Cause analysis for unsatisfactory results in proficiency testing activities: a case study of Brazilian calibration laboratories accredited under ISO/IEC 17025:2005*

M.A.F. Silva[1], M.H. Moura[1], R. Nogueira[1,**], and S.R.R. Costa[2]

[1] National Institute of Metrology, Quality and Technology (Inmetro), General Coordination for Accreditation (Cgcre), Laboratory Accreditation Division (Dicla), Xerém 25250-020, Rio de Janeiro, Brazil
[2] Fluminense Federal University; Technology, Management and Environment Laboratory, Niterói 24220-900, Rio de Janeiro, Brazil

Abstract. This work presents the results of a survey carried out among Brazilian calibration laboratories accredited under ISO/IEC 17025:2005 with the objective to identify how these laboratories investigate the root causes of unsatisfactory results in proficiency testing. The survey was coordinated by the Brazilian accreditation body, the General Coordination for Accreditation (Cgcre), of the Institute of Metrology, Quality and Technology (Inmetro).

Keywords: Laboratory accreditation; ISO/IEC 17025:2005; proficiency testing; measurement audits; corrective action

1 Introduction

Testing and calibration laboratories accredited under ISO/IEC 17025:2005 must demonstrate that they are able to generate technically valid results and to operate their management systems [1]. According to clause 5.9 of this standard, the validity of tests and calibrations shall be monitored. This monitoring may include, but not be limited to, participation in proficiency testing or interlaboratory comparisons, regular use of certified reference materials, replicate tests/calibrations using the same or different methods, and retesting/recalibration of items [1].

Both proficiency testing ("evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons" [2]) and interlaboratory comparisons ("organization, performance and evaluation of measurements or tests on the same or similar items by two or more laboratories in accordance with predetermined conditions" [2]) are essential tools to monitor, on a regular basis, the overall performance of the laboratory [3] and its technical competence, to validate uncertainty claims within the scope of accreditation [4], to evaluate the laboratory's performance against other participant laboratories, and to identify systematic problems related to testing/calibration procedures. Proficiency testing are not only essential for laboratories and their customers, but also for regulators and laboratory accreditation bodies [2]. The frequency of participation in proficiency testing is strongly influenced by costs, shipment of materials [5] and availability of programmes. Besides these difficulties, the participation in proficiency testing in developing countries is considerably increasing [6].

Other type of interlaboratory program is the measurement audit, which is conducted as part of the assessment process to assist the accreditation bodies in establishing competence of the laboratory [7]. This procedure consists in offering well-known artifacts to be calibrated by the participant laboratory and then evaluating the participant's results against the assigned values established by a reference laboratory.

According to ILAC, "proficiency testing activity" is all those activities of comparisons of tests, calibrations and inspections between laboratories/inspection bodies and used by accreditation bodies to assess performance, including proficiency tests, interlaboratory comparisons and measurement audits [8].

In order to determine whether or not a participating laboratory is proficient for a particular measurement, an evaluation of the laboratory's performance must be conducted and the most commonly used method for determining the performance of an individual calibration laboratory [9] is the "E_n numbers" [10], also called "normalized error", which is calculated according to equation (1).

$$E_n = (x - X)/\sqrt{U_{\text{lab}}^2 - U_{\text{ref}}^2} \qquad (1)$$

where x is the laboratory's measured value, X is the assigned value, U_{lab} is the laboratory's expanded uncertainty, and U_{ref} is the expanded uncertainty of the assigned value. Both expanded uncertainties (U_{lab} and
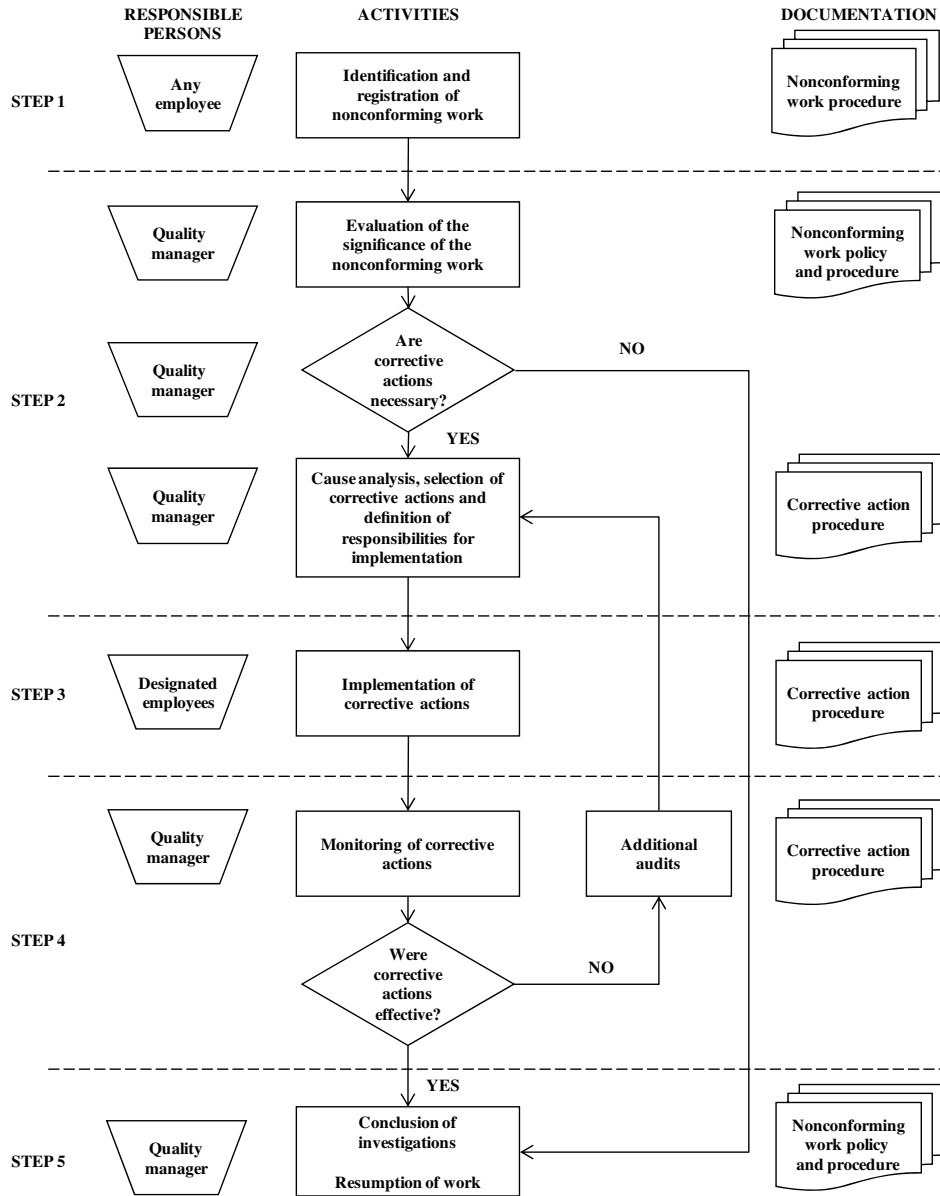

Fig. 1. Steps for treatment of nonconformities based on the ISO/IEC 17025:2005.

U_{ref}) shall be stated as the standard uncertainties of measurement multiplied by the coverage factor (k), which for a t-distribution corresponds to a coverage probability of approximately 95% [11]. For a normal distribution, the coverage factor k equals 2. The use of approximately the same coverage probability is essential whenever two results of measurement of the same quantity have to be compared [11].

In principle, if E_{n} values are between –1 and 1, the laboratory is proficient for the type of measurement performed. However, if E_{n} values are outside this range, laboratories are asked to perform a cause analysis of unsatisfactory results and to take corrective actions.

The accomplishment of ISO/IEC 17025:2005 [1] clauses 4.9 (Control of nonconforming testing and/or calibration work) and 4.11 (Corrective action) is usually carried out as shown in Figure 1. Clause 4.9 asks for policy and procedure for nonconforming work, definition of responsibilities and authorities for nonconforming work (which includes halting of work and withholding of certificates if necessary) and for resumption of work, evaluation of significance of nonconforming work, customer notification and work recall, if necessary, and adoption of corrective actions. Additionally, clause 4.11 requires policy and procedure for corrective action, definition of responsibilities, investigation of the root causes of problems, as well as selection, implementation and monitoring of corrective actions.

In Brazil, the General Coordination for Accreditation (Cgcre), which is the Brazilian accreditation body, is signatory of the mutual recognition arrangement (MRA) of the International Laboratory Accreditation Cooperation (ILAC) and makes part of the Institute of Metrology, Quality and Technology (Inmetro), but works

independently from it to ensure impartiality, as required by the ISO/IEC 17011:2004 [12]. The department called Metrological Reliability Sector (Secme), which makes part of the Laboratory Accreditation Division (Dicla), controls the participation of Brazilian testing and calibration accredited laboratories in proficiency testing, and also organizes measurement audits, usually before accreditation, but also in case of customers complaints concerning accredited laboratories and prior to revoke the suspension of accreditation. Secme activities are in accordance with the clause 7.15 of ISO/IEC 17011:2004 [12], which states that "the accreditation body shall ensure that its accredited laboratories participate in proficiency testing or other comparison programmes, where available and appropriate, and that corrective actions are carried out when necessary", and also with the ILAC policy for participation in proficiency testing activities [13].

According to Secme experience, Brazilian calibration laboratories usually tend to expend a lot of time in the cause analysis of unsatisfactory results in proficiency testing activities, but they are not always able to select the most appropriate corrective action that will really solve the problems. This situation brings consequences not only for the accredited laboratories, but also for Cgcre, since it creates demands for suspension of previously accredited services and/or reduction of accreditation scopes.

In this work, we present the results of a voluntary-response survey carried out by Cgcre among Brazilian calibration laboratories accredited under ISO/IEC 17025:2005 in order to better understand the procedures used by them to investigate root causes and to take corrective actions in case of unsatisfactory results in proficiency testing activities. The survey intended to point out the items that should be more carefully investigated by the Brazilian calibration laboratories, in order to help them to improve their management systems and to strength their technical competence.

2 Materials and methods

2.1 Structure of the survey questionnaire

A questionnaire was elaborated based on ISO/IEC 17025:2005 and consisted of six parts, as shown in Table 1. The part 3 of the questionnaire consisted of 36 questions, which can be seen in the Supplementary material. The correlation between these questions and the ISO/IEC 17025:2005 clauses are summarized in Table 2.

Prior to the survey, a pilot study was carried out. The questionnaire was sent to 10 laboratory technical managers to evaluate its adequacy. Based on the 7 replies received, some improvements were made in the document, including the use of shorter sentences.

2.2 Survey participants from Brazilian accredited laboratories

In May 2009, 271 laboratory technical managers of 242 Brazilian calibration laboratories accredited under ISO/IEC 17025:2005 were invited to take part in the survey by means of answering to the questionnaire. At the end of June of this same year, 89 questionnaires were returned, corresponding to 32.84% of responses based on the total number of invited technical managers. The confidentiality about the participant laboratories and managers was guaranteed.

Some laboratories are accredited by Cgcre in more than one metrological area (calibration service groups), e.g., electricity and temperature. Therefore, the calculations from the survey data were based on the number of calibration service groups. The 89 answered questionnaires corresponded to 141 calibration service groups, while the sum of all calibration service groups of the 242 Brazilian calibration laboratories totalized 439.

2.3 Data treatment

The answers to questions from parts 1 and 2 of the questionnaire were used to identify the profiles of the laboratories and of the laboratory's technical managers.

For the answers to questions from part 3 (items that are investigated by the laboratories and how they do it), a Likert-based scale was used to allow the quantitative data analysis based on averages and dispersions. The scale value corresponded to four values assigned according to the laboratory responses: the laboratory investigates systematically (scale value: 1), it investigates only sporadically (scale value: 3), it never investigates (scale value: 5), or not applicable (scale value: 0). The results were grouped into six categories: I. Purchase of calibration services, II. Equipment and reference standards, III. Methods, IV. Environmental conditions, V. Personnel, and VI. Measurement parameters. Additionally, an Ishikawa (cause-and-effect) diagram was constructed based on the answers from part 3 (what and how the laboratories investigate) and part 5 (root causes of nonconformity in proficiency testing activities, divided into the same six categories listed above). The free answers to questions from parts 4 and 6 were used to cross-check the answers obtained in parts 3 and 5, respectively.

Finally, Secme previous records about the participation of the calibration laboratories in proficiency testing activities were reviewed, in order to cross-check this information with those reported in the questionnaires answered by the participant laboratories.

3 Results and discussion

3.1 Purpose of the questionnaire

The survey questionnaire was designed to gather information from Brazilian calibration laboratories accredited under ISO/IEC 17025:2005 such as:

– How do laboratories investigate the root causes of unsatisfactory results in proficiency testing activities?

Table 1. Survey questionnaire main topics for cause analysis of unsatisfactory results in proficiency testing activities.

Part	Main purpose	Questions	Types of answers
1	Identification of the profile of the laboratory technical responsibles	Education Sex Years of experience as technical manager	Graduated, MBA, M.Sc., PhD Male, female Number of years
2	Identification of the laboratories' profile	Type of organization Accredited calibration service groups Number of technicians per accredited calibration service group Number of years since laboratory accreditation (per calibration service group)	Public, private Acoustics and vibrations; dimensional; electricity; chemistry; force, torque and hardness; mass; optics; pressure; radio-frequency; temperature and humidity; time and frequency; flow; viscosity; volume and specific mass Number of employees Number of years
3	Items investigated by the laboratories in case of unsatisfactory results in proficiency testing activities	36 questions according to ISO/IEC 17025:2005 (eletronic Supplementary material)	IS: the laboratory investigates systematically NS: the laboratory investigates only sporadically NI: the laboratory never investigates NA: not applicable
4	Additional items investigated by the laboratories besides those given in question (3)	Does the laboratory investigate other root causes besides those given in question (3)? Which ones?	Free answer
5	Main groups of root causes observed by the laboratories in case of unsatisfactory results in proficiency testing activities, based on Ishikawa 6M (Machine, Method, Materials, Measurement, Man, and Mother nature (environment))	Based on your experience, the main root causes of unsatisfactory results in proficiency testing activities are related to which of these factors?	I – Purchase of calibration services (3.1 to 3.4) II – Equipment and reference standards (3.5 to 3.14) III – Methods (3.15 to 3.19) IV – Environmental conditions (3.20 to 3.25) V – Personnel (3.26 to 3.29) VI – Measuring parameters (3.30 to 3.36)
6	Examples of root causes of unsatisfactory results in proficiency testing activities previously observed by the laboratories	Could you share examples of root causes of unsatisfactory results in proficiency testing activities observed in your laboratory?	Free answer

Table 2. Correlation between the ISO/IEC 17025:2005 clause and the thirty-six multiple choice questions to investigate the root causes of unsatisfactory results in proficiency testing activities (questionnaire part 3).

	Question nr.	ISO/IEC 17025:2005 clause
I - Purchase of calibration services	3.1	4.6.4
	3.2	4.6.3
	3.3	4.6.2
	3.4	4.6.2
II - Equipment and reference standards	3.5	5.5.2
	3.6	5.6.1
	3.7	5.5.2
	3.8	5.5.3
	3.9	5.5.10
	3.10	5.5.11
	3.11	5.5.12
	3.12	5.5.7
	3.13	5.5.6
	3.14	5.5.8
III - Methods	3.15	5.4.2
	3.16	4.4.1, 5.4.2
	3.17	5.4.6.3
	3.18	5.4.5
	3.19	5.4.7.2
IV - Environmental conditions	3.20	5.3.2
	3.21	5.3.2
	3.22	5.3.3
	3.23	5.3.3
	3.24	5.3.2
	3.25	5.3.4
V - Personnel	3.26	5.2.1
	3.27	5.2.2
	3.28	5.2.2
	3.29	5.2.5
VI - Measuring parameters	3.30	5.8
	3.31	4.13.2
	3.32	4.13.2
	3.33	4.13.2
	3.34	4.13.2
	3.35	4.13.2
	3.36	5.10

– Which are the main items considered in the cause analysis? How systematically are they investigated?
– Which are the main root causes of unsatisfactory results?
– Is there a correlation between the main root causes observed and what has been investigated?

3.2 Summary of answers to survey questions from parts 1 and 2: profile of laboratory technical managers and of the participant laboratories

The participant laboratory technical managers presented the following profile: (i) 91% were male and 9% female; (ii) concerning education, 82% of the participants attended at least an university graduation course, from which 47% were graduated, 20% took some specialization course (e.g.: MBA), 11% had a M. Sc. degree, and 4% were PhD; (iii) concerning professional experience, 54% had worked as technical managers for less than 5 years, while the others had experiences between 6 and 10 years (29%), 11 and 15 years (13%), and more than 15 years (4%).

Although the participation of calibration laboratories able to carry out dimensional and pressure calibration services was quite significant in the survey (23% and 22% of the total calibration service groups, respectively), the numbers of laboratories per calibration service group was considered to be balanced. The profile of this sample of participant laboratories was considered to be quite similar to that obtained for the total number of accredited calibration service groups, as shown in Figure 2, and the survey was considered to be representative.
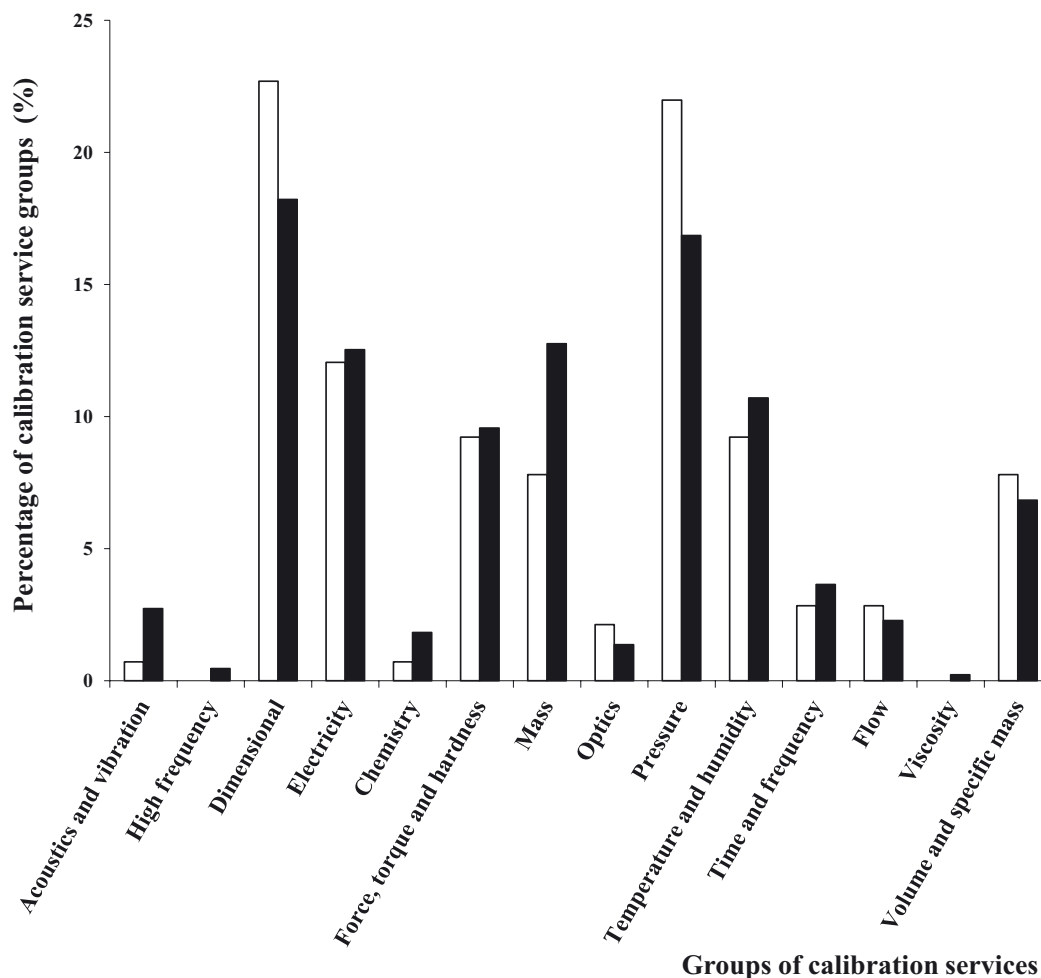
Fig. 2. Percentage of calibration service groups in the survey participant laboratories group (white bars) and in the total population of calibration laboratories accredited by Cgcre (dark bars). The chemistry group is represented by laboratories that calibrate pH/conductivity meters and portable gas analyzers (e.g., CH_4, CO_x, NO_x, SO_x, NH_3).

Summary of answers to survey questions from parts 3 and 4: systematic investigation of items by the laboratories in case of unsatisfactory results in proficiency testing activities.

The questionnaire part 3 consisted of 36 multiple choice questions based on the ISO/IEC 17025:2005 (as summarized in Tab. 2 and given in full as electronic Supplementary material) and the laboratories were asked to point out which items they used to investigate systematically, to investigate sporadically, or not to investigate. In order to allow quantification, the laboratories responses to question (3) were multiplied by Likert-based scale factors previously described (see Sect. 2.3). Therefore, the responses should be included in the range between 1 (the laboratory investigates systematically) and 5 (the laboratory never investigates), being the average note 3 given to sporadical investigation. The results are graphically presented in Figure 3.

For most of the 36 questions, the averages were around 1 (systematic investigation). For some questions there was a significant number of answers corresponding to "sporadic investigation" (note 3), including questions (1)

and (2) (purchase of calibration services), 8 (equipment operation only by authorized technicians), 10 (use of updated correction factors), 11 (protection of hardware and software), and 12 (segregation of equipment out of service). Additionally, for questions (21)–(25), which were included in the category IV (environmental conditions), the answers varied largely between 1 and 5.

The question from part 4, which was answered by 24 participants, consisted of free identification of other items (besides those from part 3) that were investigated by the laboratory in case of unsatisfactory results. The most important contributions included the participation in other proficiency testing activities, investigation of the adequacy of calibration procedures, analysis of previous calibrations, investigation of the behavior of calibration standards along the time, as well as investigation on calculation and data transcription, which were considered to be somehow related to the questions of part 3. Other answers, including the check of the proficiency testing item/artifact or evaluation of performance of other proficiency testing participants, were considered to be outside of the laboratory's responsibility and were not taken into account.
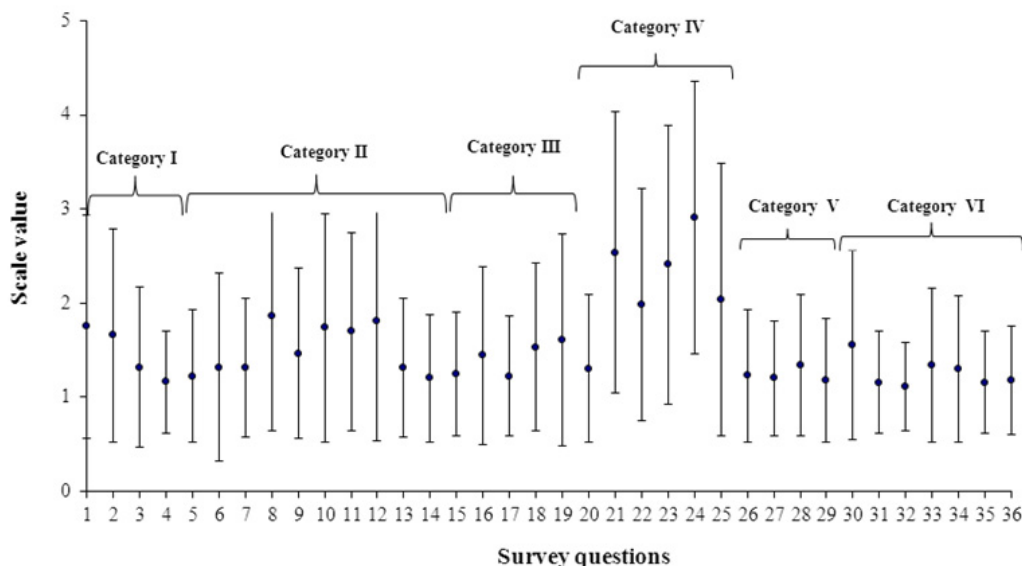
Fig. 3. Plots of averages and dispersions for answers to the 36 questions from part 3, grouped into 6 categories: I. Purchase of calibration services (3.1 to 3.4), II. Equipment and reference standards (3.5 to 3.14), III. Methods (3.15 to 3.19), IV. Environmental conditions (3.20 to 3.25), V. Personnel (3.26 to 3.29), and VI. Measuring parameters (3.30 to 3.36). The scale value corresponds to the following values, assigned according to the laboratory responses: the laboratory investigates systematically (scale value: 1), it investigates only sporadically (scale value: 3), it never investigates (scale value: 5), and not applicable (scale value: 0).
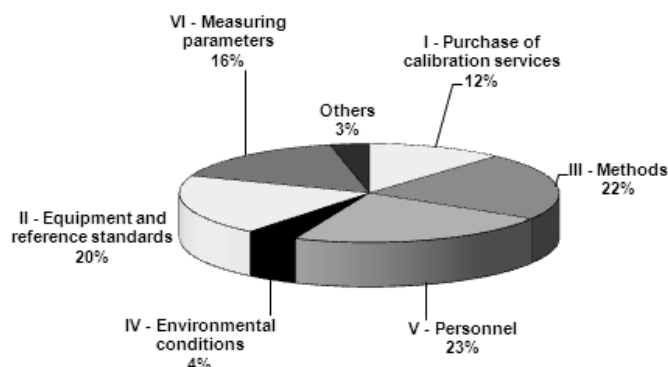

Fig. 4. Root causes of nonconformities in proficiency testing activities reported by the laboratories.

3.3 Summary of answers to survey questions from parts 5 and 6: groups and examples of root causes of unsatisfactory results in proficiency testing activities observed by the laboratories

In part 5, the laboratories indicated which of the six main categories (previously shown in Tab. 1) included the usual root causes of unsatisfactory results in proficiency testing activities. As can be seen in Figure 4, most of the laboratories associated their unsatisfactory results to personnel (23%, category V), equipment and calibration standards (20%, category II), and methods (22%, category III), while measurement parameters (16%, category VI), purchase of calibration services (12%, category I), and environmental conditions (5%, category IV) were considered to be less important. The correlation of the causes of unsatisfactory results in proficiency testing with personnel (also called human errors), equipment, calibration, and methods was already reported by Ellison and Hardcastle [14] for testing laboratories in the field of analytical chemistry.

In part 6, the laboratories were asked to point out examples of root causes of unsatisfactory results in proficiency testing. These answers, when grouped into the six categories previously described, validated the results from part 5.

The correlation between the answers to questions from part 3 (items investigated by the laboratories) and part 5 (categories of root causes usually observed by the laboratories) allowed the construction of the cause-and-effect diagram (Ishikawa diagram) shown in Figure 5, which summarizes all root causes that should be investigated by the laboratories in case of nonconforming proficiency testing results.

3.4 Survey conclusions

From the 89 received questionnaires (32.84% of the invited technical managers), the main root causes observed by the laboratories were related to personnel (category V), equipment and calibration standards (category II), and methods (category III). These three main category were reported to be systematically investigated by the Brazilian calibration laboratories in case of unsatisfactory results in proficiency testing activities, although in category II some items are still not deeply investigated, specially personnel authorization to operate equipment, use of updated calibration correction factors, protection of hardware and software, and segregation of equipment out of service.
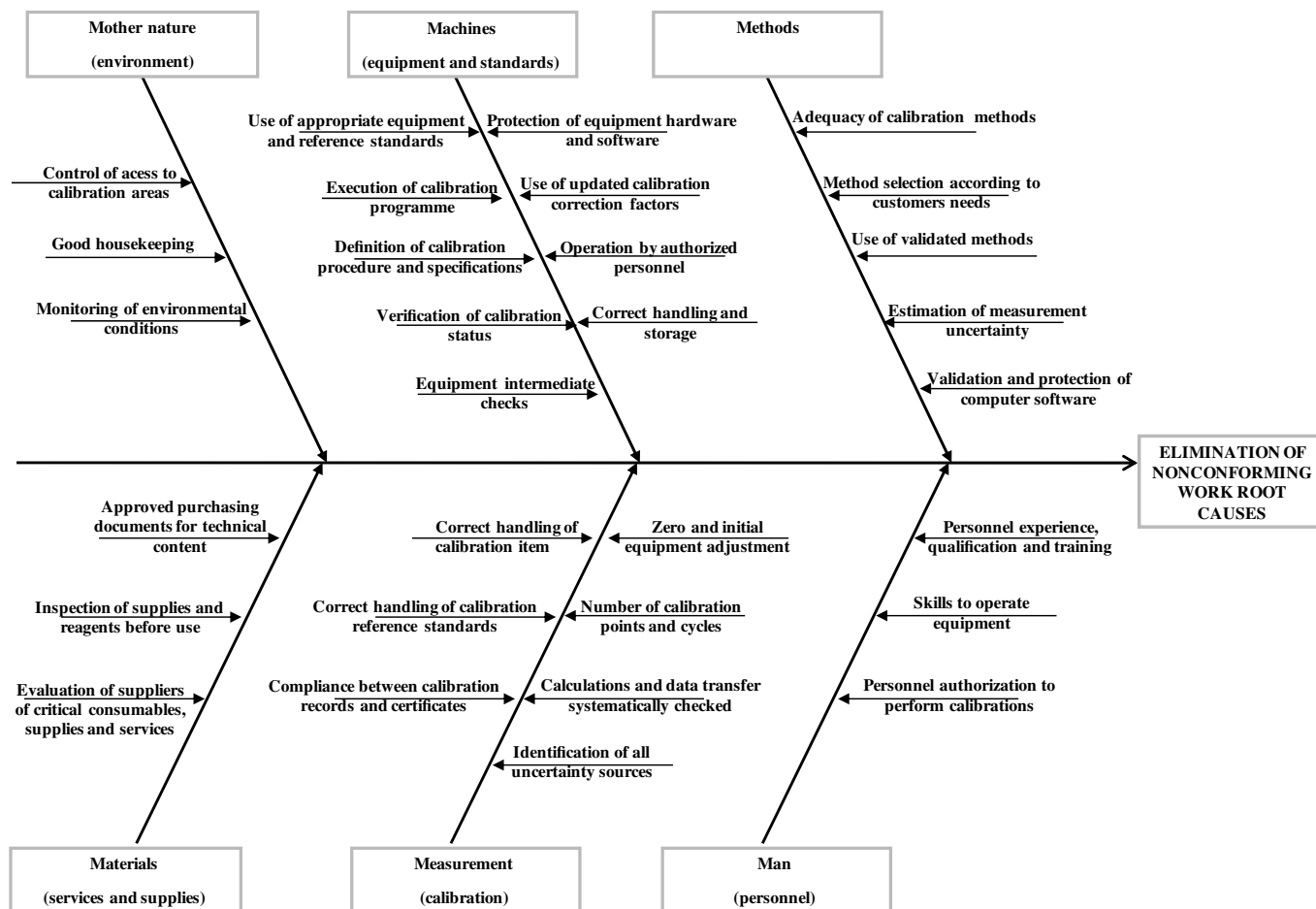
Fig. 5. Cause-and-effect (Ishikawa) diagram of root causes that shall be investigated by calibration laboratories in case of unsatisfactory results in proficiency testing activities.

Measurement parameters (category VI), purchase of calibration services (category I) and environmental conditions (category IV) were considered by the laboratories to be less important factors in case of unsatisfactory results. Concerning the investigation practices of Brazilian calibration laboratories, calibration of reference standards may deserve more attention. Additionally, environmental conditions, which have been strongly considered in the investigations of some laboratories, but are completely neglected by others, may also require attention.

4 Conclusion

A survey was carried out by the Brazilian accreditation body, Cgcre, to evaluate how Brazilian calibration laboratories accredited under ISO/IEC 17025:2005 have been investigating the root causes of unsatisfactory results in proficiency testing activities. The main root causes observed by the laboratories were related to personnel, equipment and calibration standards, and methods. Some items that were considered to need more careful investigation included authorization for equipment operation, use of updated calibration correction factors, protection of hard-

ware and software, segregation of equipment out of service, purchase of calibration services, and adequacy of environmental conditions.

References

1. ISO/IEC 17025, *General Requirements for the Competence of Testing and Calibration Laboratories* (ISO, Geneva, 2005)
2. ISO/IEC 17043, *Conformity Assessment, General Requirements for Proficiency Testing* (ISO, Geneva, 2010)
3. QUAM:2012.P1, Eurachem/Citac Guide CG 4, *Quantifying Uncertainty in Analytical Measurement*, 3rd edn. (Eurachem/Citac working group, London, 2012)
4. S. Bruce, G.F. Strouse, Proficiency testing for achieving accreditation in thermometry, Int. J. Thermophys. **30**, 351 (2009)
5. M. Belli et al., Proficiency testing in analytical chemistry, microbiology and laboratory medicine: working discussions on current practice and future directions, Accred. Qual. Assur. **14**, 507 (2009)
6. B. Brookman et al., Proficiency testing in analytical chemistry, microbiology and laboratory medicine: working discussions on current practice and future directions, Accred. Qual. Assur. **17**, 445 (2012)

7. ILAC-P13:10/2010, *Application of ISO/IEC 17011 for the Accreditation of Proficiency Testing Providers* (ILAC, Silverwater, Australia, 2010), https://www.ilac.org/documents/ILAC_P13_10-2010.pdf

8. IAF/ILAC-A2:01/2013, *Multi-Lateral Mutual Recognition Arrangements: Requirements and Procedures for Evaluation of a Single Accreditation Body* (ILAC, Silverwater, Australia, 2013), http://www.ilac.org/documents/IAF-ILAC_A2_01_2013.pdf

9. J.C. Gust, *A discussion of stability and homogeneity issues in proficiency testing for calibration laboratories* (ILAC, Vienna, 2007), https://www.ilac.org/documents/PTCG_07_8a.pdf

10. ISO 13528, *Statistical Methods for Use in Proficiency Testing by Interlaboratory* Comparisons (ISO, Geneva, 2005)

11. EA-4/02, *Expression of the Uncertainty of Measurement in Calibration* (European co-operation for Accreditation of Laboratories, Paris, 1999), http://www.european-accreditation.org/publication/ea-4-02-m

12. ISO/IEC 17011, *Conformity Assessment, General Requirements for Accreditation Bodies Accrediting Conformity Assessment Bodies* (ISO, Geneva, 2004)

13. ILAC-P9:11/2010, *ILAC Policy for Participation in Proficiency Testing Activities* (ILAC, Silverwater, Australia, 2010), http://www.ilac.org/documents/ILAC_P9_11_2010.pdf

14. S.R.L. Ellison, W.A. Hardcastle, Causes of error in analytical chemistry: results of a web-based survey of proficiency testing participants, Accred. Qual. Assur. **17**, 453 (2012)