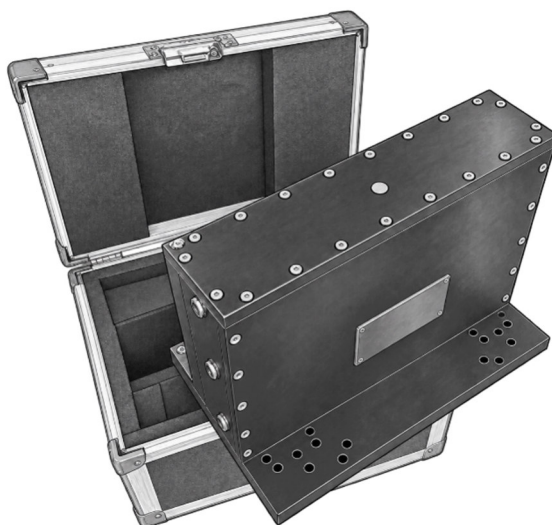


WHITE PAPER

Vibration - Proficiency Testing

Consolidated Lessons Learned — Five Years of Continuous Multi-Round Operation



**Deutsches Referenzbüro für Ringversuche und
Referenzmaterialien GmbH – Kempten, Germany**

<https://drrr.de/en/proficiency-testing/environmental-testing-emc/>

**In cooperation with
4TIERS GmbH – Hamburg, Germany**

Published in October 2026

Executive Summary

This white paper presents the consolidated findings and methodological lessons accumulated across five years of continuous operation of Proficiency Testing for Vibration testing - a comparison testing programme for environmental simulation offered by the Deutsches Referenzbüro für Ringversuche und Referenzmaterialien GmbH (DRRR GmbH) in cooperation with 4TIERS GmbH. The programme runs multiple parallel rounds per year, covering sinusoidal vibration (IEC 60068-2-6) and broadband random vibration (IEC 60068-2-64), and has by now accumulated a substantial cross-laboratory dataset spanning a few hundred of individual test executions across a broad and growing participant base. The participants in these proficiency tests come from various countries, all of which are located in Europe for organizational reasons.

The comparison test is designed to mirror commercially realistic test scenarios: one characterised specimen is subjected to defined vibration environments at independent laboratories, enabling direct quantitative assessment of result comparability. The programme serves a dual function — providing individual laboratories with peer-referenced performance benchmarks across repeated rounds while generating a growing body of evidence-based findings that directly inform the ongoing revision of internationally recognised test standards, including IEC 60068-2-64.

However, the program is explicitly and mainly designed to compare laboratory performance under realistic test conditions — providing participants with meaningful peer-referenced benchmarks for their methods, equipment, and execution. The focus here lies specifically on the use of the program within the framework of accreditation procedures under ISO/IEC 17025; accordingly, the emphasis is on laboratory assessment, reflecting the current technical level that participating laboratories are actually able to deliver. Contributions to test standard improvement are a welcome by-product of the accumulated data but are not a primary objective: advancing standardization in a rigorous sense would require dedicated measurement campaigns with additional reference instrumentation, tighter experimental controls, and substantially greater laboratory effort than is practical or commercially justifiable within a participation-fee-based proficiency testing scheme.

Key lessons consolidated from five years of programme data encompass: (i) systematic effects of specimen fixation methodology on control-loop fidelity; (ii) the need to define frequency resolution with a minimum Hz-per-line criterion, whereas we would encounter a consequent and repeatable result divergence between laboratories using, i.e. 1024, 3200, or other common line counts; (iii) sensor terminology inconsistencies between English and German language implementations; (iv) the prevalence of sensor-sensitivity unit errors detectable via measuring wedge verification; and (v) directivity ambiguities in reported measurement values. These findings are presented with supporting data, graphical evidence, and actionable recommendations directed at both test laboratories and the standardisation community.

The primary objective of this scheme is to provide participating laboratories with quantitative, peer-referenced benchmarking data — not to rank laboratory performance, but to facilitate targeted process improvement and informed investment in test infrastructure.

Programme Context: Accredited Proficiency Testing Framework

Institutional Background

The DRRR GmbH - Deutsches Referenzbüro für Ringversuche und Referenzmaterialien GmbH - is an ISO/IEC 17043 accredited proficiency testing provider and a leading supplier of certified reference materials worldwide. With more than 15 years of operation and over 700 proficiency tests conducted annually, DRRR is one of the largest providers of external quality assurance services worldwide, combining deep statistical expertise with domain-specific technical competence across a wide range of testing disciplines.

The environmental simulation proficiency testing programme — developed in cooperation with 4TIERS GmbH and launched in 2022 — has grown into a mature, multi-round programme now in its fifth year of continuous operation. Multiple parallel rounds are conducted each year, covering different test types and specimen configurations. Full programme details, the current schedule, and the registration catalogue are available at drrr.de/en/proficiency-testing/environmental-testing-emc/. The programme was conceived to close a long-standing gap in the accreditation landscape — prior to 2022, accredited environmental simulation laboratories had no dedicated, sector-specific proficiency testing scheme available for demonstrating ongoing measurement competence — and has since established itself as the primary external benchmarking reference for this sector in the DACH region and many European countries.

2026 Programme Scope - Environmental Simulation (DRRR / 4TIERS)

Environmental Testing:

- Vibration (IEC 60068-2-6 / -64)
- Shock (IEC 60068-2-27)
- Temperature Change (IEC 60068-2-14)
- Damp Heat (IEC 60068-2-30)
- Cold Testing (IEC 60068-2-1)

Registration: odin.drrr.de — DRRR Online Catalogue (ODIN)

Relevance for Accredited Laboratories

For laboratories holding or seeking accreditation under ISO/IEC 17025, participation in a recognised proficiency testing scheme is a mandatory element of the quality management system. In the environmental simulation sector — where the consequences of inadequate test severity range from undetected qualification failures to unwanted component rejections — the availability of a purpose-built, technically rigorous comparison programme represents a material advance in the accreditation infrastructure for this discipline.

Laboratories are strongly encouraged to participate in one or more DRRR vibration comparison rounds per year. With several parallel rounds available annually — covering vibration, shock, climatisation — laboratories can select participation rounds that best match their current test portfolio

and improvement priorities. Participation fulfils the ISO/IEC 17025 proficiency testing requirement while simultaneously providing a longitudinal diagnostic instrument for tracking performance trends across rounds, identifying systematic deviations that would not be detectable through internal calibration or uncertainty budgeting alone.

The random and sinusoidal vibration comparison test (IEC 60068-2-64 resp. IEC 60068-2-6) is described in detail at [4tiers.de](https://www.4tiers.de) — [Ringversuch Schwingung und Vibration IEC 60068-2-6](https://www.4tiers.de).

Test Design Philosophy

Concept and Scope

The comparison test addresses reproducibility and measurement accuracy under realistic specimen–machine interaction conditions. Rather than employing a passive, dynamically inert artefact, the programme uses a purposefully designed reference specimen whose mechanical resonances interact with the shaker system in a manner representative of typical electro-technical, mechatronic or industrial components.

For each proficiency testing round, a single reference specimen is used and circulated sequentially among all participating laboratories, rather than distributing separate specimens to each laboratory. This specimen is regularly checked for stability throughout the round to ensure that its mechanical properties remain unchanged over the course of circulation, so that any differences observed between laboratories can be attributed to the laboratories' measurement performance rather than to variation in the specimen itself. In order to keep the overall duration of a round within a reasonable timeframe, the number of participants is limited to 15–20 laboratories per round.

The framework addresses the following fundamental industrial question: when a test client dispatches nominally identical specimens to multiple independent laboratories, or when a laboratory repeats a defined test sequence in different configurations, how comparable are the resulting environmental excitations applied to the specimen? This question determines the degree of confidence that can be placed in qualification decisions based on test results from different sources.

Statistical data analysis

The statistical evaluation of the measurement data is carried out in accordance with the requirements of ISO/IEC 17043, together with the relevant test standards referenced therein. The assigned value and the standard deviation for proficiency assessment (SDPA) are calculated using the consensus value approach, i.e. derived from the data set contributed by the participating laboratories themselves. In this process, particular characteristics of the data — such as the logarithmic nature of the measurement values — are taken into account accordingly.

Specimen Design and Measurement Configuration

The reference specimen introduces controlled mechanical perturbations in both the principal excitation axis and the transverse axes, analogous to the resonance behaviour of real electro-mechanical components. The following quantities are assessed:

- Principal-axis acceleration at the control point (EN: reference point for control / DE: Bezugspunkt, Regelpunkt)
- Transverse-axis acceleration characterising cross-talk and parasitic loading
- Selected characteristic resonance frequencies of the specimen, which should not be excited by the shaker
- Coarse displacement verification at a defined frequency via the measuring wedge (Messkeil)

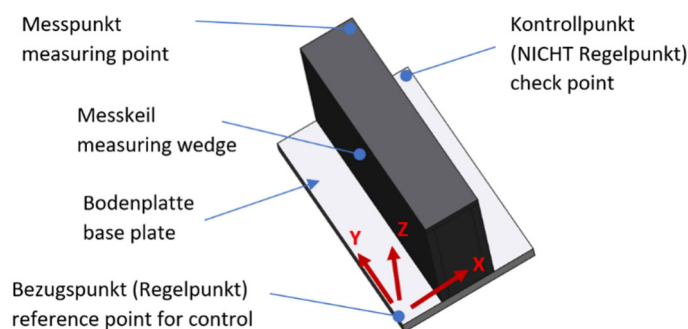


Figure 1 — Reference specimen with reference point for control (Bezugspunkt/Regelpunkt), check point (Kontrollpunkt), measuring wedge (Messkeil), and coordinate axis orientation. This diagram is provided bilingually in the test documentation to prevent terminology-related setup errors.

Technical Findings

4.1 Principal-Axis Excitation Fidelity

The qualification of principal-axis motion and the assessment of specimen-induced perturbation sensitivity constitute the primary measurement objective. The central question is: to what extent does the actual environmental stimulus applied to the specimen, under comparable specimen-induced disturbance, deviate from the target profile, and how consistently is this deviation reproduced across laboratories?

A Fatigue Damage Spectrum (FDS) analysis would represent the most rigorous basis for comparing the damage potential of applied stimuli across participants. However, FDS computation requires explicit assumptions regarding the failure mechanism of the specimen and necessitates dedicated, independently supplied measurement instrumentation to ensure comparability. To maintain an accessible entry point and limit participant cost burden in the early programme rounds, FDS analysis was deferred to a planned advanced variant of the scheme currently under development. Coarse displacement verification at a defined frequency has successfully served as a first-order consistency check without additional hardware requirements instead ever since the start.

4.2 Shaker size and fixtures

Data from several measurements in different labs shows that even shakers typically considered very small for such a large and heavy specimen (app. 25 kg) show good and well comparable results. With the force vector enquired in every round together with the results, it can be seen that shakers down to 18kN (stationary) force vector give results well comparable to the best mid-range shakers. From own measurement series it is assumed that well designed test fixtures, correct fixation and well done test execution are all necessary to reach good results even with small shakers. It must be noted that large shakers (>100kN) typically have lower main axial resonance modes, especially with large and heavy head expanders. These lowered main resonances will deteriorate the results with potentially rather high cross-axis motion and maybe distinctive local deviations in excitation axis acceleration.

The following figure describes one selected set of results from one round in one of the years since 2022 with the rounded force vector data of the laboratories taking part.

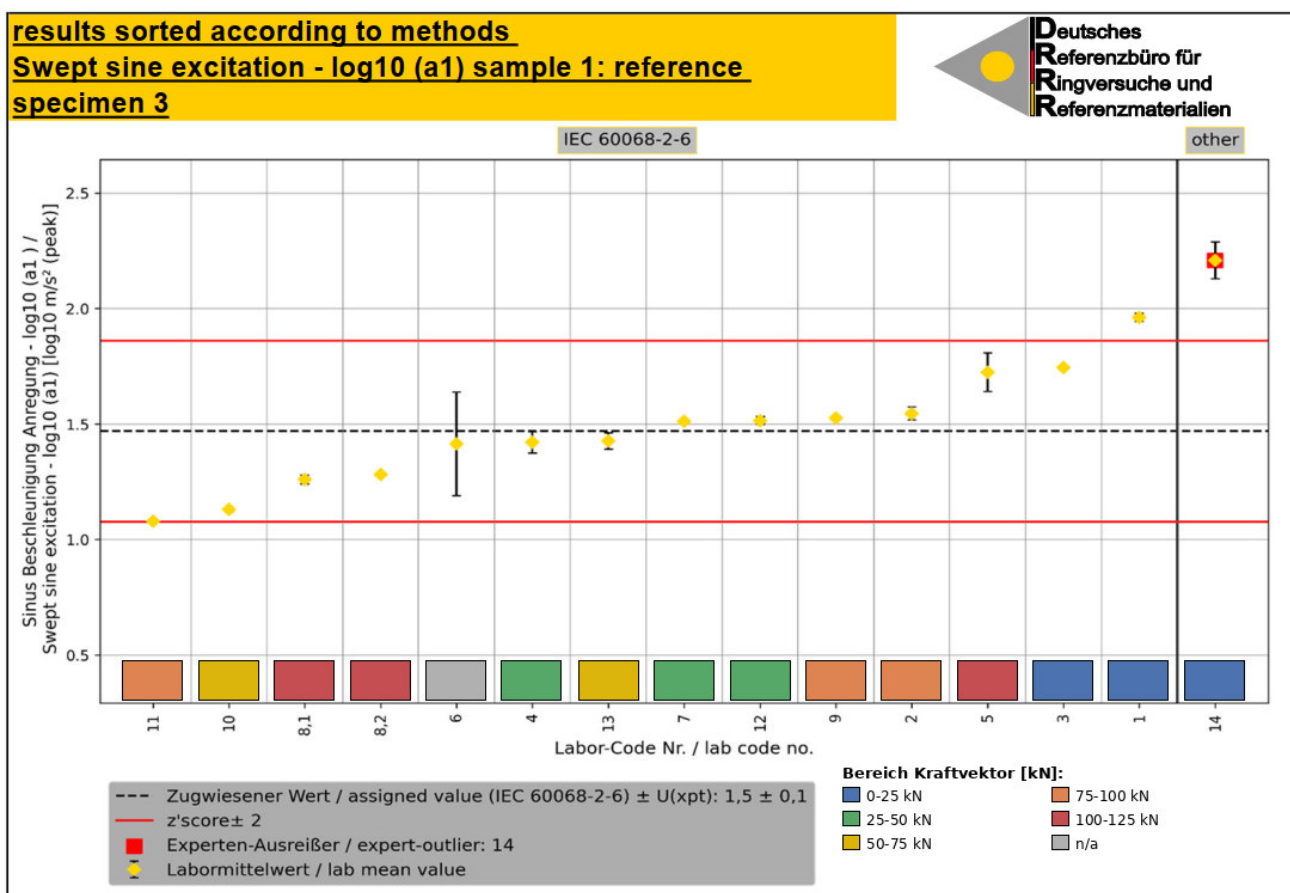


Figure 2 — deviations in main axis excitation for participating shaker force vector ranges

It can be seen that rather small shakers for the given weight and size of the specimen tend to be yield rather high deviations from the average of the data set. However, one result of a <25kN shaker is comparable to a >100kN shaker and well within the statistical limit. With all participating shaker sizes from less than <25kN up to 125kN, well agreeing results within the data set can be achieved. This is also substantiated from other round over several years. Nearly identical data for the given measurement task regarding main axis excitation deviation is generated in the range from 25kN to 100kN. The range of 100kN – 125kN provides slightly higher as well as lower deviations from the average of the data set, but both in a statistically acceptable manner.

The given test specimen with its characteristic dynamic behaviour, size and weight surely limits the range of shakers with which comparable results can be achieved, but this range seems to be very

broad – given the test is executed well. The limitation to a certain shaker size applies also to any given test specimen a customer might send to the laboratory, which then has to decide on the right shaker size to test the specimen with. Hence, the test shows, that a very broad range of shaker sizes can yield comparable results but that technical limitations below 25kN and likewise somewhere above 125kN may exist.

4.3 Specimen Fixation: Clamping versus Screwed Connection, using sufficient number of screws

A recurring and consequential source of inter-laboratory variation, consistently observed across multiple rounds over five years of programme operation, is the method employed to secure the specimen to the shaker table. Despite explicit guidance in the test specification recommending screwed fixation for the relevant frequency ranges and specimen mass, a small subset of participating laboratories has continued to employ clamping as the primary attachment method.

The mechanical consequence is directly visible in the drive signal. The effective stiffness of the coupled multi-resonator system — comprising the shaker armature, fixture, and specimen — is modified by clamping compliance. This introduces additional resonance modes and alters existing ones, yielding control-loop responses that differ systematically from those obtained with rigid screwed connections. The effect is particularly pronounced in broadband random and sinusoidal sweep tests spanning specimen resonance frequencies, where small changes in structural coupling produce large deviations in transmitted acceleration.

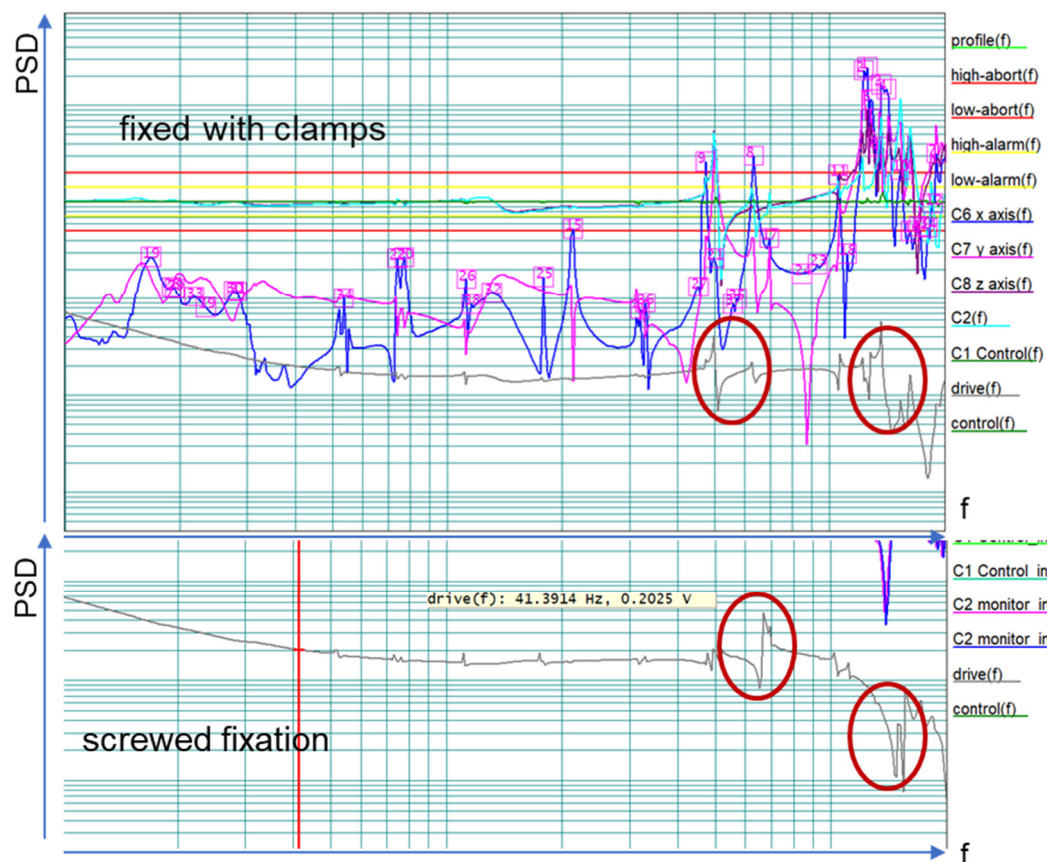


Figure 2 — Power spectral density (PSD) drive and control signals comparing clamped fixation (upper panel) against screwed fixation (lower panel). Circled regions highlight frequency bands where the control signal deviates from the target profile due to resonance effects introduced by inadequate fixation compliance. The screwed configuration demonstrates substantially improved control fidelity.

The practical consequence is evident: laboratories employing clamping report results that are not directly comparable to those of the screwed-fixation majority, even when all other test parameters are nominally identical. This finding reinforces the recommendations of IEC 60068-2-47 and underscores the need for explicit, unambiguous fixation requirements in comparison test specifications.



Figure 3 — Reference specimen mounted on a vibration test system with clamp-type fixation, illustrating a configuration identified as a primary source of inter-laboratory result divergence. The compliance at the fixture–table interface directly affects control-loop behaviour at specimen resonance frequencies.

Furthermore, feedback from laboratories and external experts have reported similar results and specimen behaviour for head expanders or fixtures with less screws in use than intended by the design of the fixation devices. With only a few screws, high frequency resonance may occur within the excitation range with the screws being the most yielding part in the load transfer path, essentially acting as springs. Most professionally designed head expanders and fixtures provide so many fastenings screw fixations that the respective resonance of the mass fixture with the screws acting as springs is sufficiently higher (typically higher than 4kHz) for most typical shaker testing applications.

4.4 Frequency Resolution in Broadband Random Testing

One of the most practically significant and consistently reproduced findings across all programme rounds concerns the effect of spectral frequency resolution on broadband random PSD results. The root cause traces to an underspecified requirement in IEC 60068-2-64, which in the version applicable at the time prescribed only a minimum of 200 spectral lines across the analysis bandwidth.

This 200-line floor is largely irrelevant in practice: it is routinely and substantially exceeded by all modern vibration controllers on the market. A survey of participating laboratory configurations found that typical implementations employ approximately 1024 lines, with others operating at i.e. 2048, 3200, or higher counts depending on controller make, preset, model, and operator configuration. The real problem is therefore not whether the 200-line minimum is met — it always is. The problem is the absence of a harmonised upper resolution requirement or a defined Hz-per-line criterion. Two laboratories, both fully compliant with the standard and both using calibrated, high-quality controllers, can operate at fundamentally different spectral resolutions and will report systematically different results even when the underlying vibration environment is nominally identical.

Numerical example: A 5 Hz–2000 Hz analysis range at 1024 lines yields approximately 1.95 Hz/line; at 3200 lines, the same range resolves to approximately 0.62 Hz/line. For a specimen resonance with a half-power bandwidth of 2 Hz — typical of lightly damped mechanical assemblies — the 1024-line configuration places only one spectral point within the resonance, while the 3200-line configuration resolves it with three or more. The resulting difference in peak amplitude characterisation and in reported RMS levels within the resonance band is measurable, repeatable, and entirely attributable to controller configuration — not to any real difference in test severity.

This finding has direct implications for inter-laboratory comparability: systematic PSD result differences between laboratories are produced not by test execution errors or equipment deficiencies, but by a configuration variable that the current standard leaves unharmonized. In response, the test specification was revised to mandate a specific frequency resolution in Hz per line, decoupled from absolute line count. This ensures consistent spectral characterisation between participating laboratories and is common practice for well-defined test specifications, especially in high quality test applications.

The forthcoming revision of IEC 60068-2-64 is expected to address this at the standard level by replacing the obsolete 200-line minimum with a criterion requiring a minimum number of spectral points per resonance bandwidth where necessary. Also, more emphasis is put on precise test specifications.

Hence, future comparison test specifications will increase the required line resolution, so that also lower damped resonances causing peak values will be measured more precisely, while with the current rather broad industry standard definition, only a value comparable value can be derived.

4.5 Transverse-Axis (Cross-Talk) Assessment

The qualification of transverse-axis motion constitutes the second principal measurement objective. Parasitic accelerations orthogonal to the nominal excitation direction can, under adverse conditions, reach a substantial fraction — or in some configurations a multiple — of the nominal test level. These loads represent an uncontrolled source of test severity variation that can significantly affect qualification outcomes and invalidate direct inter-laboratory comparisons.

The cross-talk metric presents a fundamental statistical challenge: its magnitude depends strongly on the interaction between the specific specimen and the specific shaker system, fixture geometry, and mounting configuration at each laboratory. Unlike principal-axis deviation, which is referenced against a defined target profile, cross-talk levels do not converge to a single consensus value — they are inherently site-specific measurements reflecting the particular mechanical coupling at each installation.

The scheme addresses this through a global RMS comparison in broadband random testing and in sinusoidal testing. The specification has evolved over successive rounds to prevent the reuse of historically "good" values as a performance proxy, recognising that favourable cross-talk results in a previous round may simply reflect a fortuitous specimen–shaker interaction that no longer applies. This evolutionary refinement — made possible by five years of continuous, multi-round operation — is itself a demonstration of the programme's maturity. Full statistical analysis analogous to that applied to principal-axis results remains an area of active methodological development for upcoming rounds.

4.5 Sensor Terminology: Bilingual Consistency

A recurring source of reporting inconsistency involves the assignment of sensor roles — in particular the distinction between the control point and the check point:

- Reference point for control (EN) = Bezugspunkt / Regelpunkt (DE): the primary feedback sensor regulating the drive signal to the target spectrum
- Check point (EN) = Kontrollpunkt (DE): an independent monitoring sensor verifying response at a secondary location — explicitly NOT a control input

Confusion between these roles results in fundamentally non-equivalent test executions that in real life cannot be compared even when all nominal parameters are identical. Over several years of active participation and continuing quality analysis of the round robin test results, the bilingual specimen diagram (Figure 1) has been incorporated as a reference into the test specification to eliminate this source of ambiguity since the beginning of the test programme. It has been continuously improved to maximize clarification and guidance to the user. Yet, we still receive comments that question the understanding of the correct definition.

4.6 Measuring Wedge Verification and Unit-of-Measure Errors

The measuring wedge (Messkeil) provides an independent coarse verification of sensor sensitivity entered into the data acquisition system. Its primary function is detecting unit-of-measure errors — specifically the confusion between g (standard gravity, $\approx 9.81 \text{ m/s}^2$) and m/s^2 — which produce a systematic scaling offset of approximately one order of magnitude in reported acceleration values, constituting a gross mis-test condition. Instances of this error have been identified across multiple programme rounds, validating the utility of this low-cost verification step prior to test commencement. Its inclusion in the test specification has remained a fixed element since the earliest rounds.

4.7 Sign Convention and Directivity of Reported Values

Sinusoidal and broadband random vibration test results — whether reported as PSD, zero-to-peak (0-pk) amplitudes or root-mean-square (RMS) — are inherently non-negative scalar quantities derived from quadratically averaged, stationary stochastic processes. Directivity carries physical significance in shock testing; it does not in steady-state vibration testing. Nevertheless, submissions of negative values for principal-axis and cross-axis results have been observed across several programme rounds, i.e. when asking for RMS or PSD values.

Interpretation of Performance Scores

For each measurement task in the test specification, a theoretical ideal value exists, determined by the physical conditions of the experiment and the target parameters. In most cases, the consensus value derived from participant data will not coincide with this theoretical ideal — it will typically exceed it, reflecting the systematic influence of specimen resonance interactions on the actual stimulus applied.

The z-score metric quantifies the standardised deviation of an individual result from the population consensus. The following interpretive principles are essential for correct application:

- An unusually low z-score — placing a result closer to the theoretical ideal than the consensus — may reflect exceptionally good test execution, but may equally represent a statistical artefact. Such results require critical examination in conjunction with cross-talk and resonance frequency metrics.
- Both high and low z-score magnitudes indicate reduced comparability with peer results and are a strong indicator of a divergent methodology at the reporting laboratory — irrespective of the absolute accuracy of the result.
- A large deviation from the consensus is not automatically a negative finding. A laboratory able to provide a technically explanation for its deviation — for example, deliberately characterising the performance of an undersized shaker or a simplified test setup — derives operationally valuable diagnostic information from that result.

The comparison test is a diagnostic instrument, not a pass/fail certification. Its value lies in understanding the sources of deviation and applying that understanding to improve test processes, equipment selection, and result interpretation.

Evolution of the Remarks and Guidance System

Each test report issued to participating laboratories includes a structured remarks section that contextualises statistical results and provides actionable guidance. The content and structure of these remarks has been substantially refined across five years of programme rounds, incorporating feedback from both participants and independent subject-matter experts accumulated through dozens of completed rounds.

The remarks system serves two functions. First, it enables laboratories to distinguish between deviations attributable to methodological choices — such as fixation method or resolution settings — and those attributable to equipment limitations or execution errors. Second, it communicates developments in test standardisation and best practice directly to the participant population. The longitudinal nature of the programme means that improvements introduced into the remarks in one round can be traced to measurable shifts in the result distribution of subsequent rounds, providing direct feedback on their effectiveness.

Conclusions and Recommendations

For Participating Laboratories

- Specimen fixation must employ screwed connections for all broadband random and swept-sine tests in the relevant frequency range. Clamping introduces uncontrolled compliance that degrades control and result comparability. If needed, well fixed adaptation plates bridging the screw patterns from specimen to fixture are to be utilized.
- Head expanders and fixtures should be used with all screw connection as intended by the manufacturer, all tightened to the given nominal values of the manufacturer or respective standard
- Frequency resolution for broadband random PSD analysis must be specified in the test specification depending on the test purpose, i.e. in Hz per line, not as a minimum line count — see also future revisions of IEC 60068-2-64. Modern controllers routinely exceed the 200-line standard minimum by a large margin — 1024 lines is a common default, with many more in widespread use — but the absence of a harmonised resolution criterion means different configurations produce systematically different results. Resolution setting must be documented as an explicit test parameter.
- Sensor role assignments — particularly control point versus check point — must be verified against the bilingual specimen diagrams provided in the test specification before each test execution.
- Sensor sensitivity entry must be verified via the measuring wedge procedure or other references before data collection commences, to detect unit-of-measure errors.
- All reported acceleration values for sinusoidal and broadband random tests must be non-negative scalars (0-pk or RMS) if not specified otherwise in time-domain tests, i.e. for shock testing.
- Participation in the DRRR proficiency testing programme provides accredited laboratories with formal evidence of ongoing measurement competence as required under ISO/IEC 17025. Registration: odin.drrr.de.

For Standardisation Bodies

- The revision of IEC 60068-2-64 should replace the minimum 200-line requirement with a resolution criterion specifying a minimum number of spectral points per resonance bandwidth — for example, five points per half-power bandwidth for resonances of interest where no alternative is specified. The 200-line floor is practically irrelevant given typical controller capabilities; the absence of a harmonised Hz-per-line criterion is the actual source of inter-laboratory divergence documented in this programme.
- Test standards should provide explicit guidance on specimen fixation methods as a function of specimen mass, frequency range, and expected resonance behaviour. Despite the normative reference to IEC 60068-2-47 in IEC 60068-2-6 and IEC 60068-2-64, its provisions on fixation method are evidently not being followed in practice by a small subset of laboratories — suggesting the issue is one of awareness and enforcement rather than a gap in the standard's structure.

- Bilingual terminology tables — such as the EN/DE correspondences established in the DRRR / 4TIERS programme — should be incorporated into the normative body of international standards to prevent systematic translation-induced setup errors.

The DRRR Vibration Comparison Test is now in its fifth year of continuous operation, with multiple parallel rounds conducted per year. The programme continues to evolve: FDS-based analysis and expanded cross-talk statistical methods are in development for future rounds. Participation enquiries and registration: drrr.de | odin.drrr.de.

References and Programme Links

1. IEC 60068-2-6:2007 — Environmental testing — Part 2-6: Tests — Test Fc: Vibration (sinusoidal).
2. IEC 60068-2-64:2008+AMD1:2019 — Environmental testing — Part 2-64: Tests — Test Fh: Vibration, broadband random and guidance.
3. IEC 60068-2-47:2005 — Environmental testing — Part 2-47: Mounting of specimens for vibration, impact and similar dynamic tests.
4. ISO/IEC 17025:2017 — General requirements for the competence of testing and calibration laboratories.
5. ISO/IEC 17043:2023 — Conformity assessment — General requirements for the competence of proficiency testing providers.
6. ISO 13528:2022 — Statistical methods for use in proficiency testing by interlaboratory comparison.
7. DRRR GmbH — Proficiency Testing: Environmental Testing drrr.de/en/proficiency-testing/environmental-testing-emc/
8. 4TIERS GmbH — Ringversuch Schwingung und Vibration IEC 60068-2-6. 4tiers.de/spezialisten-ringversuche-umweltsimulation/.../ringversuch-schwingung-und-vibration-iec60068-2-6/
9. DRRR Programme Test Specification: Vibration Comparison Test. Internal technical document, updated per round (available to registered participants via ODIN).

For participation enquiries and registration:

drrr.de | odin.drrr.de | 4tiers.de

DRRR GmbH | +49 (0) 831 960878-0 | info@DRRR.de