Assessing Differences in Music Perception Using Music-In-Noise Testing (MINT)

Shafaqat Rahman[1] & John Kyle Cooper[1]

[1]Department of Biomedical Engineering, University of Rochester, Rochester, NY 14623, USA

ABSTRACT

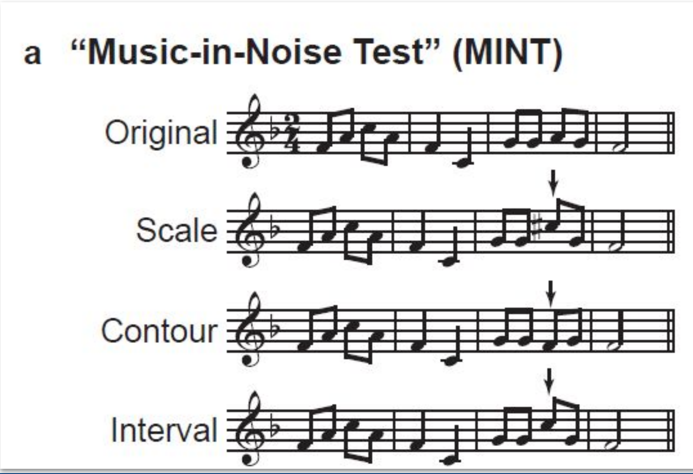
The ability to isolate target sounds from background noise is a research topic in both the clinic and computer audition. Wong et al. (2003) elucidated the issue of poor speech understanding in noise as a difficulty amongst patients with hearing impairments. However, recent literature in neuroscience suggests that people with musical training may experience less difficulty when understanding speech in noise than people without musical training. We hypothesized that musical training enhances a person's ability to make accurate pairwise comparisons through improved pitch perception. In order to validate this hypothesis, we have two goals in mind with this study: 1) use YIN pitch estimation algorithm (Cheveigné & Kawahara, 2002) and onset detection to determine spectral and temporal differences in music that are out of scale, contour, or interval and 2) to relate these differences to perceptual discernment in humans. Our results show that a person with normal hearing and no musical training can achieve a MINT accuracy of 75% at an SNR of -9 dB (n=3) , while a person with normal hearing and musical training has the same accuracy at an SNR of -13.5 dB (n =1). Our ultimate goal is to use this program to test if musical training enhances pitch perception in noise in children from 5th to 12th grade using the Music-In-Noise Test (MINT).

BACKGROUND

- The Hearing In Noise Test (HINT) is a standardized test that evaluates speech recognition in noise.
  - The level of speech is adjusted with respect to a fixed noise level.
- The test converges so that the patient obtains a 75% accuracy rate at a unique SNR level, which can then be used to compare scores across different patients and different experimental conditions (noisy and impaired).
- A recent study assessing the effects of musical training in pediatric mandarin-speaking cochlear implant users showcased results that music training can significantly improve the music and speech perception of mandarin-speaking cochlear implant users (Cheng et al., 2018).
- A modified version of the HINT test relies on using melodies with changes in scale, counter, and interval (Melodies in Noise Test, MINT). Information from MINT testing can further elucidate the benefits of using music as a therapy to improve our brain's ability to source separate stimuli from noisy environments.

Hypothesis: We are interested in knowing if musical training can sharpen a person's ability to detect subtle spectral and temporal differences in a target signal at varying levels of noise.

$$\text{SNR (dB)} = 20\log\frac{\text{rms SIGNAL}}{\text{rms NOISE}}$$



APPROACH/METHODOLOGY:

2.1 Dataset

- Synthetic piano recordings and an isolated noise source were provided by Dr. Anne Luebke's lab at the University of Rochester.
- These piano recordings were modified to have changes in pitch, scale, and contour in their melodies. Recordings were synthesized so that two melodies would be played with a fixed pause in between.
  - Each melody only appears once in a list and each list contains two subsets of equally balanced recording types (12 same, 12 different). Within each list, the subset of 12 different recordings is balanced to have 4 out of scale changes, 4 contour changes, and 4 interval changes.
- A total of 7 lists are used for testing but are adjusted to have different SNR values.
  - We applied a noise to a fixed melody recording and modulated the noise level via MATLAB. The first list is a practice list and contains four melodies in quiet. The second, third, fourth, fifth, and sixth lists contains recordings at 0, -6, -9, -12, and -15 dB SNR. The final list contains recordings at zero noise (quiet condition).

2.2 YIN Algorithm: Validating Pitch Differences Amongst Recording Types

- An aim for this project was to validate which files in the provided MINT dataset are designated as "same", "out of scale change", "contour change", "interval change".
- The YIN pitch estimation algorithm (Cheveigné & Kawahara, 2002) was used.
- Using these pitch estimates in each audio file, we detected the differences in pitch between the two music segments within each audio file to determine their type.

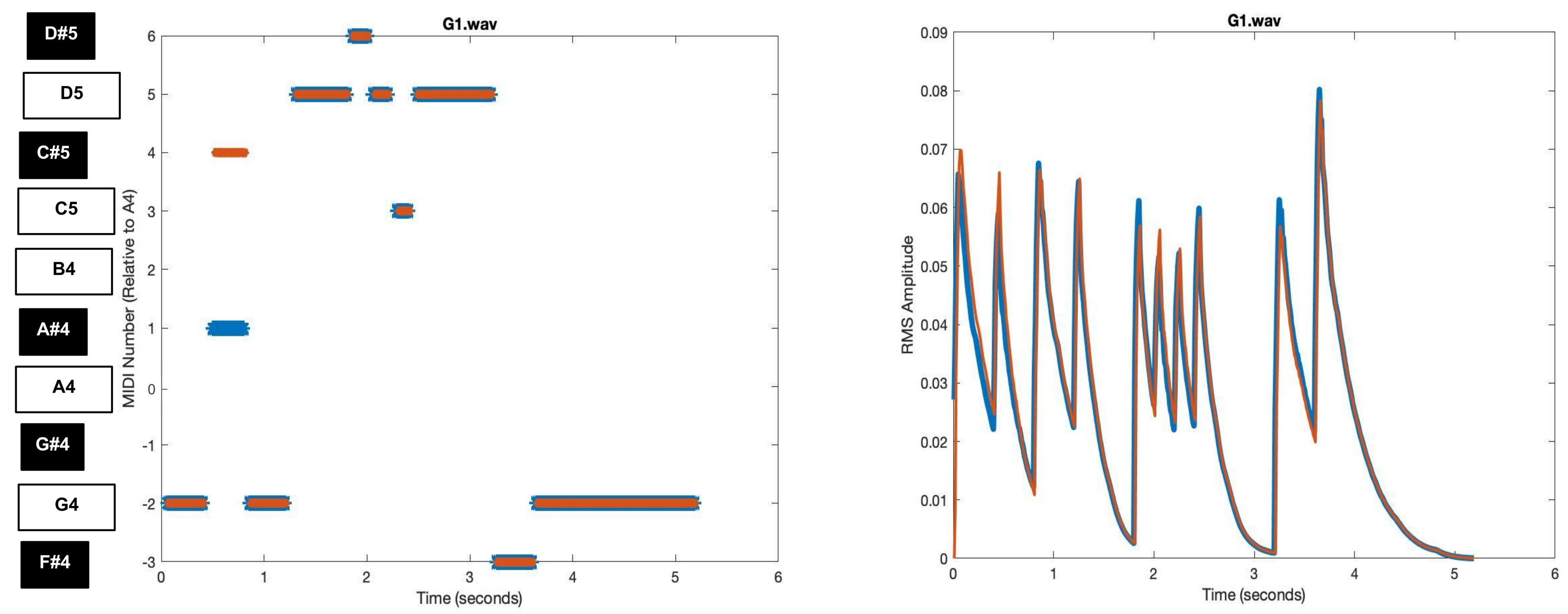


Figure 2. Pitch estimates and RMS values of MINT audio file in which the second music segment has one note that is an out of scale change.

LIST 1		
File	**Answer**	**Time (sec)**
A0.wav	Same	14.4
B2.wav	Different	10.45
E0.wav	Same	10.45
F1.wav	Different	10.45
G0.wav	Same	12.01
H0.wav	Same	10.45
I0.wav	Same	12.01
J0.wav	Same	12.01
K3.wav	Different	12.01
L1.wav	Different	11.61
M1.wav	Different	12.01
N0.wav	Same	11.62
O3.wav	Different	10.53
P1.wav	Different	14.41
Q0.wav	Same	12.01
R3.wav	Different	16.80
S0.wav	Same	12.01
T0.wav	Same	10.95
U0.wav	Same	12.01
V0.wav	Same	11.97
W3.wav	Different	10.86
X2.wav	Different	11.97
Y2.wav	Different	11.97
Z2.wav	Different	10.53

Figure 3: Paired melodic comparisons were derived from 27 original melodies from the Montreal Battery for the Evaluation of Amusia (MBEA). Test melodies were assigned the letters A-Z in the file name. For each melody, a 0 = same, 1 = out of scale change, 2 = contour change, 3 = interval change

Figure 4. The MINT GUI interface for carrying out the music-in-noise perceptual test. The application was created in MATLAB 2018b.

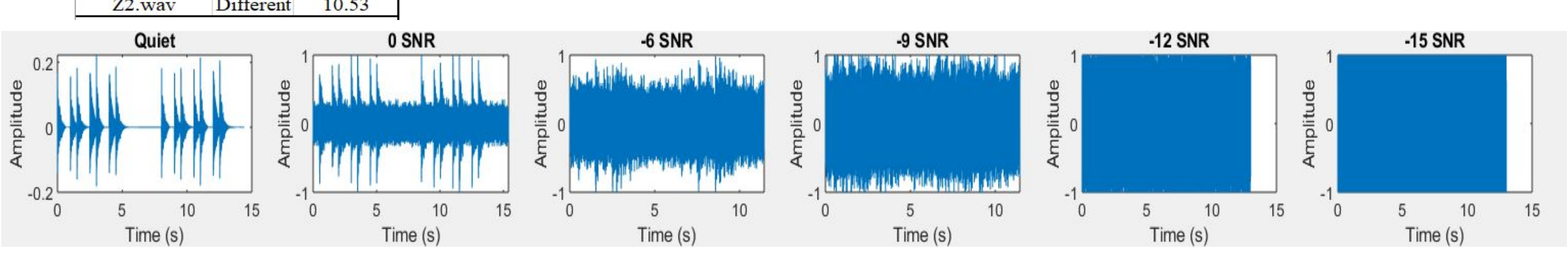


Figure 5: Audio files were separated into six lists: quiet, 0 SNR, -6 SNR, -9 SNR, -12 SNR, and -15 SNR. Each list had 12 audio files for testing.

RESULTS (NON-MUSICALLY TRAINED VS. MUSICALLY TRAINED):

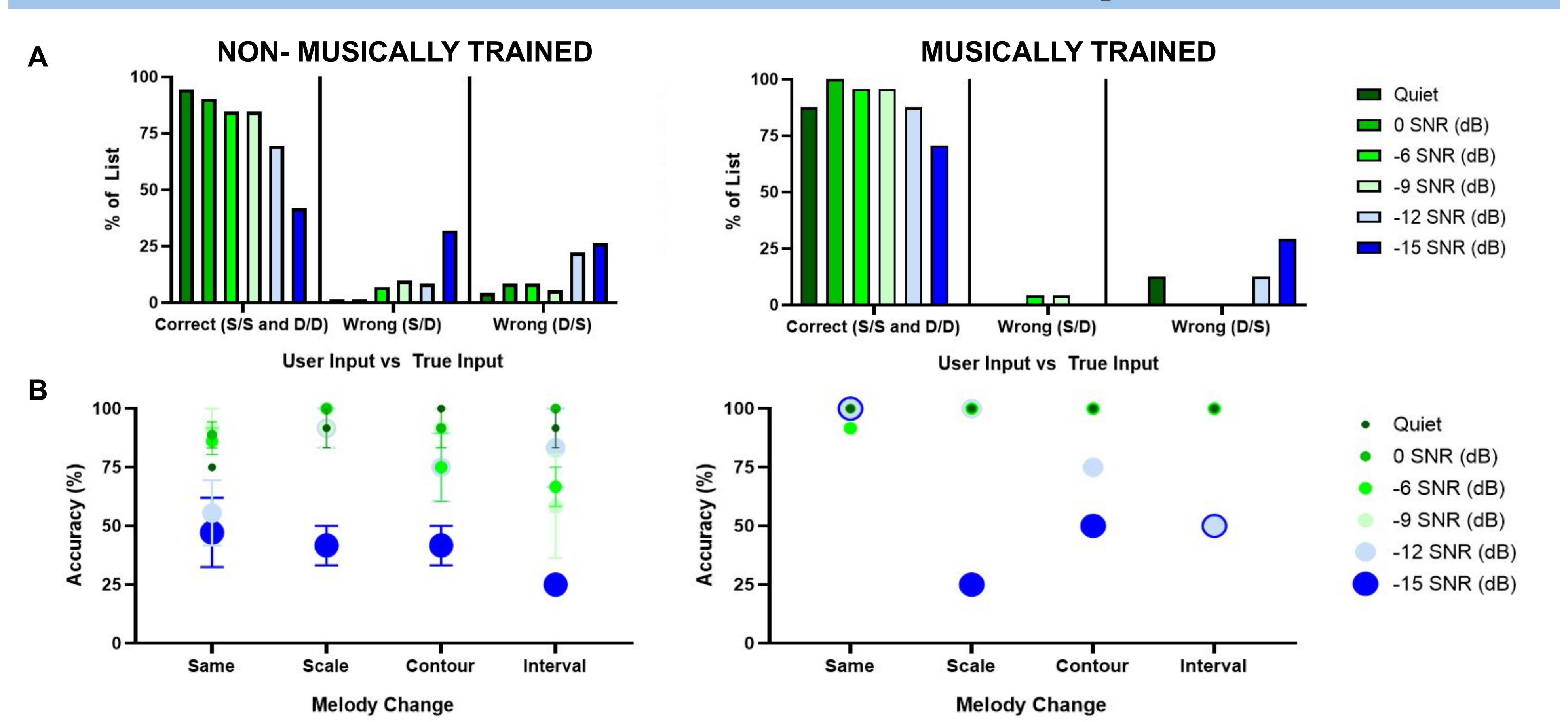


Figure 6: **(A)** At each SNR, the user input is compared to the true input. We make the distinction of separating the wrong answers as S/D and D/S (where S/D suggests the user perceived the melodies to be similar while the melodies were different) **(B)** Each melody type is plotted with respect to accuracy

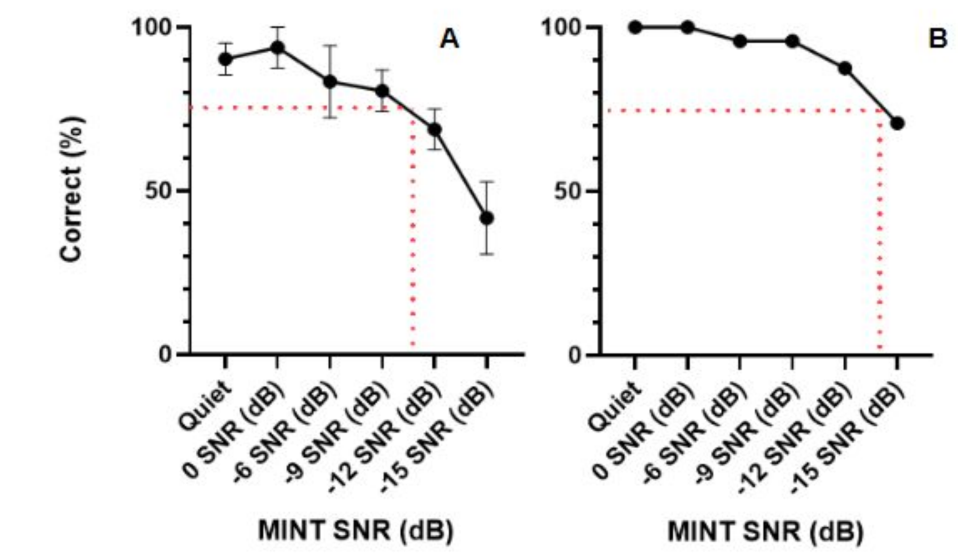


Figure 7: **(A)** MINT Plot for non-musically trained participants and **(B)** MINT plot for musically trained participants plotted with respect to accuracy

DISCUSSIONS:

- The MINT examination was carried out on 3 participants with normal hearing and no musical training and 1 participant with normal hearing and musical training. Based on the collected MINT data, the point at which a person with normal hearing and no musical training can achieve a 75% accuracy is around an SNR of -9 dB, while a person with normal hearing and musical training has the same accuracy around an SNR of -13.5 dB.

CONCLUSIONS:

- Currently, our experiment is lacking in terms of data from participants with musical training (potentially enhanced perceptual sensitivity) and participants with hearing impairments.
- Using the YIN algorithm, we generate pitch contours that can visibly showcase the pitch differences in musical melodies that are out of scale, contour, or interval.
- We also calculate the RMS of each melody segment to detect onsets, and show that the onsets do not widely differ across the melody types. Thus, the participant's decision making on whether paired melodies are the same or different must come from the pitch and not necessarily the beat.
- Despite low sample size for each cohort, we can see a clear difference in performance between those with and without musical training.
- Based on our preliminary MINT data, we believe people with normal hearing and musical training can perceive subtle differences between paired melodies better than people with normal hearing and no musical training.
- Currently, we cannot make any conclusions regarding people with impaired hearing.

FUTURE DIRECTIONS:

- Further MINT assessments are required to assess the 75% accuracy SNR for participants based on musical training or hearing impairments.
- In addition, we would like to make modifications to the MINT so that it is more adaptive and converges onto an SNR at 75% accuracy during the examination.
- We will also consider varying the melody signal in respect to a fixed noise when assessing future cohorts (rather than varying noise as done in this study).
- Would also would like to revisit studying HINT (speech in noise) with musically trained participants.

LITERATURE CITED:

1. Wong, L. L. N., Hickson, L., & McPherson, B. (2003). Hearing aid satisfaction: what does research from the past 20 years say? *Trends in amplification, 7*(4), 117-161. doi:10.1177/108471380300700402
2. Cheveigné, A., & Kawahara, H. (2002). YIN, A fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America, 111*, 1917-1930. doi:10.1121/1.1458024
3. Cheng, X., Liu, Y., Shu, Y., Tao, D. D., Wang, B., Yuan, Y., . . . Chen, B. (2018). Music Training Can Improve Music and Speech Perception in Pediatric Mandarin-Speaking Cochlear Implant Users. *Trends Hear, 22*, 2331216518759214. doi:10.1177/2331216518759214

ACKNOWLEDGEMENTS:

We would like to acknowledge Dr. Anne Luebke, Dr. Zhiyao Duan, and the ECE Department at the University of Rochester for the support of this research.