Northeastern University

College of Engineering

Department of Mechanical and Industrial Engineering

Acoustic Quality of Northeastern Lecture Halls: A Comparative Measurement Study

Submitted by: Jack Cairo (30%), Alder Nadon (40%), Jeremy Yeh (30%)

Date Submitted: April 22, 2026

Course Instructor: Prof. Bridget Smyser

Lab TA: Di Chang

Abstract

There is no public acoustic characterization of Northeastern lecturing spaces, and students perceive a wide range of acoustic quality across different rooms. This project analyzed 54 measured impulse responses across 9 rooms recorded via 10 s exponential sine sweeps and deconvolved with the Farina inverse filter. ISO 3382 parameters (T20, T30, D50, C50, D80, C80, Ts) were extracted and background noise (L_{Aeq}) was measured separately. Parameters varied significantly across rooms, with T20 ranging from 0.40 s in Snell 168 to 1.00 s in EXP 204; D50 from 0.78 to 0.93 and L_{Aeq} from 30.5 to 45.3 dBA. A blocked two-way ANOVA test found room effects highly significant for all parameters ($p < 10^{-6}$). A composite ranking driven by reverberation time, speech definition, and background noise identified ISEC 102 as the best and EXP 204 as the worst rated rooms.

Table of Contents

List of Figures

List of Tables

1 Problem Statement

Northeastern University operates dozens of lecture halls and classrooms that host spoken instruction for thousands of students each day. Anecdotally, students report that some of these rooms sound clear while others feel muddy or hard to follow, yet no publicly available acoustic characterization of the teaching spaces on campus was found during the planning phase of this project. The goal of the project was to measure and compare the acoustic quality of a representative set of these rooms using the standardized parameters defined by ISO 3382, and to determine whether the observed differences between rooms are statistically meaningful.

Nine rooms spanning small classrooms, medium lecture halls, and large auditoria were selected for measurement. Each room was characterized by six impulse response recordings taken at combinations of two source positions and three receiver positions. From the recovered impulse responses, the ISO 3382 parameters T20, T30, D50, C50, D80, C80, and centre time T_s were extracted, and A-weighted equivalent background noise L_{Aeq} was measured separately. The intended outcome was threefold: to produce an objective acoustic ranking of the measured rooms, to identify which rooms meet the ANSI S12.60 background-noise limit for core learning spaces, and to develop a reusable measurement pipeline that can be applied to additional campus spaces in future work.

2 Introduction

Room acoustic measurement has existed as a quantitative discipline since Wallace Clement Sabine established the basic relationship between reverberation time, room volume, and surface absorption at the start of the twentieth century [1]. For most of the century that followed, reverberation time was estimated by exciting a room with a pistol shot, a balloon burst, or bandlimited noise and recording the resulting decay with a sound level meter. Schroeder (1965) later showed that a single impulse response, integrated backward in time, produces a noise-free energy decay curve that is equivalent to averaging many noise-excitation decays [2]. This result reframed room acoustic measurement as a deterministic exercise rather than a statistical one, and it remains the basis of every modern parameter extraction pipeline, including the one used in this project.

Modern impulse response measurement uses deterministic excitation signals rather than impulsive sources. Of the available options, the exponential sine sweep (ESS), proposed by Farina (2000), has become the de-facto standard for room acoustic measurement [3]. An ESS is a sine wave whose instantaneous frequency increases exponentially from a start frequency f_1 to an end frequency f_2 over a fixed duration. Two properties make it attractive for this application. First, its analytical inverse can be constructed in closed form and applied as a fast convolution, avoiding the numerical instability of naive spectral division. Second, any harmonic distortion produced by the loudspeaker is pushed into negative time after deconvolution, where it can be windowed out. The second property matters whenever a non-professional loudspeaker is used, as was the case in this project.

The condenser microphone used in this project operates on the same physical principle as a parallel-plate capacitor. A thin diaphragm sits a short distance from a rigid back plate, and a polarizing voltage maintains

a constant charge on the two plates. Sound pressure displaces the diaphragm, changing the capacitance and producing a small voltage across an internal load resistance. A built-in preamplifier buffers this signal and, in a USB microphone, an integrated analog-to-digital converter delivers the sample stream directly to the host computer. The microphone used here has a cardioid polar pattern, which attenuates sound arriving from behind the diaphragm by roughly 15 to 25 dB relative to sound arriving on-axis. This directivity is convenient for speech recording but is not the pattern specified by ISO 3382-1; the implications of this substitution are discussed in Section 3.5 and Section 5.

The parameters reported in this study are those defined in ISO 3382-1 and ISO 3382-2 [4, 5] and are summarized with their governing equations and literature target ranges in Table 1. Reverberation time T_{60} is the time required for sound energy to drop by 60 dB after the source stops. Because 60 dB of usable dynamic range is rarely available in a real measurement, the standard defines T_{20} and T_{30}, which are computed by linear regression on the portion of the energy decay curve between -5 dB and -25 dB (or -35 dB) and extrapolated to a -60 dB decay. Definition D50 and clarity C50 quantify the fraction and ratio of acoustic energy arriving within the first 50 ms of the direct sound, which corresponds to the psychoacoustic window over which the ear integrates reflections as support for a talker rather than as separate echoes. D80 and C80 use an 80 ms boundary and are more relevant for music than for speech. Centre time T_s is the energy-weighted mean arrival time of the impulse response and has the advantage of not requiring an arbitrary boundary.

The speech intelligibility literature suggests specific target ranges for these parameters in lecture-style rooms. Rakerd et al. (2018) survey classroom acoustics and report that reverberation times between 0.4 s and 0.8 s are generally preferred for speech, with shorter times sounding dead and longer times blurring consonants [6]. ISO 3382-1 recommends D50 above 0.5 and C50 above 0 dB as indicators of good speech clarity. ANSI S12.60 sets a 35 dB(A) limit on continuous background noise in core learning spaces [7], and the World Health Organization's Guidelines for Community Noise echo the same figure for classrooms [8]. These target ranges are collected in the rightmost column of Table 1 and appear as shaded bands in the parameter comparison figures presented in Section 5.

The recorded microphone signal $y(t)$ is the convolution of the loudspeaker signal $x(t)$ with the room impulse response $h(t)$, plus additive noise $n(t)$:

$$y(t) = x(t) * h(t) + n(t). \tag{1}$$

For an ESS with sweep rate $R = \ln(f_2/f_1)$, the Farina analytical inverse filter is constructed by time-reversing the active sweep and multiplying it by the amplitude envelope $\exp(-nR/(N-1))$, where N is the number of samples in the active sweep. Convolving the recorded signal with this inverse filter recovers an estimate of $h(t)$. From the estimated impulse response the energy decay curve is then computed by Schroeder backward integration,

$$\mathrm{EDC}(t) = \int_t^{\infty} h^2(\tau)\, d\tau, \tag{2}$$

with a noise-energy correction due to Chu [9] subtracted point by point, and with the upper limit of integration replaced by a truncation point chosen by the iterative procedure of Lundeby et al. [10]. The ISO 3382

parameters are then computed as integrals of $h^2(t)$ over the appropriate early or late time windows, or by linear regression on the EDC in the case of T_{20} and T_{30}.

Table 1: Acoustic parameters reported in this study, with defining expressions, units, and literature target ranges for speech in lecture spaces. $h(t)$ is the room impulse response and EDC is the Schroeder energy decay curve. Targets are drawn from ISO 3382-1 [4], Rakerd et al. [6], and ANSI S12.60 [7]; an em dash indicates no established single-number target in the literature.

Symbol	Definition	Units	Target
T_{20}	Linear regression on the EDC over -5 to -25 dB, extrapolated to a -60 dB decay.	s	0.4–0.8
T_{30}	Linear regression on the EDC over -5 to -35 dB, extrapolated to a -60 dB decay.	s	0.4–0.8
D_{50}	$\int_0^{t_{50}} h^2(t)\,dt \,/\, \int_0^{\infty} h^2(t)\,dt$, where $t_{50} = 50$ ms.	—	> 0.5
C_{50}	$10 \log_{10}\left[\int_0^{t_{50}} h^2(t)\,dt \,/\, \int_{t_{50}}^{\infty} h^2(t)\,dt\right]$	dB	> 0
D_{80}	As D_{50}, with $t_{80} = 80$ ms in place of t_{50}.	—	—
C_{80}	As C_{50}, with $t_{80} = 80$ ms in place of t_{50}.	dB	-2 to 5
T_s	$\int_0^{\infty} t\,h^2(t)\,dt \,/\, \int_0^{\infty} h^2(t)\,dt$	ms	—
L_{Aeq}	A-weighted time-averaged sound pressure level, measured once per room in the empty-room condition with the measurement equipment idle.	dB(A)	≤ 35

Few published studies have applied this methodology to United States university teaching spaces directly. Rakerd et al. [6] present case studies from K-12 and small university rooms, and provide the closest available reference numbers for comparison, but no public acoustic characterization of Northeastern University's teaching spaces was found during the planning phase of this project.

3 Procedure

3.1 Rooms Measured

Nine rooms on the Northeastern Boston campus were selected to span a range of volumes and construction types. The selection is listed in Table 2 and shown in Figure 1, and includes two large auditoriums (ISEC 102, WVF 020), two large lecture halls (Snell 168, Mugar 201), one medium lecture hall (Robinson 109), three medium classrooms (EXP 204, EV 002, WVG 108), and one small classroom (Shillman 215).

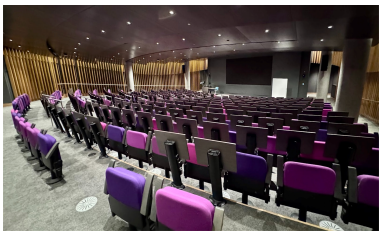
3.2 Equipment

Measurements were made with a Jounivo JV-601 USB condenser microphone with a cardioid polar pattern. The manufacturer rates the microphone at a sensitivity of -38 ± 2 dB re 1 V/Pa and a signal-to-noise ratio

Table 2: Measured rooms and their general type.

ID	Room	Seats	Type
1	EXP 204	65	Medium classroom
2	ISEC 102	275	Large auditorium
3	Snell Engineering 168	184	Large lecture hall
4	Robinson 109	74	Medium lecture hall
5	East Village 002	59	Medium classroom
6	Mugar 201	187	Large lecture hall
7	West Village G 108	73	Medium classroom
8	West Village F 020	257	Large auditorium
9	Shillman 215	52	Small classroom

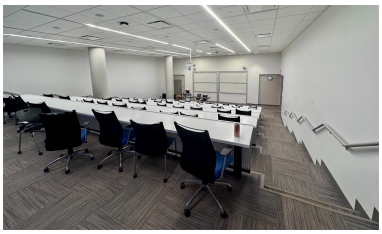
(1) EXP 204



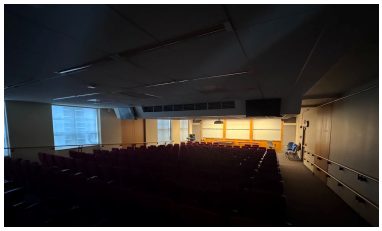
(2) ISEC 102

(3) Snell Engineering 168

(4) Robinson 109



(5) East Village 002



(6) Mugar 201

(7) West Village G 108

(8) West Village F 020

(9) Shillman 215

Figure 1: Photographs of the nine lecture halls and classrooms measured in this study, arranged in the same order as Table 2. Numbers in parentheses correspond to the room IDs in that table.

of 76 dB(A), which corresponds to an equivalent input noise of roughly 18 dB(A). The microphone was mounted on a stand with a clip and oriented toward the source for every measurement. An omnidirectional measurement microphone was originally specified for the project but could not be obtained before the scheduled measurement day; the implications of this substitution are discussed in Section 3.5 and Section 5.

The sweep was reproduced through a portable powered JBL Flip 5 loudspeaker. An ASUS Zephyrus G14 laptop running Audacity 3.7.7 was used to record the microphone signal through the JV-601's built-in USB audio interface at 48 kHz and 24-bit resolution. The same microphone, loudspeaker, and laptop were used for all 54 measurements to ensure internal comparability across rooms. Absolute sound pressure levels were not calibrated. This does not affect the extracted ISO 3382 parameters, which depend on energy ratios and decay rates rather than absolute levels.

A-weighted background noise L_{Aeq} was measured with a phone-based sound level meter application (Decibel Meter) on an iPhone 16. The reading was taken as an instantaneous value with the room empty and the measurement equipment idle. The displayed level was observed to be stable (fluctuating by roughly ± 0.5 dB) over a short observation window of 30 seconds, so time-averaging or data streaming was not performed and a single representative value was recorded for each room.

3.3 Excitation Signal

A 10-second exponential sine sweep from 20 Hz to 20 kHz was generated by the MATLAB script `generate_ess.m` at a sample rate of 48 kHz and a bit depth of 24. One second of silence was prepended to the sweep and three seconds of silence were appended, giving a total file duration of 14 seconds. The peak level was set to -3 dBFS to provide headroom against clipping at the loudspeaker input. The sweep parameters were $f_1 = 20$ Hz, $f_2 = 20{,}000$ Hz, and $R = \ln(f_2/f_1) = 6.9078$. The excitation signal, minus the lead and fade silence, is described by the equation:

$$s(t) = g\,\sin\!\left(\frac{2\pi f_1 T}{R}\left(e^{Rt/T} - 1\right)\right), \qquad R = \ln\!\left(\frac{f_2}{f_1}\right) \qquad g = \frac{\text{gain}_{\text{dB}}}{20} \tag{3}$$

3.4 Measurement Protocol

All 54 measurements were taken on April 4, 2026. Each room was measured with six source-receiver configurations: two source positions crossed with three receiver positions. The two source positions were a front-center location, corresponding to the typical lectern or instructor position, and a front-offset location. Both source positions placed the loudspeaker at roughly 1.5 m above the floor and oriented toward the seating area. The three receiver positions were located near the front, middle, and back of the seating, at approximately 1.2 m above the floor (seated ear height), at least 1 m from any wall, and at least 2 m from the nearest source. The microphone was pointed at the source for every measurement. The geometry used in every room

is illustrated in Figure 2.

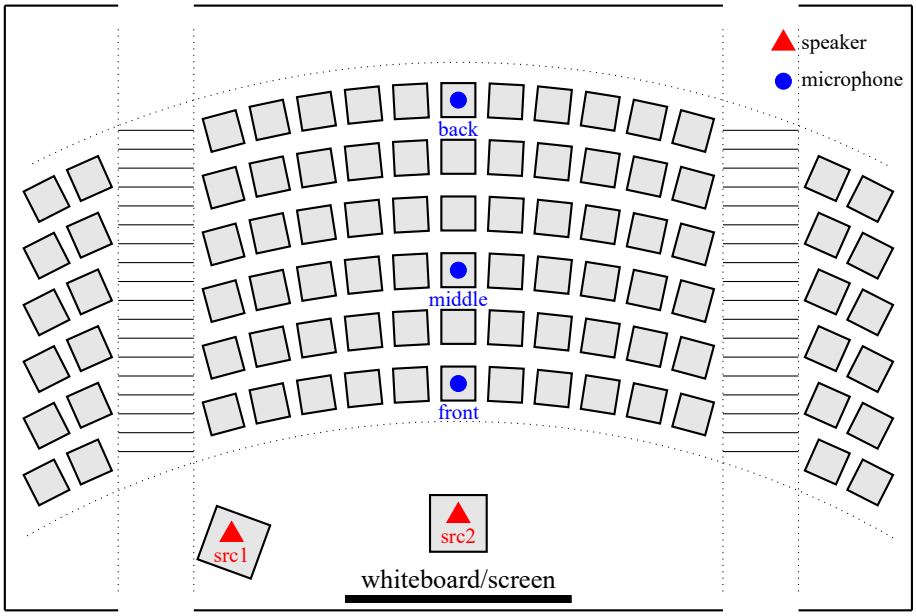


Figure 2: Source-receiver geometry used in every room. Two source positions (front-center and front-offset) were each combined with three receiver positions (front, middle, and back of the seating) for a total of six measurements per room.

The measurement session was run by a three-person team, with each member assigned a single task. One member operated the microphone and recording laptop, a second member operated the loudspeaker and managed source placement, and the third member took the L_{Aeq} readings with the phone application. Each member performed the same role in all nine rooms to reduce operator variability.

Each room was measured with all doors closed, with no occupants other than the three-person team, and with the HVAC state documented in a per-room metadata file. L_{Aeq} was recorded in the same empty-room condition but with the measurement equipment idle, before any sweep was played. For each recording, the sweep was played once and the microphone signal was captured as a single WAV file. The recorded peak level was verified to be below digital clipping and the waveform was inspected for a visible direct-sound peak and decay tail before the next position was measured.

3.5 Deviations from the Planned Procedure

Two deviations from the originally planned procedure were made before the measurement day. First, as noted above, the intended omnidirectional measurement microphone was unavailable and was replaced with the cardioid Jounivo JV-601 already on hand. The cardioid pattern attenuates reflections arriving from behind the microphone and biases the extracted D50 and C50 values upward relative to an omnidirectional measurement. Because the same microphone with the same orientation was used in every room, the bias is a consistent offset across the data set and relative comparisons between rooms remain valid. Octave-band analysis was removed

from the scope of the project for the same reason: the cardioid response and the JV-601's uncalibrated frequency response are not flat enough to support reliable band-limited parameter extraction. All reported parameters are broadband.

Second, the ISO 3382-1 reference source is a dodecahedral omnidirectional loudspeaker. No such source was available, so a single portable powered loudspeaker was used for all measurements. The loudspeaker was kept in the same orientation (facing into the seating area) in every room, so the source-directivity bias is also approximately constant across the data set.

3.6 Analysis Pipeline

Impulse response recovery and parameter extraction were implemented in MATLAB. The complete pipeline is orchestrated by `main.m`, which loops over rooms, loads room-level metadata, and calls `deconvolve.m` once per measurement. The signal flow is summarized in Figure 3.

The recorded WAV file was trimmed to the active sweep region by thresholding at 5% of the peak absolute amplitude. The Farina inverse filter was then built as shown in Figure 3 and normalized so that convolving the original sweep with the inverse produced a unit-amplitude peak. FFT-based fast convolution of the recording with the inverse filter, zero-padded to avoid circular-convolution artifacts, produced the recovered impulse response; the direct-sound peak was located in the first half of the output (the second half can contain wrapped harmonic distortion products) and a 1.0 s window starting 2 ms before the peak was retained.

The noise-floor truncation point was found with the iterative algorithm of Lundeby et al. [10]. The impulse response was partitioned into short energy windows, a regression line was fit to the decay, and the algorithm converged on the intersection of the extrapolated regression with the estimated noise floor. The energy window length was set adaptively from a coarse reverberation-time estimate so that roughly thirty windows spanned the decay region; this avoided the need to tune a window length per room. Once the truncation point was fixed, Schroeder backward integration was applied to the squared impulse response, and a Chu noise-energy correction [9] subtracted the expected contribution of the noise from each point of the EDC.

Reverberation times and early-energy parameters were then extracted from the EDC and the impulse response according to the definitions in Table 1; the coefficient of determination R^2 of each T_{20} regression was stored so that measurements with non-linear decays could be flagged. Per-room spatial means and standard deviations were computed across the six source-receiver combinations.

Cross-room statistical analysis was performed by `analyze_comparison.m`, which fit a two-way additive ANOVA (room and source-receiver position as factors) to each parameter, performed Tukey HSD post-hoc comparisons between all 36 room pairs, and computed the composite speech-quality score and its confidence interval. The statistical model and composite construction are described in Section 4.5 and Section 4.6.

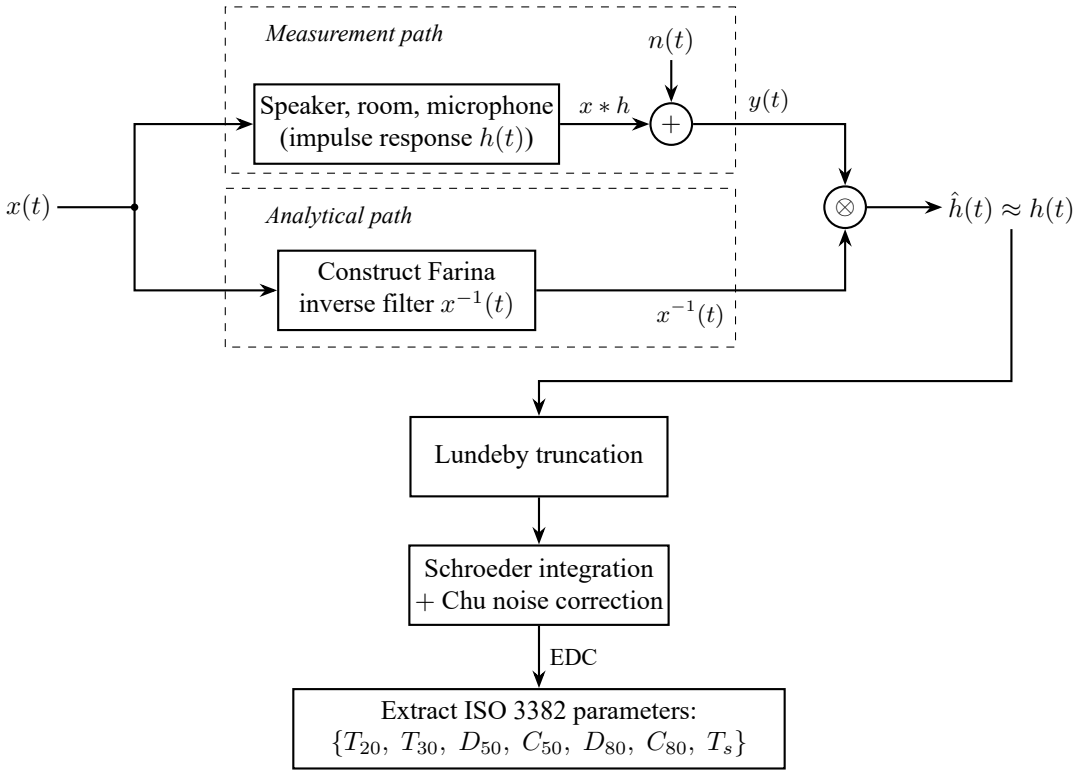


Figure 3: Block diagram of the impulse response recovery and parameter extraction pipeline. The top rail is the physical signal chain: the sweep $x(t)$ excites the speaker, room, and microphone (with room impulse response $h(t)$) and the summing junction adds ambient noise $n(t)$ to produce the recorded signal $y(t) = x(t) * h(t) + n(t)$. The bottom rail constructs the Farina analytical inverse filter $x^{-1}(t)$ from the known sweep. Convolving the recording with the inverse filter ($\otimes$) yields an estimate $\hat{h}(t)$ of the room impulse response, because $x(t) * x^{-1}(t) \approx \delta(t)$ by construction. The recovered impulse response is then truncated by the Lundeby algorithm, integrated by Schroeder backward integration with the Chu noise correction to produce the energy decay curve (EDC), and the ISO 3382 parameters are extracted from $\hat{h}(t)$ and the EDC.

3.7 Validation

Before the pipeline was applied to the campus data, its correctness was checked against a synthetic ground-truth test case implemented in `validation/ess_deconvolution.m`. A known room impulse response was convolved with the generated ESS and contaminated with additive white noise to produce a test recording. The deconvolved impulse response was then compared against the known reference by the normalized misalignment (NM),

$$\mathrm{NM} = 10 \log_{10} \frac{\|\hat{h} - h_{\mathrm{ref}}\|^2}{\|h_{\mathrm{ref}}\|^2}, \tag{4}$$

where $\hat{h}$ is the estimated impulse response and h_{ref} is the known reference. Two inverse-filter implementations were compared: the Farina analytical inverse used in the production pipeline and a smoothed-envelope spectral-division alternative. The Farina inverse achieved $\mathrm{NM} = -11.4$ dB with a sidelobe energy of 74.15 dB compared to a peak-region energy of 90.40 dB, and required no regularization parameter. The spectral-division approach reached comparable accuracy only after manual tuning of its regularization parameter and was therefore rejected as less robust for batch processing. The residual sidelobe energy in the Farina case is a known property of the discrete-time approximation [3] and lies well below the range needed for the broadband energy ratios reported here.

4 Results

4.1 Example Measurement

Figure 4 illustrates the output of the pipeline for a single measurement, taken at the `src1_back` position in EXP 204 (the largest and most reverberant room in the study). The top panel shows the deconvolved impulse response in the time domain: a sharp direct-sound peak near the start of the trace is followed by a dense series of early reflections and a slowly decaying reverberant tail. The middle panel shows the same signal in decibels with the Lundeby truncation point overlaid; the decay is visibly close to a straight line over roughly 40 dB of dynamic range before the noise floor is reached. The bottom panel shows the energy decay curve computed from the truncated impulse response, with the T_{20} regression line and its extrapolation to -60 dB drawn over it.

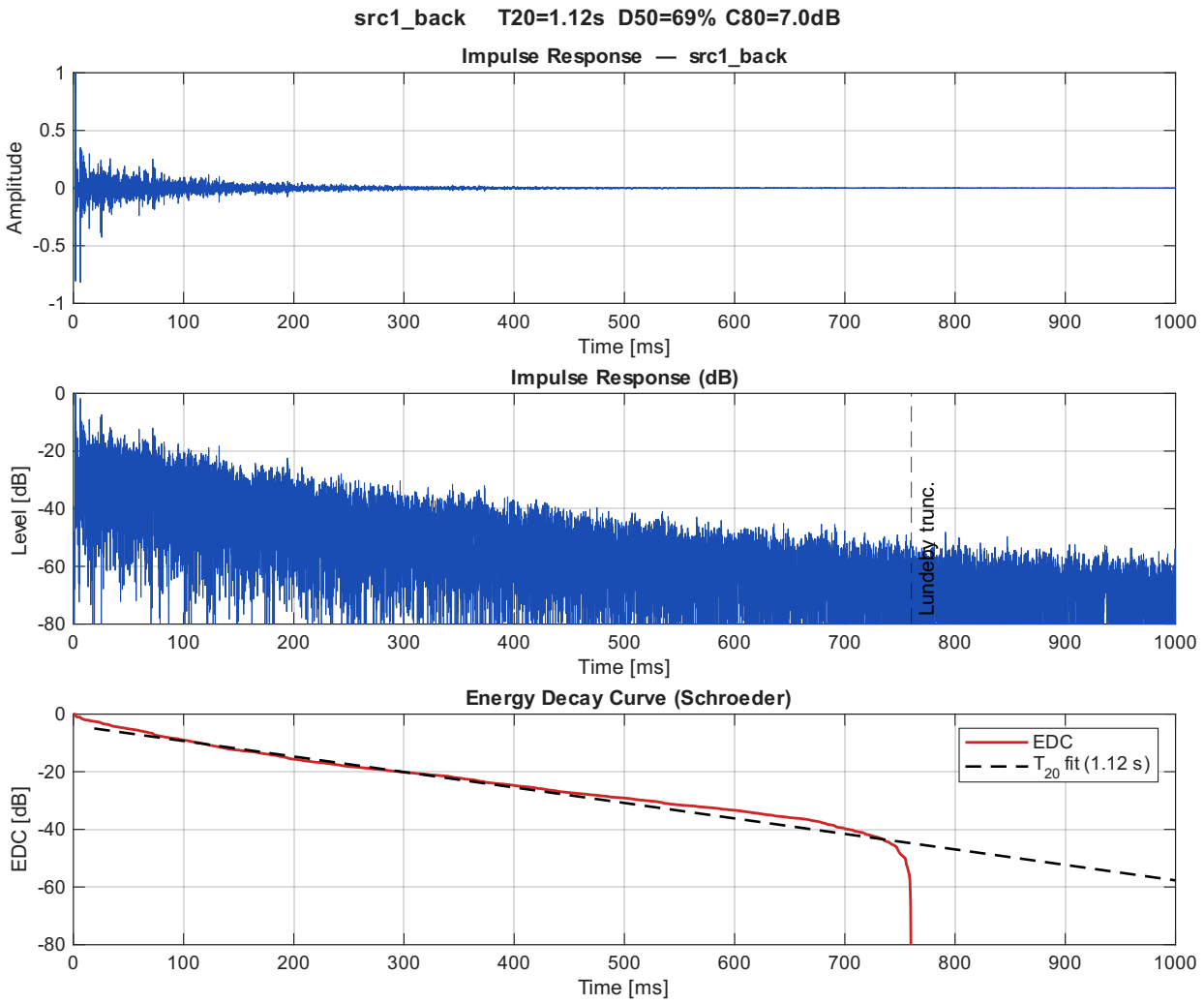


Figure 4: Three-panel impulse response figure for EXP 204 at the rear receiver position from source 1. Top: impulse response in the time domain; middle: impulse response in decibels with the Lundeby truncation point marked; bottom: energy decay curve with the T_{20} regression line.

4.2 Per-Room Spatial Behavior

The per-room summary for West Village F 020 is shown in Figure 5 as a representative example. Two features of this figure are visible in the wider data set. First, D50 and T_s show a systematic front-to-back gradient: D50 decreases and T_s increases as the microphone moves away from the source, because the reverberant energy grows relative to the direct energy with distance. Second, T_{20} is far more spatially stable than the clarity-related parameters: the six T_{20} values in this room cluster within roughly $\pm 10\%$ of their mean. This is consistent with reverberation time being, to first order, a property of the room as a whole rather than of a particular source-receiver pair. The same general behavior was observed in every room.

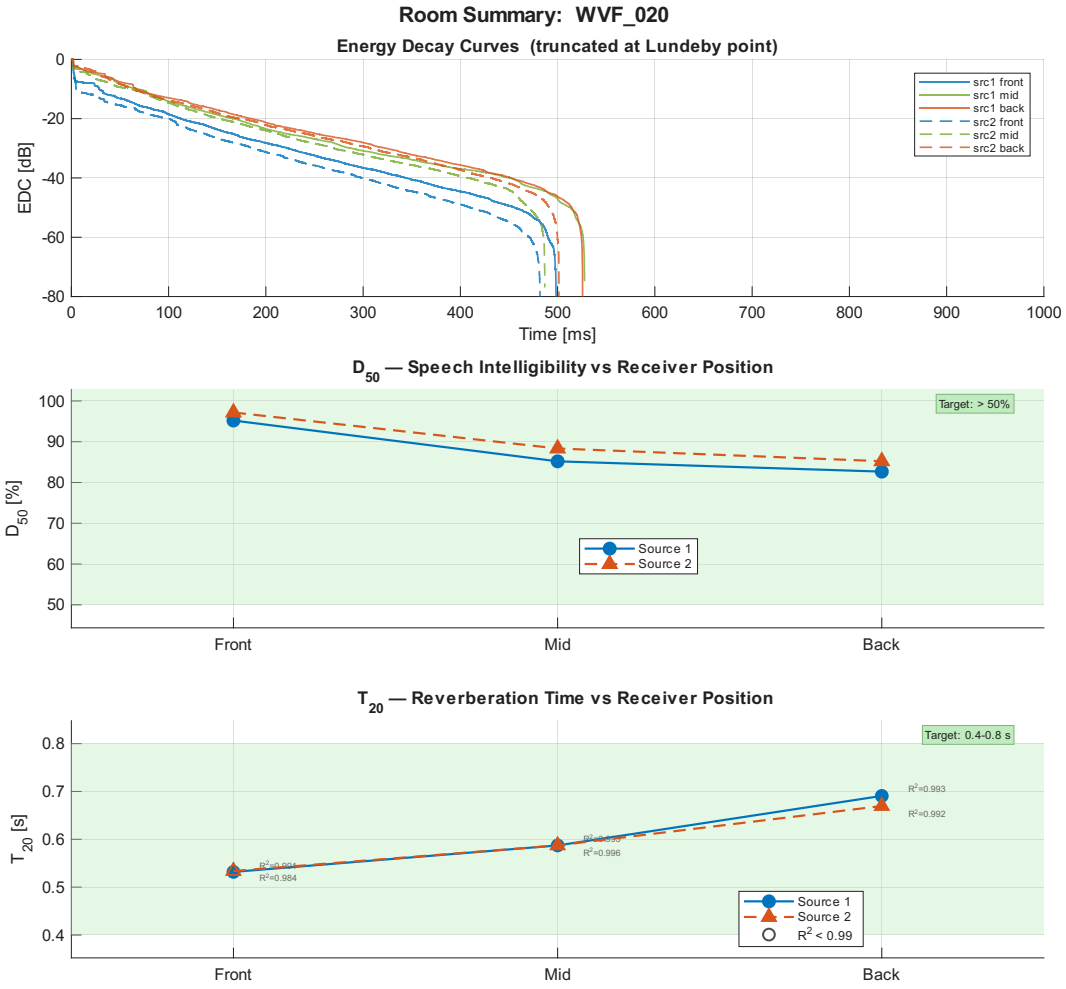


Figure 5: Per-room summary figure for West Village F 020, showing energy decay curves for all six source-receiver positions on the left and dot-strips of each ISO 3382 parameter across the six positions on the right. Error bars on the dot-strips are $\pm 1\sigma$ across positions. Open markers on the T_{20} panel indicate measurements with $R^2 < 0.99$ for the linear regression on the EDC.

4.3 Room-Level Summary

Table 3 presents the spatially averaged parameters for all nine rooms, computed across the six source-receiver configurations measured in each room. Each tabulated value is the sample mean across the six positions, with the sample standard deviation given after the $\pm$ symbol. T_s is reported in milliseconds, reverberation times in seconds, and clarity parameters in decibels.

Table 3: Spatially averaged ISO 3382 parameters for all nine rooms. Values are the sample mean $\pm$ standard deviation across the six source-receiver positions. Background noise L_{Aeq} is a single reading per room and is reported alongside the composite ranking in Table 5.

Room	T_{20} (s)	T_{30} (s)	D_{50}	C_{50} (dB)	D_{80}	C_{80} (dB)	T_s (ms)
EXP 204	1.003 ± 0.166	1.124 ± 0.122	0.776 ± 0.116	6.05 ± 3.62	0.877 ± 0.068	9.45 ± 3.97	33.9 ± 15.9
ISEC 102	0.551 ± 0.083	0.561 ± 0.062	0.887 ± 0.042	9.22 ± 1.95	0.945 ± 0.024	12.74 ± 2.19	18.8 ± 4.9
Snell 168	0.398 ± 0.016	0.412 ± 0.010	0.928 ± 0.008	11.12 ± 0.46	0.975 ± 0.005	15.89 ± 0.79	14.4 ± 1.8
Robinson 109	0.468 ± 0.096	0.534 ± 0.069	0.930 ± 0.027	11.55 ± 2.08	0.973 ± 0.012	16.00 ± 2.31	15.1 ± 3.8
EV 002	0.647 ± 0.128	0.769 ± 0.141	0.893 ± 0.052	9.70 ± 2.57	0.949 ± 0.027	13.30 ± 2.66	19.5 ± 7.6
Mugar 201	0.402 ± 0.056	0.432 ± 0.060	0.912 ± 0.049	10.69 ± 2.55	0.969 ± 0.020	15.69 ± 2.87	17.3 ± 6.5
WVG 108	0.635 ± 0.040	0.709 ± 0.050	0.887 ± 0.049	9.28 ± 2.10	0.940 ± 0.024	12.25 ± 1.78	19.3 ± 6.4
WVF 020	0.600 ± 0.067	0.669 ± 0.079	0.890 ± 0.059	9.86 ± 3.49	0.951 ± 0.024	13.54 ± 3.05	18.6 ± 8.6
Shillman 215	0.563 ± 0.040	0.603 ± 0.030	0.921 ± 0.031	11.05 ± 2.22	0.964 ± 0.014	14.71 ± 2.16	15.0 ± 4.1

EXP 204 has the longest reverberation time of the nine rooms at $T_{20} = 1.00$ s, roughly double the next-longest values and well outside the lecture range of 0.4–0.8 s recommended by ISO 3382-1 and by Rakerd et al. [6]. Snell 168 and Mugar 201 sit at the opposite end of the scale with $T_{20} \approx 0.40$ s, close to the lower edge of the recommended band. Background noise spans a wide range, from 30.5 dB(A) in ISEC 102 to 45.3 dB(A) in EV 002, with only ISEC 102 and EXP 204 meeting the 35 dB(A) ANSI S12.60 target. Every measured room had D50 above 0.75 and C50 above 0 dB, which would formally meet the ISO 3382-1 indicators for good speech clarity. As discussed in Section 5, however, these elevated early-energy values are partly attributable to the cardioid directivity of the measurement microphone and should be interpreted on a relative rather than absolute basis.

4.4 Cross-Room Comparison

Figure 6 plots each parameter across rooms, with the spatial standard deviation drawn as an error bar and the literature target ranges drawn as shaded bands. Two observations are apparent. First, EXP 204 is a clear outlier on every parameter that is sensitive to reverberation: its T_{20} and T_{30} means lie well above the target band and its D50, C50, D80, and C80 means lie below those of the rest of the rooms. Second, the remaining eight rooms cluster tightly on the clarity-related parameters, with means that fall inside or just above the target bands, and differ from each other primarily in their reverberation time and background noise level.

The same data are presented as a normalized heatmap in Figure 7. Rooms are ordered by composite ranking (Section 4.6), and each cell is coloured by the z-score of that parameter relative to the cross-room distribution.

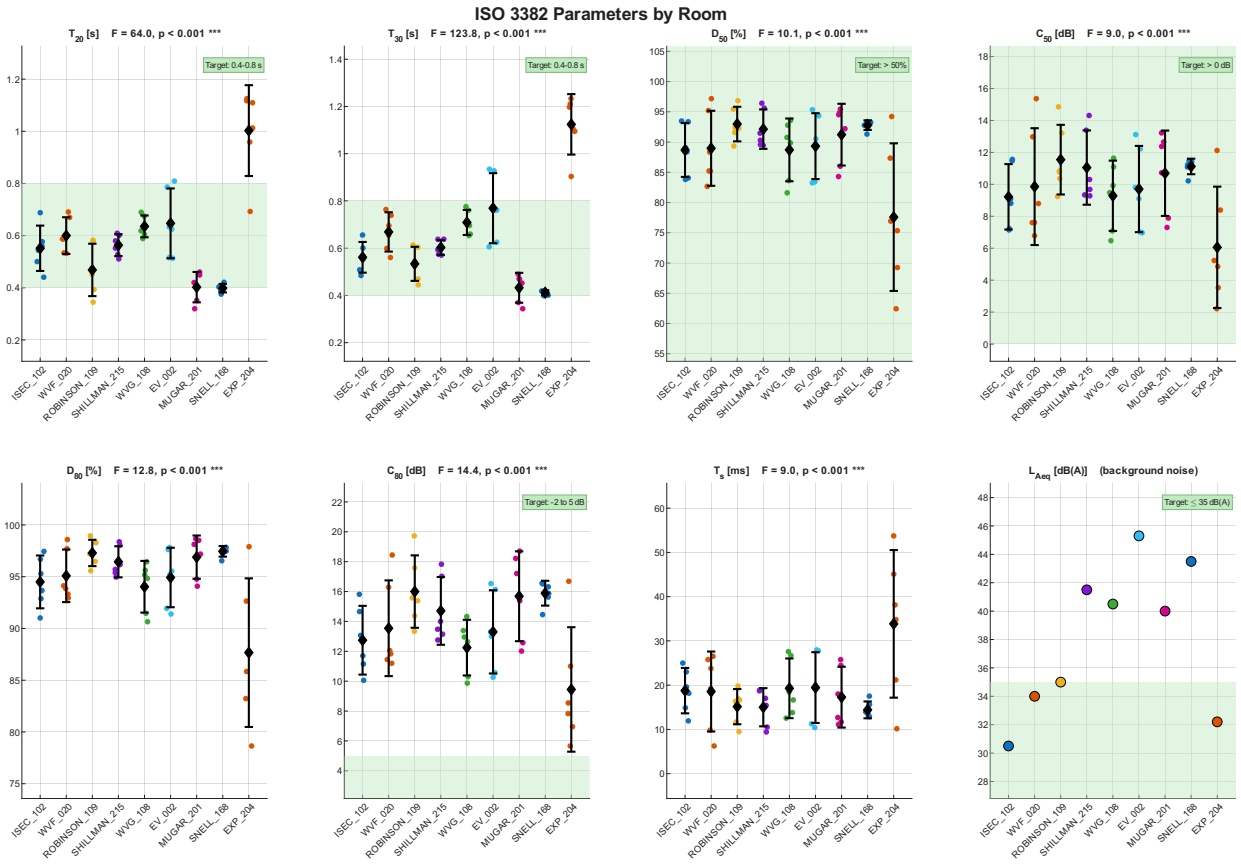


Figure 6: Cross-room comparison of ISO 3382 parameters and background noise. Points are spatial means, error bars are $\pm 1\sigma$ across the six source-receiver positions (there are no error bars on L_{Aeq}, which was a single reading per room). Shaded bands mark the literature target ranges from ISO 3382-1, Rakerd et al. [6], and ANSI S12.60 [7].

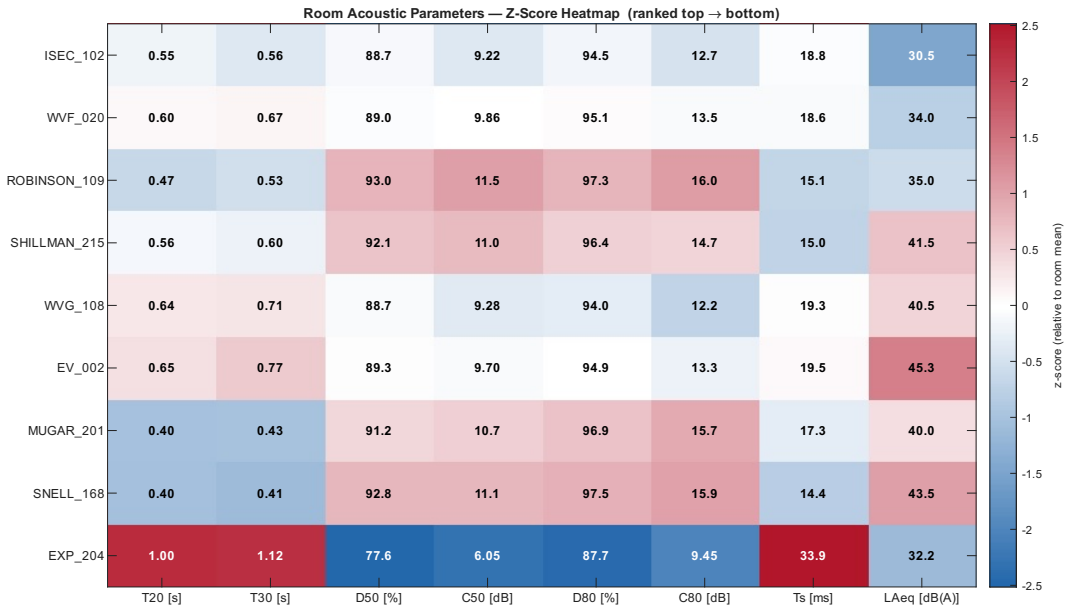


Figure 7: Heatmap of ISO 3382 parameters and background noise across all nine rooms, normalized as z-scores within each parameter. Rooms are sorted top-to-bottom by composite ranking. Blue cells are better than the cross-room average for that parameter; red cells are worse.

4.5 Statistical Significance

A two-way additive ANOVA was fit to each parameter with room and source-receiver position as factors. The additive model

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \tag{5}$$

was used because the balanced 9×6 design provides exactly one observation per cell and cannot support an interaction term. Y_{ij} is the measured parameter in room i at position j, μ is the grand mean, α_i is the room effect, β_j is the position effect, and ε_{ij} is the residual. The error degrees of freedom for this design are $(9-1)(6-1) = 40$. The blocked design was preferred over a one-way ANOVA on room alone because the source-receiver position is known to be a systematic source of within-room variance (as shown by the spatial gradients in D50 and T_s in Section 5.2); treating position as a blocking factor partitions this variance out of the error term and increases statistical power.

Table 4 reports the room-effect F statistic, room-effect p-value, position-effect F statistic, and position-effect p-value for each parameter. For all seven ISO 3382 parameters, the room effect was significant at $\alpha = 0.05$ with $p < 10^{-6}$. The F-statistics ranged from $F_{9,40} = 8.96$ (for C50) to $F_{9,40} = 123.8$ (for T_{30}). The position effect was also significant for every parameter at the same α level, which confirms that the source-receiver position carries real systematic variance and that the blocked design was justified.

Tukey HSD post-hoc comparisons were computed on each significant room effect, using the Tukey-Kramer method to control the family-wise error rate at $\alpha = 0.05$ across all 36 pairwise comparisons. Two summary

Table 4: Two-way additive ANOVA (room and position factors) for each ISO 3382 parameter. Room-effect and position-effect F-statistics and p-values are reported for an error degrees of freedom of 40.

Parameter	F_{room}	p_{room}	F_{position}	p_{position}	**Significant?**
T_{20}	64.0	2.3×10^{-20}	14.9	3.0×10^{-8}	Yes
T_{30}	123.8	9.6×10^{-26}	17.3	4.2×10^{-9}	Yes
D_{50}	10.1	1.4×10^{-7}	13.6	8.7×10^{-8}	Yes
C_{50}	8.96	6.2×10^{-7}	22.7	9.9×10^{-11}	Yes
D_{80}	12.8	6.4×10^{-9}	10.9	1.2×10^{-6}	Yes
C_{80}	14.4	1.2×10^{-9}	23.2	7.3×10^{-11}	Yes
T_s	8.98	6.0×10^{-7}	14.9	3.0×10^{-8}	Yes

observations are worth noting. EXP 204 differed significantly from every other room on T_{20}, T_{30}, D_{50}, D_{80}, C_{50}, C_{80}, and T_s. Snell 168 and Mugar 201 were statistically indistinguishable from each other on T_{20} but both differed significantly from the longer-reverberation group (ISEC 102, Shillman 215, WVG 108, WVF 020, EV 002, and EXP 204). The complete pairwise matrices are available in the raw output of `analyze_comparison.m`.

4.6 Composite Speech-Quality Score

A composite score was computed for each room from a weighted combination of three standardized parameters: D_{50} (weight 0.40), a T_{20} penalty relative to a 0.7 s target (weight 0.25), and L_{Aeq} (weight 0.35). The parameters were chosen to cover three approximately independent dimensions relevant to speech intelligibility: early-energy ratio, reverberation time, and background noise level. The weights reflect the relative importance of each dimension in the speech intelligibility literature; the slight down-weighting of L_{Aeq} relative to its importance accounts for the higher measurement uncertainty of its single phone-based reading compared to the six-replicate acoustic parameters.

Each parameter was converted to a z-score across the nine rooms so that higher scores always indicate better speech quality:

$$z_{D50,i} = \frac{D_{50,i} - \overline{D_{50}}}{s_{D_{50}}}, \qquad z_{T20,i} = \frac{-|T_{20,i} - 0.7| - \overline{-|T_{20} - 0.7|}}{s_{-|T_{20}-0.7|}}, \qquad z_{L,i} = \frac{-L_{Aeq,i} - \overline{-L_{Aeq}}}{s_{-L_{Aeq}}}. \tag{6}$$

The absolute-value construction of z_{T20} penalizes reverberation times that are either shorter or longer than 0.7 s, which is the centre of the 0.4–0.8 s range recommended for lecture spaces. The composite score is

$$C_i = 0.40\, z_{D50,i} + 0.25\, z_{T20,i} + 0.35\, z_{L,i}. \tag{7}$$

A 95% confidence interval was attached to each composite by linear error propagation from the within-cell

covariance matrix of the D_{50} and T_{20} residuals under the blocked ANOVA model, with the critical value $t_{0.975,\,40} = 2.021$. The L_{Aeq} term contributes no variance to this interval because it was measured only once per room, so the reported confidence intervals are lower bounds on the true composite uncertainty.

The resulting ranking is given in Table 5 and visualized in Figure 8. ISEC 102 has the highest composite score (+0.54), driven primarily by its very low background noise (30.5 dB(A), the lowest measured) combined with a reverberation time within the recommended 0.4–0.8 s lecture band. The next two ranks (WVF 020 and Robinson 109) are within one confidence interval of ISEC 102 and cannot be separated from each other at $\alpha = 0.05$. The middle four rooms (Shillman 215, WVG 108, EV 002, Mugar 201) also have overlapping confidence intervals and are not individually separable. EXP 204 is the lowest-ranked room with a composite score of -0.89, driven by its reverberation time of roughly 1.0 s and its comparatively low D_{50}, despite a modest L_{Aeq} of 32.2 dB(A). The confidence interval for EXP 204 is slightly larger than those of the other rooms because its T_{20} lies far above the 0.7 s target, which changes the sign of the T_{20} Jacobian relative to the cross-covariance term in the propagation formula.

Table 5: Composite speech-quality ranking of the nine rooms. T_{20} and D_{50} are room means across six source-receiver positions; L_{Aeq} is a single instantaneous reading. Composite scores are defined in Equation 7. The $\pm$ column is the 95% confidence interval from linear error propagation with $t_{0.975,\,40} = 2.021$; this interval reflects only the D_{50} and T_{20} uncertainty from the blocked design and does not include L_{Aeq} measurement error.

Rank	**Room**	T_{20} (s)	D_{50}	L_{Aeq} (dB(A))	**Composite**	$\pm$95% CI
1	ISEC 102	0.551	0.887	30.5	$+0.54$	0.21
2	WVF 020	0.600	0.890	34.0	$+0.45$	0.21
3	Robinson 109	0.468	0.930	35.0	$+0.41$	0.21
4	Shillman 215	0.563	0.921	41.5	$+0.13$	0.21
5	WVG 108	0.635	0.887	40.5	$+0.08$	0.21
6	EV 002	0.647	0.893	45.3	-0.16	0.21
7	Mugar 201	0.402	0.912	40.0	-0.23	0.21
8	Snell 168	0.398	0.928	43.5	-0.34	0.21
9	EXP 204	1.003	0.776	32.2	-0.89	0.33

5 Discussion

5.1 Error and Uncertainty

The residuals of the two-way additive ANOVA model provided the basis for all cross-room confidence intervals. Because the 9×6 design has only one observation per cell, the model assumes no room-by-position interaction. This assumption is approximate: rooms with different geometries can in principle have different spatial gradients, and any such departure inflates the residual mean square. The resulting bias is

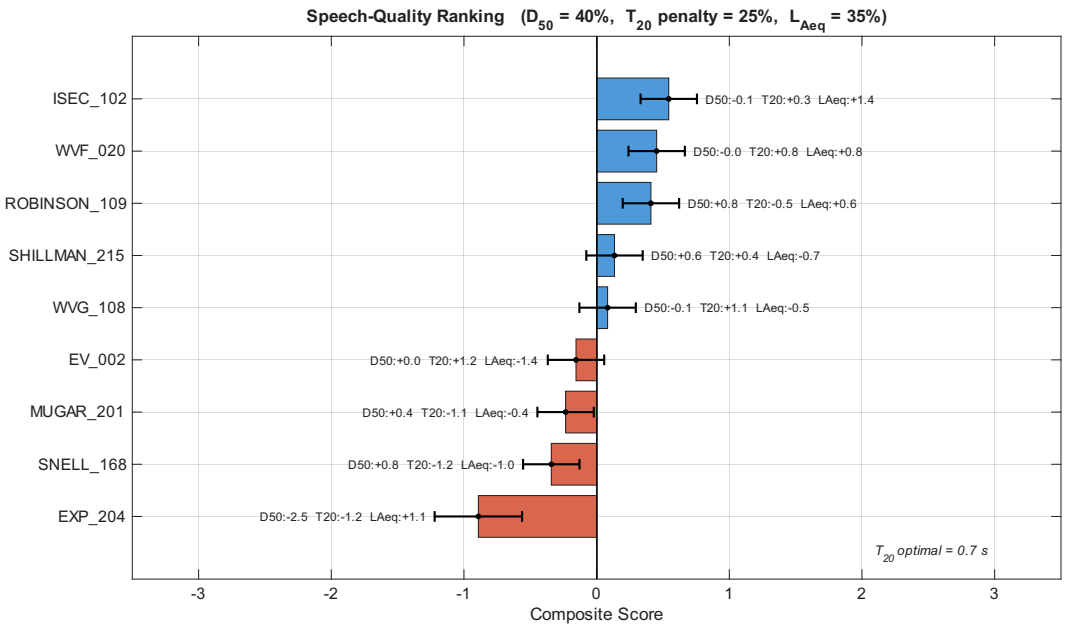


Figure 8: Composite speech-quality ranking of the nine rooms with 95% confidence intervals from linear error propagation. Rooms above the dashed line have above-average speech quality relative to the measured set; rooms below have below-average quality. Overlapping error bars in the middle of the ranking indicate room pairs that cannot be distinguished at $\alpha = 0.05$.

conservative in the sense that it widens the confidence intervals and lowers the F statistic rather than inflating the Type I error rate, so the reported p-values are not artificially small.

The quality of each T_{20} regression was checked by computing the coefficient of determination of the linear fit over the -5 to -25 dB range of the EDC. A threshold of $R^2 \geq 0.99$ was used to flag measurements in which the decay deviated meaningfully from a single exponential. 39 of the 54 measurements (72%) passed this threshold. The failures did not cluster as a dynamic-range argument would predict. The pass rate was lowest at front receiver positions (10 of 18, 56%), rose at middle positions (13 of 18, 72%), and was highest at back positions (16 of 18, 89%); and the Spearman correlation between room-mean R^2 and room L_{Aeq} was not significant ($\rho = +0.23$, $p = 0.55$, $n = 9$ rooms). In other words, failing measurements do not concentrate in the noisiest rooms or at the positions farthest from the source, which is where SNR-limited decays would be expected to break down first.

The observed front-to-back pattern is consistent with direct-sound contamination of the early decay at near-field positions. At front positions the direct sound is strong relative to the reverberant field, and the steep drop immediately after the direct arrival extends into the -5 to -25 dB evaluation window, steepening the apparent slope and pulling the regression away from a single exponential. At mid and back positions the diffuse field dominates, the direct-sound energy is a smaller fraction of the early integral, and the EDC in the evaluation window is closer to linear. Frequency-dependent absorption, which produces a superposition of exponential decays with different time constants across bands, is a secondary whole-room contributor

that would be expected to affect all receiver positions roughly equally and therefore cannot by itself explain the position gradient observed here. Separating the two effects would require octave-band analysis, which was excluded from the scope of this project because the cardioid measurement microphone's uncalibrated frequency response (Section 3.5) does not support reliable per-band parameter extraction. Measurements that failed the R^2 check were retained in the reported means but were drawn with open markers in the per-room summary figures so the reader can identify them.

The Jounivo JV-601 microphone has a manufacturer-rated SNR of 76 dB(A), which corresponds to an equivalent input noise of roughly 18 dB(A). The measured L_{Aeq} values in the nine rooms ranged from 30.5 dB(A) to 45.3 dB(A), which is 12 to 27 dB above the microphone's self-noise floor. In every room, therefore, the dynamic range of the impulse response measurement was limited by the ambient noise in the room rather than by the microphone, and a quieter measurement microphone would not have materially improved the results. This argument was also the primary reason T_{20} was preferred over T_{30} in the composite score: at the back receiver positions in noisier rooms such as EV 002, the effective SNR is marginal for the 30 dB decay window required by T_{30}, while the 20 dB window required by T_{20} remained well clear of the noise floor.

The cardioid polar pattern of the JV-601 is the most consequential limitation of the measurement chain for absolute interpretation. A cardioid microphone attenuates sound arriving from its rear hemisphere by roughly 15 to 25 dB relative to on-axis sound, which preferentially removes late-arriving reverberant reflections from the recorded signal. The net effect is that the recovered impulse responses have systematically less late energy than they would with an omnidirectional microphone, which biases D_{50} and C_{50} upward and T_{20} and T_{30} slightly downward. The magnitude of this bias cannot be quantified without a reference omnidirectional measurement in the same rooms. Two observations support the view that the bias is roughly consistent across rooms. First, the microphone was held in the same orientation relative to the source in every measurement. Second, the rank order of D_{50} across the rooms is internally consistent with the rank order of T_{20}, with longer-reverberation rooms showing lower D_{50}, as expected. Across the nine rooms, D_{50} was uniformly above 0.75, which is higher than typical omnidirectional values reported in Rakerd et al. [6] for similar lecture spaces, and is consistent with the expected cardioid bias. For this reason, the composite ranking is reported as a relative comparison among the nine measured rooms rather than as an absolute quality rating.

The residual sidelobe floor of the Farina inverse filter, quantified against a synthetic ground truth in the Validation subsection above, is an intrinsic property of the discrete-time ESS deconvolution method [3] and sets a lower bound on the usable EDC range. All parameters reported in this project were extracted from the upper 35 dB of the EDC or less, so the sidelobe floor does not affect the reported values.

5.2 Other Limitations

The single portable powered loudspeaker used in place of a dodecahedral omnidirectional source introduces an additional source-directivity bias. The loudspeaker was kept in the same orientation (facing the seating area) in every room, so this bias is approximately constant across the data set. The L_{Aeq} values were obtained from a phone-based sound level meter application rather than a calibrated Type-1 sound level meter, and only

a single reading was taken per room. The relative ranking of rooms by L_{Aeq} is likely correct because the same phone was used throughout, but the absolute numerical values should not be compared directly to published measurements. The HVAC state was documented per room but was not experimentally controlled; in rooms where the HVAC was running during the measurement, the L_{Aeq} value includes the HVAC contribution.

The composite weights (0.40 for D_{50}, 0.25 for T_{20} penalty, 0.35 for L_{Aeq}) were chosen before the data were analyzed and were not varied after the fact. A sensitivity analysis in which the weights are perturbed within plausible bounds was not performed but would be a natural extension of this work, and would identify which rooms have robust rankings and which depend sensitively on the specific weight choice.

5.3 Confidence in the Findings

The extremely small p-values returned by the ANOVA for every parameter ($p < 10^{-6}$) make it essentially impossible that the observed room differences are an artifact of random measurement variability. The ranking at the extremes of the distribution is also robust: ISEC 102 at the top and EXP 204 at the bottom have non-overlapping confidence intervals under the reported composite and cannot be interchanged under any plausible reweighting. The middle of the ranking is less firm. Four rooms (Shillman 215, WVG 108, EV 002, Mugar 201) have overlapping composite confidence intervals and cannot be individually ordered at $\alpha = 0.05$, so the specific rank number assigned to any one of these rooms should be interpreted with caution.

6 Conclusions

Nine Northeastern lecture halls and classrooms were characterized with 54 impulse response measurements, recovered by Farina analytical inverse-filter deconvolution of 10-second exponential sine sweeps. The seven ISO 3382 parameters extracted from the recovered impulse responses differed significantly across rooms under a two-way additive ANOVA with position as a blocking factor, with $p < 10^{-6}$ for every parameter. A weighted composite of D_{50}, a T_{20} penalty from a 0.7 s target, and background L_{Aeq} identified ISEC 102 as the highest-ranked room for speech and EXP 204 as the lowest. ISEC 102 combined a low background noise level (30.5 dB(A)) with a reverberation time within the recommended 0.4–0.8 s lecture band, while EXP 204 had a reverberation time of roughly 1.0 s (approximately double the recommended range) and the lowest measured D_{50}.

Relative differences between rooms are trustworthy; absolute values should be interpreted with caution because of the cardioid directivity of the measurement microphone and the uncalibrated background-noise measurement. The measurement and analysis pipeline developed for this project, including its validation against a synthetic ground-truth reference, is available in the project repository and can be applied directly to additional campus spaces in future work. A practical outcome is the identification of EXP 204 as a strong candidate for acoustic treatment on the basis of its long reverberation time, and the observation that most of the measured teaching spaces exceed the ANSI S12.60 background-noise limit of 35 dB(A) even when unoccupied.

References

[1] Wallace Clement Sabine. *Collected Papers on Acoustics*. Harvard University Press, Cambridge, Massachusetts, 1922. Published posthumously; foreword by Theodore Lyman.

[2] Manfred R. Schroeder. New method of measuring reverberation time. *The Journal of the Acoustical Society of America*, 37(3):409–412, 1965. doi: 10.1121/1.1909343.

[3] Angelo Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *Proceedings of the 108th Convention of the Audio Engineering Society*, Paris, France, February 2000. URL `https://aes.org/publications/elibrary-page/?id=10211`. AES Preprint 5093.

[4] *ISO 3382-1:2009 Acoustics — Measurement of Room Acoustic Parameters — Part 1: Performance Spaces*. International Organization for Standardization, Geneva, Switzerland, 2009. URL `https://www.iso.org/standard/40979.html`.

[5] *ISO 3382-2:2008 Acoustics — Measurement of Room Acoustic Parameters — Part 2: Reverberation Time in Ordinary Rooms*. International Organization for Standardization, Geneva, Switzerland, 2008. URL `https://www.iso.org/standard/36201.html`. Technical Corrigendum 1 issued 2009 (ISO 3382-2:2008/Cor 1:2009).

[6] Brad Rakerd, Eric J. Hunter, Mark Berardi, and Pasquale Bottalico. Assessing the acoustic characteristics of rooms: A tutorial with examples. *Perspectives of the ASHA Special Interest Groups*, 3(19): 8–24, 2018. doi: 10.1044/persp3.SIG19.8.

[7] *ANSI/ASA S12.60-2010/Part 1 American National Standard Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, Part 1: Permanent Schools*. Acoustical Society of America, Melville, New York, 2010. ANSI-accredited standard published by the Acoustical Society of America; reaffirmed 2015 and 2020.

[8] *Guidelines for Community Noise*. World Health Organization, Geneva, Switzerland, 1999. URL `https://iris.who.int/handle/10665/66217`.

[9] W. T. Chu. Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method. *The Journal of the Acoustical Society of America*, 63(5):1444–1450, 1978. doi: 10.1121/1.381889.

[10] Anders Lundeby, Tor Erik Vigran, Hannes Bietz, and Michael Vorländer. Uncertainties of measurements in room acoustics. *Acustica*, 81(4):344–355, 1995.

A Data and Code Availability

All MATLAB source code, raw WAV recordings, processed impulse responses and parameter JSONs, per-room summary figures, and cross-room comparison outputs are publicly available at `https://github.com/aldernadon/room-acoustics` (frozen version cited: release `v1.0`). The repository includes the orchestration script `main.m`, the deconvolution and parameter-extraction module `deconvolve.m`, the per-position and per-room plotting routines `plot_ir.m` and `plot_summary.m`, the cross-room statistical analysis and composite-ranking code `analyze_comparison.m` and `plot_comparison.m`, and the synthetic-ground-truth validation in `validation/ess_deconvolution.m`. Each recording's metadata (excitation bandwidth, HVAC state, background-noise level) is stored alongside the corresponding WAV file as a JSON document. Running `main.m` from the repository root regenerates all processed data and per-position figures; running `analyze_comparison.m` regenerates the cross-room statistics, composite ranking, and comparison figures reported in Section 4.5 and Section 4.6.