Simulated Hearing Loss on Speech Recognition, Flight Performance, and Workload in Aviators

Jennifer Noetzel; Paula Henry; Ryan Mackie; Kara Cave; J. R. Stefanson; J. Kyle Hale; Kevin Andres; Heath Jones

INTRODUCTION: Hearing loss can compromise U.S. Army aviators' performance, safety, and situational awareness, resulting in increasing mental workload and listening effort. This study evaluated simulated hearing loss on performance and cognitive workload among Army aviators.

METHODS: A mixed-effects linear regression study design was used. A total of 21 aviators underwent clinical audiological testing and simulated flight performance assessments. Simulated hearing loss and workload were manipulated to investigate their effects on speech recognition, flight performance, and subjective workload. Flight simulator routes included normal hearing and simulated hearing loss conditions for both high and low workloads. Task load questionnaires were administered for subjective workload assessments and compared across conditions.

RESULTS: Speech recognition scores decreased with increasing levels of hearing loss. In-flight speech intelligibility declined in high workload conditions, with a 26% decrease for mild hearing loss and a 40% decrease for severe hearing loss. High workload conditions degraded flight performance and response times to a secondary task which was exacerbated by simulated hearing loss. Workload scores validated increased workload with simulated hearing loss. No significant findings were observed on the hearing assessment.

DISCUSSION: Findings suggest hearing loss negatively impacts speech recognition and flight performance, especially under high workloads. These results support the importance of addressing hearing loss in aviators. Further research is needed to determine if the clinically adapted Modified Rhyme Test can reflect the impact of hearing loss on aviator performance.

KEYWORDS: hearing loss, workload, aviation.

Noetzel J, Henry P, Mackie R, Cave K, Stefanson JR, Hale JK, Andres K, Jones H. *Simulated hearing loss on speech recognition, flight performance, and workload in aviators*. Aerosp Med Hum Perform. 2025; 96(4):269–278.

Rotary-wing aircraft noise poses both short-term risks to the communication abilities of aviators and aircrew, as well as long-term risks to their hearing health. Aeromedical concerns regarding degraded hearing include difficulty with in-flight communications and understanding radio transmissions, as well as missing audio warnings and alerts in the cockpit. Consequently, Army aviators are required to have a level of hearing acuity that allows for effective communication in operational environments. Our study definition of operational performance, in the context of auditory performance in the aircraft, refers to the ability of a pilot to effectively hear and understand communications while simultaneously maintaining the ability to perform essential tasks such as aviating, navigating, and managing aircraft systems. An aviator's hearing ability can impact their overall cognitive workload, as it influences the level of mental and physiological resources required to complete tasks and maintain situational awareness in the aircraft.

In 2019, the Department of the Army Pamphlet 40-502[1] and Army Regulation 40-502,[2] which governs medical standards for service members in the U.S. Army, were updated to include a functional assessment of auditory performance known as the

From the U.S. Army Aeromedical Research Laboratory, Fort Novosel, AL, United States; Goldbelt Frontier, LLC, Alexandria, VA, United States; and the Oak Ridge Institute for Science and Education, Oak Ridge, TN, United States.

This manuscript was received for review in August 2024. It was accepted for publication in November 2024.

Address correspondence to: Jennifer Noetzel, Ph.D., AuD, 6901 Farrell Road, Fort Novosel, AL 36362, United States; Jennifer.m.noetzel.mil@health.mil.

DOI: https://doi.org/10.3357/AMHP.6570.2025

clinically adapted Modified Rhyme Test (MRT). The current study examined the impact mild and severe degrees of simulated hearing loss have on flight performance and cognitive workload in aviators while assessing the MRT as a predictive measure of operational performance in a flight simulator.

The current hearing standards for Army aviators are outlined in U.S. Army Aeromedical Policy Letters (APL).[3] The hearing standards for aviation are more stringent compared to other military occupations; however, these standards have traditionally been based primarily on pure tones and speech recognition test scores in quiet environments. Currently, no research exists that directly supports the predictive value of pure tone thresholds or word recognition in quiet on aviator-related performance. Hearing loss in general has been shown to be detrimental to the effectiveness of a dismounted soldier,[4,5] but individual differences due to varying degrees of hearing loss remain uncertain.[6] Although some evidence in the literature suggests a synergistic relationship between the degree of hearing loss and aviator performance during portions of flights with high workload, there is no evidence of predictive value in using pure tones for performance outcomes.[7] In pursuit of operationally relevant auditory performance data, the Army developed the clinically adapted MRT to evaluate the functional implications of hearing loss.[8,9] This speech test was adapted to evaluate the functional auditory performance of individuals with elevated audiometric thresholds. Therefore, if a service member's audiometric thresholds do not meet the Army standard, the adapted MRT is administered as part of a test battery, and functional performance dictates their disposition. For a full description and scoring summary of the Army's Military Operational Hearing Test, which includes the MRT, refer to Brungart et al.[8] A passing score on the MRT for individuals with normal hearing is 69%.

The APL outline serial Class categories 1 through 4 (C1-C4). All aviator applicants are C1. C2 represents all rated aviators. C3 and C4 refer to trained aviation personnel with a requirement for flight status and include flight surgeons, aircrew, air traffic controllers, and unmanned aerial systems operators. **Table I** outlines the audiological thresholds required by the APL. If the Class category is met with audiological thresholds alone, no further assessment is required. If an aviator does not meet the APL standard, a waiver is considered. The current audiological workup required for a waiver includes pure tone air and bone conduction testing, tympanometry, acoustic reflex testing, speech reception threshold testing, and word recognition scores (WRS) in quiet in both monaural and binaural conditions. There is a requirement for all aviators to score ≥84% on binaural WRS. If an aviator's score is <84%, the APL notes a requirement for an in-flight evaluation. The nature of this in-flight evaluation is determined by the aeromedical provider and is not standardized. The in-flight evaluation is an attempt to ensure that an aviator's hearing loss does not have a negative functional impact on their operational performance; however, in-flight evaluations can be fiscally expensive and time intensive. Therefore, the adapted MRT may provide a clinically feasible alternative to in-flight evaluations should it prove to be more predictive of functional operational performance in aviators than current word recognition scores.

Table I. APL Hearing Standards in Decibel Hearing Level (dB HL) for Army Aviation.

CLASS	500 Hz	1000 Hz	2000 Hz	3000 Hz	4000 Hz	6000 Hz
1	≤ 25	≤25	≤25	≤35	≤45	≤45
2/3/4	≤ 25	≤ 25	≤25	≤35	≤55	≤65

Cognitive workload describes the level of mental or physiological resources required by an individual to complete one or more tasks.[10] Hearing difficulties increase the cognitive workload required to understand auditory input, which is referred to as "listening effort".[11,12] Growing support in the literature suggests listening effort be considered an additional metric to complement speech intelligibility when quantifying functional impacts of various hearing conditions.[12,13] Noise and the threat of noise-induced hearing loss remains ubiquitous in Army aviation operations. As such, understanding the effects of hearing loss on cognitive workload and aircrew-related tasks remains integral to optimizing aircrew performance. Although an aviator with a hearing loss may be able to understand the same percentage of radio communications as expected with normal hearing, the amount of listening effort required to achieve the same level of performance is likely much greater and may also impact secondary task performance.[14] Previous research has shown that an increase in the amount of mental effort demanded by a primary task results in an increase in reaction time for completing a secondary task.[15,16]

The objectives of this study were to: 1) quantify the effect of hearing loss on functional performance, perceived cognitive workload, and listening effort in aviation; 2) examine the interaction effect of workload and hearing loss on functional performance; and 3) to establish correlational relationships between operational performance and MRT performance, and how that relationship changes with workload.

METHODS

Study M-10,999 was approved by the U.S. Army Medical Research and Development Command Institutional Review Board prior to execution. Using a mixed-effects design, subjects completed two sets of assessments: 1) clinical audiometric and word recognition testing, and 2) simulated flight missions in a full-motion Utility Helicopter (UH)-60 Black Hawk simulator at the U.S. Army Aeromedical Research Laboratory, Fort Novosel, AL, United States. Each subject experienced two listening conditions: normal hearing and one of two simulated hearing loss conditions (mild or severe) for both assessments. In the simulator phase, aviator performance was compared in high and low workload levels between the normal hearing and hearing loss conditions.

Subjects

Army rotary-wing aviators were recruited from Fort Novosel, AL, and were self-screened for current active flight status. Recruited were 31 aviators between ages 25 to 61 yr (39.7 ± 9.7, mean ± SD). All aviators were native English speakers and deemed fit to fly at the time of the study. There were 10 subjects excluded from the study due to having hearing thresholds >25 decibels (dB) hearing level (HL) at one or more test frequencies. A total of 21 subjects completed the study (20 men and 1 woman).

Materials

The hearing loss simulator was comprised of two USB sound cards connected to a laptop. Stereo communication earplugs (CEPs) were connected as the output of one sound card, and the input to the opposing sound card was used to deliver the audio stimulus being tested. The system programming and real-time audio was managed using a software architecture based on the commercially available Hearing Loss and Prosthesis Simulator headset (Sensimetrics Corporation, Gloucester, MA, United States).[17] Depending on the hearing loss being simulated, the algorithm passed audio from the input of the sound card at select frequencies and raised the absolute detection thresholds (i.e., dB HL) for other frequencies by attenuating the levels being sent to the CEPs and adding masking noise. That is, the system was designed to give the user a sense of loudness recruitment, which is the unusually rapid increase in perceived loudness with increases in intensity that occur concurrently with a particular hearing loss. This rendered sounds that fell below the predetermined threshold at specific frequencies inaudible, while sounds that are presented well above the threshold are as loud as they would be to a listener whose hearing is uncompromised. Refer to **Table II** for simulated hearing loss levels. One important note is that the present study examined the impacts of initial hearing loss rather than adaptation to hearing loss over time.

This method of simulating hearing loss has been used in previous studies.[6,18,19] For clinical testing, the hearing loss simulator received an audio output signal from a calibrated audiometer and routed the processed signal to stereo CEPs. For the MRT test via tablet, the hearing loss simulator received the audio output from the tablet and routed the processed signal to the CEPs. Sound pressure levels were measured on an acoustic test fixture to ensure the hearing loss simulator was applying the appropriate levels of attenuation.

Procedure

Subjects completed the study individually over the course of a 4-h visit. Following completion of the informed consent process, a brief questionnaire was used to record demographic information. Next, an audiological evaluation consisting of otoscopy and audiometry was performed in an audiometric booth. The otoscopic exam ensured there were no abnormalities that might interfere with testing or proper fitting of the CEPs. Audiometric testing was conducted to ensure participant hearing thresholds were normal [≤25 dB HL at each of the tested frequencies (500, 1000, 2000, 3000, 4000, and 6000 Hz)] and to identify the amount of attenuation to apply during the simulated hearing loss conditions. All subjects were randomly assigned to either a mild (MHL) or severe hearing loss (SHL) condition representing either a C2 or greater than C2 classification as outlined in the APL (see Table II).

Following threshold testing, word recognition testing was conducted in quiet, both monaurally and binaurally, using the Northwestern University Auditory Test Number Six (NU-6) wordlists with a normal hearing condition (NHC) and a simulated hearing loss condition (MHL or SHL). Word recognition tests using the NU-6 were conducted binaurally at each subject's preferred listening level for both hearing conditions. The clinically adapted MRT was administered next via tablet using CEPs. The MRT was administered binaurally and scored according to its test instructions. Performance outcomes were defined as percent correct on the binaural NU-6 and MRT performance. Both were included in the analyses.

Following completion of the clinical audiometry and speech recognition testing, subjects completed four flight routes in a UH-60 Black Hawk helicopter simulator. Flight simulator tasks consisted of maintaining heading, altitude, and airspeed, listening for directions from the air traffic controller, and responding to the master caution warning light whenever the aviator noticed it illuminate. Each route was a combination of workload (low workload or high workload) and hearing condition (no hearing loss or simulated hearing loss).

- NH-Low – Normal Hearing/Low Workload
- HL-Low – Hearing Loss (mild/severe)/Low Workload
- NH-High – Normal Hearing/High Workload
- HL-High – Hearing Loss (mild/severe)/High Workload

In the study, the within-subjects variable is the combination of workload (low or high) and hearing condition (normal hearing or simulated hearing loss). This means that each participant experienced all four combinations of workload and hearing

Table II. Normal Hearing and Simulated Hearing Loss Thresholds in dB HL (Hearing Level).

HEARING CONDITION	EAR	500 Hz	1000 Hz	2000 Hz	3000 Hz	4000 Hz	6000 Hz
Normal Hearing (NHC)	Both Ears	<25	<25	<25	<25	<25	<25
Mild Hearing Loss (MHL)	Better Ear	25	25	25	35	50	55
	Worse Ear	25	25	25	35	50	55
Severe Hearing Loss (SHL)	Better Ear	35	35	35	45	65	90
	Worse Ear	35	35	35	45	65	90

NHC thresholds meet the C1 Aeromedical Policy Letter (APL). MHL thresholds meets the C2 APL, SHL thresholds meets the criteria for those who do not meet the C2 APL.

condition (NH-Low, HL-Low, NH-High, and HL-High) during the flight simulator tasks. The between-subjects variable is the subjects' hearing condition, as subjects were assigned only to either a mild or severe hearing loss condition.

Workload was manipulated by the number of radio calls and master caution light instances that the subject had to respond to during each 10-min flight. Specifically, workload was increased by delivering an increased number of radio calls, both target and maskers, increasing the number of master caution warning light instances, and increasing the level of turbulence such that the aviator had to continually adjust the controls to maintain heading and elevation. Each route was flown under instrument meteorological conditions while simulating radar vectors from an air traffic controller via prerecorded voiceover broadcasts.

Instrument routes were simulations of flights in northern California and used prerecorded air traffic controller broadcasts. Subjects started each route at 4000 ft (1219 m) mean sea level and 110 kn indicated airspeed. All subjects flew NH-Low first as a baseline. Order of subsequent routes was randomized. Subjects were asked to maintain a selection of appropriate common standards from the UH-60 Series Aircrew Training Manual throughout the duration of the flight such as maintain heading (±10°), maintain altitude [±100 ft (30.5 m)], or maintain airspeed (±10 kn).

In-flight speech recognition was measured based on responses to radio calls. Subjects were allowed three attempts at responding to each target radio call within 15–20 s of the call occurring. If the subject correctly acknowledged the call and no repeat was necessary, it was counted as correct. If the subject missed the call completely, had an incorrect read back, or requested the tower to "say again," it was counted as incorrect. If the subject failed all three attempts, a research copilot, a pilot who was part of the research team, would dial the correct heading/altitude and notify the subject. In-flight speech recognition was scored based on percent correct. Flight performance in the simulator was quantified by calculating the root-mean-square deviation (RMSD) for three flight metrics: altitude, heading, and airspeed. Subjects were instructed to maintain a constant airspeed of 110 kn for all routes. Adjustments to heading and altitude requirements were provided to subjects via radio calls. Subjects were instructed to follow a standard rate of climb [500 ft $\cdot$ min^{-1} (152 m $\cdot$ min^{-1})] and a standard rate of turn (180°/min) during all flights.

Subjects completed a secondary task that required them to respond to a randomized illumination of the "Master Warning" light on the panel. This was administered by broadcasting a "Check EICAS" (engine indicating and crew alerting system) caution into the simulator at scheduled times. The Check EICAS caution would illuminate for up to 5.5 s and then extinguish regardless of the participant's response. Subjects were able to acknowledge this caution by pressing the VOX-CAUT switch on the cyclic or by pressing the Master Caution Press to Reset Button, congruent with the UH-60 Technical Manual (TM 1-1520-280-10).[20] If the subject responded to this stimulus, the Master Caution button light would extinguish.

Subjects were instructed to turn off the Master Caution warning light as quickly as possible. The warning light came on 20 times during low workload flights and 25 times during high workload flights. The warning light response times were averaged, considering only the instances where a response was made. The resulting average response time for the warning light was used for statistical analysis.

Following each route, subjects were asked to complete a subjective questionnaire, the National Aeronautics and Space Administration Task Load Index (NASA-TLX), which is a widely used, multidimensional assessment tool that asks individuals to rate their perceived workload.[21] The NASA-TLX results in unweighted workload scores derived from subjective ratings according to six subscales: mental demand, physical demand, temporal demand, effort, frustration, and performance. For each of these subscales, a 21-point scale was used with verbal anchors at the beginning and ending of the scale (e.g., Good at the beginning and Poor at the end of the scale). Subjects were asked to rate their perception for each of the subscales at the completion of each route while in the flight simulator. Within-subjects differences were calculated for each category of each route and averaged across subjects. The resulting change in subscales between workload and hearing loss were analyzed.

Statistical Analyses

All statistical analyses were performed using R (version 4.2.3, R Core Team, Vienna, Austria) and R Studio (version 2022.07.1, RStudio Team, Boston, MA) with the following packages: tidyverse, rstatix, and lmerTest.[22–24] All statistical tests were evaluated at a significance level of 0.05.

Mixed-effects linear regression models (from the lmerTest package) were used to analyze potential differences in RMSD for each flight metric while also accounting for individual differences in subjects. Each regression model contained a fixed factor for route (four levels), a fixed factor for hearing loss (two levels: MHL, SHL), an interaction between route and hearing loss, and a random intercept for each subject. Maximum likelihood was chosen as the estimation method so that the fixed and random effects could be estimated simultaneously.

Prior to the regression model, the data were evaluated to ensure a normal distribution. An appropriate transformation was applied to any data that were not approximately normally distributed. Outliers were removed from the data if the assumptions of the linear model were violated. Each regression model was checked to ensure that the residual values were normally distributed and had a constant variance. When the regression model showed a significant interaction effect between route and hearing loss, the data were split by hearing loss condition and new regression models were created as described above. When the data were split and retested, *P*-values were adjusted using the Bonferroni method.

RESULTS

Binaural NU-6 WRS in quiet for the NHC had the highest average percent correct (97%). As expected, when simulated hearing loss went from mild to severe, the percent correct decreased (MHL scored 93% and SHL scored 88%). These findings are expected as decreases in audibility should result in decreases in speech recognition. The current APL identifies ≥84% binaural WRS in quiet as the criterion for continued service in aviation. No subject in the NHC scored <92%. In the MHL condition, there were two subjects who scored <84%. In the SHL, two subjects scored <84%, (both scored 72%) and two subjects scored exactly 84%.

There were 21 subjects who completed the clinically adapted MRT in the NHC, 11 in the MHL condition, and 10 in the SHL condition. According to official test instructions, if a subject's score does not meet or exceed the 80-word list MRT scoring criteria, they complete a second MRT_{80} word list and, when this occurs, the test is referred to as MRT_{160}. Two subjects in the MHL condition did not complete the MRT according to directions and were excluded from the data below. The results of the MRT followed the same pattern as WRS in quiet with increasing hearing loss resulting in poorer performance. See **Table III** for results across all hearing conditions. The NHC MRT results demonstrated the highest average at 69%, meeting the passing criteria for normal hearing.[8,9]

In-flight speech recognition scores were calculated based on the number of correct radio calls. **Table IV** shows the mean percent correct within each workload and hearing loss condition. As shown in Table IV, increasing the workload did not drastically change the overall in-flight speech recognition in the NHC. The presence of simulated hearing loss resulted in reduced speech recognition in both low and high workload conditions. The most significant decreases were observed in the high workload condition, with a 26% decrease in speech recognition for mild hearing loss and a 40% decrease for severe hearing loss. In low workload conditions, the decrease in speech recognition was comparatively small, with a 13% decrease for mild hearing loss and a 40% decrease for severe hearing loss.

The term "ideal flight path" refers to the path the aircraft would follow if aviators perfectly maintained all the instructed headings and altitudes throughout the route. Flight performance was evaluated based on deviations from ideal flight paths across altitude, heading, and speed. In general, larger deviations occurred in the high workload routes compared to the low workload routes. Hearing loss contributed additional decreases in flight performance for altitude and heading deviations in the high workload conditions.

Table III. Average MRT_{80} and MRT_{160} Overall Percent Correct.

HEARING CONDITION	MRT_{80}	MRT_{160}
Normal Hearing	69 ± 7.8%	65 ± 5.1%
Simulated Mild Hearing Loss	65 ± 6.4%	67 ± 5.4%
Simulated Severe Hearing Loss	53 ± 7.1%	63 ± 6.5%

The benchmark criteria for the 80-word Modified Rhyme Test (MRT_{80}) are 69% for normal and mild hearing loss and 74% for severe hearing loss. The benchmark criteria for the 160-word MRT (MRT_{160}) are 65% for normal and mild hearing loss and 70% for severe hearing loss, as specified in Department of the Army Pamphlet 40-502.

Table IV. Average Percent Correct for In-Flight Speech Recognition.

WORKLOAD	NHC	MHL	SHL
Low	89%	76%	49%
High	93%	67%	48%

NHC: normal hearing condition; MHL: mild hearing loss; SHL: severe hearing loss.

The mixed-effects linear regression model for altitude RMSD showed that flight route was statistically significant [F(3, 60.8) = 67.98, P < 0.001]. Hearing level (P = 0.41) and the interaction effect (P = 0.42) were not statistically significant. Pairwise comparisons of routes were extracted directly from the regression model and showed that altitude RMSD was significantly different between routes: NH-Low and NH-High (P < 0.001), HL-Low and HL-High (P < 0.001), and NH-High and HL-High (P = 0.006). The NH-Low and HL-Low did not show a statistically significant difference (P = 0.40) (see **Fig. 1**). All these results point to a detrimental effect of hearing loss with a differential effect of workload, whereby increased workload had a greater effect on flight performance, specifically altitude, in the SHL compared to the MHL conditions.

The mixed-effects linear regression model for heading RMSD showed that flight route [F(3, 60.1) = 47.63, P < 0.001] and the interaction between route and hearing level [F(3, 60.1) = 5.62, P = 0.002] were statistically significant. Hearing level was not statistically significant (P = 0.44). Due to the significant interaction effect, the data were split into two groups based on simulated hearing loss levels (MHL and SHL).

The mixed-effects linear regression model for heading RMSD in the MHL group showed that flight route continued to be significant [F(3, 31.9) = 30.80, P < 0.001]. Pairwise comparisons of routes showed that heading RMSD for the MHL group was significantly different between NH-Low and NH-High (P < 0.001), HL-Low and HL-High (P < 0.001), and NH-High and HL-High (P = 0.010). The comparison of NH-Low and HL-Low was not significant (P = 0.42). These results show a detrimental effect of workload within the NHC and simulated HL condition, a detrimental effect of hearing loss within the high workload condition, and no effect of hearing loss within the low workload condition.

The mixed-effects linear regression model for heading RMSD in the SHL group showed that route continued to be statistically significant [F(3, 28.5) = 22.88, P < 0.001]. Pairwise comparisons of route showed that heading RMSD for the SHL group was significantly different between all conditions: NH-Low and HL-Low (P = 0.009), NH-Low and NH-High (P = 0.004), HL-Low and HL-High (P < 0.001), and NH-High and HL-High (P < 0.001). These results support a detrimental effect on heading for high workload conditions as well as a negative impact of hearing loss.

The mixed-effects linear regression model for airspeed RMSD showed that route was statistically significant

[$F(3, 63) = 3.37$, $P = 0.024$]. Hearing level ($P = 0.26$) and the interaction effect ($P = 0.20$) were not statistically significant. Pairwise comparisons of routes showed that airspeed RMSD was significantly different between HL-Low and HL-High ($P = 0.020$), indicating a significant difference between the high and low workloads when hearing loss was present. All other routes showed no significant differences (see Fig. 1 for airspeed deviations).

Response times were analyzed for acknowledgment of the warning response light. The hit rate for this task was consistently above 89% for all flight routes. The average response times to the secondary task were all less than 2 s. The mixed-effects linear regression model showed that route was significant [$F(3, 1626.79) = 4.3$, $P = 0.005$]. Hearing level ($P = 0.22$) and the interaction effect ($P = 0.18$) were not statistically significant. See **Fig. 2** for response time comparisons.

The addition of hearing loss to the low workload condition revealed a significant difference ($P = 0.0174$). Subjects responded slowest in the baseline NH-Low flight, which is likely attributed to a learning effect as it was always completed first. Furthermore, the increased workload within the hearing loss condition also showed a significant difference ($P = 0.0457$). This finding supports the hypothesis that in the high workload condition, the addition of hearing loss resulted in slower response times to a secondary task during simulated flight. None of the other flight comparisons reached significance.

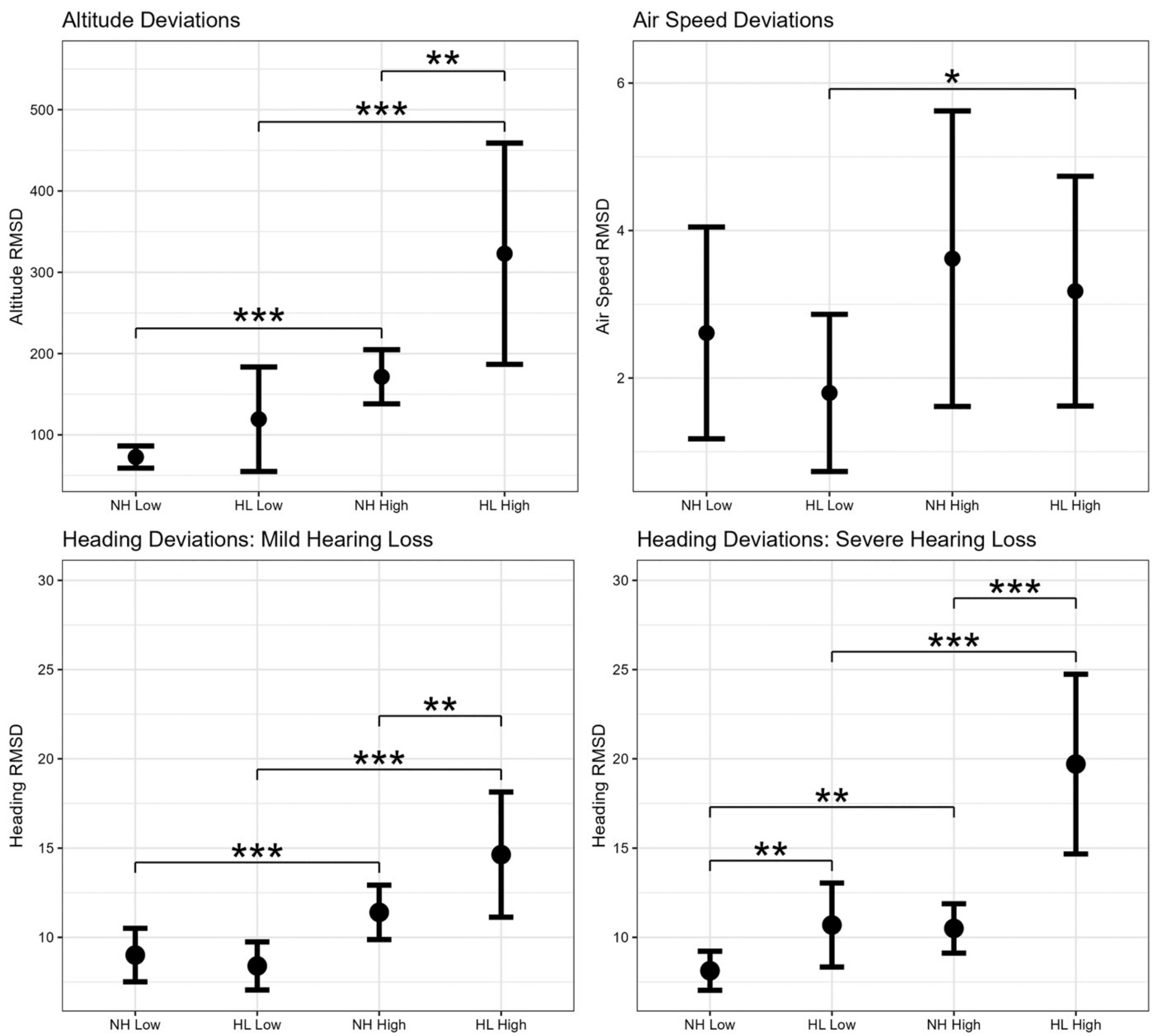

Fig. 1. Summary of flight path deviations for each listening and workload condition. Significant pairwise comparisons annotated. Due to the significant interaction in the heading category, the bottom of the figure breaks down the significance within the simulated hearing loss conditions of mild hearing loss and severe hearing loss. Significance levels: *0.05, **0.01,***0.001.

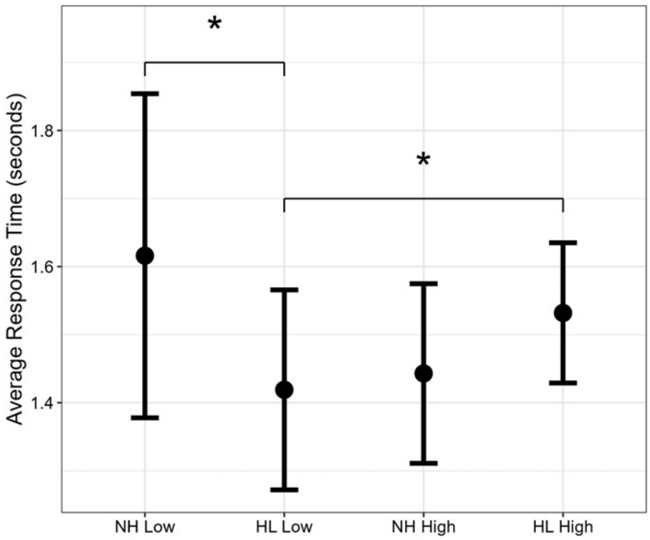

Fig. 2. Average response time and *t*-tests for the secondary task of acknowledging the master caution warning light. *$P = 0.05$.

Subjects completed the NASA-TLX following each route. Unweighted average subscale differences are plotted in **Fig. 3**. As expected, the NASA-TLX scores reflected an increase in perceived workload when either hearing loss or increases in workload were applied. Results of the NASA-TLX showed an increase in perceived workload on nearly every subscale and decrease in perceived performance when comparing high vs. low workloads.

MRT percent scores were correlated to heading and altitude deviations using the Pearson correlation coefficients for each of the hearing conditions (see **Fig. 4**). RMSD values were plotted against the subject's MRT percent score. Correlational analysis indicated that the variability in the outcome data could not be explained by the model and that there was no significant correlation between the MRT score and flight performance in the form of deviations from altitude, heading, or airspeed. However, when examining deviations in heading, there were patterns beginning to emerge, specifically in the HL-High workload condition. As performance on the MRT improved, the deviations in heading became smaller.

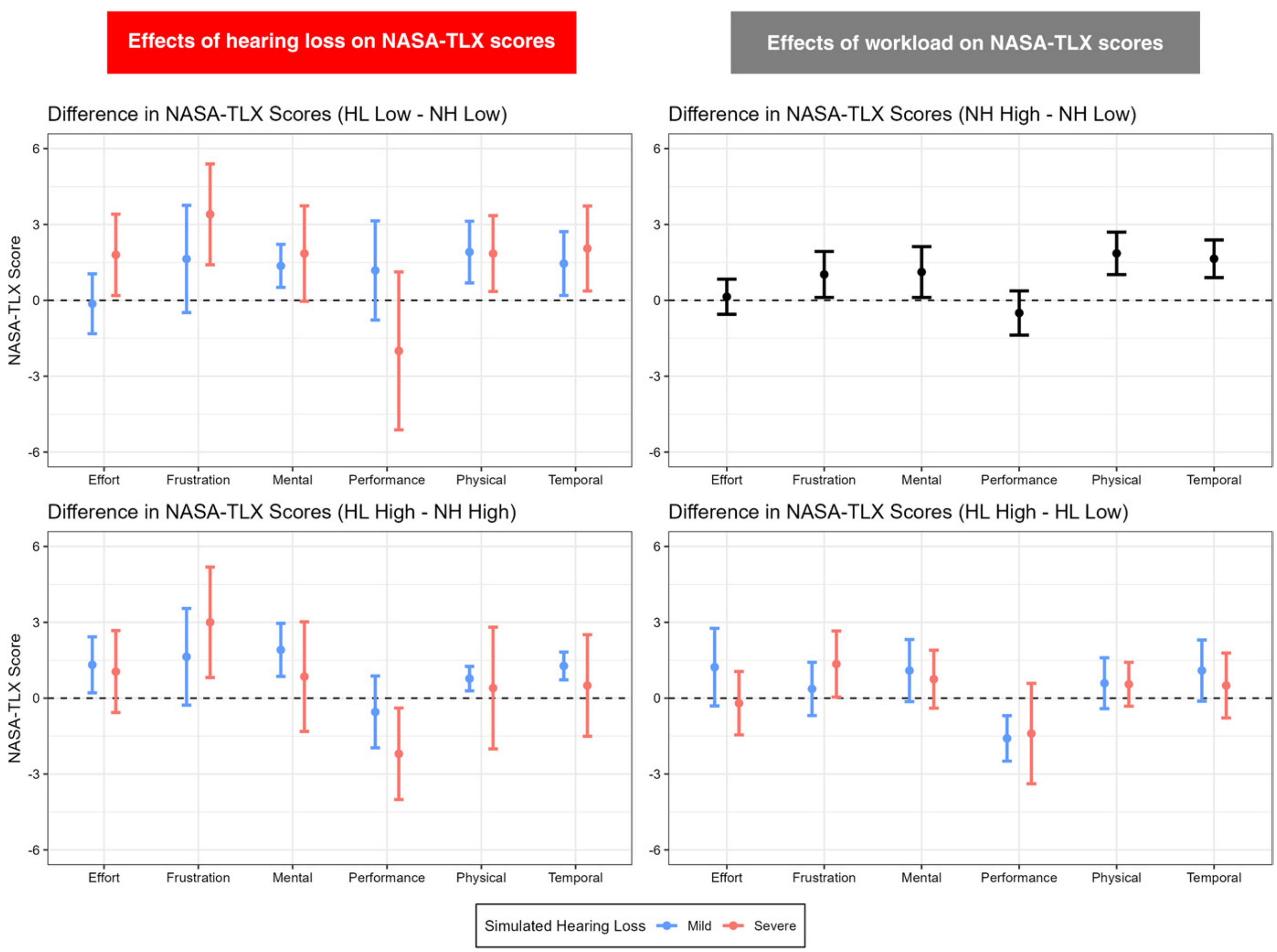

Fig. 3. Summary of all flight route NASA-TLX score comparisons. The graphs on the left display the effects of hearing loss on NASA-TLX scores with workload held constant, while the graphs on the right show the effects of workload on NASA-TLX scores with the hearing condition kept constant. Each point represents the across-participant average difference in score and standard deviation for each subscale (i.e., effort, frustration, etc.) of the survey. Positive values above the dotted line indicate an increase in the metric listed on the y-axis, whereas negative values represent a decrease in score.

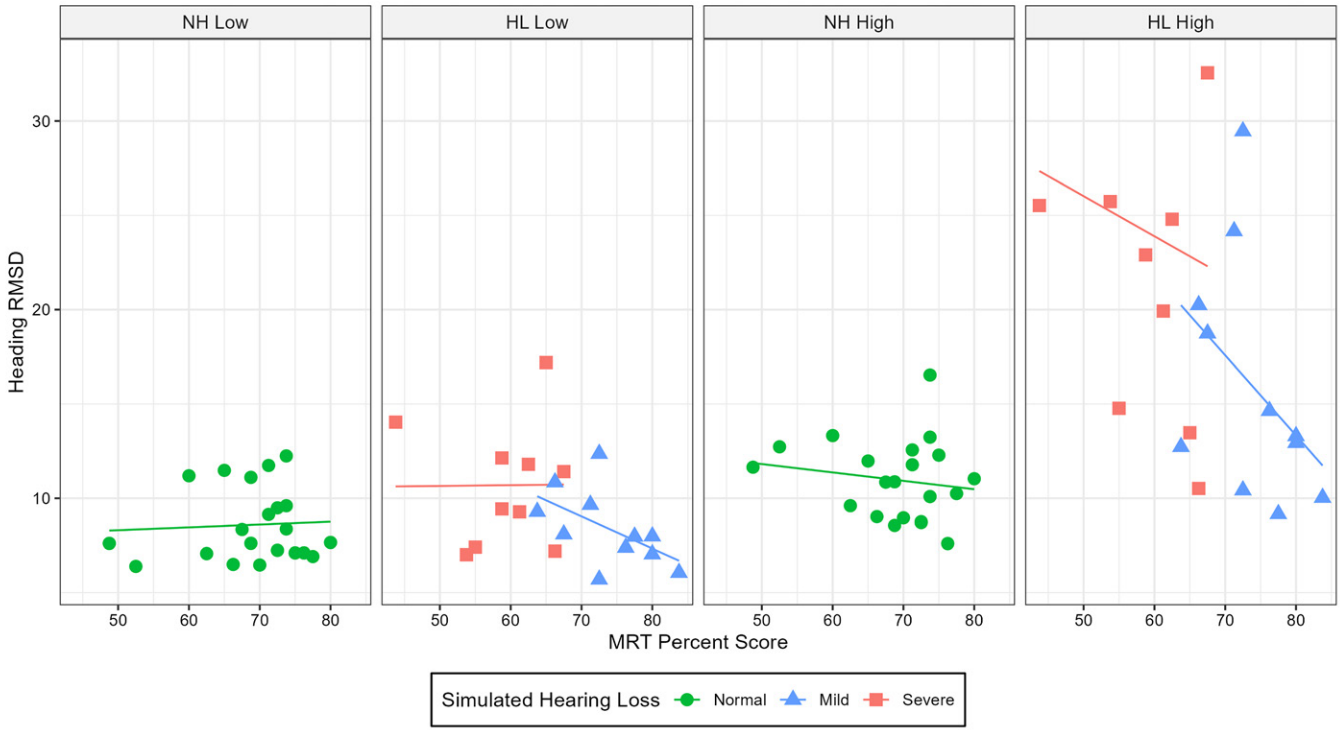

Fig. 4. Correlations between the heading deviations and MRT score. One simulated hearing loss data point in the high workload condition for SHL is not depicted on this figure.

DISCUSSION

The current study investigated the impact of simulated hearing loss and changes in workload in U.S. Army aviators on operational performance and cognitive workload during simulated flight. Subjects completed simulated flights under two listening conditions (normal hearing and simulated hearing loss) and in high and low workloads. Results of speech recognition testing showed the hearing loss simulation was successful, as indicated by decreases in WRS, MRT scores, and in-flight speech recognition across the hearing loss conditions.

NASA-TLX surveys confirmed that the aviators' perceptions of workload matched the changes in workload that were implemented by increasing the number of radio communications and the frequency of the master caution light illuminations during simulated flight. The NASA-TLX was analyzed to determine the degree to which the increases in workload were manifested and the degree to which hearing loss contributed to perceived workload. The NASA-TLX scores reported here indicated that the simulated hearing loss increased the aviators' workload. Scores on the NASA-TLX validated the change in workload as NASA-TLX scores changed with increased workload. Scores also suggested that listening effort increased with the addition of a simulated hearing loss.

It was hypothesized that higher workload and the presence of hearing loss would disrupt response times to the master caution warning light. The results showed that in the low workload condition, the addition of hearing loss led to faster response times. However, it is important to consider that the NH-Low condition was always completed first, which may have influenced the observed improvements. This suggests that the effects observed could be attributed to a learning effect rather than a change due to hearing loss or workload. Another possible explanation is that individuals with hearing loss were unable to hear radio communications, allowing them to focus more on visual information and leading to faster response times in the low workload condition.

Lastly, it may be the case that the subjects are too well-trained to respond to the warning caution light and that it would take a much larger increase in workload than was used in this study to start to see response times increase. As anticipated, in the high workload condition, the addition of hearing loss resulted in slower response times. This supports the hypothesis that hearing loss can have a detrimental effect on response times, particularly in high workload situations. However, the secondary task used in the study was unable to consistently differentiate between the hearing loss and workload conditions based on the flight profiles used.

Flight performance was assessed by examining deviations from the ideal path, specifically focusing on altitude, heading, and speed. Analyses revealed that the combination of hearing loss and high workload led to degraded performance across all flight performance metrics. However, in low workload conditions, performance was comparable regardless of the presence of hearing loss, both in terms of altitude and heading. Increasing the workload, whether in the normal hearing or simulated hearing loss condition, resulted in higher deviations in both altitude and heading. The largest deviations were observed in high workload conditions. Additionally, there was an interaction effect between hearing loss and workload on heading deviations, indicating that the impact of hearing loss on heading deviations may vary depending on the workload. Further analysis was conducted by dividing the data based on simulated hearing loss. Among individuals with simulated severe hearing loss, deviations in heading were also observed, specifically in high workload conditions. Heading deviations can have serious consequences, including navigation errors and compromised flight safety. Therefore, addressing and mitigating the functional impacts of heading, particularly in high workload conditions and for individuals with hearing loss, is of the utmost importance to ensure optimal flight performance and safety.

Sheffield et al. demonstrated with Navy watch standards that hearing loss increased perceived workload and increased the latency of the crew response time to threats and orders.[19] The current study demonstrated an increase in perceived workload with the introduction of hearing loss; however, there was only an increase in response times in the high workload condition. The current study closely aligns with Casto and Casali, who reported that flight performance and speech intelligibility decreased when aviators were operating with degraded communication signals, below 50% intelligibility, and in high workloads.[7] Additionally, Casto and Casali found that during poor signal quality conditions and high workload, the number of readbacks increased as well as flight deviations for altitude and airspeed. Casto and Casali advocated for the development of a functional hearing assessment for aviators with hearing loss.

The clinically adapted MRT was adopted by the Army as a test to identify individuals at high risk for poor performance on auditory-related tasks.[8,9] One of the analyses within the current study looked at whether the MRT, a functional hearing assessment, could be used to predict aviator performance. Correlational analysis between MRT performance and flight performance metrics showed no significant correlations. Although it did not reach statistical significance, there are some emerging patterns where better MRT scores were associated with better flight performance, particularly in terms of heading.

For heading deviations in the hearing loss and high workload condition, the data show that the MRT may be predictive of performance, although it was not found to be statistically significant. It should be noted that this could potentially be due to the small sample size. Better predictive value from the heading data were observed compared to altitude, because there were more changes in heading compared to changes in altitude during these flight plans. Consequently, there are more opportunities for aviators to miss heading calls compared to altitude calls. Further studies are needed to investigate the value of the MRT for predicting functional performance in aviators at higher workloads.

Regarding the choice between using pure tones, an in-flight evaluation, or the clinically adapted MRT, it is difficult to make a specific recommendation without more information about the characteristics and performance of these assessment tools. It is important to consider factors such as reliability, validity, and practicality when selecting an assessment tool for auditory performance in aviators. Further studies are needed to compare the effectiveness of different assessment tools and determine which one provides the most accurate and reliable predictions of functional performance in aviation. Standardization is indeed important to determine changes in performance over time, especially in individuals with hearing loss. Having standardized assessment protocols and criteria can help track and monitor performance trends, allowing for more objective evaluations and comparisons. Tracking performance over time can be particularly valuable in identifying any declines in performance and implementing appropriate interventions or accommodations.

The findings of this study showed that hearing loss resulted in decreased speech recognition, both clinically and in the simulator. Most notably, workload increases resulted in flight performance degradation with the addition of hearing loss further impacting performance. The hearing task in the experiment is likely more cognitively demanding for those with hearing loss than for those without. Although MRT comparisons to flight performance were not significant, further research should examine different methods of increasing workload along with simulations of hearing loss to further determine whether or not the MRT can be used as a predictor of flight performance. It is important to note that the present study focused on simulated hearing loss within the context of Army aviation, and further research and validation are necessary to generalize the findings to broader contexts and populations. Future studies should include aviators with hearing loss. This is important because the effect of permanent or long-term sensorineural hearing loss could potentially be more severe due to cochlear distortion, or it could be better because individuals may be more experienced in dealing with their own hearing loss compared to a simulated one. By including aviators with hearing loss, researchers can gain a better understanding of the specific challenges and impacts faced by this population.

ACKNOWLEDGMENT

The views expressed are those of the authors and do not reflect the official guidance or position of the U.S. Government, the Department of Defense (DoD), or the U.S. Army. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government. The appearance of external hyperlinks does not constitute endorsement by the DoD of the linked websites, or the information, products, or services contained therein. The DoD does not exercise any editorial, security, or other control over the information you may find at these locations. Imagery in this document are property of the U.S. Army.

Financial Disclosure Statement: The authors have no competing interests to declare.

Authors and Affiliations: Jennifer Noetzel, Ph.D., Au.D., Paula Henry, Ph.D., Ryan Mackie, M.S., Kara Cave, Ph.D., J. R. Stefanson, M.S., J. Kyle Hale, Kevin Andres, and Heath Jones, Ph.D., U.S. Army Aeromedical Research

Laboratory, Fort Novosel, AL, United States; Paula Henry, J. Kyle Hale, and Kevin Andres, Goldbelt Frontier, LLC, Alexandria, VA, United States; and Ryan Mackie, Oak Ridge Institute for Science and Education, Oak Ridge, TN, United States.

REFERENCES

1. U.S. Department of the Army. Medical readiness procedures (DA PAM 40-502). Washington (DC): U.S. Government Publishing Office; 2019.
2. U.S. Department of the Army. Medical readiness (Army Regulation 40-502). Washington (DC): U.S. Government Publishing Office; 2019.
3. U.S. Army Aeromedical Activity. 2021. Aeromedical Policy Letters and Aeromedical Technical Bulletins. [Accessed 1/30/2025]. Available from https://aero.health.mil.
4. Peters LJ, Garinther GR. The effects of speech intelligibility on crew performance in an M1A1 tank simulator. Aberdeen Proving Ground (MD): Human Engineering Lab; 1990. Report No.: HEL-TM-11-90.
5. Price GR, Hodge DC. Detection of combat sounds by the human ear. Aberdeen Proving Ground (MD): U.S. Army Human Engineering Lab; 1976. Technical Note.
6. Sheffield B, Brungart D, Tufts J, Ness J. The effects of elevated hearing thresholds on performance in a paintball simulation of individual dismounted combat. Int J Audiol. 2017; 56(suppl. 1):34–40.
7. Casto KL, Casali JG. Effects of headset, flight workload, hearing ability, and communications message quality on pilot performance. Hum Factors. 2013; 55(3):486–498.
8. Brungart DS, Makashay MJ, Sheffield BM. Development of an 80-word clinical version of the modified rhyme test (MRT80). J Acoust Soc Am. 2021; 149(5):3311–3327.
9. Brungart DS, Sheffield BM, Galloza H, Schurman JR, Russell S, et al. Developing an evidence-based military auditory fitness-for-duty standard based on the 80-word modified rhyme test. Ear Hear. 2023; 44(1): 209–222.
10. Van Acker BB, Parmentier DD, Vlerick P, Sladien J. Understanding mental workload: from a clarifying concept analysis toward an implemental framework. Cogn Technol Work. 2018; 20(3):351–365.
11. Ohlenforst B, Zekveld AA, Jansma EP, Wang Y, Naylor G, et al. Effects of hearing impairment and hearing aid amplification on listening effort: a systematic review. Ear Hear. 2017; 38(3):267–281.
12. Zekveld AA, Kramer SE, Festen JM. Pupil response as an indication of effortful listening: the influence of sentence intelligibility. Ear Hear. 2010; 31(4):480–490.
13. Winn MB, Wendt D, Koelewijn T, Kuchinsky SE. Best practices and advice for using pupillometry to measure listening effort: an introduction for those who want to get started. Trends Hear. 2018; 22:23312165 18800869.
14. Peelle JE. Listening effort: how the cognitive consequences of acoustic challenge are reflected in brain and behavior. Ear Hear. 2018; 39(2): 204–214.
15. Verwey WB, Veltman HA. Detecting short periods of elevated workload: a comparison of nine workload assessment techniques. J Exp Psychol Appl. 1996; 2(3):270–285.
16. Wickens CD. Processing resource in attention. In: Parasuraman R, Davies R, editors. Varieties of attention. Cambridge (MA): Academic Press; 1984:63–101.
17. Sensimetrics Corporation. Hearing loss and prosthesis simulator (HeLPS). Somerville (MA): Sensimetrics Corp.; 2006.
18. Sheffield B, Brungart D, Tufts J, Ness CJ. The relationship between hearing acuity and operational performance in dismounted combat. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting; Oct. 26–30, 2015; Los Angeles, CA. Vol. 59, Issue 1. Washington (DC): HFES; 2015:1346–1350.
19. Sheffield B, Ziriax J, Keller MD, Barns W, Brungart D. The impact of reduced speech intelligibility on reaction time in a naval combat environment. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting; Oct. 9–17, 2013; Austin, TX. Vol. 61, Issue 1. Washington (DC): HFES; 2017.
20. U.S. Department of the Army. Technical manual operator's manual for helicopters, utility tactical transport UH-60M (TM 1-1520-280-10). Washington (DC): U.S. Government Publishing Office; 2017.
21. Hart SG, Staveland LE. Development of the NASA-TLX (Task Load Index): results of empirical and theoretical research. Advances in Psychology. 1988; 52:139–183.
22. R Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2022. [Accessed Jan. 31, 2025]. Available from: https://www.R-project.org/.
23. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. J Open Source Softw. 2019; 4(43):1686.
24. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: tests in linear mixed effects models. J Stat Softw. 2017; 82(13):1–26.