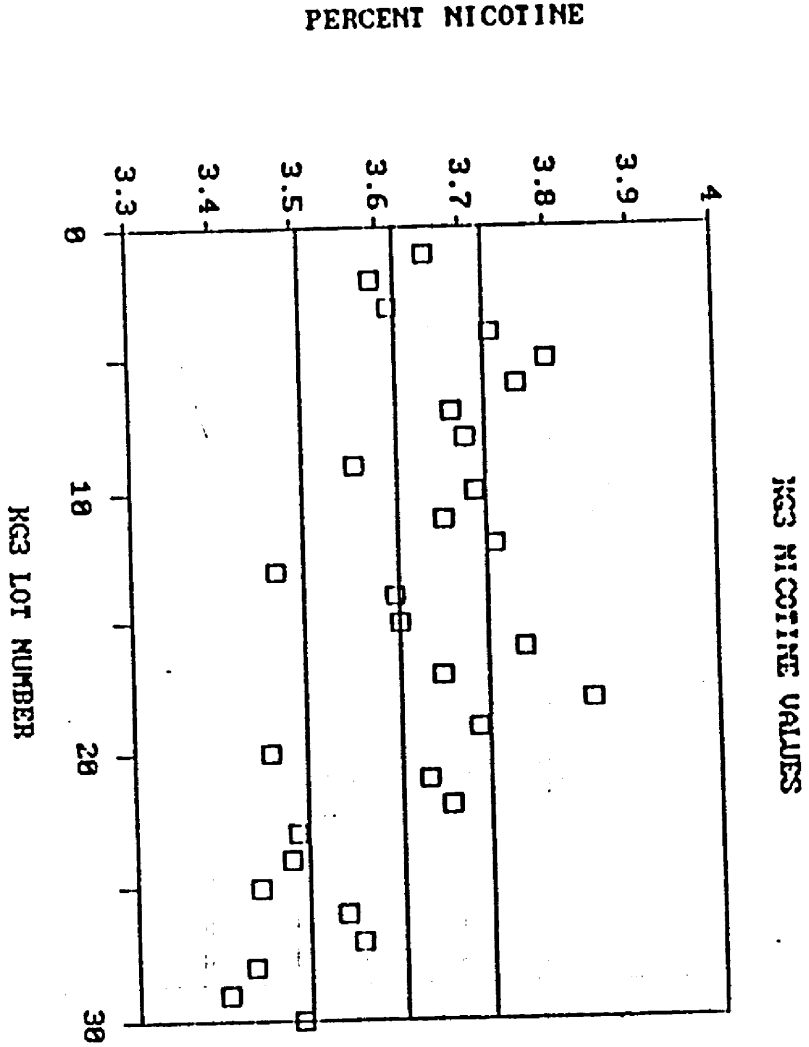


51895 8394



51895 8395





51895 8397