Measurement uncertainty of ISO CD 23351-2: acoustic performance of furniture ensembles and enclosures in situ

Applied Acoustics 238 (2025) 110773

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Measurement uncertainty of ISO CD 23351-2: acoustic performance of furniture ensembles and enclosures in situ

Valtteri Hongisto [a,*], Riccardo Caradonna [b], Jukka Keränen [a], Jarkko Hakala [a], Reijo Alakoivu [a], Arianna Astolfi [b]

[a] *Turku University of Applied Sciences, Built Environment, Acoustics, Joukahaisenkatu 3-5, FI-20520 Turku, Finland*
[b] *Politecnico di Torino, Department of Energy, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*

A R T I C L E I N F O

Keywords:
Round Robin test
Measurement uncertainty
Speech level difference
Pods
Furniture ensembles
Booths

A B S T R A C T

Enclosures (E), such as mobile pods or rooms, and partial enclosures (PE), such as sofa groups and open pods, are increasingly used in indoor environments where occupants temporarily need better speech privacy or peace of work. Such objects are usually tested in laboratory using ISO 23351-1. That standard excludes very large objects. In addition, acousticians need to measure objects in situ. The field method, ISO CD 23351-2 is based on sound power level tests according to ISO 3744 standard. The main outcome is apparent speech level difference, $D'_{S,A}$ [dB]. Since the method of ISO CD 23351-2 is novel, information about measurement uncertainty (MU) is lacking. Our purpose was to determine the uncertainty by conducting two independent studies: 1. between-laboratory study and 2. between-room study. In study 1, thirteen participants tested one PE and one E in the same room. Based on eight accepted participants, the between-laboratory standard deviations (SD) of $D'_{S,A}$ were 0.8 dB for PE and 0.7 dB for E. In study 2, a single participant tested one PE and one E in ten different rooms. The between-room SDs of $D'_{S,A}$ were 0.6 dB both for PE and E. The total SD was 1.0 dB and it considers the MUs both due to operator and room choice. The results can be useful in the finishing of ISO 23351-2 standard. Our study also provides data regarding the MU of ISO 3744 and ISO 3382-2 standards.

1. Introduction

1.1. Background

International standard ISO 23351-1 [1] is increasingly used to test the speech level reduction of enclosures (e.g., mobile workrooms, phone booths, meeting pods) and partial enclosures (furniture ensembles which are not totally enclosed). These will be called objects from now on.

The test shall be conducted in a laboratory environment (reverberation room). The method yields a single-number quantity, speech level reduction, $D_{S,A}$ [dB], which describes how much the A-weighted sound power level (SWL) radiated by a speaker (producing standard effort speech) is reduced after moving inside the object. The theory of the method is based on Hongisto et al. [2]. The measurement uncertainty of ISO 23351-1 was reported by Hongisto et al. [3] to be approximately 1.1 dB for enclosure (phone booth) and 0.6 dB for partial enclosure (open workstation).

After the publication of the ISO 23351-1 standard, it has become evident that the standard is not sufficient to serve many purposes of testing. These are listed below:

- ISO 23351-1 shall not be applied for products larger than 5 % of room volume. The typical reverberation room volume is 200–300 m^3, so the upper limit is 10–15 m^3. Examples of products exceeding this limit are meeting pods for more than 6 people. There are many such products on the market, and they should be tested in another way to avoid the violation of ISO 23351-1.
- Product developers of furniture manufacturers usually do not have reverberation rooms that meet the requirements of ISO 3741 [4] and ISO 23351-1 [1]. Prototypes are often tested several times before the final test is conducted. Manufacturers would benefit from a test method that can be used in any kind of room and that gives reasonably similar values as ISO 23351-1.
- Facility managers and buyers should have a right and means to verify that the purchased product fulfils the $D_{S,A}$ value declared in the

* Corresponding author at: Turku University of Applied Sciences, Built Environment, Joukahaisenkatu 3-5, FI-20520 Turku, Finland.
E-mail address: valtteri.hongisto@turkuamk.fi (V. Hongisto).

https://doi.org/10.1016/j.apacoust.2025.110773
Received 29 January 2025; Received in revised form 31 March 2025; Accepted 23 April 2025
Available online 12 May 2025

specifications. Therefore, acoustic consultants should be able to measure the product in field conditions using a test method that gives reasonably close value to ISO 23351-1.

These needs were anticipated already in 2018, when the ISO 23351 process was started. Therefore, the standard was finally named ISO 23351-1 (laboratory method) to give the option for a later field method. Finland proposed to ISO in 2022 that a field test method (ISO 23351-2) should be developed. This proposal was accepted in an international ballot in spring 2023. Finland accepted to take the lead of the working group WG 34 ("Speech attenuation") in ISO TC 43 SC 2. WG 34 started to prepare the committee draft (ISO CD 23351-2) in June 2023. Keränen et al. [5] studied the suitability of three alternative standardized field test methods of SWL. Based on their work, ISO 3744 [4] was chosen in WG34 as the fundamental method for measuring the SWL. Finally, ISO CD 23351-2 [6] was submitted for an international ballot in February 2024. It lacked scientific evidence about the measurement uncertainty of the test method. Our work attempts to fill this gap in knowledge.

1.2. Description of ISO CD 23351-2 method

The apparent level reduction, D' [dB], provided by the object is determined by

$$D' = L_{W,P,1} - L_{W,P,2} \tag{1}$$

where $L_{W,P,1}$ [dB] is the SWL of the sound source producing broadband noise (P) in Scenario 1 (without object), and $L_{W,P,2}$ [dB] is the SWL of the same sound source (with the same spectrum and output level) in Scenario 2, when it is placed inside the object.

In Scenario 2, the sound source is placed in the most probable location of the user. If the object is a partial enclosure, an artificial mouth (a sound source having similar directivity pattern as human mouth during speaking) shall be used as a sound source, because the directive properties of speech may coincide with the directive properties of the object (due to openings in the envelope). If the object is an enclosure, an omnidirectional loudspeaker shall be used as a sound source, because the directive properties of the object are expected to be smaller (no openings in the envelope). In addition, the SWL of artificial mouth is usually insufficient for the testing of high performing objects.

The SWLs are determined in both Scenarios by ISO 3744 standard [4] according to which the L_W [dB re 1 pW] is determined by

$$L_W = L_{p,S} - K_2 + 10 \bullet \log_{10}(S) \tag{2}$$

where $L_{p,S}$ [dB re 20 μPa] is background noise corrected energy average of sound pressure level, SPL, in the measurement surface, and S [m^2] is the measurement area surrounding the object.

In ISO CD 23351-2, the measurement surface consists of four vertical sub-surfaces reaching from floor to ceiling if the room height is 0.5–1.5 larger than the object height (Option A). If the room height is larger than that (Option B), the measurement surface is rectangular consisting of five sub-surfaces (four vertical sub-surfaces, one horizontal sub-surface above the object). In all cases, the sub-surface is positioned at 0.50 m from the outer surface of the object.

The environmental correction, K_2, is determined by

$$K_2 = 10 \bullet \log_{10}\left(1 + \frac{4 \bullet S}{A}\right) \tag{3}$$

The absorption area of the room, A [m^2], where the object is located, is determined by

$$A = \frac{0.16 \bullet V}{T} \tag{4}$$

where V [m^3] is the volume of the room and T [s] is the reverberation time of the room measured by ISO 3382-2 [7].

Eq. (4) (Sabine's formula) is valid in diffuse sound fields, which presumes that the SPL does not strongly depend on the measurement position. The formula may lead to errors in strongly non-diffuse rooms. However, ISO CD 23351-2 contains two repeated SWL tests in the same room so that the potential error of this formula is "canceled out". Anyhow, our between-room experiment (Sec. 3) involves both diffuse and non-diffuse rooms to determine, in quantitative terms, how much the deviation from diffuse field influences the test result.

The SWL radiated by an object, which contains a sound source emitting standard effort speech, is determined by

$$L_{W,S,2} = L_{W,S,1} - D' \tag{5}$$

where $L_{W,S,1}$ [dB] is the SWL of standard effort speech (Table 1).

The apparent speech level reduction, $D'_{S,A}$ [dB], is defined by:

$$D'_{S,A} = L_{W,S,A,1} - L_{W,S,A,2} \tag{6}$$

where subindex A refers to A-weighting within octave bands 125–8000 Hz.

SPL in the measurement surface is the sum of direct sound energy (also known as emission SPL, $L_{p,d}$) and reverberant sound energy (also known as diffuse field SPL, $L_{p,r}$). $L_{p,d}$ increases when the distance to the test object decreases. $L_{p,r}$ increases, when the distance to the test object increases, and room absorption area A reduces. The correction K_2 is a theoretical estimation of $L_{p,r} - L_{p,d}$.

Because there is not much research about the influence of K_2 on ISO 3744, and ISO 23351-2 should be available in any kind of room, the investigation of the effect of K_2 on L_w forms an essential part of our study.

1.3. Purpose of the study

There is limited experimental research about the measurement uncertainty of ISO CD 23351-2. In the study of Keränen et al. [5], one person tested one booth in six acoustically different rooms and found a between-site standard deviation (SD) of 0.4 dB for $D'_{S,A}$. The value was surprisingly small although the K_2 range was very large (0.6–12.0 dB) between the six rooms. They did not investigate between-operator differences. Therefore, there is a need to conduct a between-laboratory experiment to see the influence of different operators.

The measurement uncertainty of ISO 3744 mostly depends on two sources:

1. Between-laboratory differences – i.e., differences caused by different operators, measurement procedures, and devices.
2. Between-room differences – ISO 3744 can be applied in different rooms, but the room size and reverberation time may affect the result.

The driving purpose of our study was to determine the measurement uncertainty of ISO CD 23351-2 by elaborating the two abovementioned sources of uncertainty. The between-laboratory experiment is presented on Sec. 2. The between-room experiment is presented on Sec. 3. The summary of them is presented on Sec. 4.

Since ISO CD 23351-2 is based on two other standards, ISO 3744 (sound power level) and ISO 3382-2 (reverberation time), our secondary purpose was to provide information about their measurement uncertainties.

Table 1
Standard effort speech according to ISO [1] and ISO [4]. The A-weighted overall value is $L_{W,S,A,1} = 68.4$ dB.

f [Hz]	125	250	500	1000	2000	4000	8000
$L_{W,S,1}$ [dB]	60.9	65.3	69.0	63.0	55.8	49.8	44.5

2. Between-laboratory experiment

2.1. Materials and methods

2.1.1. Participants and instructions

The between-laboratory experiment was planned according to the recommendations of ISO 5725-1 [8] and ISO 5725-2 [9] given for accuracy experiments. Because ISO CD 23351-2 is unknown both among the scientific and engineering population, the suggestions of ISO 5725 standards could not be precisely followed. For example, random assignment of participants was not possible, since there are very few experts who knew the measurement principle of ISO CD 23351-2. Instead, we had to accept all participants that volunteered.

The first four authors formed the executive panel of the accuracy experiment.

According to ISO 5725-1 [8], it is common to choose 8–15 participants for an accuracy experiment. Our aim was to reach more than 10 participants. Knowing that the ISO CD 23351-2 method is not familiar to most specialists, we expected that this is a demanding target. Therefore, all Finnish parties (acoustic consulting companies, furniture manufacturers, and universities) were invited by email to join, who were known to be experts in building acoustic measurements.

The invitation letter was sent to 13 parties. It explained the drivers and purposes of the study (to provide measurement uncertainty for the next standard version ISO DIS 23351-2:2025) and the main task of measuring two objects (Objects 1 and 2) once according to ISO CD 23351-2 in a fixed room within weeks 6–9 in 2024 at a meeting room located in Turku, Finland (TUAS). Registration was requested to be sent by email to one of the authors. Participation was voluntary and the financial costs of participation were not compensated.

Seven parties registered. They undertook providing at least one participant, but the final number was still unknown. These parties were sent the following written information:

- Three-page information document about the study drivers, practical arrangements, timetables, anonymity procedures (a proxy from TUAS), and description of Objects 1 and 2.
- The empty data forms (see Fig. S1 in Supplementary material).
- ISO CD 23351-2:2023.
- Room layout drawing that showed the positioning of Objects 1 and 2.

An online meeting was held between two authors and each party separately (1–3 participants in each meeting) where the issues of the information above were discussed in more detail. Parties were informed to carry their sound level analyser, distance meter, and omnidirectional sound source. They were told that the artificial mouth (Talkbox, NTi Audio, Schaan, Liechtenstein) needed to test the Object 1 will be provided by TUAS, so that the accuracy experiment also covers the ISO 3744 [4] standard. The artificial mouth was always tested at the same volume setting (pink noise, 70 dB at 1 m free field).

Finally, thirteen participants from seven parties experimented acceptedly (i.e., conducted the measurements and returned the data forms (Fig. S1 of Supplementary material) until 23rd August 2024). The proxy gave a letter from A to M to each participant using a random procedure so that the results can not be traced to individual participants. None of the participants was contacted after returning the results: all data is based on the data forms.

The participant could make quite few individual choices, since ISO CD 23351-2 gives clear requirements to the measurement arrangements. Furthermore, the position of the Object in the room already fulfilled the requirements. The room was over 1.5 m higher than the object so that Option B had to be applied. Option B states that the measurement surface is a rectangle that consists of four vertical sub-surfaces, and the fifth sub-surface is above the ceiling of the object. Because the measurement surface shall be located 0.50 m from the object, the area of the measurement surface was expected to be the same for all participants in Scenario 2. The participant could choose between fixed measurement points and scanning. In Scenario 1, it was required that the object could not be moved away during the $L_{W,P,1}$ measurement of the bare sound source. It was allowed to apply the exception, that $L_{W,P,1}$ is measured in the same room in another position and using a smaller measurement surface S. Some statistics related to these choices and measured room volumes are given on Sec. 2.2.

We invited one participant to conduct both tests five times to get the repeatability SD (s_r). Based on Hongisto et al. [3], the s_r values were expected to be much smaller than between-laboratory SD, s_L. Therefore, we did not invite all participants to do the test for Objects 1 and 2 five repeated times since the testing of two Objects once took already 3–4 h. We would not have reached 13 participants if the measurements had taken two days.

2.1.2. Test arrangements

The measurements were conducted for two objects known to have different $D'_{S,A}$ levels:

- Object 1: partial enclosure. Booth (Fig. 1) with the door 180° open.
- Object 2: enclosure. As Object 1 but with the door closed.

Their level differences, D, and speech level differences, $D_{S,A}$, tested according to ISO 23351-1 in laboratory conditions are given in Table S1 (Supplementary material).

Two objects were tested since ISO 5725 promotes conducting accuracy experiments for objects representing two completely different levels.

Fig. 1. The booth used in the between-laboratory experiment. The outer dimensions were 1.05 × 1.41 × 2.25 m. The photograph represents Object 1 (booth with door 180° open) under the Scenario 2 test where the artificial mouth inside the object and pointing towards the sitting occupant's direction.

The outer dimensions of the custom-made booth were 1.05 × 1.41 × 2.25 m (3.33 m^3).

Object 1 was tested using an artificial mouth and Object 2 was tested using omnidirectional sound source, as ISO CD 23351-2 requires.

According to authors' measurements, the floor area of the room was 45.6 m^2 (6.12x7.45 m). Room height was 7.22 m up to concrete slab resulting in a cubical room volume of 329.2 m^3. The kitchen extension of 3.0 m^3 increased the grand room volume of 332.2 m^3. The room had 500 mm suspended ceiling made of perforated gypsum board (21.5 m^3), cabinets (3.7 m^3), concrete beam (1.9 m^3), and 70 mm perforated wall absorber field (2.3 m^3). These details reduced the acoustically meaningful room volume by 29.4 m^3 down to 302.8 m^3.

The volume of the booth was 3.3 m^3. The interior of Object 1 was part of the room volume, because its door was open. Thus, we think that the acoustically meaningful room volume was not changed when Object 1 was tested. However, during Object 2 testing, the door of the booth was closed, the interior of the booth was not part of the room volume. Thus, the acoustically meaningful room volume was 3.3 m^3 smaller. Since the measurement of room volume was time-consuming, we also asked participants to report the room volume used in Eq. (4).

2.1.3. Statistical analyses

Normal distribution of data was inspected using Shapiro-Wilk test. The null hypothesis is that the data is normally distributed. The hypothesis is accepted if $p > 0.05$. Based on normality test and holistic reasoning, the existence of outliers or different populations could be suggested.

The data were analyzed according to the statistical methods described in ISO 5725-1 [8] and ISO 5725-2 [9]. The following procedure was followed both for Objects 1 and 2.

The within-laboratory (i.e., repeatability) SD, s_r [dB], was determined by

$$s_r = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_j - x_M)^2} \tag{7}$$

where n is the number of repeated tests by a single participant, x_j [dB] is the result of test j, and x_M [dB] is the mean of the repeated tests. The between-laboratory SD, s_L [dB], was determined by

$$s_L = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - y_M)^2} \tag{8}$$

where N is the number of accepted participants, y_i [dB] is the result of participant i, and y_M [dB] is the mean of the N tests.

If data is normally distributed, standard deviation SD describes the 68 % confidence interval around the mean M, i.e., 68 % of the data is within M − SD and M + SD.

2.2. Results and discussion

2.2.1. Raw data and outlier analyses

Numerical data provided by the 13 participants (A-M) for the two ISO 3382-2 tests (reverberation time of test room during Objects 1 and 2), two ISO 3744 tests (bare artificial mouth and artificial mouth inside Object 1) and two ISO CD 23351-2 tests (Objects 1 and 2) are given in Table S2 (Supplementary material).

Shapiro-Wilk test showed that the reverberation time data (based on ISO 3382-2) was normally distributed ($p > 0.05$) for six octave bands out of eight. Normal distribution was violated at 250 Hz ($p = 0.033$) and 4000 Hz ($p = 0.020$). Since reverberation time measurement is not within our main scope, we included all 13 participants in the analyses of ISO 3382-2.

The distribution of $L_{W,A}$ in ISO 3744 test of bare artificial mouth (Scenario 1 of ISO CD 23351-2 test of Object 1) is shown in Fig. 2a. The data was not normally distributed ($p = 0.002$, N = 13). Three participants (C, E, and I) provided clearly deviant results that are beyond the expected reproducibility standard deviation of ISO 3744 (±2 dB) for steady-state broadband sound sources. After removing these three outliers, the data was normally distributed ($p = 0.219$, N = 10). Therefore, the results are based on ten participants.

The distribution of $L_{W,A}$ in ISO 3744 test of artificial mouth placed inside Object 1 (Scenario 2 of ISO CD 23351-2 test) is shown in Fig. 2b. Data violated normal distribution very strongly ($p = 0.00004$, N = 13). Two participants (C and E) clearly provided deviant results. After removing these two outliers, the data was normally distributed ($p = 0.619$, N = 11). Therefore, the results are based on eleven participants.

It is notable that participants C and E provided deviant results for both SWL tests. Holistic reasoning allows us to propose that the $D'_{S,A}$ results of participants C and E could also deviate from normal distribution since the ISO CD 23351-2 method is based on two independent SLW tests.

The distribution of thirteen $D'_{S,A}$ results for Object 1 is given in Fig. 3a. The data was normally distributed ($p = 0.088$, N = 13). Unexpectedly, there are two clusters: a larger cluster 1A with 8 participants ($D'_{S,A}$ within 5–8 dB) and a smaller cluster 1B with 5 participants ($D'_{S,A}$ within 9–13 dB). Participants C and E failed to conduct SWL test according to ISO 3744 (see above). Therefore, it is justified to remove participants C and E as outliers from cluster 1B (yellow area in Fig. 3a). After this, the data was no longer normally distributed ($p = 0.046$, N = 11). That is, the removal of two undisputed outliers worsened the situation. This raises a doubt that the two clusters might indicate two populations: population 1 (standard compliant population) with 8

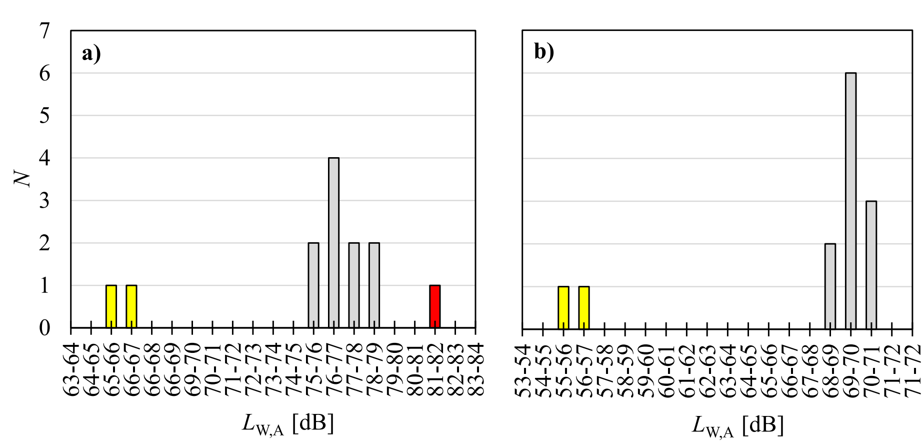

Fig. 2. The distribution of A-weighted sound power level, $L_{W,A}$, reported by 13 participants. a) Bare artificial mouth (Scenario 1 measurement dealing with Object A test). b) Scenario 2 measurement of Object A (artificial mouth inside Object A). Outlying participants C and E are indicated by yellow and participant I by red.

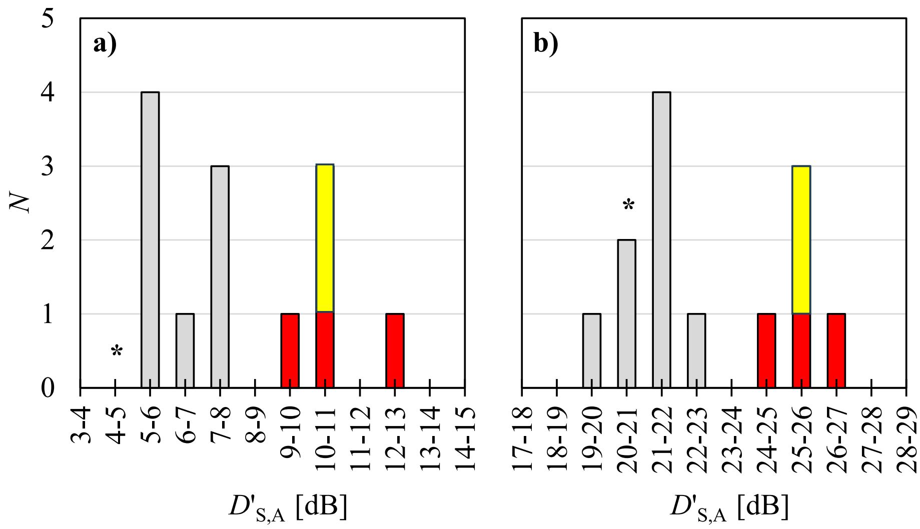

Fig. 3. The distribution of reported apparent speech level reduction, D'_{SA}, reported by 13 participants. a) Object 1. b) Object 2. The laboratory values (*) determined by ISO 23351-1 were 4.8 dB $D_{S,A}$ (Object 1) and 20.4 dB $D_{S,A}$ (Object 2). Outlying participants C and E are indicated by yellow and participants G, I, and L by red.

participants who conduct the test correctly and another population 2 (standard mis-compliant population) with 5 participants, who do not.

Proposing a standard mis-compliant population is justified since participants C, E, and I are located there. Because we know that these participants produced wrong ISO 3744 results, also participants G and L located in the same population could be outliers. Furthermore, data of cluster 1A (5–8 dB) was much closer to the laboratory value of Object 1 (4.8 dB $D_{S,A}$) than the data of cluster 1B (9–13 dB). Both Keränen et al. [5] and Sec. 3 provide evidence that ISO CD 23351-2 produces D'_{SA} values, which are very close but slightly larger than the laboratory values $D_{S,A}$ obtained by ISO 23351-1. This strengthens the claim that cluster 1B (participants C, E, G, I, L) could represent a standard mis-compliant population. Decision about this requires the parallel analysis of Object 2.

The distribution of thirteen $D'_{S,A}$ results for Object 2 is given in Fig. 3b. The data was not normally distributed (p = 0.025, N = 13). Again, two clusters exist: a larger cluster 2A with 8 participants ($D'_{S,A}$ within 19–23 dB) and a smaller cluster 2B with 5 participants ($D'_{S,A}$ within 24–27 dB). Participants C and E failed to conduct sound power level test according to ISO 3744 (see above). Therefore, it is justified to remove them again as outliers from cluster 2B (yellow area in Fig. 3a). When they were removed, the data could not reach normal distribution (p = 0.023, N = 11). Again, data of cluster 2A was much closer to the laboratory value of Object 2 (20.4 dB $D_{S,A}$) than the data of cluster 2B.

It is important to note that suspicious cluster 2B of Object 2 involved the same participants (C, E, G, I, L) as the suspicious cluster 1B of Object 1. Furthermore, data on both clusters B were approximately 4 dB higher than data of clusters A. Clusters A were much closer to the laboratory value. Therefore, it is justified to exclude the five participants of cluster 2B as outliers since they represent a standard mis-compliant population with very high probability.

After the removal of five outlying participants (C, E, G, I, L), the data of ISO CD 23351-2 data was normally distributed for Objects 1 (p = 0.336, N = 8) and 2 (p = 0.788, N = 8). Therefore, the results of Objects 1 and 2 are based on eight participants.

2.2.2. Accuracy experiment of ISO 3382-2

The statistics of reverberation time measured by 13 participants are given in Table 2 (tested during Object 1). Frequency-dependent values are given in Fig. S2 (Supplementary material).

ISO 3382-2 Annex A defines the standard uncertainty, SU, as the ratio of SD and mean. A comparison of our SU (mean SU of Objects 1 and 2) and the SU's reported from some previous accuracy experiments is

Table 2
Mean (M) and standard deviation (SD) of reverberation time reported by 13 participants measured in the same room while testing Object 1. SU is the standard uncertainty. Results with Object 2 were similar and not shown.

f	M	SD	SU
[Hz]	[s]	[s]	[%]
125	0.75	0.14	18.2
250	0.55	0.04	6.9
500	0.59	0.04	6.0
1000	0.64	0.02	3.4
2000	0.66	0.02	2.5
4000	0.61	0.02	3.9
8000	0.43	0.04	10.1

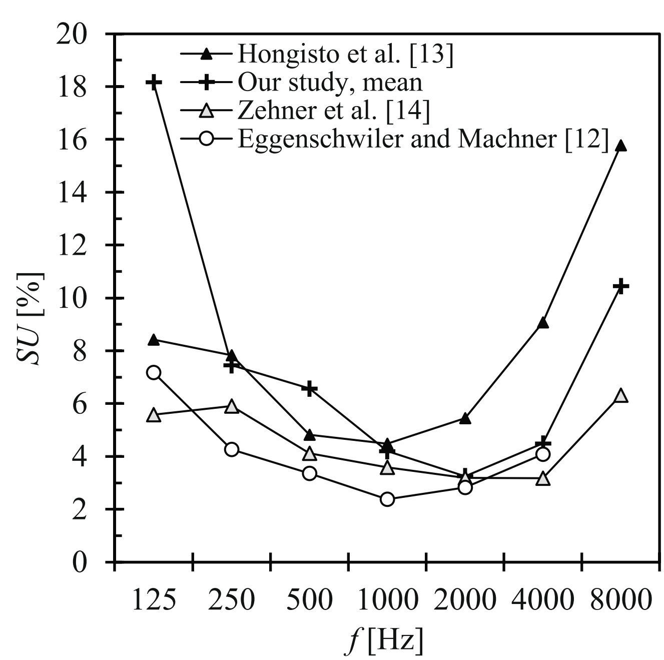

Fig. 4. Standard uncertainty, SU, of reverberation time as a function of frequency, f, or our study and three previous studies. (See above-mentioned references for further information.)

made in Fig. 4. The SU obtained in our experiment clearly exceeds those previously published at 125 Hz and within 2000–8000 Hz. Our larger SU value could not be explained by a single participant. The largest SU was observed at 125 Hz, where the measured range was 0.53–1.06 s. Despite this, the data was normally distributed ($p = 0.057$) so that we cannot easily point out any outliers.

The room was similarly furnished for every participant. The number of people in the room varied from 1 to 3 depending on the party. The room volume in our study (appr. 303 m^3) was smaller than in previous studies of Fig. 4 (museum 1500 m^3 of 4.2 m high, open-plan office 505 m^3 of 3.6 m high, multipurpose hall 1700 m^3 of 9.5 m high). It seems that reverberation time measurements in situ can involve larger uncertainties than previously thought and further research is needed to better understand potential reasons for such uncertainties.

Our primary goal deals with ISO CD 23351-2, where the reverberation time is determined both in Scenario 1 and 2. One participant measured $T = 0.53$ s and another participant measures $T = 1.06$ s at 125 Hz. The influence on K_2 values by Eq. (3) is even 1.9 dB. This difference directly reflects on the SWL value of the current scenario. For example, if the reverberation times between Scenarios 1 and 2 change by 0.53 s, as it might be theoretically possible although the room remains the same, this uncertainty also reflects to the reported $D'_{S,A}$ value. Therefore, large uncertainty in reverberation time measurement may theoretically lead to a significant uncertainty source of D' values at 125 Hz. However, large uncertainties of D' were not seen on Sec. 2.2.4. Further research is justified to better understand which factors could explain large differences in reverberation time.

2.2.3. Accuracy experiment of ISO 3744

The statistics of A-weighted sound power level measured by the accepted participants is given in Table 3. Frequency-dependent values are given in Fig. 5.

The mean reported room volume was 294.5 m^3 ranging from 280.5 to 302.0 m^3. Such volume range influences K_2 value of Eq. (3) by 0.2 dB. A detailed volume calculation was reported on Sec. 2.1.2. Although the measurement of room volume is expected to be a simple task, our data shows that differences in volume measurements are not marginal. Therefore, measurement uncertainty of room volume is also a potential source of SWL measurement uncertainty, and it may reflect to the uncertainty of D' measurements as well.

2.2.4. Accuracy experiment of ISO CD 23351-2

The statistical results concerning the apparent speech level reduction, $D'_{S,A}$ [dB], for Object 1 (booth, door open) and 2 (booth, door closed) are shown in Table 4. The distribution of frequency-dependent data is shown in Fig. 6.

The between-laboratory results of level reduction, D', and apparent speech level reduction, $D'_{S,A}$ [dB], for Object 1 (booth, door open) and 2 (booth, door closed) are shown in Table 4. The between-laboratory SDs of D', s_L, are within 0.8 and 1.8 dB depending on frequency and object. The between-laboratory SDs of $D'_{S,A}$ were nearly the same for both objects.

The within-laboratory results of level reduction, D', and apparent speech level reduction, $D'_{S,A}$ [dB], for Object 1 (booth, door open) and 2

Table 3
Mean (M) and standard deviation (SD) of the unweighted SWL determined by ISO 3744, L_W, as a function of frequency, f, reported by N participants. $L_{W,A}$ is the A-weighted SWL calculated by the authors. Scenario 1 means bare artificial mouth. Scenario 2 means artificial mouth inside Object 1.

	Scenario 1			Scenario 2		
f	M	SD	N	M	SD	N
[Hz]	[dB]	[dB]	[dB]	[dB]	[dB]	[dB]
125	72.4	1.7	10	65.2	1.6	11
250	73.6	1.2	10	65.5	1.3	11
500	73.0	1.1	10	66.0	0.8	11
1000	71.3	1.0	10	64.0	1.1	11
2000	68.9	0.9	10	61.8	0.9	11
4000	67.8	1.1	10	61.0	0.9	11
8000	61.8	2.4	10	55.2	1.9	11
L_{WA}	76.6	1.0	10	69.5	0.8	10

Table 4
Inter-laboratory test of ISO CD 23351-2 involving eight participants. Mean (M) and standard deviation (SD) of the apparent level difference, D' [dB], and apparent speech level difference, D'_{SA}. SD corresponds to between-laboratory SD, s_L. N is the number or participants reporting a value.

	Object 1			Object 2		
f	M	SD	N	M	SD	N
[Hz]	[dB]	[dB]	[dB]	[dB]	[dB]	[dB]
125	6.4	1.5	8	14.0	1.5	8
250	7.5	1.4	8	20.8	1.0	8
500	6.2	0.9	8	20.0	0.7	8
1000	6.5	1.2	8	23.1	0.7	8
2000	6.5	0.8	8	25.5	0.8	8
4000	6.1	0.9	8	28.3	1.3	8
8000	6.2	0.8	7	33.3	1.8	7
$D'_{S,A}$	6.3	0.8	8	21.1	0.7	8

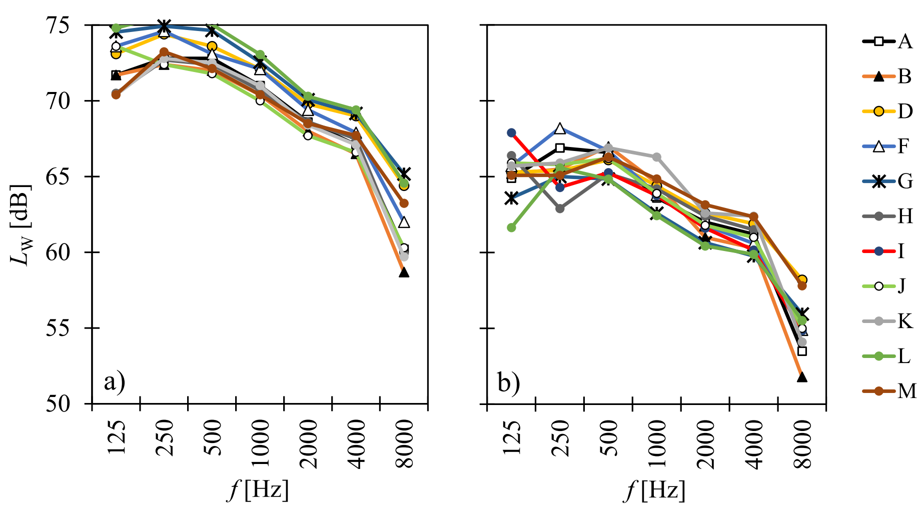

Fig. 5. Unweighted sound power level determined by ISO 3744, L_W, as a function of frequency, f. a) Scenario 1: Bare artificial mouth (N = 10). b) Scenario 2: Artificial mouth inside Object 1 (N = 11).

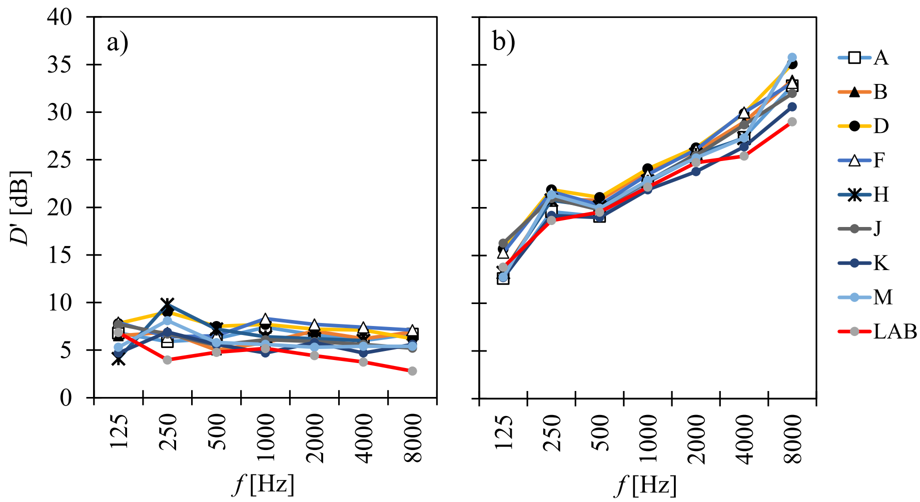

Fig. 6. Apparent level reduction determined by ISO CD 23351-2, *D*', as a function of frequency, *f*. a) Object 1 (N = 8). b) Object 2 (N = 8). LAB refers to a laboratory test result determined by ISO 23351-1.

Table 5
ISO CD 23351-2 test results conducted by one participant and the laboratory values (LAB) determined by ISO 23351-1. Mean (M) and standard deviation (SD) of the apparent level difference, *D*' [dB], and apparent speech level difference, D'_{SA}, based on five repeated measurements. Standard deviation SD corresponds to repeatability SD, s_r.

	Object 1			Object 2		
f	M	SD	LAB	M	SD	LAB
[Hz]	[dB]	[dB]	[dB]	[dB]	[dB]	[dB]
125	7.5	0.0	7.3	14.8	0.3	13.4
250	6.1	0.1	4.3	22.2	0.6	18.6
500	6.8	0.1	4.7	20.7	0.2	19.5
1000	7.9	0.3	5.3	23.7	0.4	21.9
2000	7.6	0.2	4.1	26.4	0.2	24.3
4000	7.4	0.1	3.5	30.2	0.2	25.4
8000	6.9	0.5	1.5	33.2	0.0	27.1
$D'_{S,A}$	7.1	0.1	4.7	21.8	0.2	20.2

(booth, door closed) are shown in Table 5 based on five repeated measurements conducted by one participant.

Four participants out of eight applied the scanning method in Scenario 1 and 2 for both Objects 1 and 2.

3. Between-room experiment

3.1. Materials and methods

The purpose of the between-room experiment was to determine whether the acoustic properties of the room influence the result of ISO CD 23351-2. Therefore, a single operator tested two objects in several rooms covering a large range of absorption areas, because it has strong influence on the environmental correction of Eq. (3).

The measurements were conducted for Objects 3–4 known to have different $D'_{S,A}$ levels:

- Object 3: partial enclosure (7.8 dB $D_{S,A}$). A booth with the door 180° open.
- Object 4: enclosure (21.2 dB $D_{S,A}$). As Object 3 but with the door closed (Fig. 7).

The level differences, *D*, and speech level differences, $D_{S,A}$, tested according to ISO 23351-1 in laboratory conditions are given in Table S1 (Supplementary material).

Two objects representing two completely different levels were tested since we wanted to know how the room affects the whole range of typically tested objects.

The outer dimensions of the custom-made booth were 0.97 × 1.21 × 1.93 m (2.27 m^3). The booth was not the same as Sec. 2, although the acoustic performance was almost similar.

The booth was moved as one piece between the rooms. It was not disassembled and assembled between the tests. This was necessary to minimize the potential measurement uncertainties due to changes in the physical structure. Care was taken to ensure that the booth was not damaged in any way, so that measurement uncertainty would not increase due to changes in the booth. Special attention was paid to ensuring that the floor was flat under the booth, as a floor that was too

Fig. 7. A photograph of the booth used in between-room experiment.

Table 6
Description, volume, dimensions, mean reverberation time, T_M, and mean environmental correction $K_{2,M}$ within 125–8000 Hz of the 11 studied rooms. The floor was hard except in Room 1b. The carpet was 12 mm thick and the area was 10 m^2.

Room	Description of room	V [m^3]	Length [m]	Width [m]	Height [m]	T_M [s]	$K_{2,M}$ [dB]
1a	Large empty room	200	9.9	5.7	3.6	3.9	10.6
2a	Small empty room	62	4.7	4.4	3	1.6	12.0
2b	As 2a but added wall absorption	62	4.7	4.4	3	0.58	7.9
2c	As 2b but added ceiling absorption	62	4.7	4.4	3	0.39	6.7
3a	As 3b but added wall absorption	224	8	7	4	0.39	3.0
3b	Classroom	224	8	7	4	0.63	4.3
4a	Large auditorium	1248	39	8	4	0.69	1.2
4b	As 4a + added wall absorption	1248	39	8	4	0.50	0.9
5a	Open plan office	490	24	6	3.4	0.44	1.9
6a	Very large atrium	5184	16	18	18	1.3	0.6
1b	As 1a but carpet under booth	200	9.9	5.7	3.6	3.0	9.5

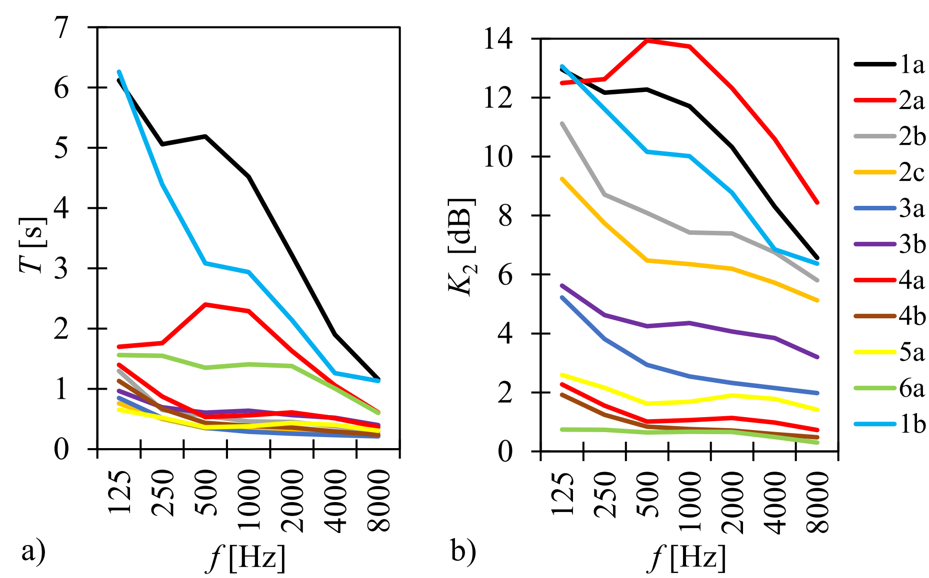

Fig. 8. a) Reverberation time, T, and b) environmental correction, K_2, as a function of frequency, f, in 11 rooms.

slanted would cause the door to be installed crooked and uncontrolled sound leakage along the door seams. Flatness was determined by measuring the thickness of the gap in the door seams, which was kept constant. Straightening was done by placing a suitable number of 2 mm thick wedges under the most hanging corner.

The rooms were chosen so that the K_2 of Eq. (3) would vary as much as possible. Therefore, we paid attention to the fact that the rooms differed as much as possible with respect to room volume and reverberation time.

The studied rooms are described in Table 6. Their T and K_2 values are given in Fig. 8. The number of physical rooms was six. Four of the rooms were acoustically modified by installing absorbers to the walls and/or ceiling so that the K_2 was significantly increased. It should be noted that Room 1a represents a diffuse field, because it is a reverberation room that fulfills ISO 3741 [4], ISO 23351-1 [1] and ISO 354 [10] standards. Therefore, we can provide a good quantitative understanding whether the deviation from diffuse sound field has an influence on the test results of ISO CD 23351-2.

The between-room SD, s_s [dB], was determined by

$$s_s = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(z_k - z_M)^2} \qquad (10)$$

where K is the number of rooms where the same product is tested, z_k [dB] is the result in room k, and z_M [dB] is the mean of results in all K rooms.

Object 3 was tested using an artificial mouth and Object 4 was tested using omnidirectional sound source, as ISO CD 23351-2 requires.

Measurements were conducted using sound level meter (XL2, NTI Audio, Liechtenstein) with Type 1 condenser microphone (M2230, NTI Audio, Liechtenstein), omnidirectional sound source (NOR 276, Norsonic Ltd., Norway) with amplified and inbuilt pink noise generator (NOR 280) and artificial mouth (Talkbox, NTi Audio, Schaan, Liechtenstein).

In every room, the volume setting of both sound sources was the same, which allowed the inspection of the effect of room on measured sound power level.

All measurements were conducted using the scanning method. The scanning speed was approximately 0.5 m/s and the distance between adjacent scanning lines was approximately 40 cm. The measurement distance was 50 cm from the outer surface of the booth walls or ceiling.

3.2. Results and discussion

The SWL measured for bare artificial mouth (directive sound source) and omnidirectional sound source according to ISO 3744 in all rooms is reported in Table 7 and Fig. 9. The room did not affect the SWL in a systematic way. The SDs within 125–8000 Hz octave bands were within 0.7–1.8 dB for a directive sound source (artificial mouth) and within 0.9–1.6 dB for an omnidirectional sound source. The SDs are close to the SDs of Table 3, which analyzed the SWL tests conducted by different participants in the same room. We expected the SDs to be larger for the

Table 7
Mean (M) and standard deviation (SD) of the sound power level ($L_{W,P,1}$) as a function of frequency, f, for a directive sound source and omnidirectional sound source measured by single operator in several rooms.

	Directive[a]		Omnidirectional[b]	
f	M	SD	M	SD
[Hz]	[dB]	[dB]	[dB]	[dB]
125	72.2	1.8	103.4	1.5
250	72.3	1.5	111.8	1.6
500	71.4	0.7	105.1	0.9
1000	70.3	1.0	102.6	1.0
2000	67.5	1.2	104.5	1.0
4000	66.0	1.3	111.8	1.1
8000	59.6	1.2	98.4	1.1

[a] Number of measured rooms was 10.
[b] Number of measured rooms was 11.

directive sound source as the SPL measurements in the most directive direction are sensitive to sampling density. It seems that the scanning line distance of 40 cm was suitable since the SWL measurement uncertainty was so small.

The ISO CD 23351-2 results for Objects 3 and 4 are shown in Table 8. The $D'_{S,A}$ values between rooms varied very little: all values were within 1.8 dB. Pearson's correlation coefficient between $D'_{S,A}$ and environmental correction, $K_{2,M}$, was negligible both for Object 3 (−0.16, p = 0.66, two-sided) and Object 4 (−0.29, p = 0.39, two-sided) indicating that there was no trend that the environmental correction would be associated with the $D'_{S,A}$ values. Thus, the upper limit of $K_2 = 4$ dB in ISO 3744 needs not to be applied in the SWL tests of ISO CD 23351-2.

D' values obtained for Objects 3 and 4 in all studied rooms are illustrated in Fig. 10. Despite the small correlation coefficients reported above, there is a minor issue worth mentioning. The results of room 1a were systematically 1.1—1.5 dB smaller within 1000–4000 Hz than the results of any other room. The result could not be explained by K_2, since room 2a had larger K_2 than room 1a.

Keränen et al. [5] is the only study so far that contains similar research design as our between-room study. They conducted a similar between-site experiment in six rooms. The measurement technician was different in our study (Caradonna) and in Keränen et al. [5] (Laukka) and the measurements were conducted in different years so that these two studies are independent. Keränen et al. [5] did not report the environmental correction values as we did. They used measurement distance of 0.25 m, while we used the 0.50 m distance later specified in ISO CD 23351-2. Since our study involved a larger number of rooms and room absorption manipulations, our study provides a larger coverage of environmental corrections. In addition, Keränen et al. [5] studied only an enclosure while we studied both enclosure (Object 4) and partial enclosure (Object 3). Therefore, our study provides a significant progress in this field. Keränen et al. [5] obtained $s_s = 0.36$ dB for $D'_{S,A}$ for an enclosure, while we obtained a slightly larger value, i.e., $s_s = 0.63$ dB. The reason may be the larger range of environmental corrections among the eleven rooms. Although the difference is almost twofold, our result supports the view that the uncertainty due to room choice is reasonable.

Object 4 was tested in the same room 1 in two ways: on hard floor (room 1a) and on top of a 12 mm thick carpet (room 1b). The carpet (10 m^2) extended over the measurements surface and it was installed in both Scenarios 1 and 2. The addition of the carpet did not significantly affect the test result. Therefore, the ISO CD 23351-2 seems to be applicable also in rooms having a sound-absorbing carpet, provided that the carpet thickness is not more than 12 mm.

Table 8
Apparent speech level reduction, $D'_{S,A}$ [dB], measured in 10 or 11 rooms for Objects 3 and 4. M = mean. SD = standard deviation. SD corresponds to between-site SD, s_s.

Room	Object 3	Object 4
1a	7.2	21.2
2a	8.5	22.0
2b	9.0	22.3
2c	8.0	21.9
3a	8.7	22.4
3b	8.8	22.3
4a	7.7	21.9
4b	8.7	22.6
5a	9.0	20.8
6a	7.8	21.9
1b	–	20.8
M	8.3	21.8
SD	0.62	0.63

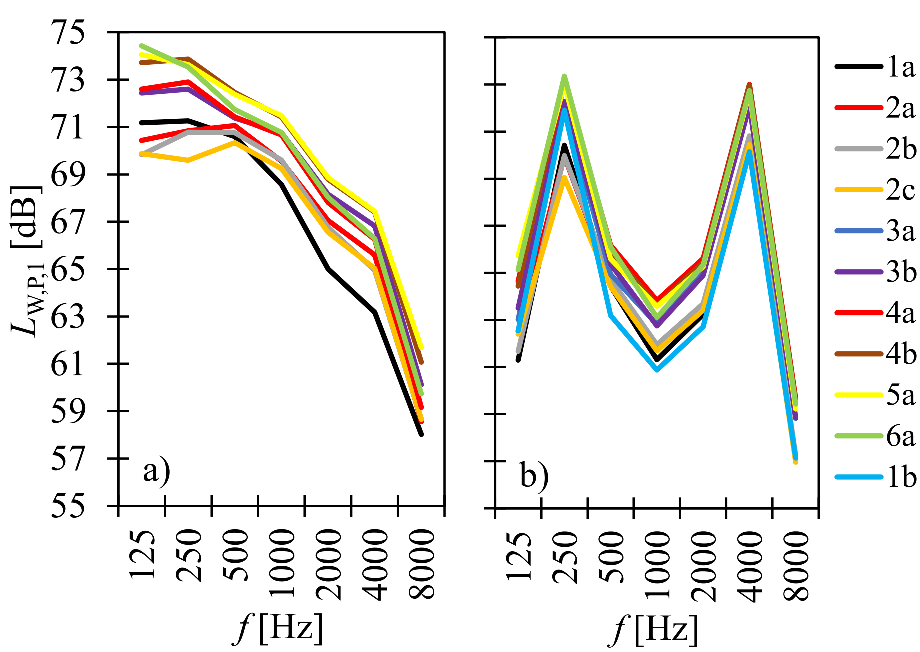

Fig. 9. Sound power level determined by ISO 3744, $L_{W,P,1}$, as a function of frequency, f, measured by a single operator. a) Bare artificial mouth measured in ten rooms. b) Bare omnidirectional sound source measured in eleven rooms.

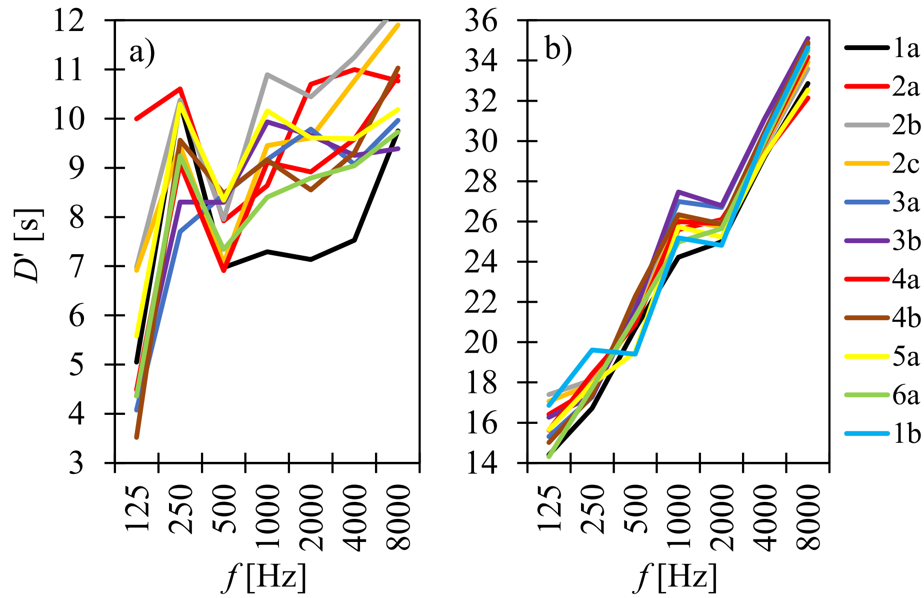

Fig. 10. Apparent level difference determined by ISO CD 23351-2, D', as a function of frequency, f, for a) Object 3 in ten rooms, and b) Object 4 in eleven rooms.

Table 9
Main results.

Object type	s_r [dB]	s_L [dB]	s_R [dB]	s_s [dB]	s_{tot} [dB]
Partially enclosed, $D'_{S,A} < 8$ dB	0.14	0.82	0.83	0.62	1.03
Enclosed, $D'_{S,A} > 20$ dB	0.22	0.68	0.72	0.63	0.93

4. Summary analysis of two experiments

Table 9 summarizes the SDs obtained on Secs. 2-3. The reproducibility SD, s_R, was calculated from the results of between-laboratory experiment by

$$s_R = \sqrt{s_L^2 + s_r^2} \tag{11}$$

Our study went beyond conventional between-laboratory experiment, since we also conducted the independent between-room experiment. We suggest that the total SD, s_{tot}, which best explains the measurement uncertainty caused by different operators, repeatability, and room choice, is calculated by

$$s_{tot} = \sqrt{s_L^2 + s_r^2 + s_s^2} \tag{12}$$

The values of s_R and s_{tot} are given in Table 9.

5. Discussion

Our study represents the first systematic research regarding the between-laboratory measurement uncertainty of ISO CD 23351-2 method.

Hongisto et al. [3] reported an accuracy experiment of the laboratory test method ISO 23351-1, which involved eight participants from six countries. In their study, each participant tested one partial enclosure ($D_{S,A} = 4.2$ dB) and one enclosure ($D_{S,A} = 28.7$ dB) in the reverberation room. The corresponding s_R values were 0.6 dB and 1.1 dB. The values were very similar to ours, although ISO 23351-1 is a laboratory method, which requires a reverberation room. The agreement between their and our data proposes that ISO CD 23351-2 is a robust field test method. However, the large number of outliers found in the between-laboratory experiment suggests that the method is complex and requires training and individual qualification.

ISO CD 23351-2 has two measurement options related to the choice of measurement surface. Option A is used in low rooms, where the distance d between the object ceiling and room ceiling is within 0.5–1.5 m. In this option, the ceiling shall be reflecting and the measurement surface consists of four vertical sub-surfaces which extend from floor to room ceiling. Option B is used in rooms where $d > 1.5$ m. In this option, there is no requirement for the ceiling to be reflective, and the measurement surface consists of four vertical sub-surfaces and one horizontal sub-surface, which is above the object. Option B was applied both in the between-laboratory experiment by all participants and in the between-room experiment in all rooms. Therefore, our study is limited to Option B. Complementary research is needed regarding the measurement uncertainty of Option A.

Finally, the between-laboratory test showed that measured room volumes reported by participants differed almost by 10 %. This was unexpected. ISO CD 23351-2 does not give further advice about the measurement of room volume, whether the volume of fixed cabinets and suspended ceilings should be subtracted or not. Furthermore, the subtraction of the volume of the test object from empty room volume is logical, if the object is an enclosure and the interiors are not connected to the room. The between-participant differences in room volume assessments can be even 100 % in cases when the test object volume is 50 % of room volume. Such situations can exist when very large objects (such as pods for 10 persons) are tested in rooms, which just fulfill the volume requirements of ISO CD 23351-2 for the test environment (room walls at least 1.5 away from object, room ceiling at least 0.5 m away from object). This error source should be considered when ISO 23351-2 is finished. The final standard should give clear advice on how the room volume is calculated. We propose that the volume of the object is subtracted from the room volume if the object is enclosed.

The design of our between-laboratory experiment was strong, because the participants tested two independent objects having two completely different levels of test result, and the between-laboratory experiment was possible to conduct for ISO 3744 for two different sound sources. If we had tested only one object, we would not have been able to identify the five outliers of ISO CD 23351-2 test. Instead, we should have accepted the lack of normal distribution, and the between-laboratory experiment would not be satisfactory. Based on this experience, it is strongly recommended that the next between-laboratory experiments should also involve at least two different objects, and the sound power level test of at least one known sound source.

The laboratory values ($D_{S,A}$ by ISO 23351-1) of Objects 1–4 were 0.5–2.1 dB lower than the field values ($D'_{S,A}$ by ISO CD 23351-2). The values are reported in Table S1 (Supplementary material). The difference may be real or statistical coincidence. This analysis is not perfect since the data is based on only two operators. Anyhow, this finding agrees with Keränen et al. [5]. Systematic research is needed about the differences of laboratory and field values.

The between-laboratory experiment of Object 1 suffered from one clear weakness. Based on the survey among registered participants, none of the participants owned an artificial mouth. This is understandable since ISO 23351-2 is not yet an accepted standard and there is no commercial activity related to this method. Furthermore, artificial mouth is not needed in any other building acoustic field test standard. Therefore, we decided to provide our artificial mouth for all participants. However, all participants should carry their own measurement apparatus in between-laboratory experiments. We violated this principle. The quantitative bias caused by the fixed artificial mouth for the test results of Object 1 is assumed to be small, because ISO CD 23351-2 is based on two subsequent sound power level tests for the same source: scenario 1 for bare sound source and Scenario 2 for the sound source inside Object 1. The benefit of using a fixed artificial mouth was that we could conduct two between-laboratory experiments of ISO 3744 and point out probable outliers. Anyhow, there is a possibility that the directivities of different artificial mouths would differ. Our study did not cover this potential source of bias. The same choice was also made by Hongisto et al. [3], who conducted the between-laboratory experiment of ISO 23351-1. The next between-laboratory experiments should no longer follow this example, but all apparatus shall be provided by the participants.

Finally, our between-laboratory experiment has also another limitation. Although we emphasized during the recruitment, that all participants shall conduct the measurements using different apparatus, we did not control the apparatus during the measurement session. Measurement apparatus had to be reported by all participants. The analysis of these reports showed that four participants used the same apparatus as another participant on the same day. These occurrences are due to two or three participants from the same company. Therefore, our data may slightly underestimate the true measurement uncertainty since the measurement uncertainty due to measurement apparatus was slightly biased.

The newer version of the method, ISO DIS 23351[11]does not contain such technical differences from ISO CD 23351-2[6], which would have affected the measurements in our study. Therefore, our results are valid also for ISO DIS 23351-2.

6. Conclusions

Our purpose was to determine the measurement uncertainty of ISO CD 23351-2:2024 method by two experiments: 1. between-laboratory experiment and 2. between-room experiment. In study 1, thirteen participants tested one partial enclosure and one enclosure in the same room. Based on eight accepted participants, the between-laboratory standard deviations of $D'_{S,A}$ were 0.8 dB for the partial enclosure and 0.7 dB for the enclosure. Five outliers were identified because of obvious errors in ISO 3744 tests. In study 2, a single participant tested one partial enclosure and one enclosure in ten different rooms. The between-room standard deviations of $D'_{S,A}$ were 0.6 dB both for the partial enclosure and the enclosure. The total SD aggregating the between-laboratory and between-room SDs was 1.0 dB. It considers the measurement uncertainties both due to operator, repeatability, and room choice. Our study was limited to Option B of ISO CD 23351-2, which considers rooms which are at least 1.5 m higher than the test object. Option A considers rooms, whose ceiling is at 0.5–1.5 m distance from the test object. Further research is needed for Option A. In addition, the ISO 23351-2 should give advice that the volume of the test object is subtracted from the room volume if the object is an enclosure. Our results can be useful in the finishing of ISO 23351-2 standard. Our study also provides useful results regarding the measurement uncertainty of ISO 3744 and ISO 3382-2 standards.

CRediT authorship contribution statement

Valtteri Hongisto: Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization, Writing – review & editing. **Riccardo Caradonna:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Jukka Keränen:** Writing – original draft, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Jarkko Hakala:** Resources, Methodology, Investigation, Data curation. **Reijo Alakoivu:** Supervision, Resources, Methodology, Investigation. **Arianna Astolfi:** Supervision, Resources, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Valtteri Hongisto, grantholder on behalf of Turku University of Applied Sciences reports financial support was provided by Finnish Furniture Association. Arianna Astolfi, grantholder on behalf of Polytechnic of Turin reports financial support was provided by Polytechnic of Turin. Dr. Hongisto is the convenor of working group ISO TC 43 SC2 WG34 which is responsible for developing ISO 23351-2 standard However, the convenor position is not financially compensated but it is voluntary work. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Finnish Furniture Association (Huonekalusäätiö sr, Finland) and Turku University of Applied Sciences Ltd. (TUAS, Turun ammattikorkeakoulu Oy) for funding this research. We also thank Saint-Gobain Ecophon AB (Sweden) for supporting the scholarship of Mr. Riccardo Caradonna. We also thank Markus Forsten for providing a meeting room during the between-laboratory experiment. Finally, we sincerely thank all external parties (Karelia University of Applied Sciences Ltd, Promethor Ltd, Akukon Ltd, Sitowise Ltd, Framery Ltd) and their experts for attending the between-laboratory experiment.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apacoust.2025.110773.

Data availability

All relevant data is given in Supplementary material.

References

[1] ISO. International standard ISO 23351-1:2020 Acoustics — Measurement of speech level reduction of furniture ensembles and enclosures —Part 1: Laboratory method. Genève, Switzerland: International Organization for Standardization; 2020.

[2] Hongisto V, Keränen J, Virjonen P, Hakala J. Test method for determining sound reduction of furniture ensembles. Acta Acust Acust 2016;102:67–79.

[3] Hongisto V, Keränen J, Hakala J. Accuracy experiment of ISO DIS 23351-1 – speech level reduction of furniture ensembles and enclosures. Appl Acoust 2020;164:107249.

[4] ISO. International standard ISO 3744:2022 Acoustics Determination of sound power levels and sound energy levels of noise sources using sound pressure. Genève, Switzerland: International Organization for Standardization; 2022.

[5] Keränen J, Hakala J, Laukka J, Hongisto V. Alternative measurement methods for speech level reduction of booths in field conditions. Forum Acusticum 2023, paper

392, 11-15 Sep, Turin, Italy, 2023. Open access: https://appfa2023.silsystem.solutions/atti/000392.pdf.

[6] ISO. Standardization committee draft ISO CD 23351-2. Genève, Switzerland: International Organization for Standardization; 2024.

[7] ISO. International standard ISO 3382-2:2008 Acoustics — Measurement of room acoustic parameters — Part 2: Reverberation time in ordinary rooms; 2008.

[8] ISO. International standard ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions. Genève, Switzerland: International Organization for Standardization; 1994.

[9] ISO. International standard ISO 5725-2:1994 Accuracy (trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. Genève, Switzerland: International Organization for Standardization; 1994.

[10] ISO. ISO 354:2003 Acoustics — Measurement of sound absorption in a reverberation room; 2003.

[11] ISO. International standard draft ISO DIS 23351-2. Genève, Switzerland: International Organization for Standardization; 2025.

[12] Eggenschwiler K, Machner R. Intercomparison measurements of room acoustical parameters and measures for speech intelligibility in a room with a sound system. J Audio Eng Soc 2005;53(3):193–204.

[13] Hongisto V, Keränen J, Labia L, Alakoivu R. Precision of ISO 3382-2 and ISO 3382-3 – A Round-Robin test in an open-plan office. Appl Acoust 2021;175:107846.

[14] Zehner M, Zurwerra D, Goldberg A. Round Robin to Determine Reverberation Time. Paper 10320. 148th Convention of AES (Audio Engineering Society), 10 pp., 2–5 June, Virtual Vienna (Online); 2020.