Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
National Metrology Institute

INTERNATIONAL COOPERATION

Guide for the Implementation of Proficiency Tests at a Bilateral or Regional Level

Experiences from the MERCOSUR-PTB project

german cooperation
DEUTSCHE ZUSAMMENARBEIT

On behalf of

Federal Ministry
for Economic Cooperation
and Development

On behalf of the Federal Government of Germany, the Physikalisch-Technische Bundesanstalt promotes the improvement of the framework conditions for economic, social and environmentally friendly action and thus supports the development of quality infrastructure.

Authors
Carlos Bindé, Anett Matbadal

Here you can find this publication online:
https://www.ic.ptb.de/media/fileadmin/Publications/ptb_study_guide_proficiency_testing_EN.pdf

Objective

This guide aims to present the main challenges to consider when planning and executing proficiency tests so that they can be developed as programs that effectively assess laboratory performance, develop participant competencies, and promote measures to minimize the likelihood of risks to the activities or delays that may compromise the schedule.

The recommendations given in this guide are based on experience made during the organization and implementation of proficiency testing (PT) programs for refrigerators and room air conditioners within the framework of the MERCOSUR-PTB project *Strengthening Quality Infrastructure for the promotion of energy efficiency in MERCOSUR countries*.

© PTB/Fotografie

What are Proficiency Tests and Why are they Important?

Proficiency tests (PT) are a crucial tool in the world of metrology and laboratory quality. They play a key role in ensuring the accuracy and reliability of results produced by laboratories, especially those seeking to achieve or to maintain accreditation according to the ISO/IEC 17025[1] standard.

The assessments of a laboratory's proficiency is the determination of its technical competence and accuracy in performing specific methods within its scope of accreditation. Similarly, for laboratories seeking accreditation, it is a fundamental tool to demonstrate that they are capable of performing the calibration or test within the targeted scope of accreditation. Proficiency testing in the field of metrology is often done as interlaboratory comparison (or short *intercomparison*), or at the level of NMI as key or supplementary comparison. In the area of testing, the term *proficiency testing* is commonly used for this kind of assessment.

Proficiency tests are developed and conducted in accordance with ISO/IEC 17043[2] and involve sending artifacts (equipment in the case of calibration or samples in the case of testing or sampling) to participating laboratories that must analyze them following their standard procedures. The results are then compared with known reference values, thus allowing the accuracy and precision of the measurements made by the laboratory to be assessed.

As part of PTB projects, partners often express the wish to receive support in performing comparative measurements on specific products with peers. Comparative measurements are often very expensive to perform and partner institutes often do not have the capacity to carry them out on their own. It is important to note that carrying out comparative measurements is also very time-consuming. It is therefore recommended that the organization of such comparative measurements be prioritized from the start of the project.

When the PTB, through the German government, invests resources or co-finances the implementation of proficiency testing (PT), it is essential that the participating laboratories and entities involved fully benefit from the exercise. Therefore, we recommend that these laboratories commit to openly sharing their experiences with the other laboratories sponsored by the PT. Such measures effectively exchange knowledge and transfer skills between laboratories, providers and stakeholders contribute to the construction of a set of good practices that can act beyond the scope of the specific exercise carried out within the technical cooperation project.

1 ISO/IEC 17025:2017: General requirements for the competence of testing and calibration laboratories

2 ISO/IEC 17043:2023: Conformity assessment – General requirements for the competence of proficiency testing providers

Challenges for Accredited Laboratories in Meeting the ISO/IEC 17025 Requirements for Proficiency Testing

Below, a set of key points for successful implementation of PT programs, or risks that need to be mitigated will be presented. This document provides recommendations that may or may not be feasible to execute and should be considered on a case-by-case basis. A cost-benefit relationship among all the factors presented below should be taken into consideration in the decision-making process for the elaboration of proficiency tests. The factors are segmented into three groups relating to the stages of program implementation: planning, execution and results.

The MERCOSUR-PTB project faced several setbacks, including for the identification and contracting of the PT provider. For both PT it took more than two years until the final PT provider was contracted. The duration of the PT, i.e. the realization of the measurements by the laboratories and the processing of the results was one year for air conditioners and one and a half (1.5) years for refrigerators. The difference in duration is not subject-related but mainly caused by the different number of participants in both PT.

1. Planning

1.1. Definition of the proficiency testing provider

Generally, the request for a proficiency test originates from a demand by a national entity responsible for accreditation or by the regulation of the national service. Additionally, it may come from one or more accredited laboratories or accreditation applicants.

As the quality infrastructure environment of a country or region evolves, accredited PT providers for the scope or group of equipment or product to be evaluated can be found. The use of these accredited local or regional providers should be considered a priority.

In cases where there is no PT provider with competence for this assessment, it should be evaluated whether to use a provider that is accredited in another scope and establish a technical advisory committee involving laboratories and professionals with technical knowledge on the subject. The lower the technical level and experience of the potential members of this group, the more the use of international expertise is recommended to support the various stages of the program.

In the MERCOSUR-PTB project, PT providers were contracted which both were accredited according to ISO/IEC 17043, but in another scope. The PT provider used for the PT in refrigerators is ANAB U.S.[3] accredited and the other one holding an accreditation certificate from INMETRO Brazil. Both were supported by experienced and accredited European testing laboratories, VDE Germany[4] (for refrigerators) and CEIS España[5] (for room air conditioners). CEIS España also provided the reference value for the PT in air conditioners.

Another alternative is the use of one of the participating laboratories as the program organizer (PT provider). The risks regarding privileged access to technical information

3 ANAB – ANSI National Accreditation Board

4 VDE – Verband der Elektrotechnik Elektronik Informationstechnik e.V.

5 CEIS – Centro de Ensayos Innovación y Servicios

by this in relation to the gains of technical reliability necessary for the organization and conduct of the program should be evaluated. It is recommended that this laboratory has prior knowledge of the requirements of ISO/IEC 17043 and applies them in the different stages of the program. A way to avoid undue advantages of this laboratory is to consider it only as a reference laboratory (in this qualitative case its performance is not evaluated) or to establish measures so that PTB can receive and gather the results of this participating laboratory (and PT provider) and the other participants and allow access to the other results when the necessary stages of the schedule have been executed.

If a PT is going to be based on a reference value (see chapter 1.5), it is critical that the PT provider is contracting a reference laboratory which is technically capable and which meets the requirements in terms of measurement accuracy.

In the MERCOSUR-PTB project the previously selected accredited provider for the air conditioner PT had to be ruled out, as it turned out (very late in the project) that the laboratory with which the PT provider collaborated, used standard measurement equipment (e.g., psychrometric calorimeter test chamber) which had a lower accuracy of that of the participating laboratories, which used a balanced ambient calorimeter test chamber.

The previously selected accredited PT provider for the refrigerator PT on the other hand, also had to be changed, because it was not possible for this provider to contract a reference laboratory in country.

The PT provider must fully coordinate the schedule with the project team and the laboratories, which includes agreement on and follow up of dates for the measurements and shipments, as well as reporting (taking into account the stability tests and shipping of the devices this often takes longer than expected).

Contracting the PT provider

Formalizing the contracting of the PT provider is a process that should be prioritized at the beginning of a program in order to start the practical actions of organizing and documenting effectively. It is recommended to consider intermediate payments to the provider for him to cover the execution costs of the program.

It is most effective, considering PTB's lengthy administration processes, that the contracted provider makes payments to third parties (reference laboratories suppliers, PT consultants, insurances, customs and transport services, etc.) directly related to the activity in such a way that it guarantees its responsibility for the execution of the program. In this sense, partial payments can be linked to the completion of stages such as planning, execution of the schedule (here partial conditions can be defined in relation to time and associated costs), and issuance of the final report or participation in the closing workshop.

In the MERCOSUR-PTB project the provider, despite already having international experience in carrying out PT in the area of calibration, for the air conditioner program there were many difficulties in getting the start of the measurement round operational due to customs complexities and the difficulty of hiring a company trained in exporting and importing this type of product, which led to significant delays in getting the product out of Argentina. The PTB's constant monitoring of the supplier was a decisive factor in the progress of this process.

It is up to the PTB to evaluate its participation with the use of its team and consultants who can support the stages of the program, as well as participate in the proposed events and the technical committees that are part of the program.

1.2. Definition of technical requirements and calibration or testing methods

When organizing proficiency testing programs, defining technical requirements and calibration or testing methods is a critical aspect that can present various risks, especially when involving laboratories from different countries with distinct local regulations and standards.

Different countries may have divergent calibration or testing standards and methods, which can lead to inconsistencies in the results and not allow for the proper

evaluation of participant performance. Promoting the use of standardized and internationally recognized methods, such as for example those specified by ISO, IEC, Eurachem, US Pharmacopeia and OIML to ensure uniformity. Participating in international standardization forums can also help in harmonizing practices.

However, even if laboratories of different countries apply the same international standard for products, they sometimes work based on different versions of said standards. This often happens for products that are subject to technical regulation as regulatory work is a lengthy process and does not always keep pace with the international standards work and adaptation of regulations to new standard versions does not happen equally quickly in all countries.

A clear definition of the method, terms and conditions is very important. Moreover, some testing methods can be extremely complex and time-consuming, which can affect the efficiency of the proficiency testing program. Evaluating and selecting methods that balance precision and operational efficiency can be options to determine the main characteristics. Where possible, Technologies should be incorporated that accelerate the process without compromising the integrity of the results.

The PT provider should prepare a Technical Protocol, which clearly defines the scope of the PT and introduces its objective. It furthermore should contain the description of the artifact(s) and its/their packaging, the participating laboratories and the measurement schedule, any necessary preparatory activities, transport instructions, the calibration procedure/test method to be applied and the method for the evaluation of the results.

The PT provider should encourage the participating laboratories to use standardized formats for the submission of results as this way it can be ensured that data and results are submitted in a way that enable the PT provider to evaluate and to compare them and it can be made sure that all information are collected, and a later inquiry is not necessary. The recommended formats can be introduced in the Technical Protocol.

Many parameters and test results depend on other magnitudes and test conditions that can be as relevant as the final value (the same concept can be extrapolated for qualitative tests). In some proficiency testing programs, it is important to monitor these intermediate data or even evaluate their performance given their relevance to the product or magnitude in question. In any case it is very important to agree ahead of the PT on certain conditions, including measurement points.

For example, it is critical for tests of energy efficiency of air conditioners to clarify that measurements are carried out under the same conditions (e.g., 100 %, 47 % load) and to define the various test temperatures for the inside and outside test chamber). A set of electrical power supply conditions for the product were previously established, even though the standard allows for an acceptable range, in order to reduce potential deviations inherent to the product's operation and increase the comparability of results.

In the PT for refrigerators, it turned out that the laboratories had used different number and location of thermocouples for a specific test. This fact could be at the expense of comparability. A lesson learnt was, that this situation could have been prevented by discussing and agreeing this at the beginning (e.g., in the opening workshop). The relationship between providing too precise information and allowing the laboratory to correctly interpret the regulatory requirement must always be assessed.

Finally, another factor that should be considered in the analysis of the methods to be evaluated are potential disparities in experience and technical competence among the participating laboratories since these can lead to significant variations in the results. The previous offer of training and workshops for the participating laboratories, focusing on the standardization of techniques and deep understanding of the methods, can be an alternative to be considered, if the project timeline permits.

It has also proven to be very effective to organize an alignment workshop, as a kick-off event, ahead of the PT (see chapter 1.8 Promoting Alignment Workshop). This provides a platform for the PT provider to introduce the proceedings and methods to be applied and to clarify open issues; and for the participants to address questions and doubts.

1.3. Choice and preparation of the artifact

In some cases, the sample may be sensitive and may not withstand long periods of waiting or transport conditions, this can alter its original (initial) characteristics. Implementing rigorous handling, packaging, and transport protocols for the samples, and using special packaging and boxes that maintain the necessary conditions is essential. The main factors to be considered in these cases, for example, can be temperature, humidity, vibration and shock, usage time, storage time.

Local laws may impose restrictions on the import and export of certain samples or on the disclosure of results. Consulting the applicable regulations and customs rules in advance and seeking legal alternatives such as partnerships with local laboratories for testing in the national territory are important factors to avoid the abrupt termination of the program or significant delays to the schedule. The greatest restrictions are found in the biological, chemical, forensic areas of legal metrology, and for certified products. It is important to discuss in advance with the PT provider exactly who will *create* and pay for the transport packaging and that the provision of the transport packaging can also take some time, depending on the size of the device. It should also be noted that different countries have different customs requirements/entry regulations for certain products. The PT provider must be fully aware of this in advance so that delays in the implementation of the PT can be minimized.

> The PT provider who organized the PT for air conditioners had much experience in organizing PT within Brazil but had not organized a regional PT yet. For example, shipping abroad was very difficult because, they had not thought about the fact that international shipping is more time-consuming and costly than national shipping before subcontracting. The learning effect for the PT provider was immense and only with great efforts the PT could be finalized on time.

When the choice of the artifact in proficiency tests depends on the power supply voltage or frequency, there are specific challenges due to differences in electrical infrastructures between countries or regions. These differences can significantly affect the performance and results of the tests. It is important to verify IF the participating laboratories have sources or frequency inverters that allow these measurements to be made traceable and reliably. Developing test protocols that include the standardization of operating conditions or using test equipment that automatically adjusts power inputs are also alternatives.

Certain tests require the destruction of the sample or alteration of its characteristic when performing the test method. Develop or opt for non-destructive test methods whenever possible. Alternatively, using simulated or reproducible samples that represent the characteristics of the original material is an alternative. The conducts of stability and homogeneity studies before the start of the programs and often during are essential factors to ensure the validity of the results.

> The choice for a product with a digital thermostat was mandatory so that adjustment errors and parallax in analog controllers could not interfere with the refrigerator's internal temperatures, causing unexpected variations in energy consumption.

Providing sufficiently detailed information (for example, photos and videos of how the device is to be packaged) to participants about the handling, receipt, inspection, and preparation of the artifact for transport is a critical factor and must be planned and documented appropriately.

1.4. Circulation of the artifact

Depending on the type of artifact and the time frame a suitable method for circulating the artifact should be found. There are two typical shapes, round and radial (star-shaped).

In a round-shaped PT the item is circulated successively from one participating laboratory to the next. The reference value can be determined at the beginning or at the end or in both instances (Fig. 1).

In a star-shaped PT (Fig. 2) the item is referred back to the reference laboratory after each calibration or testing by the respective participating laboratory or each partic-

ipant receives a batch or sample that first has been calibrated/tested by the reference laboratory. In this case homogeneity must be ensured (see chapter 1.5.2).

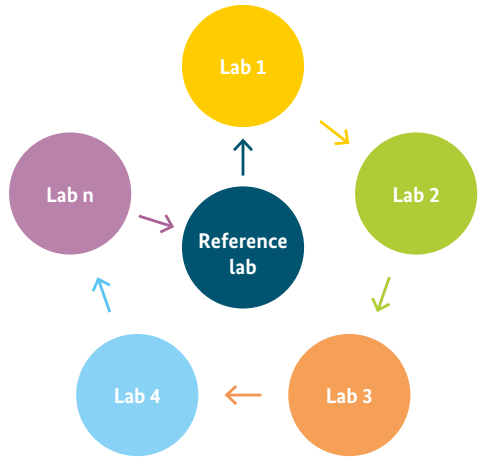

Figure 1: Round PT

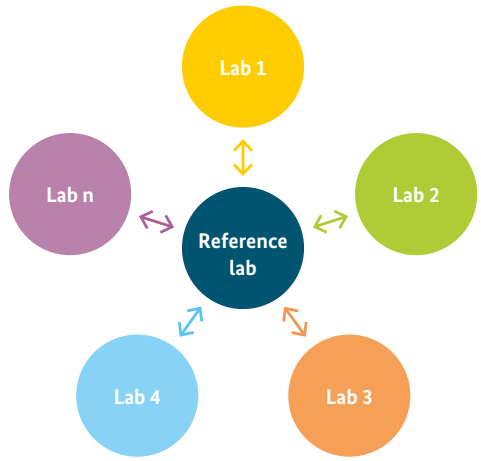

Figure 2: Star-shaped PT

Both PTs in the MERCOSUR-PTB project used a combination of the two models. The decision was based on the relationship between several factors such as program time, the lack of previous stability and homogeneity data and the probability of damage due to transport. The option of a hybrid system made it possible to establish check points to verify that results were maintained within acceptable values and prevent an error from being identified at the end of the program and the entire program being lost. It must be noted that, depending on the testing method, more time must be planned, as the device must always return to the reference laboratory.

1.5. Statistical tools and reference values

Choosing the appropriate statistical tools is essential to correctly analyze the data obtained in proficiency tests. The ISO 13528[6] standard provides guidelines on the recommended statistical techniques for these analyses.

1.5.1. Quantitative results

Option 1: The **consensus value** is determined based on the results collected from all participants in the proficiency test and serves as a benchmark to assess each laboratory (Fig. 3). This value is generally calculated as the average or median of the results, adjusted by robust statistical methods to minimize the impact of extremely high or low values (outliers). The method of using the consensus value is to be preferred when all laboratories are at the same or similar levels of capacity.

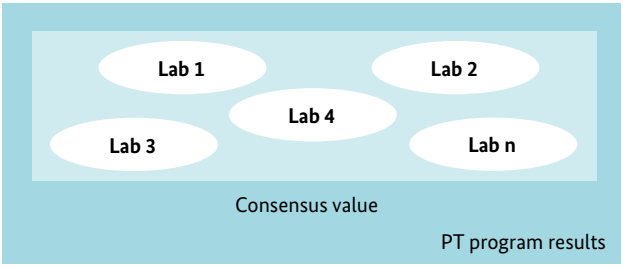

Figure 3: Consensus value – all participants

In some situations, the consensus value can be determined by a part of the participating laboratories that stands out, for example, because they are accredited, because they have demonstrated greater precision in carrying out a method or because they have lower measurement uncertainty (see *Lab 1* and *Lab 3* in Fig. 4). In this case, it is important that this criterion is clearly defined, and that the determination of this group is documented, in order to guarantee the quality of the results.

6 ISO 13528:2022 – Statistical methods for use in proficiency testing by interlaboratory comparison

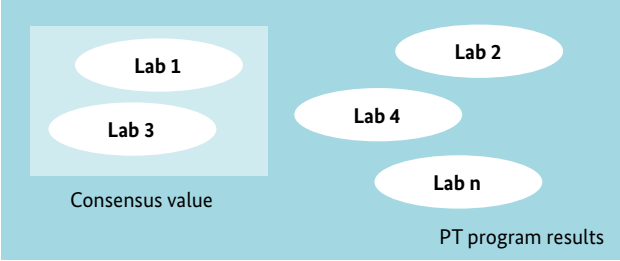

Figure 4: Consensus value – determined group

Option 2: Unlike the consensus value, the **reference value** is a theoretically *correct* or accepted value, often obtained under controlled conditions or provided by a reference laboratory like an NMI using primary standards with CMCs published in the KCDB/BIPM, or by certificated reference materials or by accredited laboratories for the quantity or method involved. The reference value may be used instead of a consensus value to assess the accuracy of the participating laboratories, especially when the properties of the sample are well known and controlled. If the competence level of the participants is still initial, the use of a reference laboratory with demonstrated competence is recommended. The feasibility of using a reference laboratory outside the country or region should be assessed.

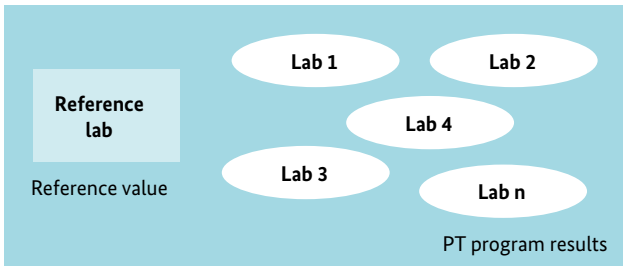

Figure 5: Reference value

For both PTs in the MERCOSUR-PTB project, the possibility of using a consensus value was studied. Only in the refrigerator program could the use of consensus be determined through accredited laboratories, while for the air conditioner program it was necessary to look for a laboratory with better capacity to provide the reference value.

For both, the consensus value and the reference value, understanding the laboratory or group of laboratories that will determine the consensus value is an important factor in bringing reliability to the program. Whenever possible, the option for accredited laboratories or NMI with their measurement uncertainties published to the KCDB for the respective quantity is very important. If such a laboratory or NMI is not available another way is to identify higher order laboratories, i.e. laboratories which use methods of higher order,or which present a lower expanded measurement uncertainty or variability (e.g., standard deviation).

It is also worth noting that proficiency testing at the level of the National Metrology Institutes is carried out under the International Bureau of Weights and Measures through Key or Supplementary Comparisons. The specific details and rules can be further investigated at *https://www.bipm.org/kcdb/*.

1.5.2. Stability and homogeneity studies

Stability and homogeneity studies are vital to ensure that the samples used in proficiency tests maintain consistent characteristics over time and/or among different batches.

Homogeneity: It is necessary to verify that the samples are uniform in all containers or batches so that all laboratories test equivalent materials. Generally, homogeneity is determined for samples that are fractionated and separated, in order to distribute them among the various participating test laboratories. For a single or the same artifact(s) measured by all participants no homogeneity tests are necessary.

Stability: the artifacts must remain stable under recommended storage and transport conditions until the conclusion of the test or calibration.

1.5.3. Qualitative results

Several specific precautions need to be taken to ensure the validity and reliability of qualitative results. These usually involve identifying the presence/absence of a component, classifying samples, or determining characteristics not directly quantifiable by numerical measures.

It can be useful to have a panel of experts review ambiguous or discrepant results. They can provide a second opinion on the interpretation of the results, increasing the reliability of the PT. Incorporate control samples or

duplicates to check the consistency of results within and between laboratories. This helps to identify any variation due to subjective interpretation or execution technique. Although the data is qualitative, statistical methods such as Cohen's kappa, concordance analysis or sensitivity and specificity tests can be applied to assess the agreement between different laboratories and the accuracy of the test methods.

1.6. Logistics

When organizing the logistics for proficiency testing programs, several critical aspects must be meticulously planned and managed to ensure the integrity and effectiveness of the program.

Logistical organization must consider the time required for transportation, customs processing, and sample analysis. Delays at any of these stages can significantly affect the overall schedule.

Often clearance times at customs are strongly underestimated in the planning, which causes delays concerning the main schedule. It is essential to proactively identify and consider customs regulations that may affect the import and export of samples. This includes tariffs, restrictions, and the need for specific documentation for the cross-border movement of regulated, sensitive, or hazardous materials.

Ensuring that the physical and chemical characteristics of the samples are maintained during transportation is essential. This includes controlling factors such as temperature, humidity and vibration, which can alter the properties of the samples. Choosing the right packaging is crucial for protecting the samples during transportation. Appropriate packaging should be selected based on the sample type and the expected conditions during transportation.

The costs associated with transportation, packaging, insurance, and customs processing must be carefully estimated and managed. These costs can significantly impact the budget of the proficiency testing program.

It must be clearly discussed who will provide the transport, who will pay for it and who is responsible.

- Does the PT provider take care of this?
- What role do the laboratories play?
- How can the PTB team provide support?
- What happens if the transport causes additional costs?
- If the test object breaks, etc.?

Such things must be precisely agreed and recorded in advance (e.g., in the Technical Protocol or in the contract with the PT provider) so that the responsibilities are clarified from the outset and there are no negative surprises or additional costs at the end.

> For the refrigerator PT in the MERCOSUR-PTB project we added a 25 % Risk Deposit for any risks that could occur during the PT (such a PT can stretch over two years and then it becomes difficult to renegotiate with the PTB administration in case of additional costs).

For many artifacts the possibility to use an ATA Carnet should be evaluated in advance. The ATA Carnet is an international customs and temporary export-import document, used to clear customs without paying duties or import taxes. The number of *CARNET Countries* is raising and it is worth checking if the countries participating in a PT belong to those countries that use ATA Carnet.

Additionally, often the laboratory schedule is another complicating factor, as it often needs to meet the speed required by clients in dynamic markets. Constant communication between the provider and participants is vital so that schedule rearrangements bring the least impact to the schedule.

Deciding whether to use one or more companies responsible for transportation or manual transport (carry by hand) may depend on various factors, including the speed required, the safety of the samples, and the cost. Manual transport may be preferable for extremely sensitive or valuable samples, while specialized companies may offer more robust and efficient solutions for most situations.

Discussing the need to return samples to the reference laboratory for stability and homogeneity control tests is crucial. This may be necessary to ensure that the samples have not undergone changes over time or during trans-

portation, which could affect the results of the proficiency tests.

The stability and homogeneity of the artifact should be strongly evaluated at the end of the program when any of these conditions occur: higher sample sensitivity, high standard deviation of results, lower measurement uncertainty declared by the participating laboratories, long distances between participants, the number of freights carried out and the duration of the program.

1.7. Confidentiality vs. learning opportunities

Confidentiality and information exchange in proficiency testing programs are crucial aspects that require a careful balance. On one hand, confidentiality protects the intellectual property, sensitive data of participating laboratories, and the possibility of collusion; on the other hand, sharing information can provide valuable learning opportunities and continuous improvement.

A common practice in PT and a fundamental requirement of ISO/IEC 17043 is to encode the results and data of participants, ensuring that the identity of the laboratories is not revealed without their explicit consent. This allows for objective analysis and discussion of results without compromising confidentiality.

Organizers of PT programs can create reports that combine data from all participants anonymously. These reports allow laboratories to compare their performance against industry benchmarks without exposing individual information. Often technical committees that support the provider receive, analyze, and debate the data without knowing the participant's identification. Developing and clearly communicating the confidentiality policies governing the PT is essential. All participants must understand how their information will be used, who will have access to it, and how it will be protected.

As mentioned before, confidentiality requirements are imposed by ISO/IEC 17043 in order to avoid collusion between participants. However, in view of the objectives of PTB's cooperation projects, i. e. developing the technical capabilities of the participating laboratories, it is important to formalize with the provider that PTB will have access to the results in order to monitor the performance of the activity and compliance with the objectives established in each project. In this same sense, it is mandatory that PTB collaborators and consultants do not disclose or share this data without the comparison report being formally issued.

On the other hand, when organizing PT in PTB projects it is advisable to remove the confidentiality of results for the purpose of learning and research, of course with the consensus of all participants. This can bring several advantages. The openness of results can increase transparency among participants, fostering an atmosphere of trust and collaboration. This can lead to a greater willingness to share practices, problems, and solutions, improving the overall level of competence among laboratories.

> In the first phase of the MERCOSUR-PTB project proficiency testing in the area of relative humidity was planned and executed very successfully by means of an interlaboratory comparison between all participating NMI. The entire group of participating NMI and the supporting reference (pilot) laboratory, INMRTRO Brazi, was openly working together in a joint training session at the alignment workshop, supported by one participating laboratory (INTI Argentina), and in the joint closing workshop where the preliminary results were presented by the pilot laboratory and openly discussed. The measurements, however, every laboratory realized on their own, obviously. The closing workshop was held at another participating laboratory (LATU Uruguay). Such an approach, however, works best at NMI level. It must be clearly said, that if private laboratories are involved, such an approach may not be possible.

When results are openly shared, laboratories can learn from each other about how different methods and techniques yield varied results. This allows participants to identify areas of improvement and implement practices that have been successful in other laboratories. With transparency, errors or inconsistencies in the data can be quickly identified and corrected. Collaboration can help solve complex problems that would be more difficult to manage in isolation.

In the MERCOSUR-PTB project, the development of the proficiency testing program for refrigerators was extensively discussed and supported by the PTB-Brazil project, in which a working group was established with the regulator, proficiency provider, testing laboratories and manufacturers in order to assess the potential impacts of the normative interpretation, the variabilities permitted by the standard, the transportation and use of the sample in order to provide a set of instructions and recommendations that increase and qualify the laboratory capacity on the aspects of energy efficiency throughout the quality infrastructure chain.

1.8. Promoting alignment workshop

Promoting an alignment workshop before the start of a proficiency testing program offers several important strategic and operational advantages and is a must-have. It is highly advisable to organize such an alignment workshop to formally open the PT. The event can serve to ensure that all participants have a uniform understanding of the objectives, processes, and expectations of the proficiency testing program. This helps to minimize misunderstandings and misalignments that could affect the quality of the results. Annex C presents a suggestion for the event agenda.

A session for alignment allows for a detailed discussion of the methods and procedures that will be used during the tests. This is crucial to ensure that all participating laboratories follow the same guidelines and techniques, which is essential for the comparability of the results.

By openly discussing all aspects of the program, including evaluation criteria and the use of results, the workshops promote a transparent environment and facilitate the identification and discussion of potential challenges or technical problems that may arise during the execution of the program. Addressing these issues in advance can prevent delays and complications.

It also creates a basis of trust between the laboratories. It can also be used to see whether the laboratories need further support to participate in the PT (training on standards, testing, etc.).

2. Execution and Monitoring

Hereunder, we address aspects related to the execution phase of proficiency tests and outline strategies that can be adopted to mitigate problems that may arise during the implementation of these programs.

Effective communication is crucial to prevent misunderstanding and misalignments. Establishing clear and efficient communication channels with the provider and monitoring the progress of the program is important to avoid long delays or rescheduling.

Regular feedback from providers can be requested for monitoring. The tasks and, above all, the responsibilities between PTB, the PT provider and the participating laboratories must be clarified from the outset. It is particularly important that everyone adheres to previously discussed agreements and deadlines. The PT provider handles all communication and agreements with the laboratories. PTB is only involved in the event of problems. Annex B summarizes the crucial steps and agreements when organizing a PT.

© iStock

Maintaining the schedule is the most challenging aspect of proficiency testing execution. Delays due to logistical problems, customs issues, and collision with laboratories' schedules can occur, and thus constant communication of the provider with each of the parties involved in these processes is fundamental to ensure appropriate deadlines. In addition, when the contract is drawn up with the PT provider, it should be stipulated what happens if a laboratory can no longer participate, the test object breaks, gets damaged or compromised or the process is delayed due to transport difficulties.

When it comes to export and import of the artifact, it is recommended that the provider monitors this closely so that bureaucratic and documentary issues are initiated before the actual release date of the artifact by the predecessor laboratory.

Local rules often require special declarations about the artifact or approval by some health or trade entity, which takes time. It is important that the PT provider familiarizes himself with the different requirements/customs regulations of the participating countries beforehand and, if necessary, consults with the laboratories again to find out how transport can be simplified so that there are no problems or delays later on.

Long proficiency testing programs or those in which the artifact is sensitive to conditions that affect its stability and/or homogeneity often require monitoring of these parameters. The execution of this monitoring, already stipulated in the planning phase, must be followed up with the provider so that any corrective action can be taken to maintain confidence in the results and ensure that any undue deviation is attributed to the participating laboratories. Remember that small deviations can be identified and added to the provider's mathematical studies.

Monitoring these points along with the provider should minimize any deviations from the expected program and thus ensure that subsequent actions can be taken throughout the project.

3. Results and Subsequent Actions

In the final phase of a proficiency test, effective data sharing with participants is essential. This phase not only represents the culmination of the entire process but also plays a key role in guiding laboratories on the actions needed to improve or maintain their competencies. On the other hand, the conclusion of a proficiency test is also an opportunity to bring the participants back together to deal with the lessons learned, with their points of satisfaction and improvement, from the point of view of both the laboratory and the provider and PTB.

3.1. Sharing results

The proficiency testing reports should include a detailed analysis of the results of each participant compared to the reference or consensus values. It is important that these reports are clear, precise, and contain practical recommendations for improvement actions. Each laboratory should receive information detailing how their results compare with those of other participants, in addition to identifying any trends or patterns.

It is an inherent process to the provider's rules, but it is always recommended that they issue a preliminary version of the document so that those involved in the program can review it and suggest any improvements to the document or clarifications. It may happen that laboratories request the sending of new results; however, this practice is not recommended, and potential errors should be treated internally by them.

Here, the aspects of confidentiality established since the initial phase of the program (see item 1.7) remain relevant. However, as mentioned under section 1.7, PTB should seek authorization from the supported laboratories to disclose their results to PTB in order to find the best strategy for continued technical support after the PT.

3.2. Promoting a closing workshop

Just as recommended for the initial phase, holding a closing event to discuss the achieved results is extremely important and a must-have for the conclusion of proficiency testing programs. Annex D presents a suggestion for the event agenda.

This session allows aligning the interpretation of the results and debating the potential sources that led to the agreement or disagreement of the results, discussing opportunities for improvement for the laboratories, as well as indicating improvements and control points in future programs. Listening to participants' experiences is fundamental to learning and evaluating the process. Annex A gives details of these results for the two programs developed in the project with MERCOSUR.

This set of outputs shared by the provider, technical committee, reference laboratory, and participants, in addition to the PTB, can constitute a list of opportunities to strengthen this very important part of the Quality Infrastructure and can be a reference for a new action plan.

If there is still enough time left in the project, challenges or difficulties that have come to light during the PT can also be addressed, support and training can be offered for relevant laboratories as required, and accreditation processes can be further supported.

3.3. Lessons learned

Based on the results and feedback from the participants, the organizers can assist each laboratory in developing a specific action plan. This plan may include measures such as additional training, recalibration of equipment, or revision of operational procedures. The idea is that each laboratory has a clear route to enhance its precision and technical competence.

Consider publishing case studies or best practices identified during the test. This not only increases the general knowledge within the quality infrastructure environment but also serves as recognition for laboratories that demonstrate excellence or significant improvements.

Offering ongoing support to the laboratories after the completion of the proficiency test is vital. This can include follow-up consultations, additional technical support, or even inclusion in future proficiency tests to measure progress over time.

For the PTB to have the opportunity to provide further support to laboratories based on their results, it is vital for the PT to be finalized at least 6 months before the end of the PTB project.

4. Updates to this Guide

© iStock

To ensure that the guide on proficiency testing remains relevant and effective, it is essential to incorporate a process of continuous review based on the lessons learned during its application. Contribute by updating this document whenever necessary.

Annex A – Feedback from Participants

The following are the results of the feedback collected from the MERCOSUR-PTB proficiency programs, in order to present the participating laboratories' perception of the actions taken.

1. Given the complexity of organising this complex programme, what challenges do you think were faced and what lessons could your lab share with others?

- Particularly for our laboratory, we did not have extensive experience with the 2020 version of the standard, so the challenge was to develop the calculations and test methods considering the new standard. As lessons learned, we look forward to the outcome of the final report and the course to detect errors or deviations we have made in the testing process. The product reception and dispatch control procedures were very enlightening and positive.
- The organisation was well planned, the timing was a bit delayed, but we understand that this is a likely possibility for such a complex study with so many logistics and different border systems.
- The selection of the sample and the design of the procedure is always a challenge. This time with the sample there were no problems, but the procedure was not clear and underwent several modifications throughout the programme.
- Nothing to declare.
- As I was the one who initiated the interlaboratory round, the main challenge was to apply the test protocol given the differences we found with respect to the reference standard. This caused us to have doubts during the execution of the tests, and we had to turn to the organiser several times, unfortunately without finding the answers we were looking for.
- As a main challenge we see the coordination with all the laboratories to take this forward. As this is not a routine test for us, there is not much we can share with other laboratories. What we have learned and the results will bring us closer to the goal of accrediting this test.
- The challenge was to perform the test while maintaining the physical integrity of the circulating equipment.
- I think the biggest complexity is precisely related to the organisation of the programme and the logistics involved in moving the sample. As a laboratory, we just need to perform the test as we are used to, only taking special care in relation to the execution instructions received from the proficiency testing provider. The most demanding aspect of the laboratory is to reconcile the arrival of samples with the availability of space in the climatic chamber, considering some small delays in the arrival of the sample.

2. Do you have any suggestions for improvement (organisation, communication, logistics, implementation, etc.)?

- Communication and execution were good. Regarding organisation and logistics, there were several delays in receiving the sample in our laboratory and the time allotted ended up being limited, especially considering that we were simultaneously developing the method and the spreadsheets.
- The organisation of this type of exercise is highly complex, so the deviations that there were are understandable, I think it was good in general terms, in our case we had to ask not to be called every day of Q work at a certain time because it was impossible to work that way, they understood and we were able to do the rehearsals without pressure. Organisation: Very good, Communication: Very good, Logistics: Had problems, Execution: Executed as planned.
- Everything went perfectly and according to plan.
- No! The programme was very well organised.
- Organisation: Correct, we would have liked to repeat one of the tests, but we ran out of time and had to send the sample to the next laboratory. Communication: Good, although at the beginning there was no fluency with those who could answer our technical questions. Improved, but coming to the end; Logistics: Correct; Execution: Correct.
- We in particular were the last laboratory to receive the cooler. This meant that due to the various setbacks and difficulties that arose, our planned date for receiving the sample changed several times. We understand that the difficulty and the different customs regulations (even between MERCOSUR) was part of

the obstacle, therefore, we believe that it was not a shortcoming of the supplier of the interlaboratory itself.
- We would like to highlight positively the constant communication from the programme manager, who was always there to clarify any doubts.
- Ideally, we would like to strictly adhere to the arrival dates of samples for testing; however, we understand the difficulty associated with this and that sometimes the situation is beyond the supplier's control.

3. What advantages and disadvantages do you see in using a proficiency testing provider in such a complex programme?

- A complex programme is actually more representative of reality and very rich in the results and improvements detected for the laboratories. The use of a supplier simplifies the presentation of results by an independent body, a key function of these activities. The drawbacks are mainly logistical and cost-related, but thanks to the support of PTB, these are reduced to a minimum (testing times and use of the chamber).
- Proficiency testing is one of the most powerful tools for demonstrating technical competence, which makes this type of exercise very valuable for laboratories. The possibility that PTB gives us to do them free of charge is very important as it would be impossible for a small laboratory like ours to buy them. We participate in many rounds of interlaboratory exercises during the year, but not for this kind of tests they are very expensive.
- Such a supplier is mandatory and necessary. The recurrent problem with these suppliers is the lack of experience in the field, as it is always their first experience. It would be ideal if they are previously accredited and with several interlaboratory tests carried out in the specific field.
- Advantages: Important to be aligned with laboratories outside Brazil, plus perhaps opportunities to test for MERCOSUR; Disadvantages: Costs that may be involved.
- I consider that whoever provides or organises a proficiency test should have extensive knowledge of the standard being applied.
- The advantages are precisely that the complexity of these tasks means that without a dedicated supplier to carry out this proficiency programme, the laboratories would have to carry out these tasks (coordination, customs clearance, transport, etc.). We see no disadvantages.
- The advantage lies in having an autonomous and independent organisation to manage the programme, ensuring compliance in the treatment of results.
- One of the advantages is the convenience for participants of not having to worry about logistical problems, especially when the sample has to be transported between different countries. The laboratory will only have to make the sample available for collection on the agreed date. Another aspect is that there will be an independent company (the supplier) to evaluate the results and able (hopefully) to propose the statistical evaluation method adapted to the comparison programme and make the necessary adjustments based on the data obtained.

4. Compared to the previous situation, how has your participation in this programme contributed to the development of your technical skills?

- We have managed to run the test in the chamber where we usually test for the previous version of the standard. We have managed to understand a little more about the application of the new standard and to make the first calculations and spreadsheets for its implementation. It has been very enriching for us in that sense, and we are looking forward to a second round of this same test in a few years' time.
- We have not yet tested with the new standard because it is not yet in force in our country. However, the experience was very good, we have identified many weaknesses and opportunities for improvement in our process. We are looking forward to the results.
- Participation in this programme was a great learning experience. Skills development is very much needed in the region and being able to have these programmes in the future makes us improve as a laboratory. This one in particular made us improve and strengthen our technical capacity. Let's wait for the final results to understand more about the measurements.
- Nothing to declare.
- Beyond the above answers, it was enriching in every technical aspect, especially to be able to discuss with colleagues our differences in interpretation when executing it.
- For us, the results of this interlaboratory will provide us with a lot. As LATU is at a stage prior to final implementation, carrying out these tests with comparative

results will force us to improve internal procedures in order to carry out the tests in the best possible way once the test has been implemented.

- Participation in inter-laboratory comparison programmes always brings benefits to the participating laboratory to verify the conformity of its results and how it compares to other participating laboratories.
- As for the development of technical skills, although we already have a fairly consolidated knowledge of this type of testing, when participating in a benchmarking programme we always carry out a review exercise of the normative references and procedures adopted, serving as an exercise in critical analysis. of what we do on a daily basis. However, in doing so, we rely solely on our own knowledge and desire to do so. This is not really a training that allows us to incorporate external knowledge, but an opportunity to compare ourselves with other laboratories, which is also very important.

Annex B – Checklist for Organizing a Proficiency Test Within a Project

Activities	Details
Defining the scope of the interlaboratory comparison	Scope should reflect the project goals and is selected as to support the accreditation efforts of participating laboratories. The scope should be selected to allow most laboratories to participate. Note: It is possible that a laboratory only participates in a part of the defined scope.
Defining the potential participating laboratories	It can be decided that only accredited laboratories participate or that the PT is open to all interested laboratories in the partner countries, provided that they have the technical means to do the measurements.
Defining the hiring of experts or partner laboratories	It can be valuable to involve a mentoring expert/laboratory to coach the pilot laboratory or to guide PTB in technically overseeing the exercise.
Mapping the potential artifact	An artifact can be an object that is already available (at the PT provider) or that has to be procured. Consider the time for purchasing a suitable item. Consider possible transport and power setup (voltage/frequency) issues in the participating countries. Check the necessary transport preparations, including import/export restrictions imposed by local laws. Depending on the artifact its circulation should be properly selected (star-shaped or round). Check the possibility to use ATA Carnet to accompany the artifact.
Hiring the provider	Necessity is depending on the objective of the PT: ■ It may be necessary to involve a PT provider which is accredited according to ISO/IEC 17043 for the respective scope. ■ It may be sufficient to involve a PT provider which is accredited in another scope or not accredited, but experienced. If registration of the PT in the KCDB is intended, involvement of a suitable pilot laboratory (NMI) is required, which provides the reference value and the link to a Key Comparison (a key comparison reference value, in particular). If a PT provider is involved, this PT provider handles all communication with the laboratories. However, timelines and responsibilities need to be clarified early: ■ Who communicates with whom? ■ Who coordinates the logistics/transport of artifact? ■ Overall approach to disclose results of participants at the end? ■ Disclosure of individual results to PTB in case of coding the results and non-disclosure to participants. ■ How often feedback has to be provided to PTB? ■ Who follows up on what?
Approve the Technical Protocol	The Technical Protocol (TP) is usually prepared by the PT provider or the pilot laboratory. PTB should review the TP (with the help of its expert/coach) and approve it.
Laboratories register and agree to the program rules, including those on impartiality and confidentiality	Impartiality and confidentiality is a MUST for all laboratories. No collusion amongst the laboratories can be accepted. If there is such an agreement, and if at the end results are openly shared amongst the participating laboratories, this does not affect the impartiality and confidentiality requirements that all laboratories have to meet throughout the exercise.

Organize the opening workshop	PTB organizes the alignment workshop in coordination with the PT provider or pilot laboratory. See Annex C for a suggested agenda.
Conduct the opening workshop	The PT provider or pilot laboratory is leading the event. Can be done virtually or presential.
Monitor the schedule of measurements	The PT provider or pilot laboratory should monitor the schedule and inform PTB on a regular basis. PTB should follow up frequently with the PT provider.
Monitor the analysis of the results	This is to be done by the PT provider or pilot laboratory. PTB should follow up frequently with the PT provider.
Organize the closing workshop	PTB organizes the alignment workshop in coordination with the PT provider or pilot laboratory. See Annex D for a suggested agenda.
Closing workshop	The PT provider or pilot laboratory is leading the event. Can be done virtually or presential.
Record improvement actions	Follow-up activities and measures should be derived from the results of the closing workshop. Measures should be individually agreed between PTB and laboratories in the form of an action plan. Could be general activities such as e.g. a subject-related training on measurement uncertainty.
Update this guide	Enrich the guide with additional experience and guidelines.

Annex C – Suggested Agenda for the Opening Workshop

* The host is either the PT provider or the pilot laboratory.

The following agenda is most suitable for a virtual event. Items 3 to 6 are presented by the host.

1. Opening remarks by PTB (or its coach) and the host
2. Introduction of participating laboratories (every laboratory)
3. Presentation of the program protocol
4. Presentation of the artifact and service to be performed
5. Presentation of the schedule
6. Technical presentation on methods, standards and exchange with participants
7. Space for questions and clarifications about the program and next steps

The opening (alignment) workshop can be organized virtually or presential. A presential event is useful and often possible in particular for interlaboratory comparisons between NMIs. In this case it is an option to include a laboratory visit and a practical session in the opening workshop.

Annex D – Suggested Agenda for the Closing Workshop

* The host is either the PT provider or the pilot laboratory.

1. Opening remarks by PTB (or its coach) and the host
2. Overall summary by the host (general course of the program and any major issues)
3. Presentation of results by the host (normally, the preliminary report or Draft A Report is being presented)
4. Space for questions and clarifications about the results
5. Space for discussion of lessons learned and necessary future actions
6. Feedback from participants about the PT – See questions in Annex A
7. Agreement on next steps (e. g. deadline for final report, follow-up activities for improvement)

The closing workshop can be organized virtually or pre-sential.

Imprint

Published by
Physikalisch-Technische Bundesanstalt
Bundesallee 100
38116 Braunschweig
Germany

Responsible
Dr. Marion Stoldt
+49 531 592-9300
marion.stoldt@ptb.de
www.ptb.de/q.3/en

Text
Carlos Bindé, Anett Matbadal

Editing
Ann-Kathrin Flamming

As of October 2024

Contact

Physikalisch-Technische Bundesanstalt
International Cooperation
Dr. Marion Stoldt
Phone +49 531 592-9300
marion.stoldt@ptb.de
www.ptb.de/q.3/en

Here you can find this publication online: