Supplemental to ILC/PT Reports

A guide for interpretation of reports issued by NAPT

The purpose of this document is to assist our participants, accreditation bodies, and assessors, as they interpret and analyze reports issued by the National Association for Proficiency Testing. The more focus and understanding we can provide, the more improvement opportunities will become evident and value we will be able to provide all who participate in ILC/PTs. It is a privilege to partner with you in this effort. Please direct any questions regarding the content of our reports to the NAPT technical manager.

At a minimum, NAPT issues two reports to all participants during their enrollment in an ILC/PT; a preliminary and a final report. Each report has a separate purpose. Explanations for each report are described in detail below. Each ILC/PT is designed to provide participants with real value. Careful consideration is placed on artifact integrity along with current measurement practices and available technology. A test scheme is developed in accordance with the established NAPT quality system, which meets all international quality standards and the requirements of accreditation bodies. Once participants enroll, artifacts are distributed, monitoring of ILC occurs, and reports are issued.

Quality System

NAPT's Quality System not only meets, but exceeds international requirements for ILC providers. For example, our analysis provides much more depth than required by current guidelines. Accreditation bodies also have requirements that are based on their interpretation of the international standards, plus their own additional set of requirements. We actively work with all accreditation bodies to assure your participation with NAPT will fulfill those requirements. Of equal importance are the specific needs of our members. Our full time staff understands not only the accreditation requirements imposed on you, but also how they can best be met.

All ILC schemes are designed per our rigorous NAPT Quality Procedure 303-1 *Development of ILC/PT Schemes*. This procedure requires: 1) a quality artifact will be used 2) the range of measurement uncertainty that will be expected 3) a determination of the type of participants who would benefit 4) the selection of the pivot lab who will characterize the artifact 5) the establishment of the reference values.

Technical Analysis

During the course of each ILC/PT, comprehensive technical reviews are conducted before, during, and after an artifact is put into distribution. This is done to assure test integrity, establish and/or validate reference data, and check for trends and/or anomalies in the data. By performing these reviews on a scheduled basis, we are better able to assist you in the event prompt corrective action is necessary. The last thing we want is for you to miss an opportunity for improvement or, worse yet, spend valuable time searching for cause that may be unfounded.

NAPT's statistical review process is so robust that many of our advisors feel we have gone overboard in the analysis. It is their opinion that our technical review is unmatched by any ILC Provider. These requirements are described in NAPT Quality Procedure 304-1 *Data and Statistical Analysis Procedure*.

NAPT's process for establishing reference values was designed with the help of lead statisticians in the metrology community and technical oversight by statisticians at NIST. Using standard technical practices, an analysis is performed and an appropriate reference value is assigned to the data set. We ensure that reference values assigned are stable and worthy of being compared against. To further analyze the results we also perform the following statistical reviews: **Two Sigma, Three Sigma, Chauvnet Criterion, Sample Median, Trimmed Mean, Interquartile, Q-Test, and Thompson Technique**. This is done in all cases to assure sound and meaningful results are published. An example of this thorough analysis is attached to this paper as an appendix.

Only after a careful review of the data does NAPT assign an established reference value. To prematurely assign a reference value could be inaccurate and may result in a value that would not pass a robust analysis. Doing so would not ensure confidence in the reference value assigned. Making an assumption that a single measurement is the correct measurement is not a technically sound process for ensuring the validity of the data (no single laboratory is infallible). That is why it is only after a thorough review of all data that a reference value should be assigned.

Another very important differentiator is the support we receive from our established base of Technical Advisors. All ILC/PTs have a technical advisor assigned to oversee the process and critique the results. These representatives from the metrology community specialize in the measurement being performed. These individuals are actively involved in developing the ILCs, assisting in the analysis, and being available to provide technical counsel if needed. The data submitted by a participant is shared, but not the name associated with the data. At no time does the advisor know the relationship between name of participant and data submitted by the participant. NAPT ensures that confidentially of participant's results is kept throughout the ILC process, in accordance with NAPT's quality system.

Explanation of Preliminary Reports

The main purpose of the preliminary report is to provide a means to share with the participant the data that they have initially submitted to NAPT. Participants are given 10 days to inform NAPT of any errors reported in the data. If changes are made, NAPT verifies all re-submitted data and requires objective evidence before any modifications are made to the original data submitted. Objective evidence will include original observations, data collection sheets, uncertainty budgets, and procedures used. These are just a few of the details/documents NAPT would need to review before data submitted might be revised to reflect actual data obtained by participants in the original data collection process.

The preliminary report includes limited analysis of the results received up to that point in the ILC/PT. While in some cases this may be revealing, this will not necessarily be the same as the final report. This can be for several reasons, the most significant is the fact that not all data has been reported either due to the number of participants or the time it takes to make a complete cycle. Another reason may be the robust analysis that will be done for the final report to assure the integrity of the comparison reference

values. Accreditation bodies and *ISO17025* assessors understand the purpose of preliminary reports and do not expect you to act on the results. This doesn't necessarily diminish the value of these reports; opportunities may exist. For example, several laboratory technicians may have completed the same test and the preliminary results may show a disparity in the data. Corrective action and further review of root cause would certainty add value to the process.

Explanation of Final Reports

The final report presents the measurement results reported to NAPT along with associated uncertainties. Two comparisons are presented. First, the reported results for each measurement data point are compared against the established reference value, including the associated uncertainties. Secondly, the reported result for each measurement data point is compared to reported results from all other test participants along with a comparison of their reported uncertainties. To add clarity and simplification in understanding the results, several additional indicators are included for each reported data point. These include:

E_n

Normalized Error designated as E_n, is described in ISO Guide 43-1 *Proficiency Testing by Interlaboratory Comparison*. It is the ratio of the deviation between the reported value and the reference value to the root sum square of their associated uncertainties. Putting this into perspective, as the reported measurement uncertainty approaches the uncertainty reported by the reference laboratory; the closer the reported value must be to the reference value. Satisfactory results will yield an E_n <=1. To further add clarity, the E_n is charted for all reported data points showing the dispersion around the ±1 acceptable range.

Satisfactory/Unsatisfactory

The S/U indicator is provided as an immediate indicator of the results. Given the performance as indicated by the E_n, analysis, a satisfactory or unsatisfactory note is made. Satisfactory performance is described as an E_n <=1. Unsatisfactory performance is an E_n >1. Most accreditation bodies look for a satisfactory performance in order for the ILC to meet their proficiency testing requirements. Most accreditation bodies require successful participation before they will grant accreditation. It is our advice that you check with your accreditation body for acceptance requirements.

Z Score

This statistical evaluation also comes from ISO Guide 43-1. In reviewing the test results for all participants and all data points, there may be times when an outlier is identified. This analysis allows the exclusion of these data points and the potential impact it could have on the final results. Satisfactory performance is indicated as Z <=2, unsatisfactory performance is indicated as Z >=3 and questionable performance is anything between Z <= 2 and Z >= 3.

I-W-O

This concept was first implemented by NAPT in 1997. Since then many organizations have used this concept as a way to understand the reporting of measurements in interlaboratory comparisons against a reference value.

I – In

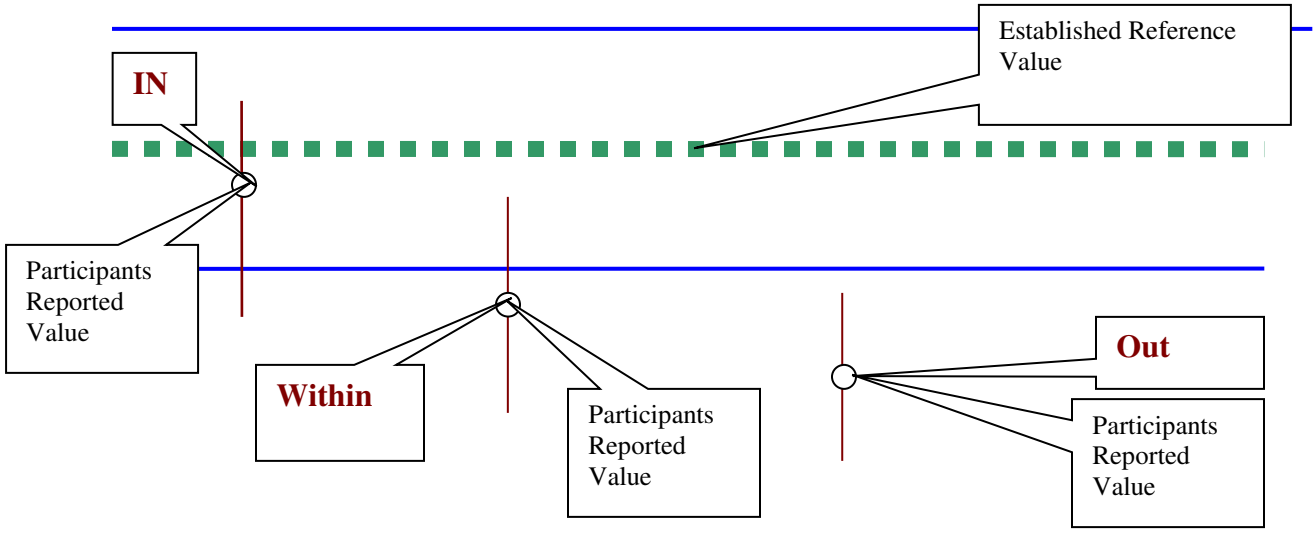
This status is achieved when the participants reported value falls within the uncertainty limits of the established reference value. The uncertainty of the participant has no bearing on this evaluation. See chart below.

W- Within

This status is achieved when the participants reported value falls outside the uncertainty limits of the established reference value, but the uncertainty of the participants overlaps the uncertainty of the established reference value. See chart below.

O – Out

This status is achieved when the participants reported value and reported uncertainty falls outside the uncertainty limits of the reference value. See chart below.



The final report indicates the performance and capability of multiple laboratories performing like calibrations of a single type of artifact. Results also give each participant a basis for evaluation and perhaps improvement of calibration performance. In many cases, this is the first time the laboratory is able to compare itself to others. While the 'tightness' of the reported values has always been the driving force, participants are now able to see how their measurement uncertainties compare to others.

OK!..... What's next?

One of the questions we get quite often-"First I pass……and now I get my final report, and I fail!!!! How can that be? Let's begin by going through the steps of developing an ILC and maybe this will provide some clarity.

First let's clear up one small misunderstanding, NAPT does not pass or fail any participant. Our job is to provide you with a technically sound ILC, one that is properly managed, properly analyzed and is accepted by your accreditation body. Of equal importance is one that complies with international requirements. Only through a comprehensive process are we able to provide this service to you. What we provide back to you is the information that allows you to make sound technical decisions.

Satisfactory or unsatisfactory rating, per the En analysis, should not be your only concern. A full review of the report issued to you is essential. Because you performed satisfactorily in all measurements does not mean you should be content with your data submittal. Maybe the En value was close to 1, maybe your uncertainty is not what others (making like measurements) are reporting. The reasons are numerous and each ILC participation should be reviewed to not only make sure your expectations were met but for continuous improvement in your measurements process. It is between you and your accreditation body to determine your performance rating in any ILC.

With that being said, let's discuss when we have a reference value from a national or accredited laboratory. At some point during the test we need to ensure that the reference values have not changed. This is particularly true for artifacts that may drift, a platinum resistance thermometer, for example. The artifact is recalibrated and the reference value has changed, maybe a very small change, but it has changed. The big question then becomes, when did the artifact change? We review the test data and identify some change or drift, but the approach is to predict the reference values based on a sound technical analysis. Because you have received your preliminary report, the final report may be different because it's compared to the new established reference value based on the revised reference value. We manage this by performing technical reviews, selecting quality artifacts and placing them on a recalibration schedule to minimize the impact of drift and monitoring the data submitted by participants.

Next let's discuss the case when we need to satisfy your immediate needs and provide an ILC without having the benefit of pre-establishing reference values. Developing the ILC goes through the same process as above except we are not able to establish reference values from a calibration/test laboratory. This is where we use pivot laboratories to assure artifact stability and integrity, but not necessarily the reference value. We begin the test and for comparison purposes we use the average of the participants to begin calculating the reference values. Preliminary reports are sent to the laboratories as they participate. The test continues providing us additional measurement data. When the test is complete we use the statistical techniques described earlier to calculate the reference values and associated uncertainties and publish the final report. In this case it's almost certain that the final report will be different than the preliminary reports.

Corrective Action

This brings us to the purpose of participating in ILC/PT's. What do the results show and what opportunities are there for improvement? What should I expect from my accreditation body and how can I be prepared for my next assessment? The most important factor here is an understanding of the results. This is the primary purpose of this document. Be knowledgeable of the process that was used to provide the test. Understand the source of the data and how calculations such as the measurement uncertainty were determined. Next, do the results show any En >1? This should initiate a formal Corrective Action and at least identify the issue and begin looking for root cause.

The question that should be asked and answered is, what does your accreditation body require of your organization as you participate and review your results from that participation? Participation and review of ILC performance should be done in accordance with an established procedure. In most cases a thorough investigation needs to occur to determine why an organization did not perform per their expectations.

Participation in proficiency testing is not a requirement of ISO/IEC 17025, though the Standard does require laboratories to "assure the quality of test and calibrations results" one option being through participation in ILC's. Participation in proficiency testing is a requirement imposed by accreditation bodies /AB's. While many of our clients participate to prove their measurement processes are valid, many also participate to comply with their AB's proficiency testing requirements. Most AB's have specific requirements for participation and additional specific requirements for evaluation of your participation. It is our recommendation that you check with your AB and ensure you are meeting the AB's requirements.

Most AB's require that your ILC/PT participation comply with international requirements for interlaboratory comparisons, that is why NAPT's ILC's not only meets but exceeds these requirements. NAPT's ILC's have been audited by and approved by most of the accreditation bodies now granting accreditation to the metrology community. By participating in a NAPT sponsored ILC, you are assured of not only meeting your accreditation body's ILC requirements, but participating in a NAPT ILC is recognized by many as participating in the best proficiency testing program available.

NAPT was formed as a nonprofit organization to assist you with your measurement processes. The NAPT staff is here to help you and your organization with your measurement needs. Please remember, NAPT is the only nonprofit ILC provider providing this service to the test & measurement community. We are here for you!

If there are any additional questions you, your accreditation body or your assessors have, please do not hesitate to contact us. We look forward to working with you in not only helping you comply with accreditation requirements but improving your measurement process.

Appendix A[1]

Example of Data Analysis Performed By NAPT

Set Point:	Example	**Reporting Units:**	Volts
Discipline & Parameter:	Example	**Resolution of Artifact/Reading:**	0.0001
Established Value		**50.0029**	
Established Uncertainty		**0.0015**	

NAPT Process for Analysis for Data Submitted

Reported Mean All:	**50.0031**	**Uncertainty All: At 95%**	**0.0015**
Median:	**50.0030**	**Median: At 95%**	**0.0015**
Min Reported:	**50.0010**	**Min Reported: At 95%**	**0.0010**
Max Reported:	**50.0080**	**Max Reported: At 95%**	**0.0020**
Standard Deviation all:	**0.0017**	**Standard Deviation all: At 95%**	**0.0005**
Average Mean/Median:	**50.0030**	**Average Mean/Median: At 95%**	**0.0015**
Mean 80%:	**50.0029**	**Mean Reduced Sample:**	**0.0015**
Median 80%:	**50.0030**	**Median Reduced Sample:**	**0.0015**
Min Reported 80%:	**50.0010**	**Min Reported Reduced Sample:**	**0.0010**
Max Reported 80%:	**50.0040**	**Max Reported Reduced Sample:**	**0.0020**
Standard Deviation 80%:	**0.0009**	**Standard Deviation Reduced Sample:**	**0.0005**
Average 80% Mean/Median:	**50.0029**	**Average RS Mean/Median:**	**0.0015**
Number of Participants:	**34**	**RS Mean-Median OK?**	**Yes**
Participants Satisfactory:	**31**	**RS Range within 2 STDEV?**	**Yes**

Weighted Analysis

Weighted Value	**50.0030000**	**Weighted Unc Value**	**0.0016667**

HPC Analysis

Reported Values		**Uncertainties**	
Mean	**50.0030750**	**Mean**	**0.0015000**
StdDev	**0.0016391**	**StdDev**	**0.0005064**
Mean 5-95%	**50.0029487**	**Mean 5-95%**	**0.0015000**
StdDev 5-95%	**0.0014500**	**StdDev 5-95%**	**0.0005064**
Mean Interquartile	**50.0029487**	**Mean Interquartile**	**0.0015000**
StdDev Interquartile	**0.0014500**	**StdDev Interquartile**	**0.0014500**
Trimmed Mean	**50.0027297**	**Trimmed Mean**	**0.0015000**
Trimmed Mean StdDev	**0.0011217**	**Trimmed Mean StdDev**	**0.0005064**
Q Test Mean	**50.0030750**	**Q Test Mean**	**0.0014872**
Q Test StdDev	**0.0016391**	**Q Test StdDev**	**0.0005064**
Thompson Mean	**50.0027297**	**Thompson Mean**	**0.0015000**
Thompson StdDev	**0.0011217**	**Thompson StdDev**	**0.0005064**

Potential Reference / Pivot Lab Analysis

	Date of Measurements	**Pivot Reported Values**	**Uncertainties**
Original observations	1-Jan-04	**50.000000**	**0.00000012**
2nd set of observations	1-Mar-04	**50.000000**	**0.00000012**
3rd set of observations	1-Jun-04	**50.000000**	**0.00000016**
Mean		**50.000000**	**0.00000013**
StdDev		**0.000000**	**0.00000003**
Max		**50.000000**	**0.00000016**
Min		**50.000000**	**0.00000009**

[1] Values used in this appendix are for representation purposes only, not actual data from an ILC.