GE
Water & Process Technologies
Analytical Instruments

# Comparison of the Capability of TOC Analyzers for the System Suitability Test


imagination at work

## Introduction

A number of statistical techniques have been developed to assess the capability of a process or measurement system. While widely used in process control and manufacturing, these techniques are not widely used in assessing the capability of analytical instruments. In this paper, we present the results from a study of Total Organic Carbon (TOC) analyzers and their capability for the USP system suitability test.

## System Suitability Test

United States Pharmacopoeia (USP) published a method for measurement of total organic carbon (TOC) in water for injection and purified water in 1997 as a replacement for the oxidizable substance test. As part of this method, a test was described to determine the "suitability" of TOC analyzers to perform the measurements. The System Suitability Test comprises the measurement of three test solutions: a Reagent Water Control ($R_W$), a Standard Solution comprised of sucrose at a concentration of 500 µg C/L prepared in the reagent water ($R_S$), and a System Suitability Solution ($R_{SS}$) comprised of a solution of 1,4-benzoquinone also at 500 µg C/L prepared in the reagent water. It was viewed by USP that 1,4-benzoquinone represented a more difficult-to-oxidize compound than sucrose and therefore the test calculation is the response efficiency defined as:

$$\text{Response Efficiency} = 100\ \frac{(R_S - R_W)}{(R_{SS} - R_W)}$$

Upper and lower specification limits were established as not less the 85% and not more than 115% of the theoretical response. If the analyzer is suitable, then a response limit (maximum allowable TOC) for the water samples is established as $R_L = R_S - R_W$.

## Statistical Methods

Two widely used statistical methods for assessing process and measurement systems are Gage R&R Studies and Process Capability Analysis. The Gage R&R attempts to determine the different sources of variation in a measurement process. Specifically, it isolates the variation due to the equipment or gage (Repeatability), the variation due to the operator (Reproducibility), the variation due to the "parts," the variation due to the interaction between the operator and the parts, and the overall measurement variation (Gage R&R), which combines the repeatability, reproducibility, and the interaction terms and the total variation. The Gage R&R is a measure of the precision of the measurement system, and the accuracy of the measurement can also be estimated in this analysis if the actual value of the parts is known. The accuracy is normally expressed in terms of the bias (due to operator or parts) and linearity, which is a measure of the accuracy of the equipment across the range.

Capability Analysis is a measure of the variation in the measurements compared to specification limits. While several different indexes have been developed to measure process capability, the most widely used are the process capability ratio, $C_p$, and the process capability index, $C_{pk}$. The capability ratio is defined as (USL-LSL)/6σ, where USL is the upper specification limit, LSL the lower specification limit, and σ is the population standard deviation estimated from the sample standard deviation. Since the $C_p$ does not not account for processes that are not centered between the limits, the process capability index is often used and is defined as minimum(USL-x, x-LSL)/3σ, where x is the sample mean. For a capable process or measurement system, values of $C_p$ and $C_{pk}$ should be >1, and a value of 1.33 is often used as a minimum acceptable value.

## Experimental

The results obtained from five analyzers are included in this report. The analyzers were two Sievers 500 RL's (one equipped with a standard **iOS** (Integrated On-Line Sampling System),* the other with the Super **iOS**), two Sievers 900's, and a Thornton 5000TOC. Originally analyzers from two other manufacturers were included in this study (Anatel A643 and Shimadzu $TOC_v$), but the data from the A643 showed poor precision and accuracy and the $TOC_v$ had low recovery for benzoquinone. These results suggest that these analyzers were not operating correctly and they were excluded from the study.

The measurements were conducted in two phases. For the first phase, the standards laboratory at GE Analytical Instruments prepared large batches of reagent water, 500 ppb sucrose, and 500 ppb benzoquinone. These were transferred to plastic containers suitable for each analyzer (500-mL bottles for the Thornton, 30-mL bottles for the Sievers analyzers) and stored at 4 °C. Prior to measurement, the samples were transferred to a constant temperature water bath at 25 °C to warm the samples. Three operators participated in the study, performing one system suitability measurement on each analyzer every day for five days. Since

the Gage analysis requires some variability in the "parts," in the second phase, standard solutions of sucrose and benzoquinone were prepared each day at different concentrations ranging from 425 to 550 ppb. Stock solutions of sucrose and benzoquinone at 1,000 ppm were prepared and stored at 4 °C. These stock solutions were diluted to prepare a 5 ppm solution, which was then diluted to prepare the final ppb-level standards. The concentration of the sucrose and benzoquinone standards were selected to yield response efficiencies of 85%, 91%, 100%, 106%, and 115%, thus covering the full specification range of the test. Again, the three operators performed a measurement of the reagent water, sucrose, and benzoquinone standards on each analyzer every day for five days.

Measurements on the Thornton 5000TOC were performed using the System Suitability Kit and the results recorded manually on the forms from the instrument's manual, while measurements on the Sievers analyzers were performed using the **iOS** and the data exported to a USB flash drive each day.

Between the measurements by the different operators and overnight, the analyzers were connected on-line to a low-TOC, deionized water system for on-line measurements.

After the first operator's measurement of the ninth test on the 5000TOC, while operating in the on-line mode, the analyzer's fault light came on and the instrument was no longer recognized by the 770Max controller. Cycling the power restored communication between the analyzer and the controller, however, the 5000TOC reported TOC and conductivity values ten times higher than before. The technical support group at Thornton suggested the problem may be with the 770Max controller, but installation of a second controller produced identical results. In order to complete the study, the multiplier for the TOC measurements on the 5000TOC was manually reset to 1, which decreased the reported TOC value by a factor of 10. However, subsequent measurements of the standards gave lower than expected TOC values. Therefore, for the Gage R&R study for this analyzer, the last two sets of standards were excluded.

## Results

**Figure 1** shows variability plots of the blank-corrected TOC values for the sucrose and benzoquinone standards from this test for each analyzer, including the range and mean value along with a plot of the standard deviation of the three replicate measurements. These plots show the complete data set and, as noted above, measurements of the last two sets of standards (RS9, RS10, RSS9, and RSS10) were excluded for the Gage R&R calculation for the 5000TOC.

### Gage R&R Results

The results from the Gage R&R study on the five analyzers are summarized in T**ables 1 and 2. Table 1** shows the individual sources of variation for each analyzer,

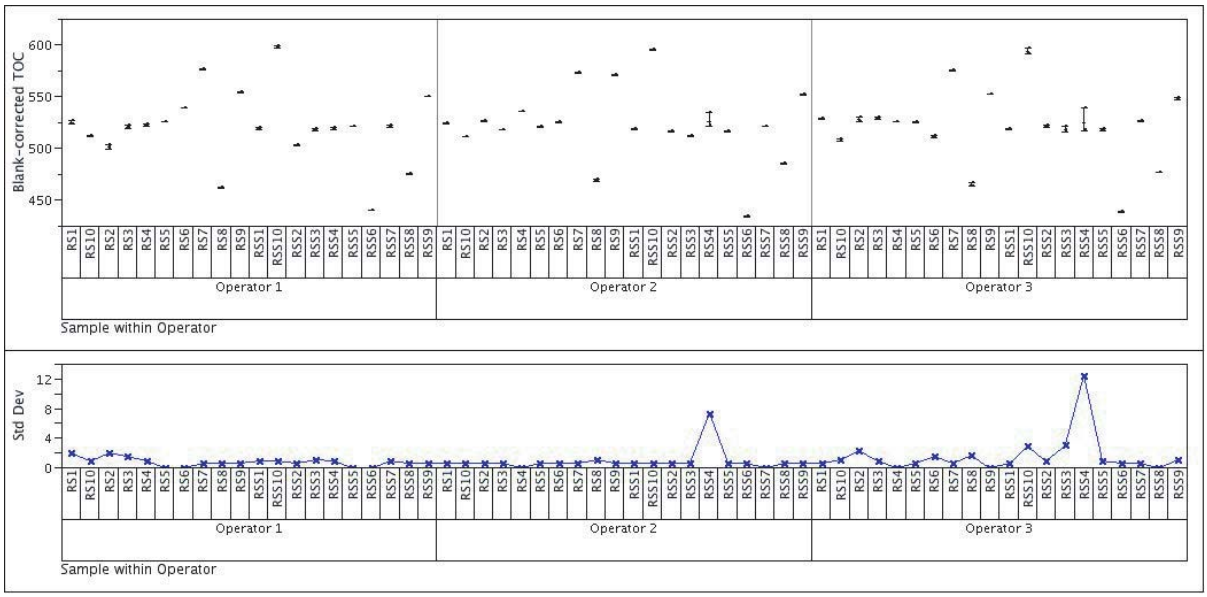


**Figure 1a.** Variability and standard deviation plots from Sievers 500 RL, serial number 21.

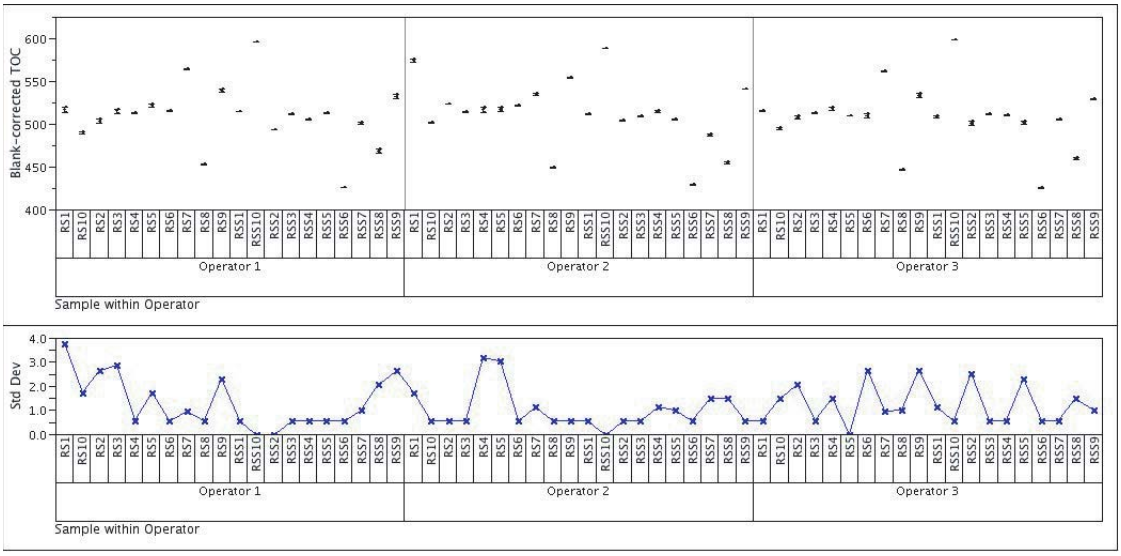


**Figure 1b.** Variability and standard deviation plots from Sievers 500 RL, serial number 24.

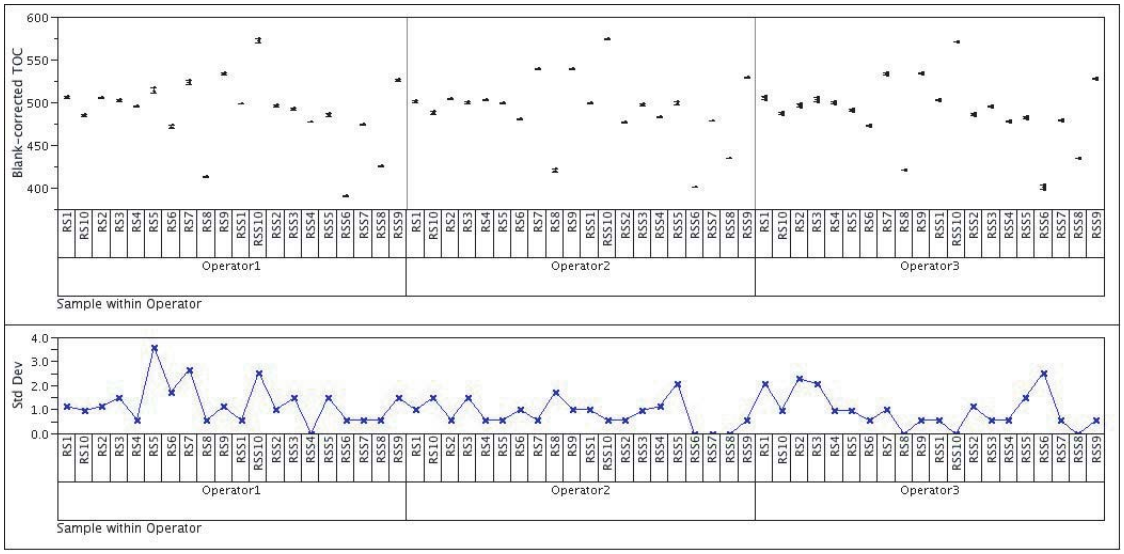


**Figure 1c.** Variability and standard deviation plots from Sievers 900, serial number 405

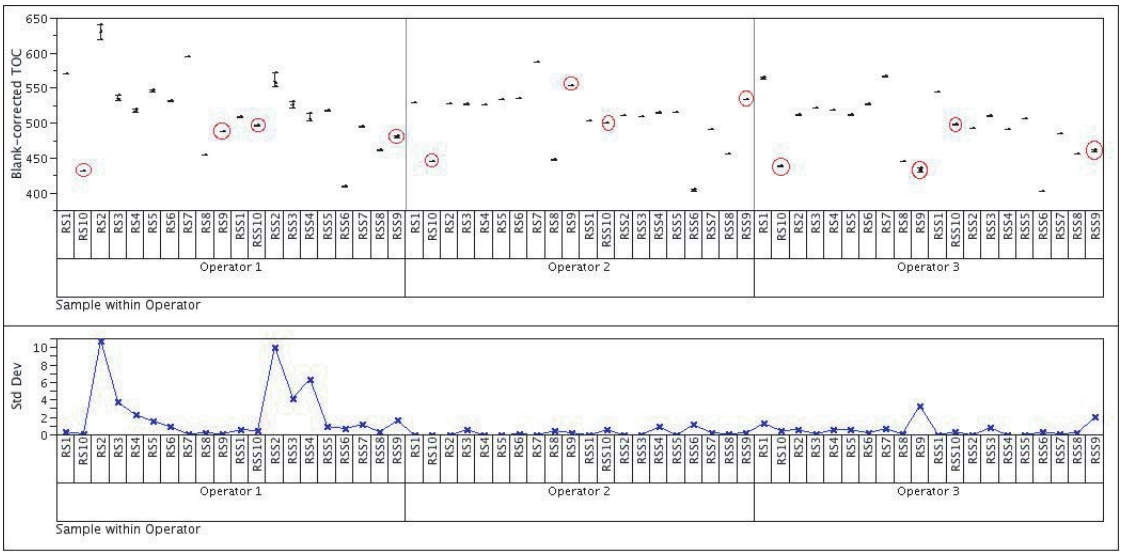


**Figure 1d.** Variability and standard deviation plots from Thornton 5000TOC. Circled data were excluded from study due to instrument malfunction.

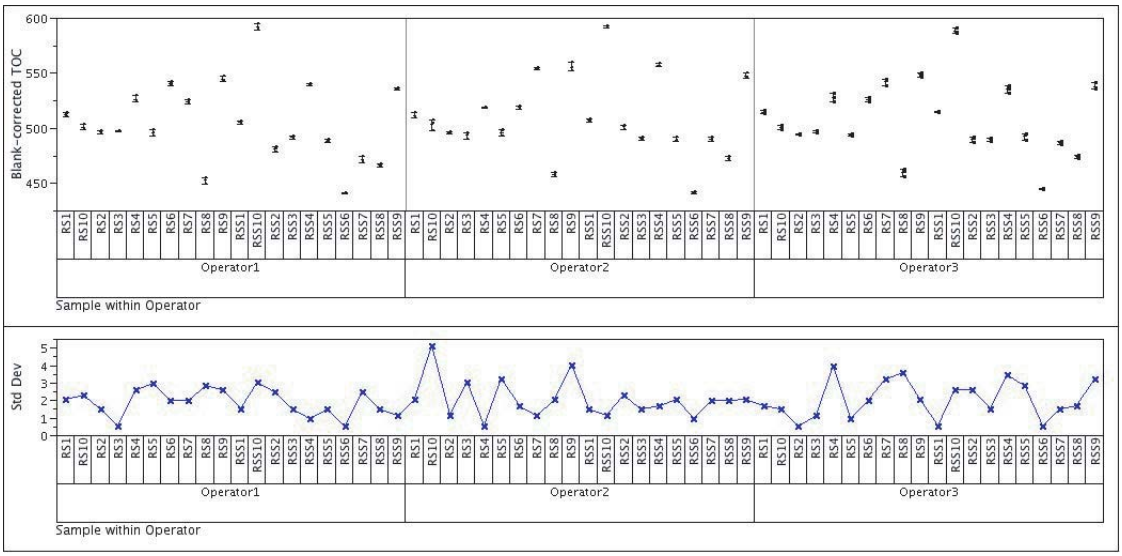


**Figure 1e.** Variability and standard deviation plots from Sievers 900, serial number 531.

**Table 1**
Sources of Variation from Gage R&R Study on Five TOC Analyzers

| Analyzer | Repeatability | Reproducibility | Operator Sample | Gage R&R | Part Variation | Total Variation |
|---|---|---|---|---|---|---|
| 500-24 | 7.16 | 0 | 32.6 | 33.4 | 177.8 | 180.9 |
| 500-23 | 7.8 | 0 | 35.2 | 36.1 | 183.0 | 186.5 |
| 900-405 | 6.6 | 4.7 | 26.9 | 28.1 | 205.8 | 207.7 |
| 900-531 | 11.6 | 9.6 | 31.5 | 34.9 | 178.2 | 181.6 |
| 5000TOC | 13.0 | 47.1 | 99.7 | 111.1 | 218.4 | 245.1 |

**Table 2**
% Gage R&R for the Five TOC Analyzers

| Analyzer | % Gage R&R | Precision to Part Variation | Barrentine's Characterization |
|---|---|---|---|
| 500-24 | 18.4 | 0.19 | adequate |
| 500-23 | 19.3 | 0.20 | adequate |
| 900-405 | 13.5 | 0.14 | adequate |
| 900-531 | 19.2 | 0.20 | adequate |
| 5000TOC | 45.3 | 0.51 | unacceptable |

while **Table 2** shows the % Gage R&R along with Barrentine's characterization of the variation.

As shown in **Table 1**, the Sievers 500's and 900's showed only small variations due to the operator (reproductivity) and the equipment (repeatability), while much greater variability from the operator and equipment was observed with the 5000TOC. The reason for the greater variability is due, in part, to the different procedures for measurement of the samples. In the Sievers instruments, the operator simply inserts the samples into the vial port of the **iOS** and presses a button. On the 5000TOC, the operator connects the sam-

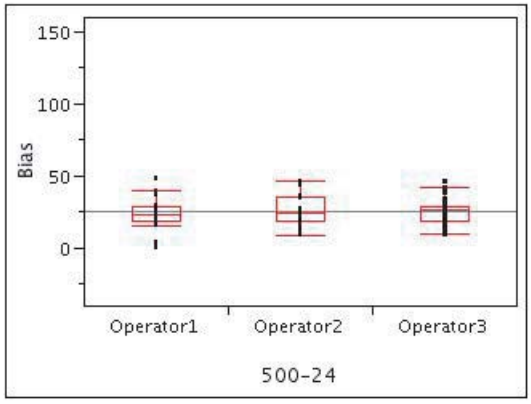


**Figure 2a.** Box plots of the bias report for Operator from Sievers 500 RL, serial number 24.

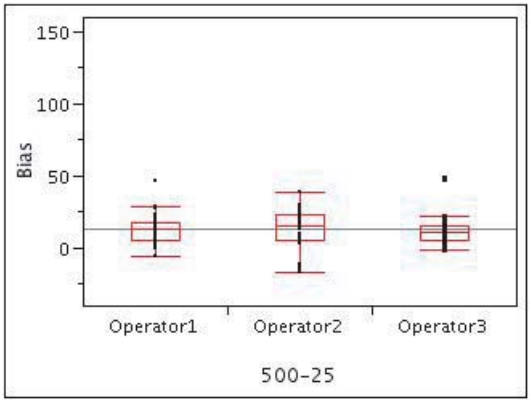


**Figure 2b.** Box plots of the bias report for operator from Sievers 500 RL, serial number 25.

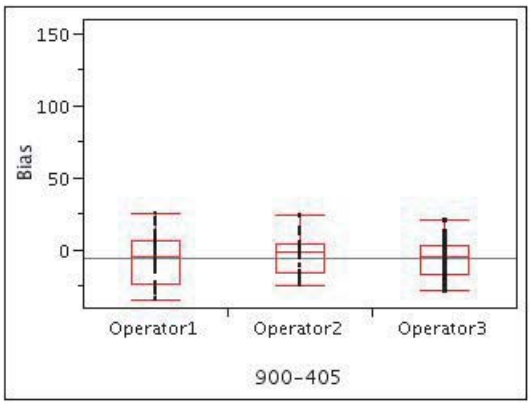


**Figure 2c.** Box plots of the bias report for operator from Sievers 900, serial number 405.

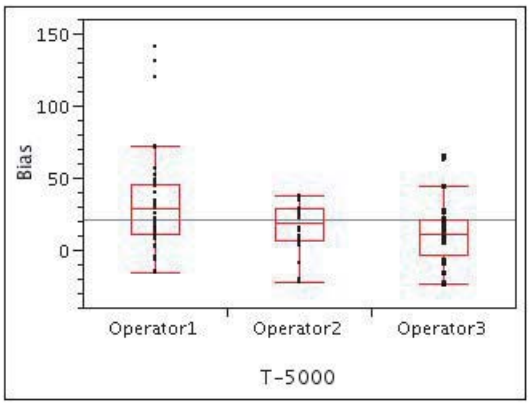


**Figure 2d.** Box plots of the bias report for operator from Thornton 5000TOC.

ple bottle to the external sample pump of the System Suitability Kit, configures the 5000TOC to measure sample flow rate, adjusts the pressure regulator to set the sample flow rate to 20 mL/min, reconfigures the 5000TOC to report TOC, and then after waiting for the TOC values to stabilize (5 minutes is recommended), records three values, ten seconds apart. The sample pump is then turned off, the next sample bottle installed on the device, and the process repeated.

As shown in **Table 2**, both models of the Sievers TOC analyzers gave acceptable results from the Gage R&R measurements of 30 reagent water, sucrose and benzoquinone standards. The results from 24 measurements of these samples on the 5000TOC showed this analyzer to be unacceptable for the system suitability test.

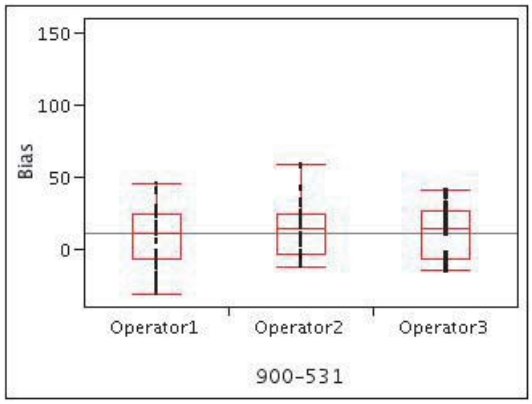


**Figure 2e.** Box plots of the bias report for operator from Sievers 900, serial number 531.

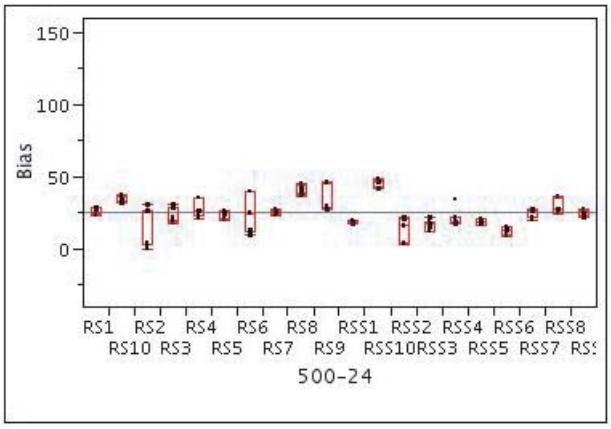


**Figure 3a.** Box plots of the bias report for Sample from Sievers 500 RL, serial number 24.

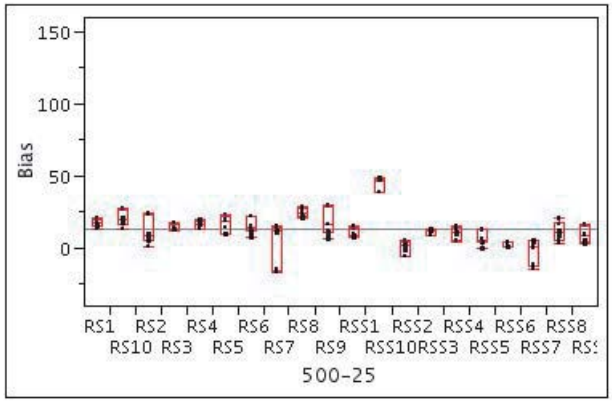


**Figure 3b.** Box plots of the bias report for Sample from Sievers 500 RL, serial number 25.

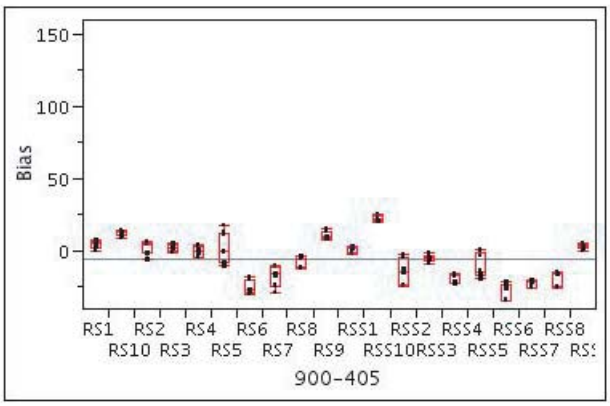


**Figure 3c.** Box plots of the bias report for Sample from Sievers 900, serial number 405.

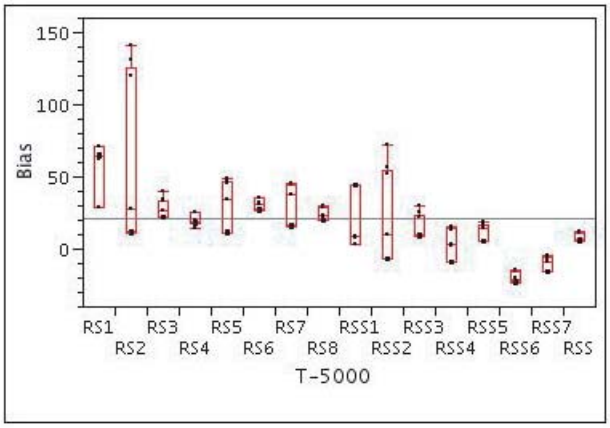


**Figure 3d.** Box plots of the bias report for Sample from Thornton 5000TOC.

## Bias and Linearity

Box plots of the bias report for Operator is shown in **Figure 2**. For the 500's and 900's, there was no significant difference in the means of the bias by operator (Tukey-Kramer Honestly Significantly Different) nor in the variances of the bias (Bartlett's Test) between operators. In contrast, for the 5000TOC, there was a significant difference between the mean bias and the variance of the bias between operators, with Operator 1 showing a larger positive bias and greater variability than the other operators.

Box plots of the bias report for Sample is shown in **Figure 3**. For the 500's and 900's, sample RSS10 showed the largest positive bias (along with RSS4 for 900-531 and RS8 for 500-24). The mean bias for this sample was significantly different for most of the analyzers and may reflect changes in the benzoquinone stock solution or an error in the preparation of this solution.

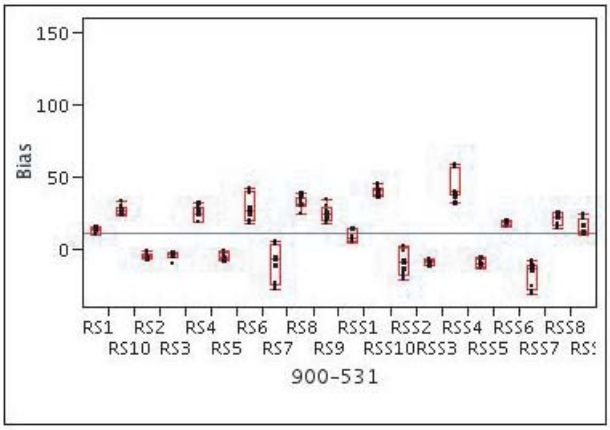


**Figure 3e.** Box plots of the bias report for Sample from Sievers 900, serial number 531.

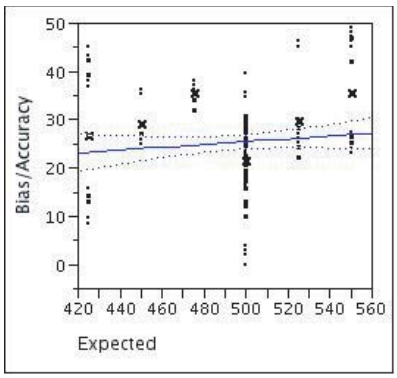


**Figure 4a.** Linearity plots from Sievers 500 RL, serial number 24.

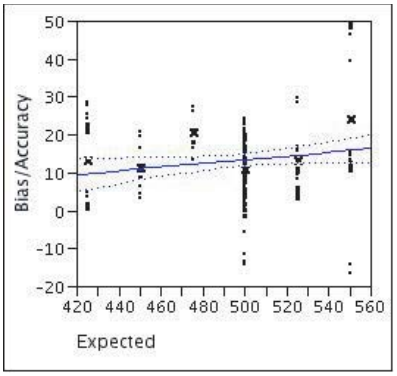


**Figure 4b.** Linearity plots from Sievers 500 RL, serial number 25.

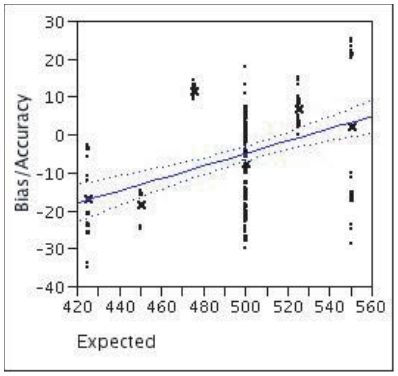


**Figure 4c.** Linearity plots from Sievers 900, serial number 405.

Linearity plots for the analyzers is shown in **Figure 4** using a process variation of 0.25. There was no significant difference in the bias versus expected concentration for the two Sievers 500 RL analyzers (slope = 0), while the two Sievers 900's showed a bias versus concentration. For 900–405, this was due to a small negative bias in the measurement of the lowest concentration combined with the before-mentioned positive bias at the highest concentration (RSS10). For 900-531, a small negative slope was observed. The 5000TOC also displayed a significant bias versus concentration, even though only four of the measurements were at a concentration other than 500 ppb.

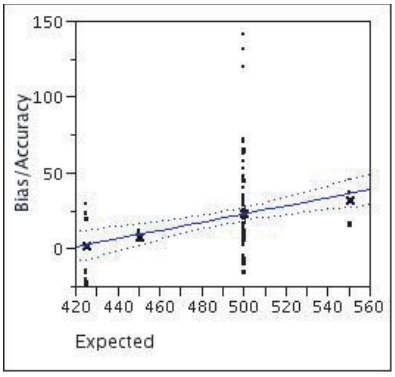


**Figure 4d.** Linearity plots from Thornton 5000TOC.

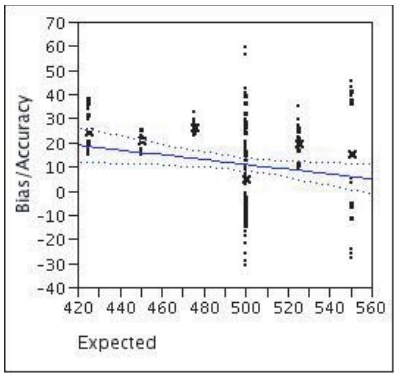


**Figure 4e.** Linearity plots from Sievers 900, serial number 531.

**Table 3**
Summary Statistics from 15 Measurements of Response Efficiency

| Analyzer | Mean | Standard Deviation | Range |
|---|---|---|---|
| 500-24 | 98.9% | 0.7% | 87.9 – 100.3% |
| 500-25 | 98.6% | 0.8% | 96.6 – 99.7% |
| 900-405 | 97.6% | 1.8% | 94.3 – 99.9% |
| 900-531 | 99.1% | 1.1% | 97.0 – 101.4% |
| 5000TOC | 95.7% | 3.0% | 88.9 – 98.8% |

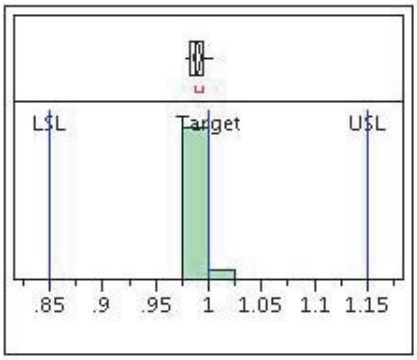


**Figure 5a.** Calculated response efficiencies, and specification limits for Sievers 500 RL, serial number 24.

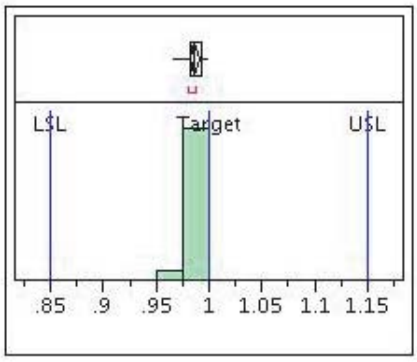


**Figure 5b.** Calculated response efficiencies, and specification limits for Sievers 500 RL, serial number 25.

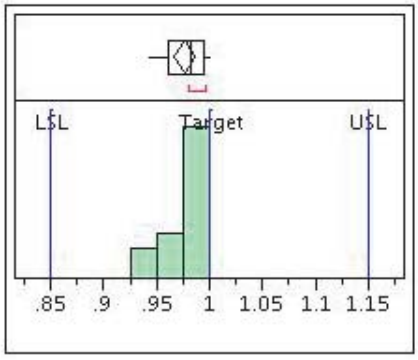


**Figure 5c.** Calculated response efficiencies, and specification limits for Sievers 900, serial number 405.

## Capability Analysis

The first five sets of samples from the Gage R&R test were all at a concentration of 500 ppb and therefore can be used to calculate the response efficiency from the system suitability test. The distribution of these response efficiencies is used, along with the specification limits, to calculate capability ratio and capability index. **Figure 5** shows a histogram and box plot for the 15 calculated response efficiencies for the five analyzers, along with the specification limits. Summary statistics for this data set are shown in **Table 4**. For the Sievers analyzers, the distribution of response efficiencies was tightly grouped around the tar-

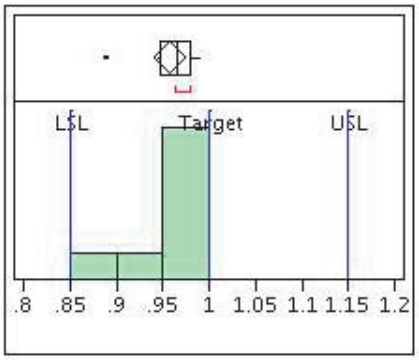


**Figure 5d.** Calculated response efficiencies, and specification limits for Thornton 5000TOC.

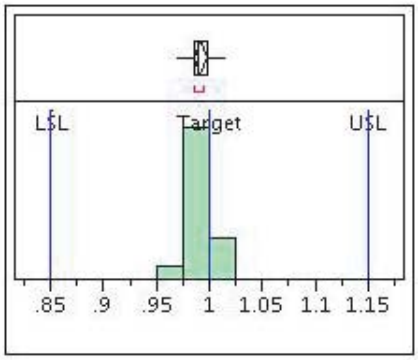


**Figure 5e.** Calculated response efficiencies, and specification limits for Sievers 900, serial number 531.

**Table 4**
Capability Analysis of Analyzers from Measurement Response Efficiency

| Analyzer | $C_p$ | $C_{pk}$ | Total Outside Limits |
|---|---|---|---|
| 500-24 | 7.42 | 6.88 | 0 |
| 500-25 | 6.11 | 5.54 | 0 |
| 900-405 | 2.73 | 2.30 | 0 |
| 900-531 | 4.35 | 4.10 | 0 |
| 5000TOC | 1.65 | 1.18 | 193 |

get value of 1 (equal response for sucrose and benzoquinone). In contrast, the distribution for the 5000TOC was skewed (skewness = -1.61) towards the lower specification limit. The mean response efficiency for the 5000TOC was significantly lower than the other analyzers.

**Table 4** lists the capability ratio and capability index for the analyzers. All the analyzers had $C_p$ and $C_{pk} > 1$ which indicates that the mean ± 3σ was inside the specification limits of the USP test. A $C_{pk}$ value of 1.33 is widely used as a minimum to indicate a capable system (mean ± 4σ within specification limits) since this allows for some shift in the mean while remaining within the limits. All the Sievers analyzers had $C_{pk} > 1.33$, while the 5000TOC was between 1 and 1.33, a range generally considered as marginal. The capability analysis can also be used to calculate how many measurements can be expected to be outside of the specification limits, usually expressed in parts per million. The Sievers analyzers would not be expected to fail any system suitability tests (0 ppm), while the 5000TOC would be expected to fail this test (197 ppm).

## Conclusions

The results from a Gage R&R study of the System Suitability test showed that two Sievers TOC analyzer models gave acceptable results, while the Thornton 5000TOC gave unacceptable results. There was no operator-dependent bias observed for the Sievers analyzers, while the 5000TOC showed significant operator influence. A capability analysis of response efficiency shows the Sievers-brand analyzers to be capable, while the 5000TOC is only marginal.



For more information, visit www.geinstruments.com. Find a sales partner near you through the "Contact Us" Section.

**USA**
GE Analytical Instruments
6060 Spine Road
Boulder, CO 80301-3687 USA
T +1 800 255 6964
T +1 303 444 2009
F +1 303 444 9543
geai@ge.com
www.geinstruments.com

**Europe**
Unit 3 Mercury Way
Urmston, Manchester, M41 7LY
United Kingdom
T +44 (0) 161 866 9337
F +44 (0) 161 866 9630
generaluk.instruments@ge.com



300 00120 Rev A
MC06-020