CTS Industry Programs
Guide to Interpreting Performance Analysis Reports

Your **Industry Program Performance Analysis Report,** commonly referred to as your individual report, is designed to serve as the primary tool for evaluation of your results for each test. For most Industry Programs the Performance Analysis Report is a two page per analysis document consisting of a Trend Chart and a Control Ellipse.

The reports have been developed to present a large amount of data in ways that are both easy to interpret and actionable by the participating labs. This guide provides an explanation of the information presented in the Performance Analysis Reports and how they may be used to improve your lab's performance.

This document, along with the *Key to Individual Reports* will provide a framework for understanding the data returned to you by CTS. However, given the broad scope of CTS testing programs, not all interlaboratory statistical analyses can be covered in this document. Please contact CTS if you have additional questions after reviewing this guide.

Table of Contents:

For further information, contact:
Collaborative Testing Services, Inc.
Phone: 571-434-1925
info@cts-interlab.com

An Introduction to the CTS Analyses and CPV Calculations

CTS employs comparative statistics for many of the tests that we offer. For most Programs two sample materials are tested by participants, with the most agreed upon (nominally) consensus value of all participating labs determining the best value for that property.

This involves the calculation of grand means: usually the mean or median of results for all laboratories for a given sample. Standard Deviations both between and within laboratory results are also calculated. CTS uses these statistics to calculate a Comparative Performance Value (CPV) as well.

CPV = (lab mean – grand mean)/between-lab standard deviation

CTS uses the CPV to allow labs to evaluate measurement performance over time. Because small differences in sample/property means and variation are not of critical importance when using the CPV to evaluate performance, laboratory results can be compared from cycle to cycle, even though the samples used may be different. When comparing data among test cycles, remember that such comparisons may be limited if there have been changes to equipment, test procedures, or technicians. Despite the limitations, labs that choose to maintain a continuous approach to the interlaboratory program should find that the Trend Charts provide more than just historical data; the Trend Charts should have diagnostic and/or predictive value.

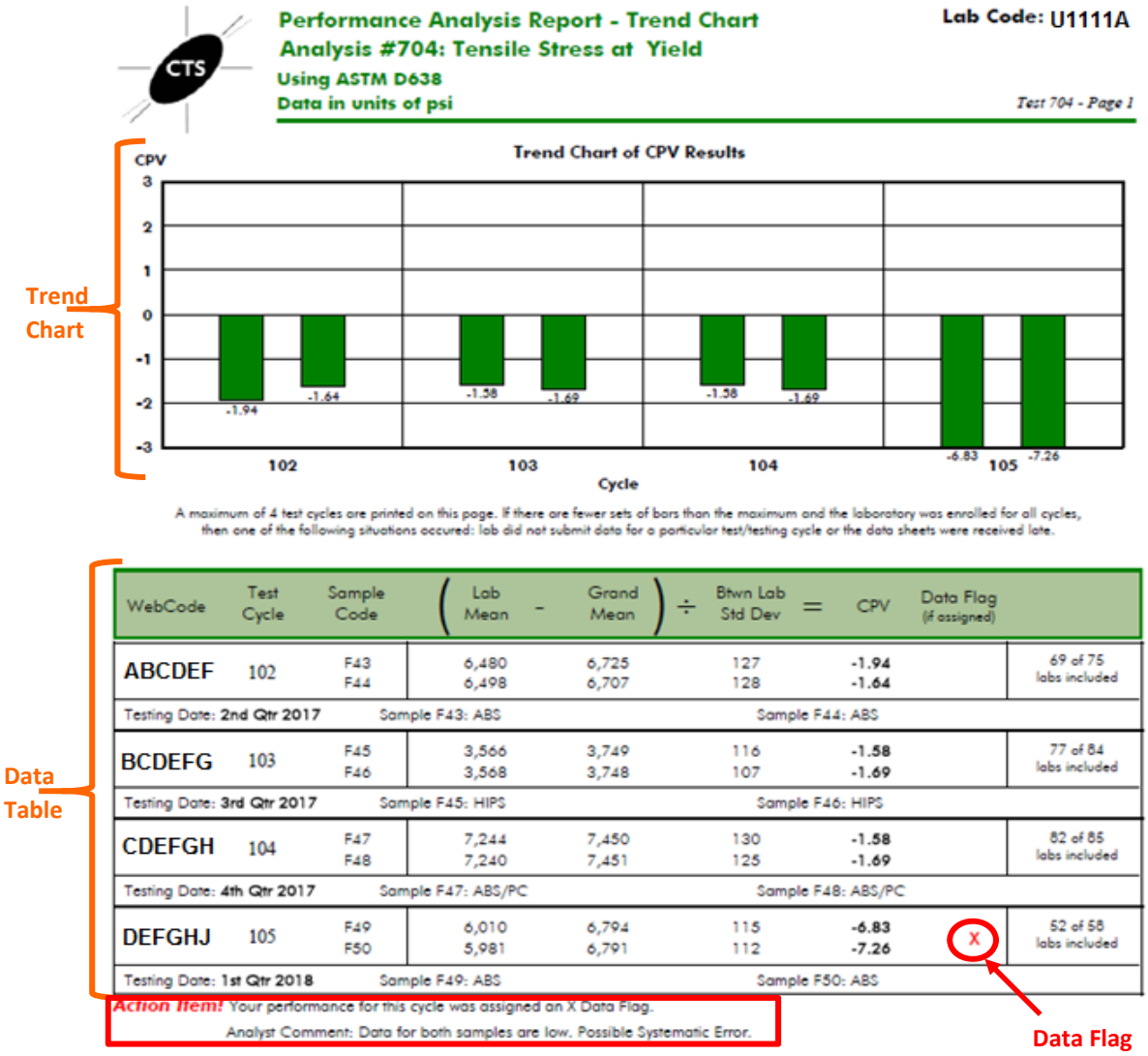
Because similar materials are chosen for both samples, there should be a correlation of measurement performance between the two samples. For most tests CTS uses a bi-variate analysis technique (represented by the ellipse) to judge measurement performance for both samples simultaneously. Quite often measurement performance that differs from the group can be classified as either a **systematic** difference (means for both samples are similarly offset from the group means) or a **consistency** difference (measurements for both samples were not as correlated as other laboratories). If your results received a Data Flag, the Action Item may include our characterization of the error as a systemic or consistency variation. The examples on the following pages illustrate how the information presented in the Performance Analysis Report could be interpreted.

Trend Chart and Data Table Overview

The first page for each analysis is the **Trend Chart and Data Table** page. The example below shows four cycles of testing in the Plastics Program. Within each Cycle the samples are shown with their corresponding CPV value. Up to a year's worth of testing may be shown in the Trend Chart. Any cycle for which data was not returned to CTS will simply be left blank for that cycle on the trend chart.

The Data Table below the trend chart lists the CPVs shown in the trend chart and also provides supporting information. If your results have been flagged in the two-sample analysis and excluded from the statistics, an **Action Item** will be posted below the data, and an analyst comment for the flag will appear. The Action Item should serve as notice of a problem that requires immediate attention. Other types of data flags are discussed in the *Key to Individual Reports*.

Note: Participants in the Color & Color Difference tests will see a separate Trend Chart for all three color spaces (CIE L, CIE a* and CIE b*).*



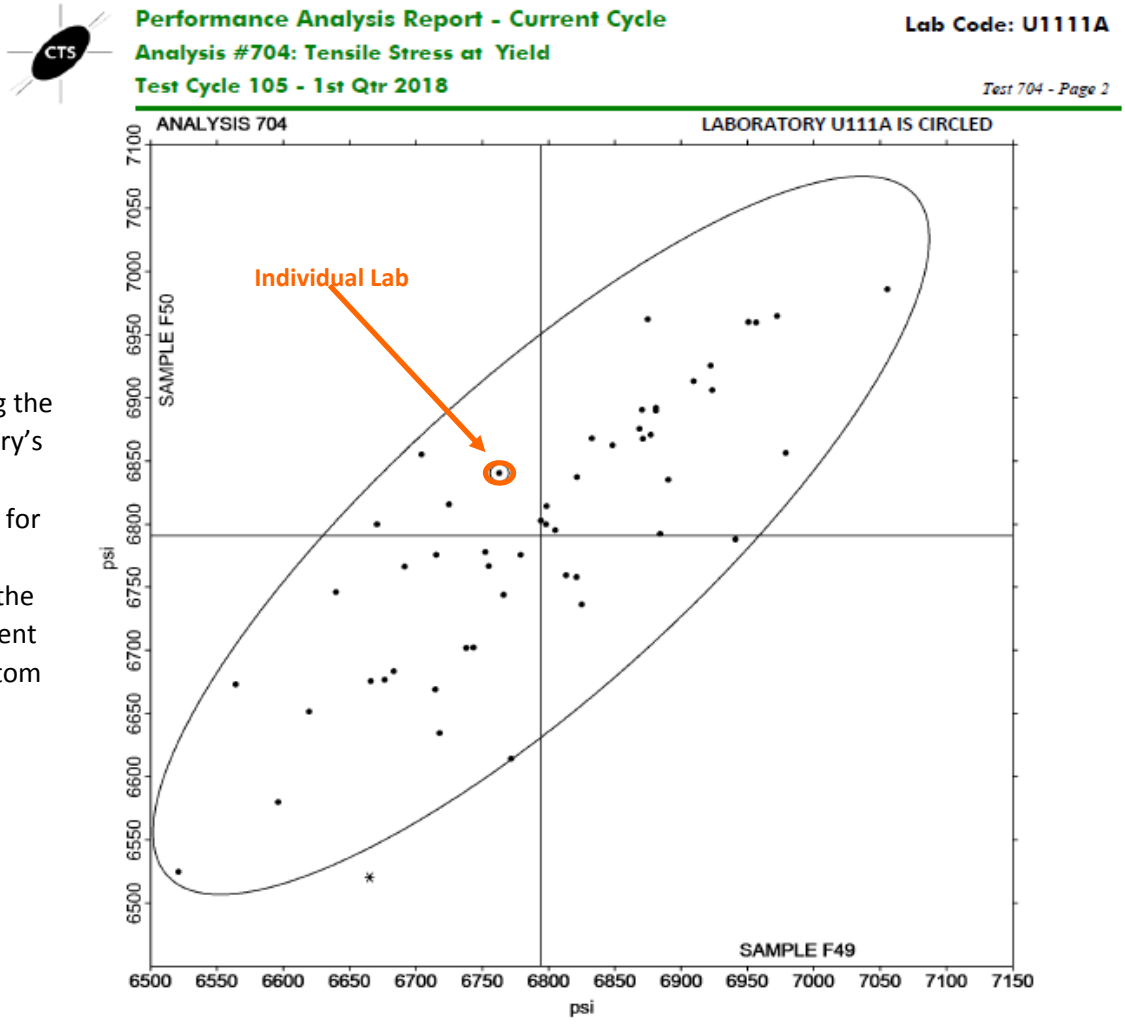
CPV Data Flags: In laboratory testing, we always expect our results to vary from test to test and from lab to lab; but how much variation is too much? That is the question that this report seeks to answer. If your CPV exceeds the warning limit, defined by the 5% significance level, a warning statement will be issued and an "*" data flag will be assigned. CTS advises you to closely monitor tests that trigger a warning signal. If your results exceed the critical limit, defined by a 0.5% significance level, an Action Item will be posted to the data table and an "**X**" flag will be assigned. The Action Item should serve notice of a problem that requires immediate attention.

CTS Control Ellipse Overview

The second page shown for most tests is **The Performance Analysis Report – Current Cycle** page which presents the two-sample plot and control ellipse. The Lab Means for the first sample are plotted on the x-axis and the Lab Means for the second sample are plotted on the y-axis. The cross hairs in the plot represent the Grand Means for both samples. Each participant's data is represented as a data point on the graph. Your data point for the analysis is circled for easy identification. However, if one or both of your lab means is extremely high or low, you may "fall off the plot".

Once the data is plotted, a control ellipse is drawn such that 95% of the time, a randomly selected lab is included inside the ellipse. This control ellipse is a graphical representation of our bi-variate analysis technique. A 99% ellipse is also calculated but not drawn. Results falling between the 95% to 99% ellipses are assigned a "*" warning flag, and data falling outside the 99% ellipse are assigned an "**X**" data flag.

Current Cycle:
The two-sample plot showing the control ellipse. Your laboratory's data point is located at the intersection of the Lab Mean for each sample.
Additional tables containing the statistical results for the current cycle are included at the bottom of the page.



WebCode	Data Flag	Sample F49 Lab Mean	Sample F49 Diff from Grand Mean	Sample F49 CPV	Sample F50 Lab Mean	Sample F50 Diff from Grand Mean	Sample F50 CPV
WEEU6K		6,763	-31	-0.27	6,840	49	0.44

Consensus Results (all laboratories)

Grand Means	6,794 psi	6,791 psi
Btwn Lab Standard Deviation	115 psi	112 psi

Consensus statistics based on 52 of 58 reporting participants

Printed March 07, 2018

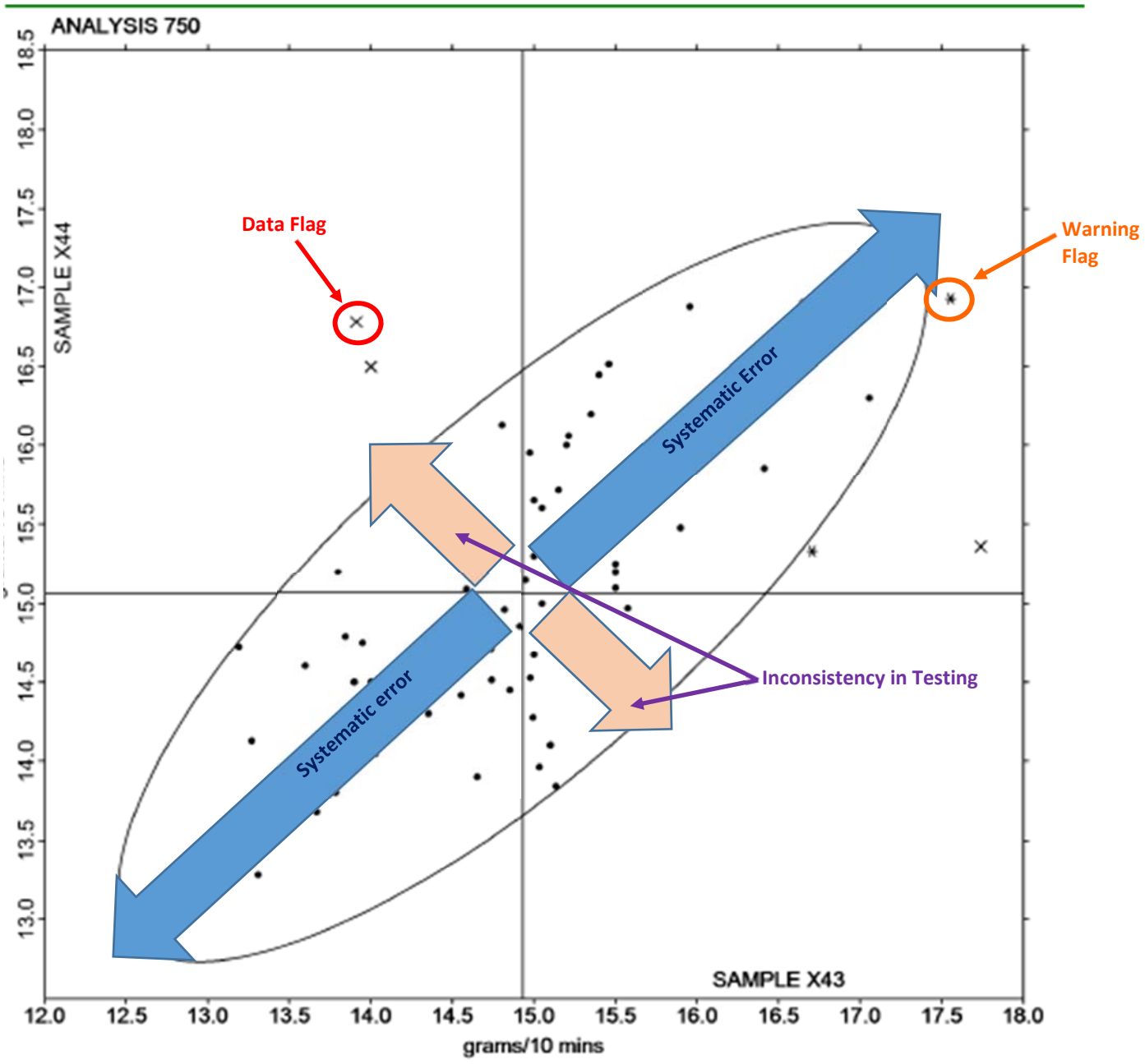
Copyright @ 2018 CTS, Inc.

Interpreting the Control Ellipse

When considering your lab's position on the plot relative to the control ellipse, remember that, generally speaking, if a lab's plotted point falls on the major axis of the ellipse, the lab is consistent in its measurements between the two samples but exhibits an offset from the grand mean (systematic bias/error). If a plotted point falls to the side of the ellipse, it indicates possible differences in the way that the lab tested the two samples or differences in lab sample behavior (inconsistency in testing). The two-sample plot enables you to see which sample, if either, is out of control and to ascertain the nature of the out of control situation.

A robust analysis will result in a narrow ellipse oriented approximately a 45 angle. If a particular analysis/sample combination did not show bias, the control ellipse would become a circle.

You will notice a correlation between your Trend Chart for the cycle and your position on the plot. For example, if both CPV bars are above or below the zero-line, you will find your lab in the upper right or lower left quadrant, respectively. If your lab falls in the lower right or upper left quadrant of the plot, your CPV bars go in opposite directions.

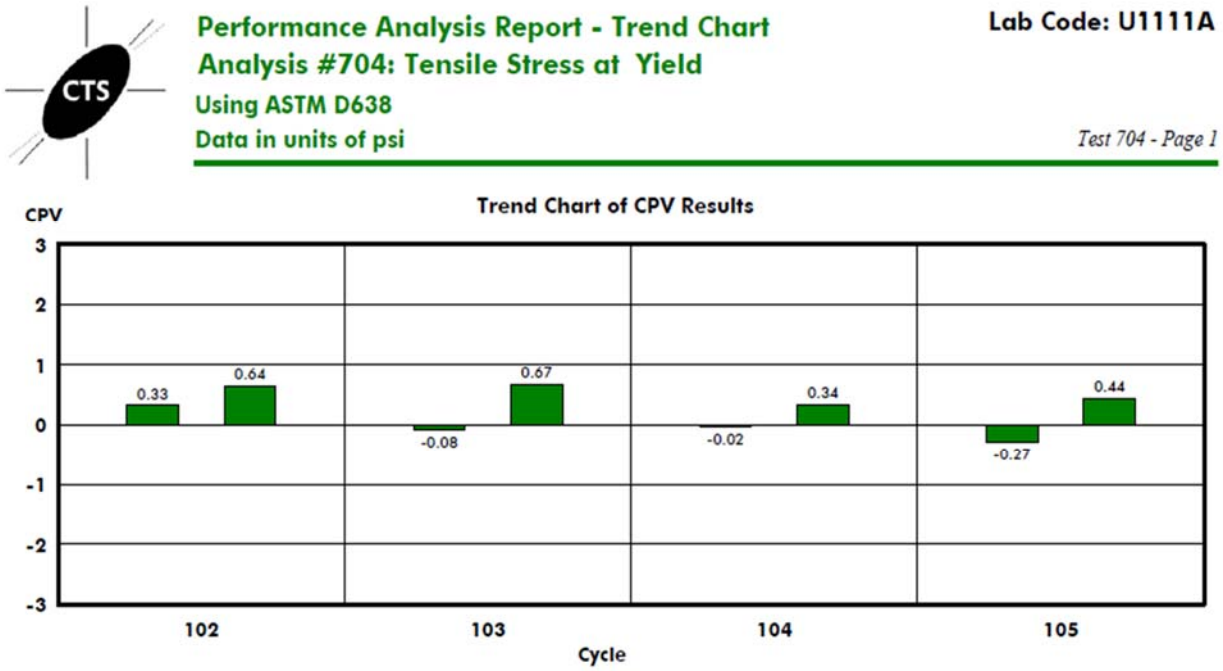


Interpreting the Trend Chart

Keeping in mind the limitations if factors such as instrument, testing procedure, or technician have changed between cycles, the following examples show how the results presented in the Trend Chart portion of the Performance Analysis Report can be interpreted.

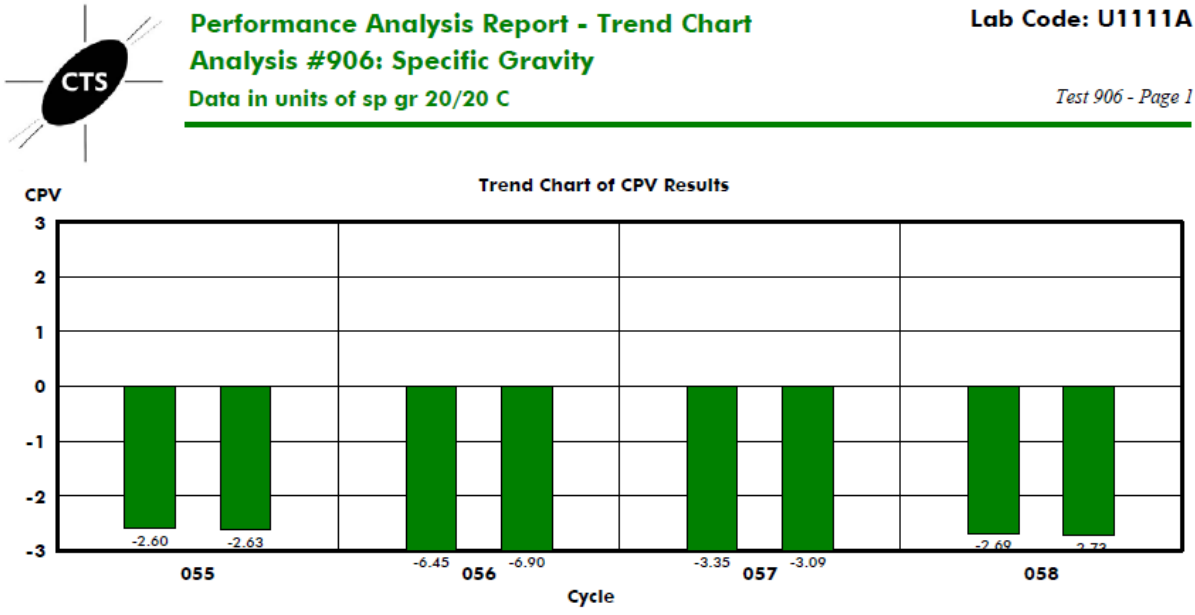
Consistency

The results below are consistent from round to round showing only normal and acceptable flutter around the Grand Mean, with all CPVs between -1.00 and +1.00. This should give the lab greater confidence in its measurements.



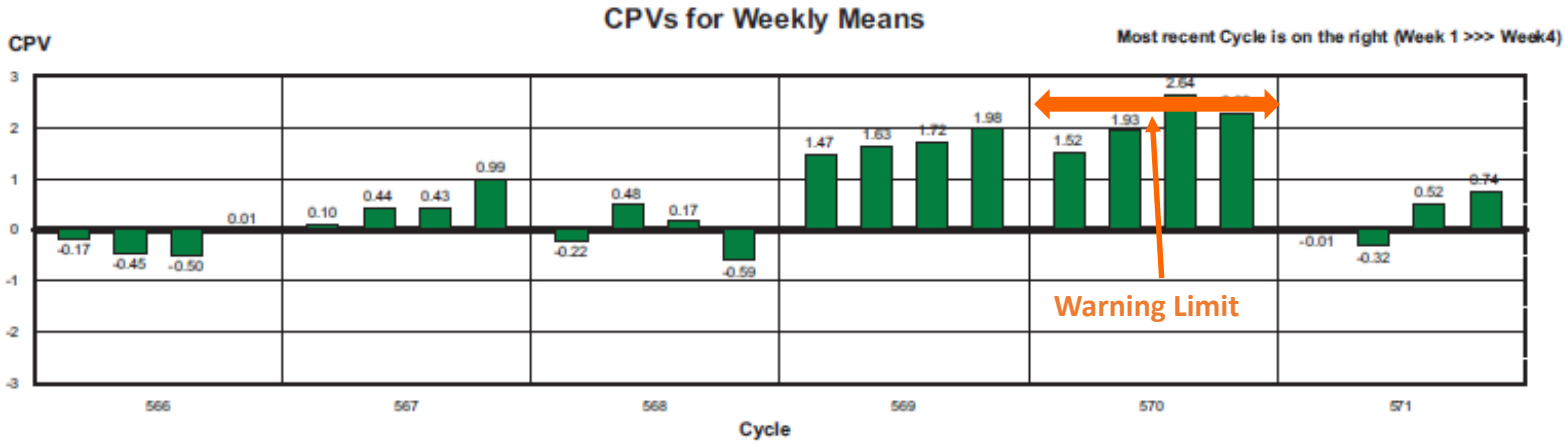
Systematic Consistency

The results below show data with a different type of consistency. The results from this laboratory are consistently lower than the Grand Mean. Depending on the laboratory's results, action may be taken to bring the test in closer agreement with the consensus.



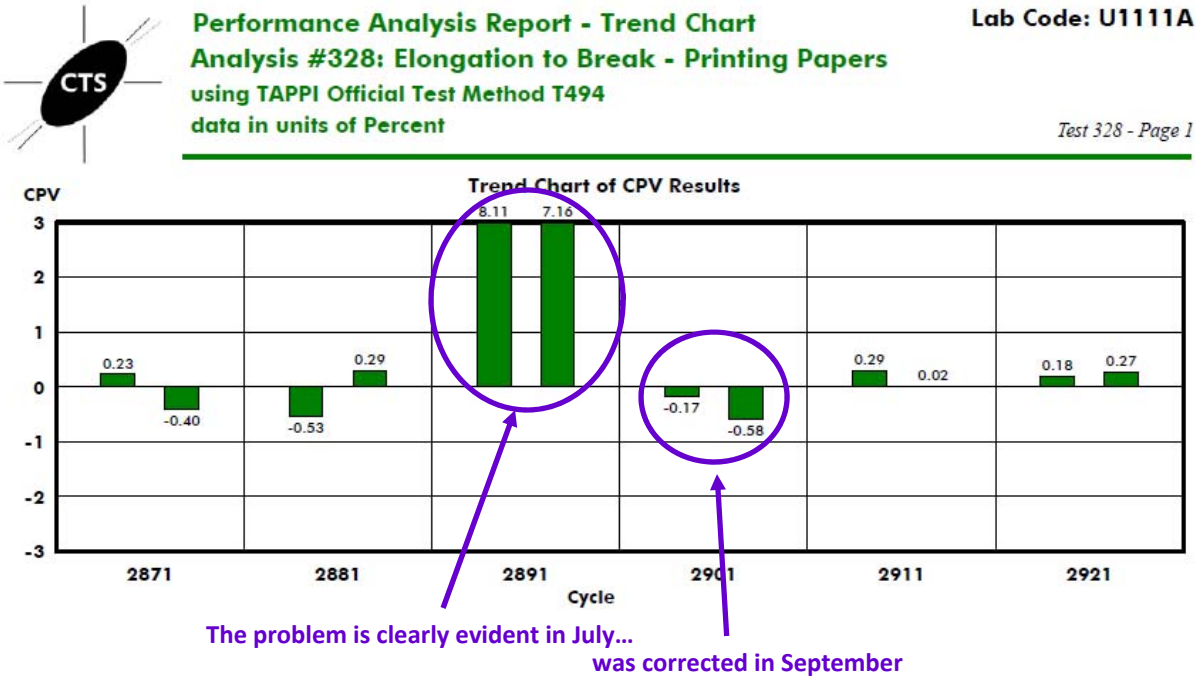
Trending Towards Extremes

Consistency is good. A bar graph that is growing consistently longer is not. This lab was alerted to an impending problem by the Trend Chart showing results that were trending increasingly higher than the consensus values. This resulted in a warning "*" flag for Cycle 570. Investigation of testing procedures and equipment were conducted at this point and the Trend Chart reflects changes effected by the lab to yield results that now agree quite well with the Grand Means.



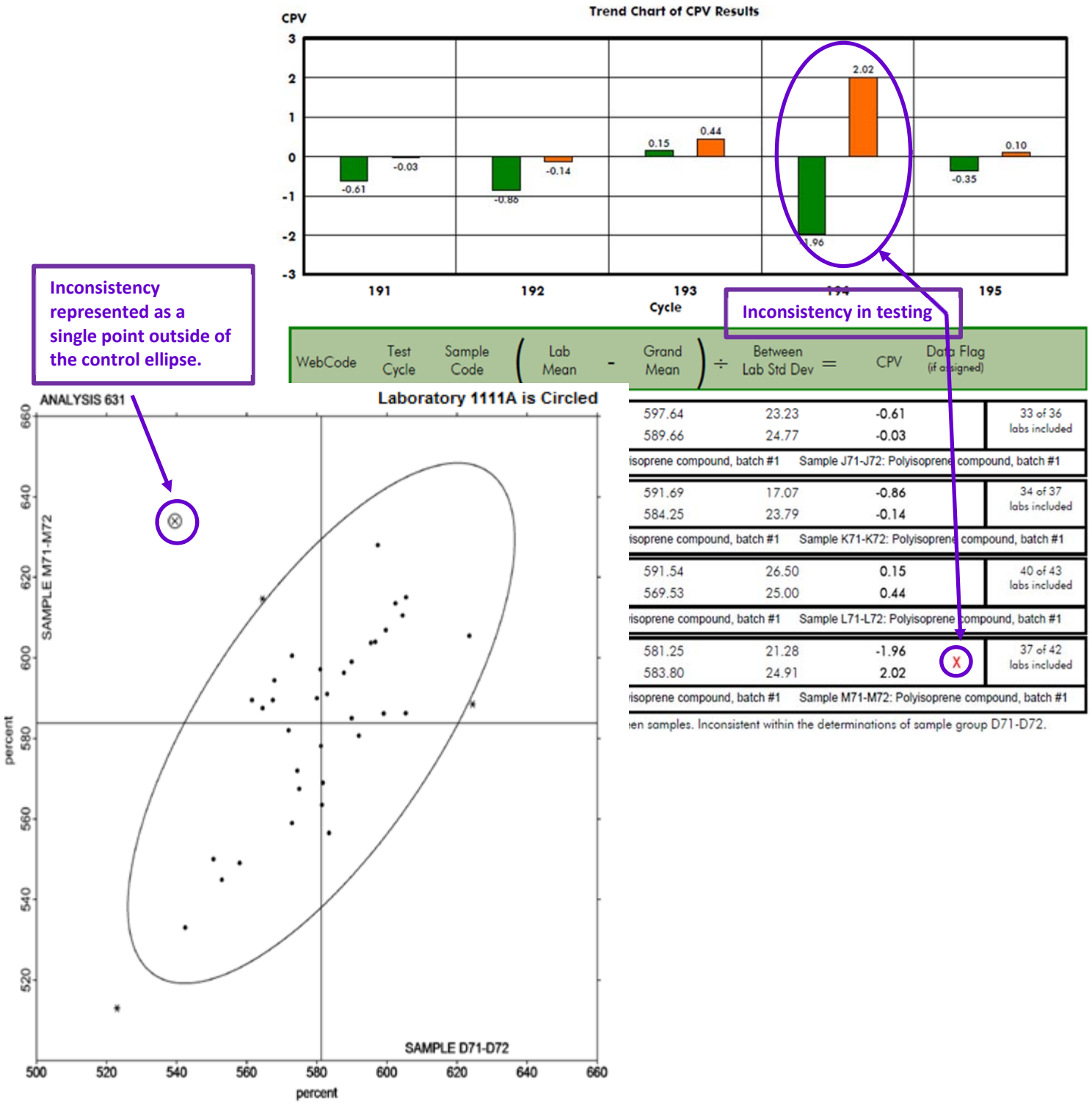
Sudden Blip

The trend chart can also reveal a one-time deviation from the usual performance. These deviations happen, even to the best of labs, and cannot be predicted; they may or may not result in exclusion from group statistics. This requires a review of the testing process; was there a change in the instruments? Improper calibration? A departure from procedure? New technician?



Inconsistency in testing

Inconsistencies that do not involve extreme data may be the most difficult error for labs to understand and to identify a cause. A lab's first instinct often is to conclude that each lab mean does not exceed a reasonable limit, so there is no problem. But because the samples provided are similar to each other, there is an expectation that there should be a correlation between the measurement results for the two samples. This correlation is clearly shown by the control ellipse. The test results for all labs are compared against each other, thereby determining an "acceptable" level of inconsistency, which is illustrated in part by the width of the control ellipse. A lab flagged for *Inconsistency in testing* has exceeded what the results from all labs have determined is a reasonable correlation between the means for the samples.

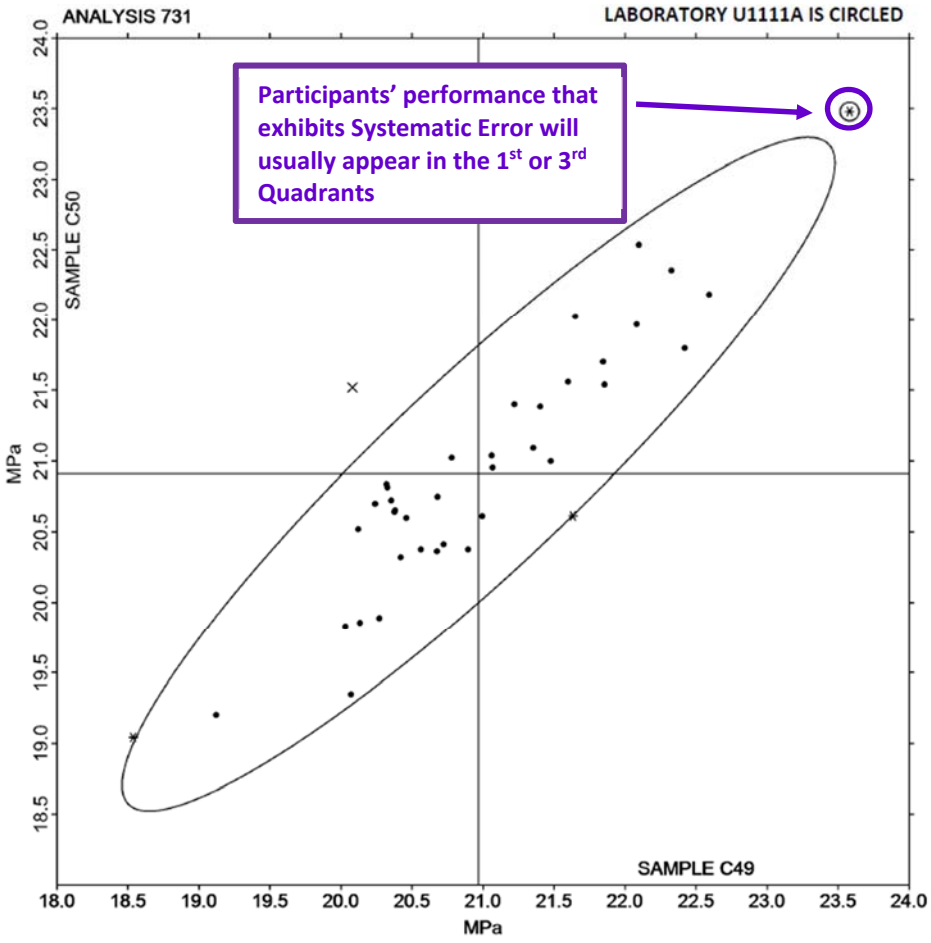


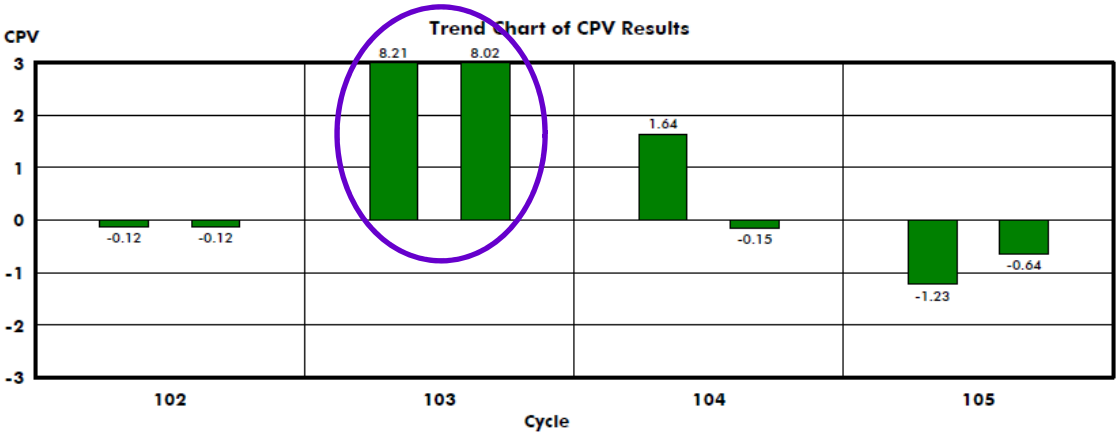
Systematic variations

Bias is an unavoidable fact of life in laboratory testing. The best illustration of bias is the control ellipse on the two sample plot.

Differences in procedures, conditions, instrumentation and sample preparation all contribute to the bias of a laboratory. When these differences become too large a laboratory may receive a Data Flag for a *Systematic Error*. When the test results for both samples are either high or low compared to the group, a laboratory has a fixed set of factors to focus on to identify a cause.

Furthermore, since additional testing on similar samples should produce similarly high or low results, it is easier to determine that a systematic error has been successfully corrected.





A maximum of 4 test cycles are printed on this page. If there are fewer sets of bars than the maximum and the laboratory was enrolled for all cycles, then one of the following situations occured: lab did not submit data for a particular test/testing cycle or the data sheets were received late.

The lab shown here indicates high systematic bias for Cycle 103.

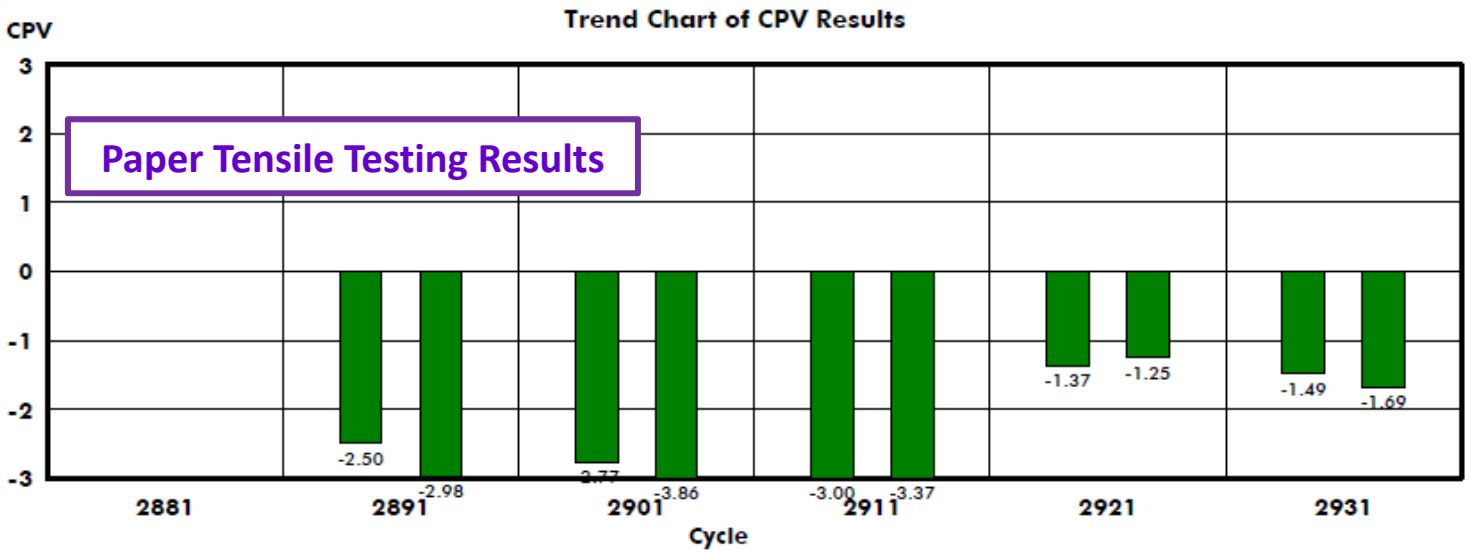
WebCode	Test Cycle	Sample Code	Lab Mean	Grand Mean	Btwn Lab Std Dev	CPV	Data Flag (if assigned)	
92NF4N	102	X43	14.82	14.93	0.98	-0.12		75 of 83 labs included
		X44	14.96	15.07	0.93	-0.12		
Testing Date: 2nd Qtr 2017		Sample X43: PP			Sample X44: PP			
9DQTNP	103	X45	6.97	5.42	0.19	8.21	X	81 of 91 labs included
		X46	8.39	6.40	0.25	8.02		
Testing Date: 3rd Qtr 2017		Sample X45: LDPE			Sample X46: LDPE			
Analyst Comment: Data for both samples are high.								
E2U9U6	104	X47	16.32	14.88	0.87	1.64		78 of 89 labs included
		X48	11.26	11.34	0.55	-0.15		
Testing Date: 4th Qtr 2017		Sample X47: PP			Sample X48: PP			

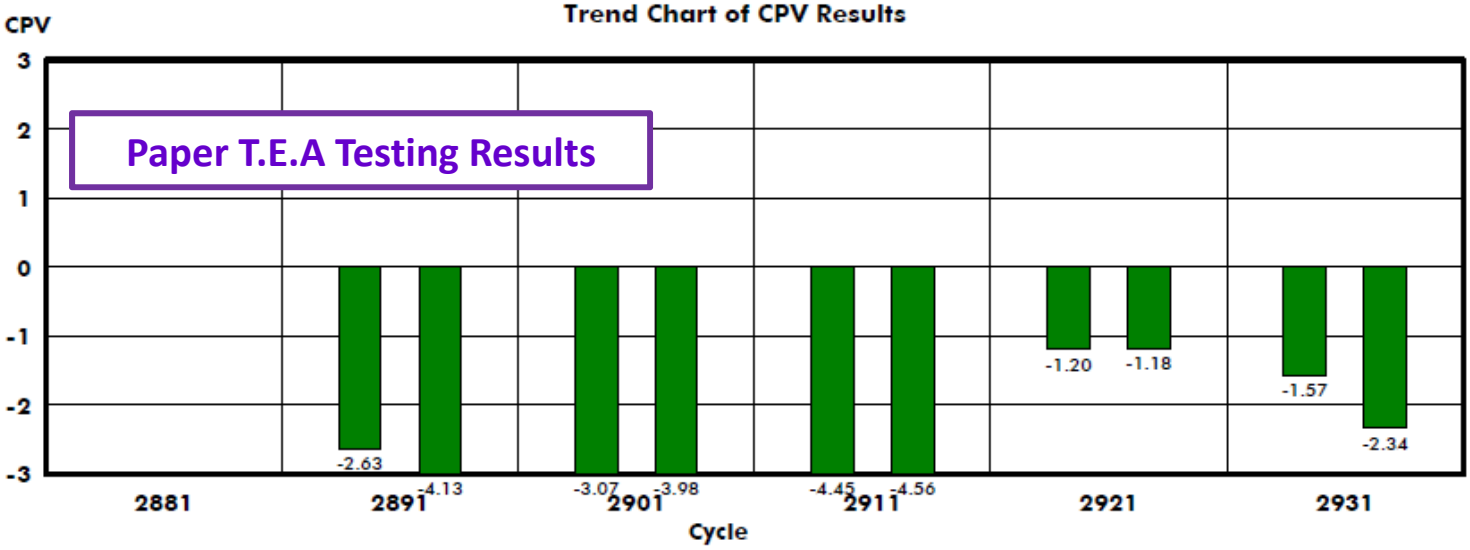
Program Specific Report Variations

Trending in Linked Properties

Although each analysis is presented on its own two pages, it is important to remember that tests are not completely independent. All of the testing in a laboratory is linked by factors such as training, maintenance, and conditioning; analyses that use the same instrument or examine linked properties are even more closely related. Tensile properties and color spaces (L*, a*, b*) are examples of the close linkage between some analyses. The trend towards lower results seen in this Paper Tensile Strength test is further confirmed by the Tensile Energy Absorption (T.E.A.) results.

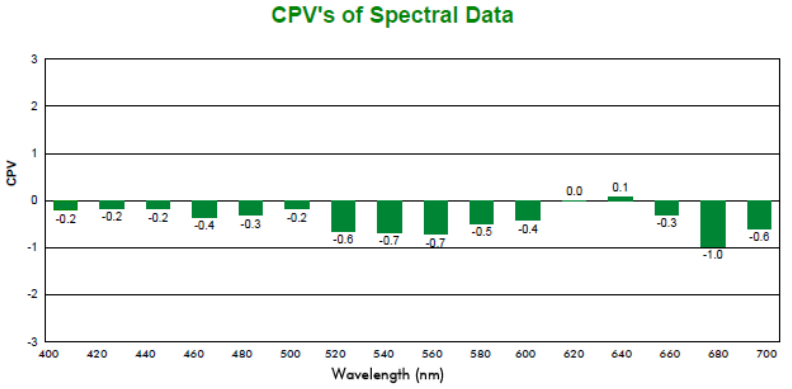
Participants can compare these results to gain additional insight into which factors may be affecting their performance.





Spectrophotometric Test (Color Program Analysis 411)

The Spectrophotometric Performance Analysis Report contains two distinct graphical presentations of the analysis conducted on a laboratory's measurements. The CPV Chart and the Spectral Reflectance Plot enable you to fully evaluate your performance and allow greater insight into the Spectrophotometric Analysis.



Overall Analysis for this Cycle

WebCode

Sample D171

Instrument X-Rite Color i7

Data Flag Assigned: None. Your spectral reflectance data were consistent with the other laboratories.

Your Data for this Cycle

Spectrophotometric Reflectance values (as %) and CPV's at selected wavelengths

	400	420	440	460	480	500	520	540	560	580	600	620	640	660	680	700
Lab Mean	15.92	19.20	23.16	25.95	27.00	30.16	35.71	38.61	31.95	24.70	21.04	19.34	19.22	20.32	19.85	17.98
CPV	-0.20	-0.17	-0.17	-0.36	-0.31	-0.17	-0.65	-0.67	-0.72	-0.48	-0.41	0.00	0.08	-0.30	-0.97	-0.59

Summary Statistics

	400	420	440	460	480	500	520	540	560	580	600	620	640	660	680	700
Grand Mean	16.04	19.24	23.20	26.05	27.07	30.21	35.92	38.90	32.23	24.85	21.15	19.33	19.20	20.43	20.13	18.16
STD	0.59	0.25	0.27	0.28	0.24	0.28	0.32	0.43	0.40	0.29	0.27	0.29	0.23	0.37	0.29	0.31

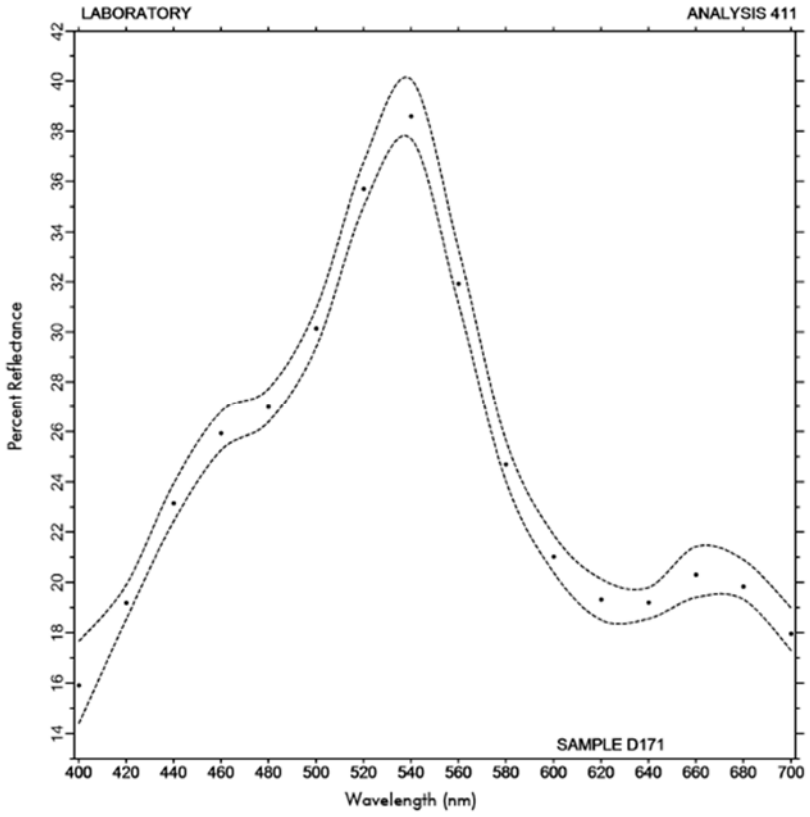
Consensus statistics based on 105 of 107 reporting participants

The first page of the Performance Analysis Report shows your **CPV's of Spectral Data Chart**. Your lab's CPV at each wavelength is shown as a bar. Immediately below the chart an **Overall Analysis for This Cycle** containing your data and consensus data for the cycle is shown.

Action Items – CPVs that exceed the critical limit calculated by CTS will be indicated by an X Data Flag. If seven or more reflectance values are flagged, an overall X Data Flag is assigned and an Action Item will appear.

The second page of the Performance Analysis Report is wholly devoted to the I**ndividual Spectral Reflectance Plot**. These plots show your data as points at the 16 specified wavelengths and the consensus data as curves*. The distance between the two curves represent the range of acceptable Lab Means at each wavelength. Data falling outside that range is assigned an X Data Flag and is shown as an X on the plot.

* The upper and lower limits are interpolated between the 16 wavelengths and do not represent an actual spectrophotometric curve.

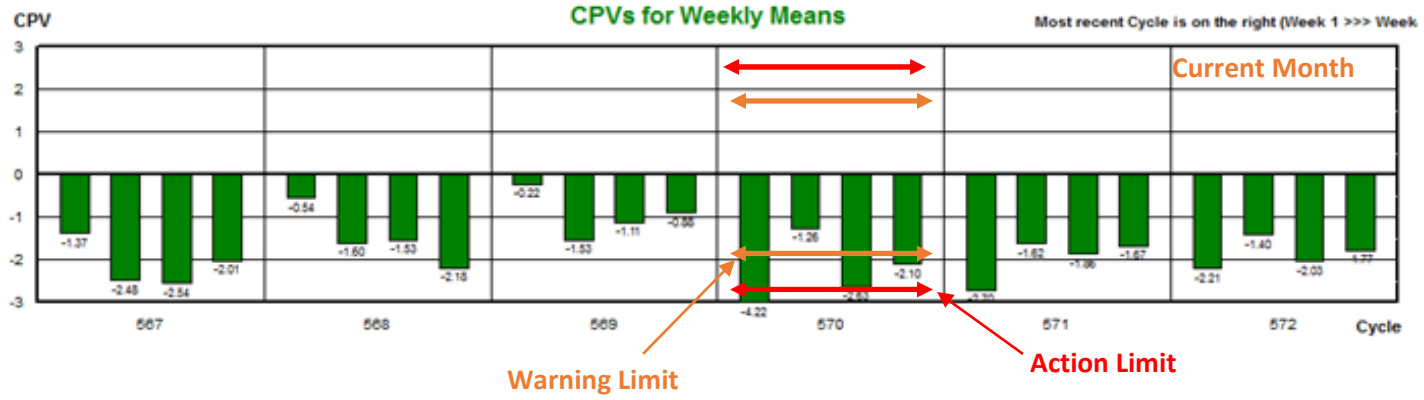


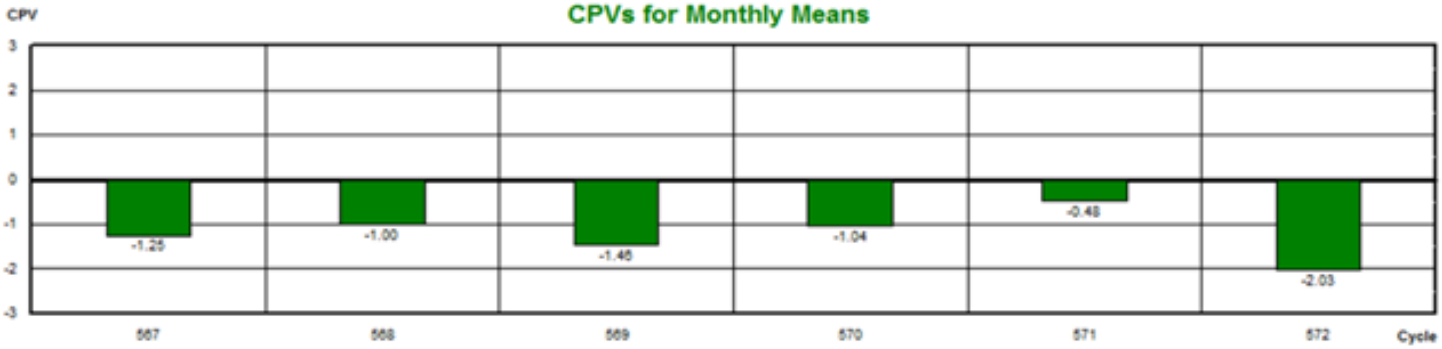
Containerboard Reports

The Containerboard Program is unique among the CTS programs as it was designed as a collaborative reference program. The two sample analysis is replaced with samples from one lot and the frequency of testing is increased to monthly or even weekly intervals. This provides participants a means of continuous measurement assessment, allowing them to monitor and if necessary take quick corrective action.

The first page of the Containerboard Performance Analysis Report contains three trend charts. The first of which displays the **Comparative Performance Value** (CPV) which is calculated and assigned the same way for all Industry Programs, but now represents either weekly or monthly testing.

Each Cycle shown on the Trend Chart shows the CPVs for each frequency of testing. Weekly CPVs have four bars with week 1 on the left through week 4 on the right.





Additional Containerboard Comparative Statistics

In addition to the CPV data, two additional comparative statistics are provided that allow you to determine, at a glance, how your results compare to those of the other participants for each interval of testing (weekly, monthly, cumulative)

The **Consistency** (k) statistic is the ratio of your within-lab standard deviation and the consensus within-lab standard deviation. This shows how high or low the variation within your lab is compared to that of other labs. A value of one indicates an average within-lab variation. For Monthly results, the k statistic is based on the 5 or 10 test determinations reported for that month of testing. Therefore, the meaning and interpretation of the k statistic and any corresponding flag is the same as it is for a single week in the Weekly tests.

Warning Limits for the k statistics are set to a 0.5% significance level and the exact value will vary based on both the number of included laboratories and the number of replicates in testing. If your within-week variation was high compared to other laboratories, an "H" data flag will be placed next to the corresponding k statistic and a warning statement will be posted in the Analyst's Comments section of the report. Low within-week variation will be assigned an "L" data flag; however, there will be no warning statement.

The **Constancy** (c) statistic is the ratio of your between-testing interval standard deviation and the consensus between-testing interval standard deviation. This shows how much or how little your testing interval means vary over time relative to that of other labs. A value of one indicates an average variation over time.

Warning Limits for the c statistic have been chosen such that they correspond to a significance level of no more than 0.5%. As with Consistency, if your test interval-to-test interval variation was high compared to other laboratories, an "H" data flag is assigned with an Analyst's Comment. Low test interval-to-test interval variation is assigned an "L" data flag with no warning statement.

These charts will contain as many as six consecutive months of historical data so that you may monitor how your results shift over time.

Containerboard Cumulative Statistics

The second page of the report contains the **cumulative statistics.** These are determined by the data of both the current cycle and that of previous cycles. The for weekly testing the cumulative data includes up to 16 weeks of data for tests that occur every month or up to 12 weeks of data for tests that occur every second month. For monthly testing the time span for the cumulative data is based on up to 4 months. It is important to note that the time span for the cumulative data may not be the same as the time span covered by the trend charts or the data table. The time span over which the cumulative statistics are determined is indicated in the "Wks Incld" or "Months" column of the data table.

Cumulative results will reset when the sample lot is changed. How often this occurs depends on the analysis type and the material used.

The SD Months statistic and c statistic will be displayed as zero whenever the cumulative results reset as there is no month-to-month variation to account for. Therefore, data flags will only be assigned to c statistics when three or four months are included in the cumulative results.