Gauge repeatability and reproducibility studies of field testing of airborne and impact insulation in multifamily residences

LoVerde, John[a]
Dong, Wayland[b]
Veneklasen Associates
1711 16th Street, Santa Monica, CA 90404 USA

ABSTRACT

A gauge repeatability and reproducibility study (GRR) uses analysis of variations (ANOVA) on an appropriately designed experiment to separate and quantify the components of the overall uncertainty. The experimental design can be tailored to extract information on the variance attributable to various components of the measurement process. The authors have previously presented results of GRR studies of the airborne and impact insulation of floor-ceiling assemblies in several apartment buildings (ICSV 2018, Internoise 2018), in which the uncertainty in the measurement method and the variability of the nominally-identical assemblies were compared. The results are potentially instructive in modifying or optimizing test procedures and methods, evaluating the number of measurements required to accurately evaluate assemblies, and for evaluation, design, and quality control of construction and workmanship. Results and evaluation of additional GRR studies on in situ testing of airborne and impact insulation of walls and floor-ceiling systems are presented.

Keywords: Repeatability, Reproducibility, Gauge, Uncertainty
I-INCE Classification of Subject Number: 77

1. INTRODUCTION

1.1. Measurement Uncertainty

In the United States, measurement uncertainties for field acoustical testing are determined by inter-laboratory studies (ILS) as described in ASTM E691[1] and E177[2]. In such a study, a single specimen is tested by the various laboratories or testing agencies, and the assembly may also be tested multiple times by a single operator. The total uncertainty is broken into repeatability, which is the variation when a test is repeated by the same test personnel, and the reproducibility, which is the variation between different test agencies. The total uncertainty is given by the sum of these two terms,

$$\sigma^2 = \sigma_{repeatability}^2 + \sigma_{reproducibility}^2 \qquad (1)$$

[a] jloverde@veneklasen.com
[b] wdong@veneklasen.com

Repeatability is defined as the precision under repeatability conditions, i.e., "the closeness of agreement between independent test results obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time."[2] It is essentially the same for laboratory and field measurements. Reproducibility in the laboratory is defined as "the closeness of agreement between independent test results obtained with the same method on identical test items in different laboratories with different operators using different equipment."[2] As it is not generally possible for assemblies to be transported between laboratories, this requires the same separating construction be built into different laboratories. This raises questions regarding the homogeneity of the materials and the identicalness of the construction methods. In field tests, the "different laboratories" refers to different testing organizations traveling to and testing the same assembly.

A different type of study, a gauge repeatability and reproducibility ("Gauge R&R" or GRR) study, uses analysis of variations (ANOVA) techniques along with a suitable experiment design. This type of study not only provides information on the repeatability and reproducibility of the measurement process but is capable of separating the uncertainty in the part under test from the uncertainty associated with the measurement method itself. Gauge R&R studies are commonplace in manufacturing, where they are often used to evaluate the repeatability and reproducibility of a literal gauge used to measure a part during the process. In this context, the repeatability is generally defined as the variation when the same part is re-measured by the same person, and reproducibility defined as when the same part is re-measured by a different person.

Whitfield and Gibbs[3,4] have applied this concept to building acoustics testing, and here we follow their interpretation. The "part" in this case refers to the assembly under test. Repeatability remains as the variation when the same test personnel re-measure the same assembly, and reproducibility is the variation attributable to different test personnel measuring the same part. That "gauge" in this case refers to the entire process of measuring the sound insulation of an assembly.

Note that in the traditional ILS, the reproducibility is based on measurements by different testing laboratories on specimens constructed on their premises in a similar manner but not actually identical specimens, while a GRR is traditionally based on measurements by different personnel but within the same company and operating under the same instructions and interpretation of the relevant standards. The GRR methodology could be expanded to include different test companies (with different equipment and practices) by increasing the number of factors in the analysis. This was not investigated here.

Including a large number of assemblies or parts in the GRR further allows an assessment of the uncertainty of the part itself. Traditional ILS are of a single assembly. Field measurements of a significant number of assemblies using the traditional methods will provide an estimate of the overall variation but does not distinguish between measurement uncertainty and part variation.

1.2. GRR Analysis

A two-factor balanced analysis of variation (ANOVA) with interaction was performed on the data set. In this model, the factors are the operators and the assemblies (parts), and the results can be written as

$$X_{ijk} = \mu + O_i + P_j + (OP)_{ij} + r_{ijk} \qquad (2)$$

where X is the measured value, μ is the overall mean, O is the random variation associated with operator i, P is the random variation associated with assembly or part j, OP is the variation associated with the interaction of operator i with part j, and r is the variation associated with the repeat k. The standard deviation (square root of the variance) of each term is labelled in the same manner as σ_O, σ_P, σ_{OP} and σ_r. The variables are assumed to be random and normally distributed.

The terms used in Eq. (1) in terms of the model in Eq. (2) are

$$\sigma_{repeatability}^2 = \sigma_r^2$$
$$\sigma_{reproducibility}^2 = \sigma_O^2 + \sigma_{OP}^2 \tag{3}$$

although, again, the reproducibility is defined between the test personnel of the same company, not between testing companies. The combined variance of the measurement method (our "gauge') is

$$\sigma_{gauge}^2 = \sigma_o^2 + \sigma_{OP}^2 + \sigma_r^2 \tag{4}$$

and the total variance of the measurement is therefore

$$\sigma^2 = \sigma_P^2 + \sigma_{gauge}^2, \tag{5}$$

where σ_P^2, the variance attributable to the part or assembly, can be separated from σ_{gauge}^2, the variation due to the measurement.

This is clearly different from the ASTM definitions that assume that the total variance is due to the summation of variance due to repeatability and reproducibility.

2. GRR STUDIES

2.1. Testing

The results of several GRR studies have been published previously[5,6] in addition to the study described in this paper. In all cases, testing was performed by Western Electro-Acoustic Laboratory, which is accredited for the test methods by the National Voluntary Laboratory Accreditation Program (NVLAP) of National Institute of Standards and Technology (NIST). WEAL is a subsidiary corporation of Veneklasen Associates, Inc., the authors' institution. The test personnel had previously been trained and qualified for the test method and had significant experience in performing such tests.

Personnel were grouped into teams of two people each. Each team used a Bruel & Kjaer type 2250 or 2270 sound level meter running the same building acoustics software for the test. The loudspeakers and noise generators were of identical models, and two models of tapping machines were used. All equipment calibrations were current and sound level meters were calibrated before and after the testing. Each team used the same test equipment during the study.

Teams were instructed to perform the measurements as they usually would. No direction was given regarding the division of labor between the teams, the order of testing, or such details as loudspeakers and tapping machine positions. All variance due to differences in equipment, details of procedure, personnel, etc., are assigned to the operator category.

2.2. Previous Results

The previous studies were described in Ref. 5 and Ref. 6, and the results are summarized here. The first study (GRR1) was a multifamily residential project in which the parts under test were the floor-ceiling assemblies in two adjacent stacks of four units each. There were therefore six "parts" or assemblies that were nominally identical. The floor plans were identical with receiving room volumes of 50.9 m^3. The rooms were finished but unfurnished. The six parts were tested by three teams and all tests were repeated once (two repeats).

The second study (GRR2) was a recently-completed multifamily residential project. Floor-ceiling assemblies in one stack of five units were tested, providing four parts. The floor plans were identical with receiving room volumes of 35.4 m^3. The rooms were finished but unfurnished. The four parts were tested by three teams with two repeats.

The assemblies for GRR1 and GRR2 were both wood-framed assemblies, with dimensional lumber joists and wood I-joists, respectively. Both had hard-surfaced flooring, a gypsum concrete screed over sound mat, and gypsum board ceilings hung on resilient channel.

For both GRR1 and GRR2, one study was performed for airborne noise isolation testing per ASTM E336[7], and another for impact insulation testing per ASTM E1007[8]. The results for the single-number ratings are shown in Table 1. Airborne ratings are Normalized Noise Isolation Class (NNIC), which is similar to $D_{nT,w}$. Impact ratings are Normalized Impact Sound Rating (NISR), which is similar to $L'_{nT,w}$, except inverted so that higher ratings represent lower sound pressure levels; the approximate relationship is $110 - L'_{nT,w}$. The analysis was repeated for each third-octave band.

Table 1: GRR1, GRR2, GRR3, and GRR4 Results (standard deviations) for Single-Number Ratings

Rating	μ	σ	σ_P	σ_{gauge}	σ_O	σ_{OP}	σ_r
GRR1 Airborne, NNIC	57.5	1.01	0.84	0.54	0.27	0.00	0.47
GRR1 Impact, NISR	57.9	1.78	1.41	1.09	0.48	0.69	0.69
GRR2 Airborne, NNIC	55.2	1.31	1.21	0.49	0.00	0.39	.029
GRR2 Impact, NISR	49.0	1.47	0.88	1.17	0.24	0.53	1.02
GRR3 Airborne, NNIC	56.6	1.21	1.00	0.67	0.00	0.36	0.56
GRR3 Impact, NISR	57.6	1.80	1.25	1.29	0.00	0.00	1.29
GRR4 Airborne, NNIC	60.8	2.60	2.44	0.91	0.67	0.44	0.41

2.3. GRR3

The third GRR study was also performed for airborne and impact noise isolation of floor-ceiling assemblies. The assemblies tested were two stacks of four units, yielding six parts. The floor plans were identical with receiving room volumes of 85.8 m^3. The rooms were finished but unfurnished. The six parts were tested by three teams with two repeats.

The floor-ceiling assembly was similar to the previous studies, with luxury vinyl plank finish flooring over 25 mm (1 inch) of gypsum concrete over 6 mm (1/4-inch) sound mat, over 16 mm (5/8-inch) plywood floor sheathing on wood joists with 150 mm (6-inch) R19 batt insulation in the stud cavities, and 1 layer of 16 mm (5/8-inch) type X gypsum board on 25-gauge resilient channels.

The results in terms of single number ratings are shown in Table 1.

2.4. GRR4

The fourth GRR study was performed on demising walls between units. Six walls were tested, two on each of three floors, of the same building used in GRR3. The floor plans on both sides of the wall were mirrored, with volumes of 84.8 m^3. The rooms were finished but unfurnished. The six parts were tested by three teams with two repeats.

The wall was constructed with double row of wood studs separated by a nominal 50 mm (2-inch) airspace, with batt insulation in both cavities, and two layers of 16-mm (5/8-inch) type X gypsum board on each side.

3. DISCUSSION

3.1. Wall Variation within Building

GRR4 revealed larger part-to-part variances than the previous studies. On review of the data, it appeared that there was a trend of higher ratings on higher floors in the building. The data was analysed with the floor as a factor instead of the part. The results are shown in Table 2 and graphed in Figure 1. The differences between the floors are statistically significant.

Table 2: Mean and Standard Deviation of NNIC rating as function of floor within the building for GRR4 study.

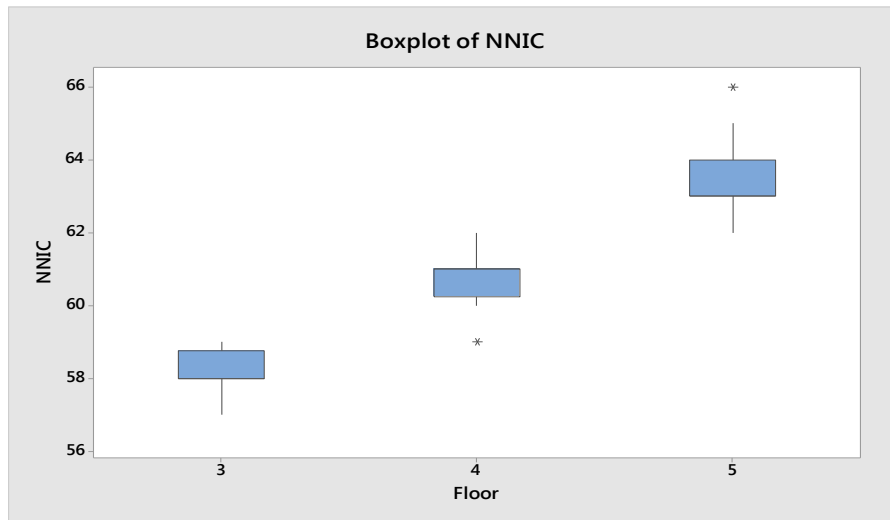| Location	μ	σ
Floor 3	58.1	0.67
Floor 4	60.8	0.75
Floor 5	63.5	1.17

Figure 1: Box plot of NNIC rating of walls as a function of floor.

There is no immediate explanation for such large variation between floors of nominally identical assemblies. The construction crews and material suppliers were not changed. Similar variations have been reported for prefabricated buildings, where the issue was attributed to over-compression of the gaskets between the pre-fabricated sections with increasing load.[9] This explanation is not applicable to this condition.

The average normalized noise reduction values in third-octave bands by floor is shown in Figure 2. The effect is largest around 250 Hz but is seen at all frequencies.

It is typical for wood-framed construction in California that the required shear strength increases at lower floors. Methods for increasing shear strength include increasing the number of shear panels, increasing this thickness of shear panels, increasing the number of nails attaching the shear panels to the studs, increasing the number and size of studs and hold-downs. However, none of the explanations is satisfactory. For example, if a shear panel was installed between the stud rows, previous testing indicates that this can reduce the transmission loss at low frequencies (below about 200 Hz); however, the additional mass tends to increase the transmission loss at higher frequencies. Similarly, if the number of layers of shear panel remained but increased in thickness, the sound insulation should be higher at lower floors.

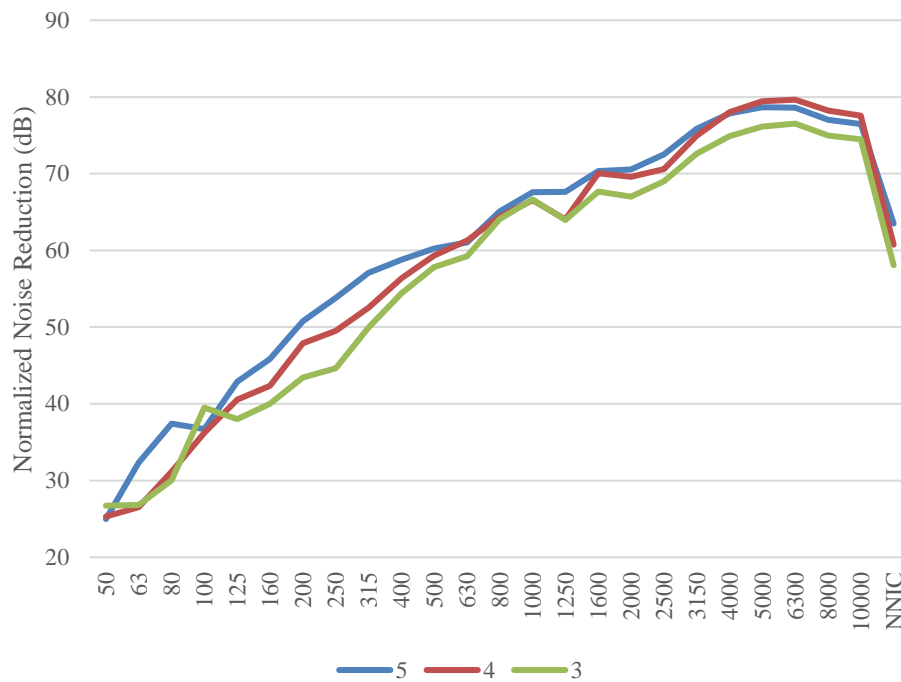The reason for this variation therefore remains unknown.

Figure 2: Averaged airborne isolation spectra by floor for GRR4.

3.2. Part Uncertainty

The part variation for the GRR1, GRR2 and GRR3 is shown in Figure 3. It is reasonable to compare these three studies as all had similar floor-ceiling assemblies.

Several operators for GRR3 had noted that there was the possibility of flanking sound transmission at high frequencies due to the lack of seals on some of the doors of the units. From Figure 3, this appears to be the case. The variation attributable to the assembly was considerably higher for the high frequencies for GRR3 than for GRR1 and GRR2, for both airborne and impact insulation measurements.

At the remaining frequencies, the part-to-part variation of all three studies was very similar. It appears that that absent particular issues or errors, this type of assembly can be constructed with a standard deviation in the range of 1-2 dB at all frequencies.
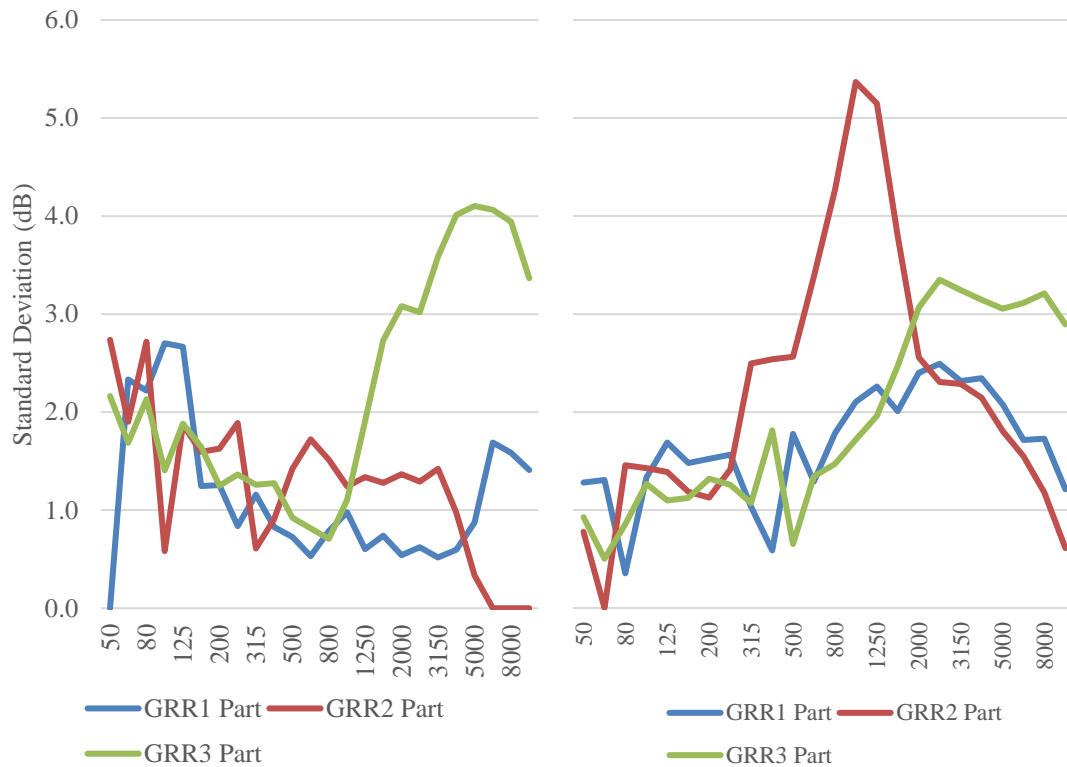
Figure 3: Part-to-Part Component of Variance for GRR1, GRR2, and GRR3 in third-octave bands. (left) Airborne noise isolation. (right) Impact noise isolation.

3.3. Measurement Uncertainty

Figure 4 plots the Gauge Repeatability and Reproducibility for the three studies on floor-ceiling assemblies. The results are broadly similar. Recall that this is the combined variance of the measurement procedure, including that attributable to the variance between operators, to operator-part interaction, and to the repeatability (see Eq. 4). This does not include the variance between testing organizations (traditional reproducibility), which would add additional variability.

The measurement uncertainty is lowest in the mid frequencies and increases at both low and high frequencies. The increase in low-frequency variance is expected, as modal effects increase spatial variation and the bandwidth-time product is reduced. The reasons for the increase in high-frequency measurement uncertainty are not known. One possible reason is differences in self-generated noise due to different scanning procedures, clothing and shoe type between operators and between repeats. Also, changes in background noise between tests were observed due to sources such as crickets, which would predominantly affect the high frequencies.
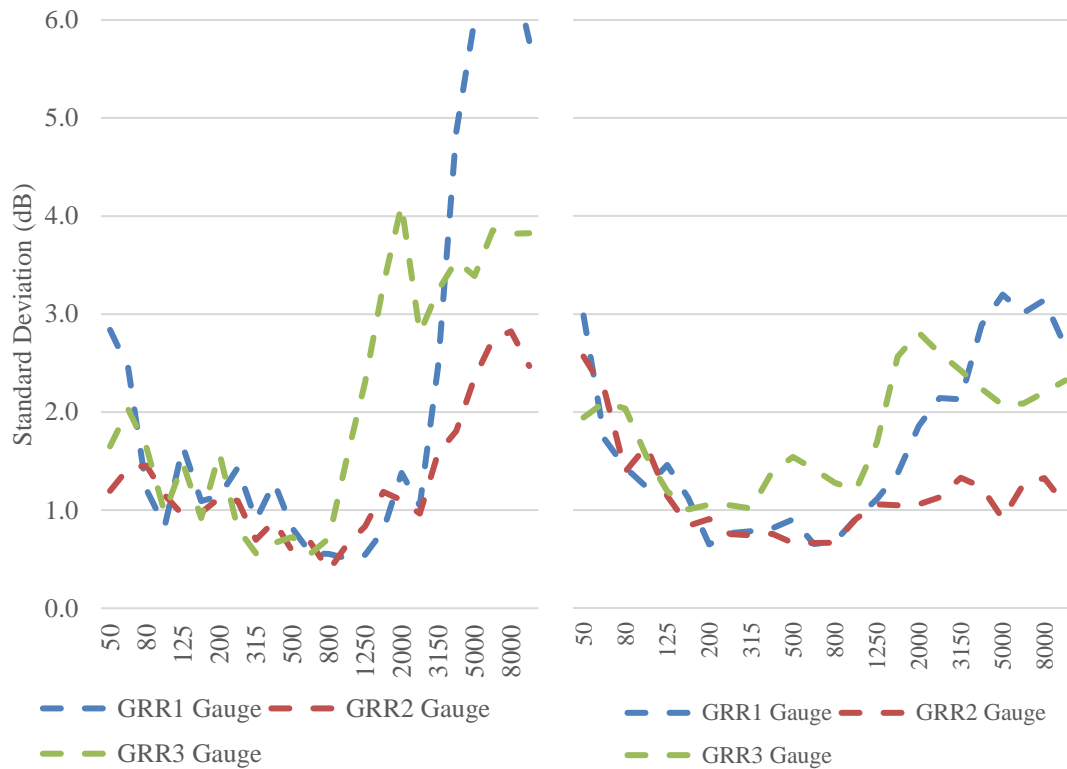
Figure 4: Gauge Repeatability and Reproducibility for GRR1, GRR2, and GRR3 in third-octave bands. (left) Airborne noise isolation. (right) Impact noise isolation.

4. CONCLUSIONS

This paper continues previous work by reporting the results of additional gauge R&R studies. After testing several buildings with similar constructions, the part-to-part variation is observed to be largely consistent and therefore usable as the expected level of variation. Larger-than-expected variations provide quality control feedback to the contractor. In this case, the larger variations are attributable to a mundane cause (flanking noise due to lack of seals). In GRR2, the larger variations in impact noise pointed to a previously unidentified difference in the assemblies.

A firm understanding of the assembly variation will also enable the developer, design team, and acoustical consultant to predict the acceptability of the design. Knowing the variance allows prediction of the likelihood of complaints due to an assembly that happens to be on the low side of the distribution.

Information on part variation is not available in laboratory situations, because it requires testing a large number of nominally identical parts. When an assembly is retested in the laboratory, it is never known whether the variation in the results is due to the measurement process or to the assemblies under test. We say "the field is the new lab" because field testing allows precise measurement of assembly variation, removing much of the guesswork that is too common in the industry.

The methodology allows the total uncertainty of the measurement procedure (within the testing organization) to be extracted even in the presence of large part variation. This information is not available from traditional ILS and other studies. The results can point to aspects of the measurement process that may require improvement, particularly at the low and high portions of the frequency range. The results show that the precision of the

measurements is limited, which in turn limits the ability to discriminate between assemblies.

While additional measurements and effort are required, the GRR experiments and analysis methodology provide otherwise unobtainable information that is valuable to prediction, analysis, quality control, and fundamental understanding of building acoustical performance.

6. REFERENCES

1. *ASTM E691-11, Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method.* (ASTM International, West Conshohocken, PA, 2016).
2. *ASTM E177-14, Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods.* (ASTM International, West Conshohocken, PA, 2014).
3. Whitfield, W. A. & Gibbs, B. M. Measurement Uncertainty in Airborne Sound Insulation: Uncertainty Components in Field Measurement. in *Proc. ICSV 22* (2015).
4. Whitfield, W. A. & Gibbs, B. M. Measurement Uncertainty in Sound Insulation Testing: Comparison of Field Test Data for Separating Floors. in *Proc ICSV24* (2017).
5. Dong, W. & LoVerde, J. Gauge Repeatability and Reproducibility Study of Field Testing of Floor-Ceiling Assemblies. in *Proc ICSV25*
6. Dong, W. & LoVerde, J. Gauge Repeatability and Reproducibility Study of Airborne and Impact Insulation of Floor-Ceiling Assemblies. in *Proc INTER-NOISE 2018*
7. *ASTM E336-16a Standard Test Method for Measurement of Airborne Sound Attenuation between Rooms in Buildings.* (ASTM International, West Conshohocken, PA, 2016).
8. *ASTM Standard E1007, 'Standard Test Method for Field Measurement of Tapping Machine Impact Sound Transmission Through Floor-Ceiling Assemblies and Associated Support Structures'.* (ASTM International, 2013).
9. Öqvist, R., Ljunggren, F. & Ågren, A. Variations in Sound Insulation in Nominally Identical Prefabricated Lightweight Timber Constructions: *Build. Acoust.* (2010).