Audio Engineering Society

Convention Paper 10241

Presented at the AES 159th Convention
2025 October 23–25, Long Beach, CA, USA

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Evaluating Microphone Choice and Placement for Tom Drum Recording: A Perceptual Study

Matthew Cheshire[1] and Samuel Bullock[1]

[1]*Birmingham City University*

Correspondence should be addressed to Matthew Cheshire (matthew.cheshire@bcu.ac.uk)

ABSTRACT

This study investigates how microphone choice and placement affect the perceptual quality of tom drum recordings. Six widely used microphones—AKG C414, AKG D40, Audix D2, Audix D6, Sennheiser MD421, and Shure SM57—were evaluated across three tom sizes (10", 12", and 16") and two microphone positions (Centre and Rim). Recordings were captured under controlled studio conditions, utilising a robotic drum beater for consistent performance, and rated by 25 listeners with audio engineering backgrounds. Statistical testing was used to evaluate significant ratings of microphones, as well as if ratings differed between tom sizes and between microphone position. Friedman tests indicated significant differences in preference across microphones. Subsequent pairwise comparisons using the Wilcoxon signed-rank test revealed context-dependent perceptual differences between microphones, with several models showing strong performance only for specific tom sizes and positions. Results also showed that the same microphone could be rated significantly different depending on placement or drum size, highlighting the importance of context in microphone selection. These findings suggest that tailoring microphone choice and position for each tom can improve perceptual outcomes.

1 Introduction

The toms or "Tom–Toms" are an integral part of the drum kit and used across many genres [1]. Both microphone selection and placement remain critical factors for tom drum recording. While close miking has long been a standard approach in contemporary drum recording [2], relatively few controlled studies have focused specifically on toms — particularly regarding how microphone choice and position affect listener preference. While practical guidance on tom miking is available [3, 4, 5], these emphasise workflow efficiency and anecdotal experience over scientific evaluation.

Previous research on snare drums has shown that perceptual differences between microphones are linked to frequency response variations and can be corrected through equalisation [6], yet these findings have not been extended to toms. Other work has explored microphone bleed [7, 8], and drumhead variations [9, 10], all of which are relevant in understanding the complex interaction between source, microphone, and placement.

The tom drum's role in popular music genres—particularly rock, where fills are common—place additional emphasis on transient detail and low-mid clarity [11]. While the application of tom miking has evolved over time [12], commercial recordings still typically favour dynamic microphones [13], but engineering decisions are often based on experience rather than empirical evidence [14, 15, 2].

Make	Model	Type	Polar Pattern
AKG	C414	C	Cardioid
AKG	D40	D	Cardioid
Audix	D2	D	Hypercardioid
Audix	D6	D	Cardioid
Sennheiser	MD421	D	Cardioid
Shure	SM57	D	Cardioid

Table 1: Microphones used for tom drum recording, showing Make, Model, Type Condenser - (C) or Dynamic (D), and Polar Pattern.

This study presents a comparison of tom recordings using multiple microphones, three distinct toms, and two different mic placements, evaluated through the use of a subjective preference listening test. The aim of which is to determine which configurations yield subjectively preferred outcomes, and if preference for one microphone is tom- and position-dependent.

2 Methods

2.1 Recording

Six commercially available microphones were selected for evaluation: AKG C414, AKG D40, Audix D2, Audix D6, Sennheiser MD421, and Shure SM57, shown in Table 1. Microphone selection was based on recommendations by recording engineers, as well as availability. These were used to capture three toms — High Tom (10" x 7"), Mid Tom (12" x 8") and Floor Tom (16" x 16") — with recordings taken from both centre and rim positions shown in Figure 1. This resulted in 36 unique tom drum recordings (6 microphones 3 toms 2 positions). The high pass filter on both the C414 and MD421 were both disengaged and the polar pattern of the C414 was set to Cardioid.

To ensure consistency across performances, all toms were struck using a POLYEND PERC[1] robotic beater triggered by MIDI from Logic Pro, striking the centre of the batter head, as seen in Figure 1. Each tom was recorded playing a sequence of single hits, steady quarter notes, short rhythmic phrases, and a brief fill. The microphones were positioned 5cm from the drumhead and their on-axis angles were matched as closely as possible. All recordings were made in an acoustically treated studio live room, tracked through an SSL AWS 924 δelta console and an Antelope Orion 32+ ADC at a sample rate of 48kHz and a bit depth of 24-bit.

[1] https://polyend.com/legacy/polyend-perc/



Fig. 1: Centre (top) and Rim (bottom) microphone positions shown with the Sennheiser MD421, and the Polyend Perc Robotic drum beater.

2.2 Evaluation

A listening test was conducted using the Web Audio Evaluation Tool (WAET) [16] with the APE interface [17], including six test pages — one for each tom/position combination (e.g., floor/centre, mid/rim, etc.). Each page included six playback sliders (one per microphone) which changed from green to red to indicate which sample was playing. These sliders were initialised on a continuous 0–1 scale, with the starting positions, slider labels, and page order all randomised per participant to minimise bias. Stimuli were loudness normalised to -23 LUFS and looped continuously, playhead position was maintained when switching between samples for uninterrupted playback. Participants were instructed to rate the six stimuli per page based on their own subjective preference, least preferred on the left and most preferred on the right of the scale. They were

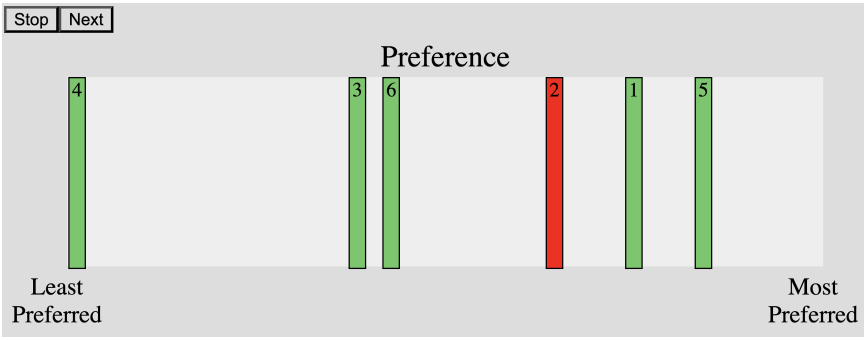


Fig. 2: Interfaced used for the subjective listening test.

not able to progress to the next test page until every sample had been played at least once, and every slider had been moved from its initial position. Participants were instructed to use the full range of the rating scale, placing at least one slider at both extremes of the scale, the interface used is shown in Figure 2.

In total, 25 participants, aged between 19 and 23 (mean age = 20.9), completed the listening test using Beyerdynamic DT990 PRO headphones in a quiet, controlled environment. As part of the pre-test questionnaire, participants were asked how many years of critical listening experience they had; responses ranged from 2 to 8 years (mean = 4.6). They were additionally asked which instruments they played — with responses including drums, guitar, bass, piano, and synthesiser — and to identify their favourite genre of music, with the most common responses being rock, metal, pop, electronic, and hip hop.

2.3 Spectral Features

To explore whether perceptual preference related to measurable spectral properties, five spectral features were extracted. The recordings were first trimmed to remove silence (40 dB threshold) prior to feature extraction. Features were then computed frame-by-frame using a short-time Fourier transform and averaged across frames to produce one value per recording. The spectral features considered were: spectral centroid (for perceived brightness), spectral bandwidth (spread of spectral energy), 95% spectral roll-off (balance of low vs. high frequency energy), spectral flatness (tonal vs. noise-like content), and peak frequency (dominant spectral component, often the fundamental). Correlations between mean preference scores and the features were assessed using Pearson's r, a standard parametric measure of linear association. This test is easily interpretable, but is limited by its sensitivity to outliers. With only six microphones per condition, statistical power was expected to be weak.

3 Results

Firstly, boxplots were generated to visualise the subjective preference ratings across all microphone and tom combinations, seen in Figure 3. This serves to highlight patterns in listener responses. While useful as an exploratory tool, visual trends alone are insufficient to draw robust conclusions regarding microphone preference. Formal statistical analysis is therefore required to assess whether any microphone is statistically more preferred than another, preferred in one position over another, or more preferred for a given tom size.

In order to determine an appropriate statistical method for analysis, the distribution of subjective rating data was first assessed using the Shapiro–Wilk test. Results indicated that 14 out of 36 microphone/position/tom combinations met the assumptions of normality. This suggests that a non-parametric approach is more appropriate for the majority of comparisons.

A Friedman test was conducted for each tom and position combination to evaluate whether there were statistically significant differences in listener ratings across the microphones. All six Friedman tests returned highly significant results, with p-values below 1×10^{-8}, indicating strong evidence that listener preference for at least one microphone differed significantly from another within each test condition. This justifies proceeding with post-hoc pairwise comparisons to determine which microphones ratings are significantly different.

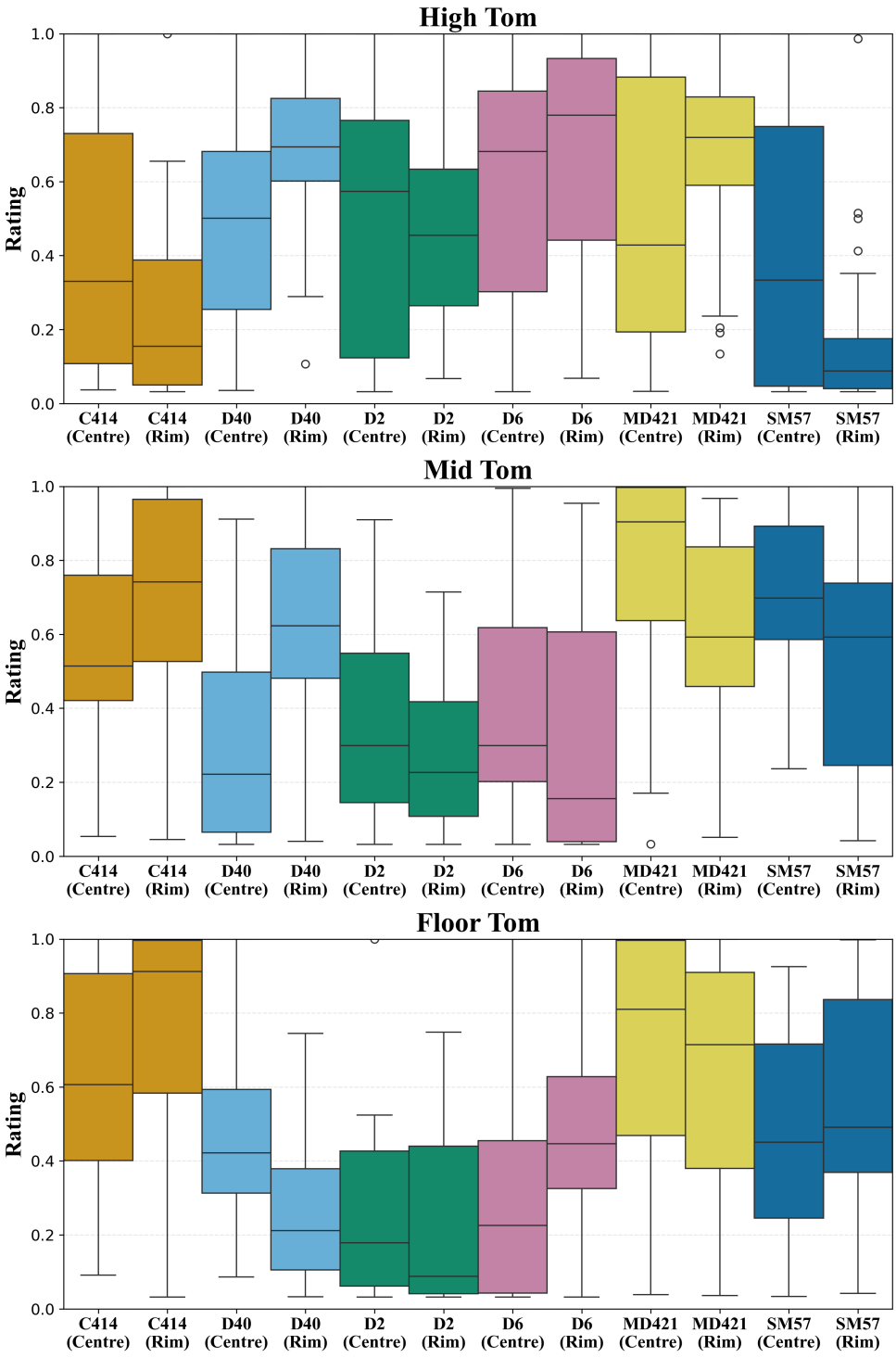


Fig. 3: Listener preference ratings for six microphones across three toms (high, mid, floor) and two microphone positions (Centre and Rim).

Mic	High	Mid	Floor
C414	*ns*	*ns*	*ns*
D40	0.0025 (↑)	0.0001 (↑)	0.0002 (↓)
D2	*ns*	*ns*	*ns*
D6	*ns*	*ns*	0.0236 (↑)
MD421	0.0136 (↑)	0.0422 (↓)	*ns*
SM57	0.0342 (↓)	0.0236 (↓)	*ns*

Table 2: Wilcoxon signed-rank test results comparing preference ratings between Centre and Rim microphone positions for each tom. Only statistically significant comparisons ($p < 0.05$) are shown, *ns* indicate no significant difference, (↑) shows if Rim position is higher rated than Centre, (↓) shows if it is lower.

The Wilcoxon signed-rank test was chosen for pairwise comparisons as it is well-suited to the within-subject design of the study, does not assume normally distributed data, and accounts for both the direction and magnitude of differences in listener ratings between pairs of microphones within each tom and position condition. A paired t-test was not appropriate, as normality assumptions were violated in the majority of these conditions.

Firstly, in order to observe which microphones had significantly different scores from one another, a pairwise comparison was conducted for each tom-position condition using the Wilcoxon signed-rank test with Bonferroni correction. Figure 5 presents the results of these comparisons, showing microphone-versus-microphone significance matrices. Each matrix shows which microphone pairs were rated significantly differently ($p < 0.05$) within that condition, microphones are ordered by their median preference rating.

Although different microphone placements (Centre vs. Rim) and tom sizes (High, Mid, and Floor) were not directly compared within the same test page, each configuration was rated by the same group of listeners under identical playback conditions. As such, while the evaluations were collected across separate listening tasks, the use of a consistent participant pool and controlled listening context allows for valid indirect comparisons between placement and size conditions for a given microphone. The results of the Wilcoxon signed-rank test between Centre vs. Rim are shown in Table 2, where only pairs that were rated significantly different ($p < 0.05$) from one another are shown via their p values.

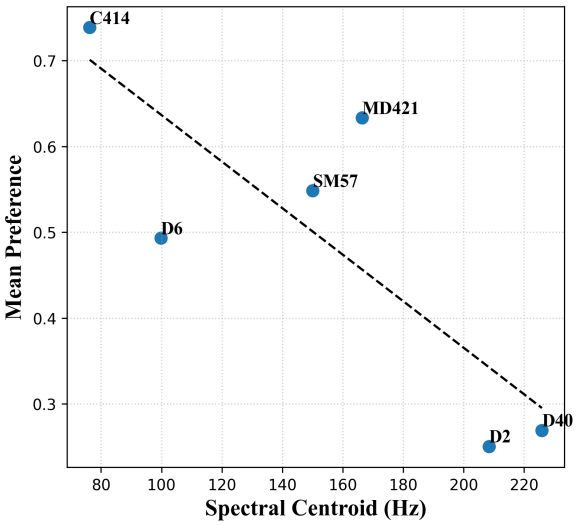


Fig. 4: Pearson correlation for floor tom rim: Centroid vs Preference, r = –0.814, p = 0.049.

Additional pairwise comparisons were performed to evaluate whether listener preference for each microphone changed significantly when used on different tom sizes. This analysis was again conducted using the Wilcoxon signed-rank test with Bonferroni correction, comparing ratings for each microphone across the three toms (High, Mid, and Floor), separately for Centre and Rim positions. Results are presented in Table 3, which shows all microphone/tom comparisons that yielded statistically significant differences ($p < 0.05$), along with arrows indicating the direction of median preference change.

Finally, Pearson correlations between preference and spectral features were investigated. While several conditions showed moderate-to-strong correlations ($|r| > 0.7$), these were inconsistent across drums and positions. Of the 30 correlations tested (6 conditions x 5 features), roughly one-third showed negligible association ($|r| < 0.3$). Notably, the Floor Tom Rim condition displayed strong negative correlations between mean preference and all five features, with two reaching significance: centroid (r = –0.814, p = 0.0489) and flatness (r = –0.867, p = 0.0253). Figure 4 illustrates the correlation for centroid. No other drum or position presented significant correlation between mean preference score and any of the five spectral features.

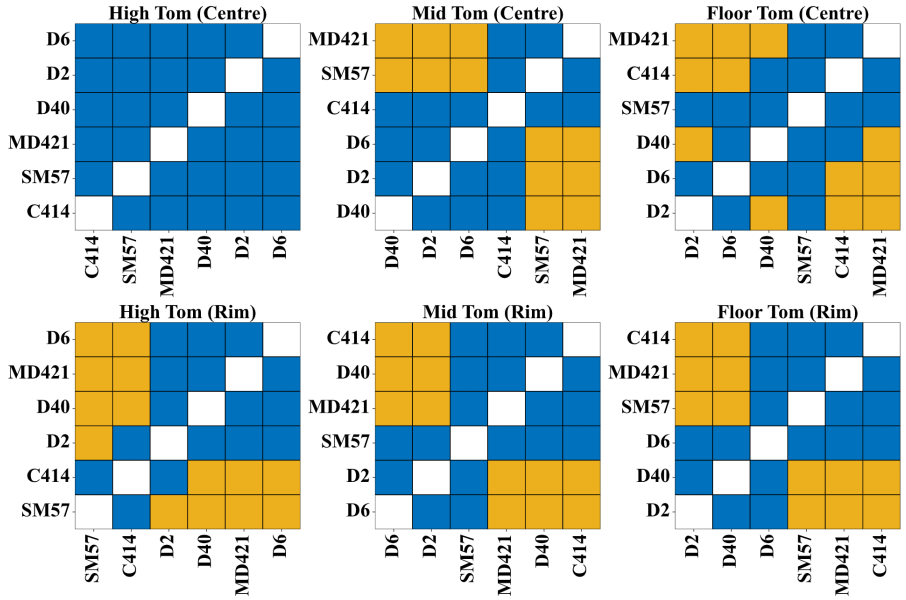


Fig. 5: Pairwise comparisons between microphones based on Wilcoxon signed-rank tests with Bonferroni correction. Yellow cells indicate a statistically significant difference in listener ratings between microphones ($p < 0.05$), while blue cells indicate no significant difference ($p \geq 0.05$). Each subplot corresponds to a different tom size and microphone placement condition. Microphones are ordered within each subplot from lowest (bottom left) to highest (top right) median preference rating.

Microphone	High vs Mid	High vs Floor	Mid vs Floor
C414 (Centre)	*ns*	*ns*	*ns*
C414 (Rim)	0.0001 (↑)	0.0001 (↑)	*ns*
D40 (Centre)	0.0105 (↓)	*ns*	*ns*
D40 (Rim)	*ns*	0.0000 (↓)	0.0001 (↓)
D2 (Centre)	*ns*	0.0147 (↓)	*ns*
D2 (Rim)	0.0034 (↓)	0.0008 (↓)	*ns*
D6 (Centre)	*ns*	0.0207 (↓)	*ns*
D6 (Rim)	0.0001 (↓)	0.0422 (↓)	*ns*
MD421 (Centre)	0.0035 (↑)	0.0027 (↑)	*ns*
MD421 (Rim)	*ns*	*ns*	*ns*
SM57 (Centre)	0.0187 (↑)	*ns*	0.0074 (↓)
SM57 (Rim)	0.0001 (↑)	0.0000 (↑)	*ns*

Table 3: Wilcoxon signed-rank test results comparing listener preference ratings across tom sizes for each microphone, shown separately for Centre and Rim positions. Statistically significant differences ($p < 0.05$) are reported alongside arrows indicating the direction of change: (↑) indicates the median rating increased in the second tom, (↑) indicates it decreased. Non-significant results are marked as *ns*.

4 Discussion

Several noteworthy trends emerge upon inspection of the boxplots in Figure 3. First, there is evidence that listener preference for certain microphones varies not only with tom size, but also with microphone position. For the high tom, the SM57 shows a noticeable difference between positions: the Centre placement receives moderate ratings, while the Rim placement is considerably lower. In contrast, the D40 and MD421 appears to exhibit the opposite trend on the high tom, with the Rim position more highly preferred than the Centre.

From the results in Figure 5, we can see that no individual microphone received significantly different ratings from any other in the High Tom (Centre) condition. While the Friedman test indicated that some microphones might be rated differently within this group, none of the pairwise comparisons remained significant following Bonferroni correction—likely a reflection of the correction's conservative nature. Significant differences, however, were observed in the remaining five conditions. In the Mid Tom (Centre) condition, both the SM57 and MD421 were rated significantly higher than the D40, D2, and D6. For the Mid Tom (Rim), the MD421, D40, and C414 all outperformed the D2 and D6. Similarly, in the Floor Tom (Rim) condition, the SM57, MD421, and C414 were also rated significantly higher than both the D2 and D40.

The position comparison results in Table 2 show that the D40 was the most sensitive to placement, exhibiting significant differences between Centre and Rim positions across all three toms. While its Rim placement was preferred on the high and mid toms, the Centre position was significantly favoured on floor tom. Both the C414 and D2 did not exhibit any change in preference across the two placements. The D6 showed a significant difference only on floor tom, where Rim placement was preferred. The SM57 also showed significant differences in position for both the high and mid toms, with Rim placement resulting in lower preference ratings in both cases. Lastly, the MD421 was significantly preferred in the Rim position on high tom, but less preferred at the Rim on mid tom. These findings, when considered alongside the distribution of scores in Figure 3, indicate that the perceptual impact of microphone placement is both microphone- and tom-dependent, and that a consistent positioning strategy may not yield uniformly optimal results across the drum kit.

The results in Table 3 highlight that tom size produced a statistically significant influence on microphone preference for several microphones. Notably, the SM57 was most sensitive to tom drum size, with four tom-dependent results from six comparisons across both positions. While the D40, D2, and D6 had three tom-dependent scores, and the C414 and MD421 only have two. For AKG (Centre) and MD421 (Rim) the size of the tom did not significantly change subjective ratings, meaning that in these positions their preference was consistent across the different toms. This indicates that for certain microphones and positions, preference can significantly change depending on the size of the tom.

These findings have practical implications for recording engineers and educators, suggesting that optimal microphone choice and placement on toms may vary not only by mic model but also by tom size. Rather than defaulting to a single microphone type or uniform positioning across the toms, a mixed configuration strategy with tailored tom-specific positions may yield perceptually better outcomes.

The limited number of microphones per condition (n = 6) likely restricted the ability of Pearson correlation to reveal robust relationships between spectral features and listener preference. Even where strong effect sizes were observed, significance was only achieved in two of the 30 tests, reflecting the low statistical power of this analysis. Additionally, preference may depend on perceptual factors that are not captured by global spectral descriptors alone, such as temporal envelope and attack characteristics.

5 Summary

This study examined the perceptual impact of microphone choice and placement for recording tom drums, using a controlled recording setup and listening test. Six commonly used microphones were evaluated across three tom sizes (10", 12", and 16") and two microphone positions (Centre and Rim), with subjective preference ratings from 25 participants. The analysis revealed that microphone preference is influenced not only by model, but also by tom size and mic placement. Statistical testing confirmed that listener ratings differed significantly across microphones, with certain models outperforming others for specific tom size and mic position combinations. Additionally, position-based comparisons showed that several microphones exhibited significant shifts in preference between Centre and

Rim placements, and tom-size comparisons highlighted that perceptual suitability is not fixed across the drum kit. These findings suggest that microphone performance is highly context-dependent, and that more nuanced, tom-specific mic selection strategies may yield improved perceptual outcomes over uniform mic deployment. Moreover, the results indicate that a given microphone's effectiveness can vary significantly with placement, suggesting that different microphones may have distinct positional "sweet spots" that interact with drum size and tuning. As such, both microphone choice and positioning should be carefully tailored to each individual tom to achieve optimal recording results.

One limitation of this study is that all recordings were captured in isolation, meaning the perceptual impact of bleed from surrounding drums and cymbals was not accounted for. In real-world recording contexts, microphone placement decisions are influenced not only by tonal quality but also by strategies to minimise bleed, particularly in live environments. While the use of noise gates may be deployed to reduce bleed, future work could explore how microphone and placement choices perform under more realistic conditions with full kit context. Additionally, examining the interaction between mic choice, placement, and gating or other processing strategies would offer valuable insights into optimising tom drum recording in applied studio settings. Further studies could also utilise a broader range of microphones, as well as investigate how listener preference is influenced by variations in tom tuning or the application of dampening techniques.

References

[1] Nicholls, G., *The Drum Book: A History of the Rock Drum Kit*, G - Reference, Information and Interdisciplinary Subjects Series, Backbeat, 2008.

[2] Major, M., *Recording Drums – The Complete Guide*, Cengage Learning, 2013.

[3] Shure Incorporated, "Microphone Techniques for Drums," https://content-files.shure.com/Pubs/microphone-techniques-for-drums/microphone_techniques_for_drums_english.pdf, 2014.

[4] DRUM! Magazine, "Studio Secrets: How To Record Toms," https://drummagazine.com/studio-secrets-how-to-record-toms/, 2025.

[5] Sound On Sound, "Recording Toms When You're In A Hurry," https://www.soundonsound.com/techniques/recording-toms-when-youre-hurry, 2025.

[6] Cheshire, M., Stables, R., and Hockman, J., "Microphone Comparison: Spectral Feature Mapping for Snare Drum Recording," in *AES Convention Paper 10263*, 2019.

[7] Clifford, A., Lindsay, H., and Reiss, J., "Simulating Microphone Bleed and Tom-Tom Resonance in Multisampled Drum Workstations," in *AES Convention Paper*, 2012.

[8] Greenwood, A. and Chafe, S., "Controlling Drum Bleed with Laser Vibrometry," in *AES Engineering Brief 109*, 2013.

[9] Quiroga, A. F., Garcia, J. D., and Páez, D., "Tom–Tom Drumheads Miking Analysis," in *Audio Engineering Society Convention 138*, 2015.

[10] D'Virgilio, N., "How to Choose the Best Drumheads for You," https://www.sweetwater.com/insync/how-to-choose-the-best-drumheads/, 2022.

[11] Modern Drummer Magazine, "Rock Perspectives – Tom-Toms," https://www.moderndrummer.com, 2017.

[12] Stamler, P. J., "Recording Drums—Then and Now," https://www.recordingmag.com/resources/recording-info/mics-miking/recording-drums-then-and-now/, 2018.

[13] Waxman, G., "6 Best Tom Mics For Live And Studio," https://drumhelper.com/audio/best-tom-mics/, 2025.

[14] Huber, D. and Runstein, R., *Modern Recording Techniques (6th Ed.)*, Focal Press, 2005.

[15] LANDR, "How to Mic a Drum Kit for Recording," https://blog.landr.com/how-to-mic-drums/, 2025.

[16] Jillings, N., De Man, B., Moffat, D., and Reiss, J. D., "Web Audio Evaluation Tool: A Browser-based Listening Test Environment," in *12th Sound and Music Computing Conference*, 2015.

[17] De Man, B. and Reiss, J., "APE: Audio Perceptual Evaluation Toolbox for MATLAB," in *136th Convention of the Audio Engineering Society*, 2014.