

Agreement and Test Duration of a Revised Modified Hughson–Westlake Down-Up Procedure for Pure-Tone Audiometry: A Method-Comparison Study

Maureen Joy M. Tanoy, RSLP, MCLinAud¹ and Maria Rina T. Reyes-Quintos, MD, MCLinAud, PhD^{2,3}

¹University of the Philippines Manila

²Philippine National Ear Institute, National Institutes of Health, University of the Philippines Manila

³Department of Otolaryngology–Head and Neck Surgery, Philippine General Hospital, University of the Philippines Manila

ABSTRACT

Background. The Modified Hughson–Westlake Down-Up Procedure (MHW) is the manual method for conducting pure-tone audiometry to determine auditory thresholds in adults. This method requires a repeat of the 10 dB descent in the second ascending run followed by two (2) successive 5 dB incremental ascending steps. It can be argued that this repeated step is unnecessary, since anecdotally and from the authors' experience a 5 dB descent (instead of 10 dB) yields similar ATs while reducing test duration. A Revised MHW (Rev MHW) Procedure eliminates the repeat step by reducing the second descending step from 10 dB to 5 dB, with the aim of improving efficiency while maintaining comparable hearing-threshold measurements.

Objective. The study aimed to determine whether the Rev MHW produced hearing threshold measurements similar to those obtained in the MHW procedure. The secondary objective was to determine whether the Rev MHW reduced the total test duration.

Methods. This was a cross-sectional method-comparison study with a within-subject repeated-measures design. Thirty adults were recruited using convenience sampling from the community, and each participant underwent both testing procedures in randomized order. Inclusion criteria were adults aged 18–60 years with clear ear canals and intact tympanic membranes. Exclusion criteria included an active ear infection, ear malformations, and physical disability that may affect the participants' ability to respond to sound. Bland-Altman analysis was used to determine the strength of agreement between MHW and the Rev MHW. A paired t-test was performed between the total test times of the two procedures to determine the mean difference and its statistical significance.

Results. Bland-Altman analysis demonstrated comparable threshold measurements between MHW and the Rev MHW. Across 600 paired hearing-threshold measurements, the mean difference (Rev MHW – MHW) was 0.55 dB, with 95% limits of agreement ranging from –11.65 to 12.75 dB, indicating minimal systematic bias between procedures. The Rev MHW significantly reduced test duration compared with MHW, with a mean reduction of 1.50 minutes (95% CI, 1.12–1.87 minutes; $p < 0.001$), representing approximately a 22% decrease in administration time.

Conclusion. Results suggested that Rev MHW yields auditory thresholds comparable to those obtained using MHW. The minimal systematic bias observed between methods suggests that the revision did not meaningfully affect threshold estimation while providing a measur-



eISSN 2094-9278 (Online)
Published: October 31, 2026
<https://doi.org/10.47895/amp.v60i20.13767>
Copyright: The Author(s) 2026

Corresponding author:
Maureen Joy M. Tanoy, RSLP, MCLinAud
University of the Philippines Manila
Taft Avenue, Ermita, Manila 1000, Philippines
Email: mmtanoy@up.edu.ph

able reduction in administration time. Rev MHW may be a more efficient alternative for determining ATs in adults while maintaining comparable hearing-threshold measurements.

Keywords: *Modified Hughson-Westlake procedure, pure-tone audiometry, hearing-threshold, audiometric threshold, method-comparison study, Bland-Altman analysis, test duration*

INTRODUCTION

Pure-tone audiometry (PTA) is a test that determines auditory threshold (AT) across frequencies ranging from 250 Hz to 8000 Hz. Testing included determining air conduction (AC) and bone conduction (BC) thresholds and the application of masking noise as needed. The Modified Hughson-Westlake (MHW) Down-Up procedure is the preferred manual method for PTA determination in adult patients.

Walter Hughson and Harold Westlake proposed the “ascending method” procedure, which determines the AT by presenting a stimulus where intensity is increased from an inaudible sound level until a response is elicited.¹ That is, as soon as the stimulus becomes inaudible from the first presented sound stimulus, an ascent in the intensity level is initiated.¹ This ascending method was accepted in 1944 by the Committee on Conservation of Hearing of the American Academy of Ophthalmology and Otolaryngology and was further revised and advanced by Jerger and Carhart in 1959.¹ The MHW adapted the ascending method and elaborated on ascending and descending protocols which included the significance of the 5 dB step size.¹ Since then, it has been accepted by audiologists as a key element of the audiological evaluation and has, essentially, remained the same.²

In the current protocol of the MHW procedure, there are two important steps to follow. The first step is familiarization, which ensures that the patient understands and is able to perform the task required by the test. A clearly perceived stimulus is presented, usually at 30–40 dB at 1000 Hz, where a response is expected.¹ A definitive response is obtained when the patient indicates that they have perceived the stimulus by notifying the tester through a variety of means, e.g., by pressing a signal switch, raising a finger, hand, or arm, or by verbalizing that they have heard the stimulus.³ The stimulus tone may be raised up to 50–70 dB if hearing impairment is present. The key element is to elicit a definitive response to a clearly audible stimulus. The second step is threshold determination, which is performed immediately after the familiarization step. The stimulus level is decreased by 10 dB until inaudibility is reached. From the inaudible intensity level, the stimulus tone is increased in 5 dB steps until a response is obtained. This marks a complete first ascent, which is then followed by a second decrease in intensity level to return to an inaudible level. Again, the key element here

is to start from an inaudible level to begin the second ascent, thus the name ascending technique. The threshold is obtained when a response is observed in two out of three ascending runs, which is the lowest intensity level at which 50% of responses occurred. The audiologist must remember that tone presentation must have a one- to two- second duration, and that intervals between stimulus presentations must be varied.³ Vencovsky and Rund illustrate a flowchart of the MHW.⁴

In Carhart and Jerger’s study, the 5 dB increment/decrement has been established as a reliable interval resistant to errors caused by moment-to-moment response variability.¹ Specifically, the interval allows a high consistency of responses and at the same time restricts the possibility of over-responding or under-responding at some intensity levels, which affects the precision of determining the threshold. Prior attempts were made to reduce the interval to 2dB to obtain a more precise threshold.⁵ However, 2dB intervals were found to be less time-efficient and were not superior to 5dB intervals in test-retest reliability. Particularly, the 2 dB step size resulted in an increased number of threshold crossings.

Testing method efficiency is linked to increased patient cooperation and greater patient capacity while maintaining result accuracy.⁶ Since pure-tone audiometry is considered the “gold standard” in determining ATs in adults, testing duration can significantly impact a clinician’s workload and overall efficiency.^{7,8} The present MHW protocol’s second ascending run has potential for improvement,

In our institution, while performing the MHW, we noticed that the threshold did not change when reducing the intensity from 10dB to 5 dB in many, if not all, of our patients. Therefore, this study was conceptualized to determine whether such modification will have similar diagnostic outcomes to the standard MHW. If similar outcomes are obtained, the Revised MHW (Rev MHW) procedure, with a reduced 10db to 5dB descent in the second ascending run, will result in significantly shorter testing time without sacrificing diagnostic accuracy.

The primary objective of this study was to determine whether the Rev MHW produced hearing threshold measurements that agreed with those obtained using the standard MHW procedure. The secondary objective was to determine whether the revised procedure reduced total test duration. We hypothesized that the revised procedure would demonstrate clinically acceptable agreement with the conventional procedure while improving testing efficiency.

METHODS

This was a cross-sectional, within-subject method-comparison agreement. The study was conducted at the Ear Unit of the Philippine General Hospital, University of the Philippines Manila. Recruitment occurred between March and May 2025 via convenience sampling, with data collection completed between April and June 2025. Because this investigation was a cross-sectional method-comparison

agreement study, it was designed and reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement and the Guidelines for Reporting Reliability and Agreement Studies (GRRAS).^{9,10}

Ethical approval was obtained from the University of the Philippines Manila Research Ethics Board (UPMREB 2024-0155-01), and written informed consent was obtained from all participants before enrollment.

The primary outcome was agreement between hearing thresholds obtained using the MHW and Rev MHW procedures. The secondary outcome was total administration time required to complete each threshold-determination procedure.

This study was conducted in four sequential phases: (1) pre-data collection, during which examiner reliability was established, and study personnel were trained; (2) participant recruitment and eligibility screening; (3) participant scheduling; and (4) data collection, which included randomization, otoscopic examination, audiometric testing using both the MHW and the Rev MHW procedures, and data management. Figure 1 provides an overview of the study workflow and participant progression through each phase of the investigation.

Pre-data Collection Phase

Interrater Reliability

Two trained testers from the University of the Philippines Clinical Audiology graduate program who have completed formal training in audiometry served as data collectors. Both testers underwent standardized protocol training before

Table 1. Intra-class Correlation Coefficient (ICC) Coefficients and 95% Confidence Intervals for Inter-Rater Reliability

		ICC	95% CI		Interpretation
			Lower band	Upper band	
Right Ear	Individual	0.78	0.68	0.86	Excellent reliability
	Average	0.91	0.86	0.94	Excellent reliability
Left Ear	Individual	0.87	0.81	0.92	Excellent reliability
	Average	0.95	0.92	0.97	Excellent reliability

participant enrollment. Interrater reliability (IRR) between the data collectors and a clinical audiologist, who served as a reference, was established using intra-class correlation (ICC) on five pretest participants who underwent MHW.¹¹

IRR was determined using a two-way random-effects model, calculating ICCs for both individual and average measurements on the same target. For individual measurements, the resulting ICCs on the right and left ears were in the excellent range (ICC = 0.78 and 0.87, respectively). For average measurements, the ICCs were also excellent (right ear ICC = 0.91; left ear ICC = 0.95). The complete results are shown in Table 1.

ICCs indicate that all three testers had a high degree of agreement on threshold determination.

Recruitment of the Timer/Interviewer

A research assistant served as a patient interviewer and timekeeper, tasked with collecting pertinent demographic and medical information as well as managing the randomization process through a computer-generated coin toss.

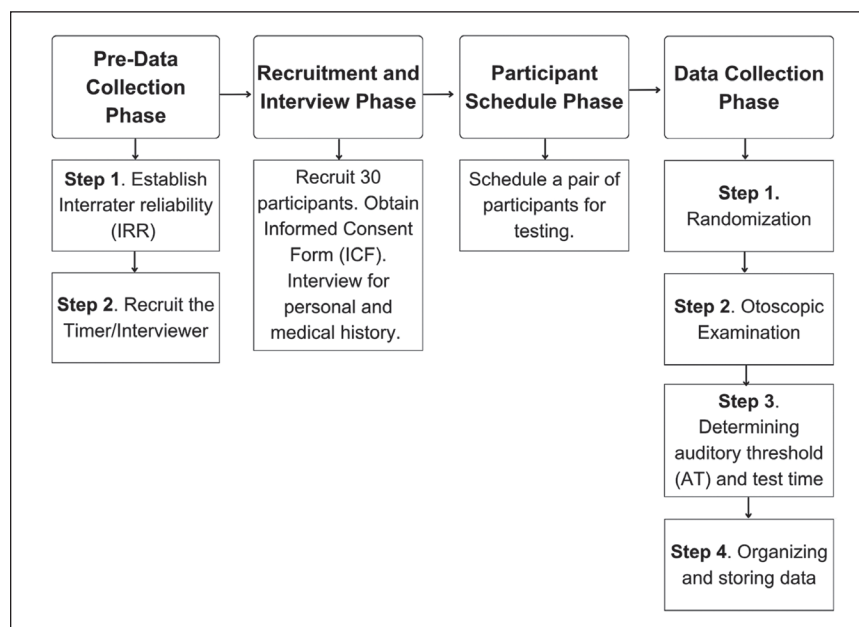


Figure 1. Data collection procedure flowchart.

Recruitment and Interview Phase

Because this investigation was an exploratory method-comparison agreement study, a formal sample size calculation was not performed. A target sample of 30 participants was selected based on study feasibility, available resources, and consistency with previously published agreement and reliability studies in audiology. This sample yielded 600 paired hearing-threshold measurements across all test frequencies and ears for the agreement analyses.

The eligibility criteria were selected to ensure that participants could reliably complete behavioral pure-tone audiometry while minimizing sources of variability unrelated to the threshold-determination procedures. Adults aged 18–60 years were included because they are generally capable of providing consistent behavioral responses and are less likely to demonstrate age-related hearing changes that could confound threshold measurements. An intact tympanic membrane and a clear external auditory canal were required to minimize the influence of conductive pathology on hearing-threshold determination. Participants with active ear infection, ear malformations, or physical disabilities that could interfere with responding to auditory stimuli were excluded to reduce measurement bias and ensure that observed differences reflected the threshold-seeking procedures rather than underlying otologic or physical conditions.

Participants were recruited through convenience sampling using institutional Facebook and Instagram advertisements.

Thirty-three (33) participants responded and shared their demographic and contact information. Respondents were initially contacted by the researcher via their mobile phone number through text messaging and an audio call. During the call, the Informed Consent Form (ICF) was thoroughly discussed, and they were instructed to spend time reviewing the ICF prior to deciding to join the study. After two to three days, respondents contacted the researcher through text message and audio call, providing informed consent to join. At this point, recruited participants were interviewed for approximately 10–15 minutes. Following the interview, three participants were excluded due to an active ear infection and a physical disability that may affect the participant's ability to respond to sound. The confirmed thirty recruited participants were assigned participant code numbers by the Timer/Interviewer. A signed ICF was obtained on the day of the test procedure, and a copy was given to the recruited participant. Figure 2 shows a participant flow diagram.

Participant Schedule Phase

A pair of participants was scheduled per time block for the data collection. Recruited participants were informed of their scheduled study procedure via text message and phone call, and their attendance was confirmed. Reminder messages were sent through text messages the day before their confirmed schedule.

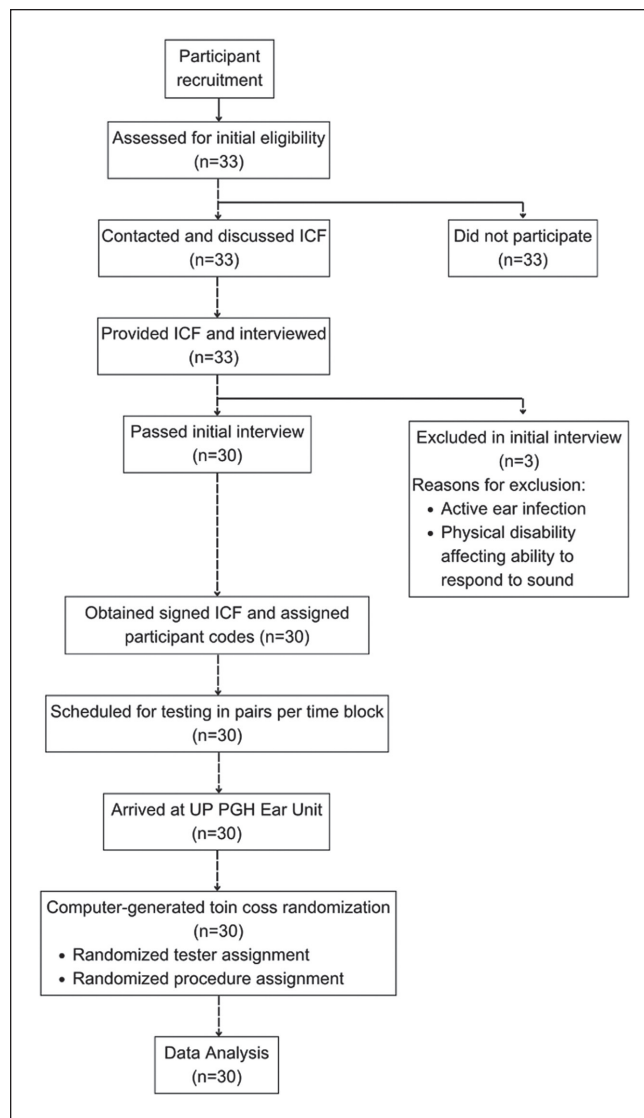


Figure 2. Participant flow diagram.

Data Collection Phase

Randomization

Testing took place at the UP PGH Ear Unit. On the day of the testing, Timer/Interviewer randomized tester and procedure assignments through a computer-generated coin toss. Tester and procedure assigned to each participant were indicated in the Participant Data Sheet (PDS). Several measures were implemented to minimize measurement bias. The order of procedures and tester assignments were randomized. Testers were trained before data collection and demonstrated excellent inter-rater reliability. Participants served as their own controls, thereby minimizing between-subject variability. Testers followed standardized testing instructions throughout all examinations.

Otoscopic Examination

An otoscopic examination was performed by the data collectors prior to performing threshold determination, and ear canal and tympanic membrane observations were recorded in the PDS.

Determining AT and Total Test Time

The audiometers used were two Otometrics Madsen Astera 2. This is a computer-controlled, two-way channel audiometer which performs standard audiometric tests such as tone, speech, and other diagnostic hearing tests. Three sound booths were prepared. Supra-aural headphones for air conduction testing and the bone oscillator transducer for bone conduction testing were prepared. Both audiometers were calibrated according to manufacturer recommendations before study initiation.

For both Procedures A (MHW) and B (Rev MHW), the following were performed: (1) AT was determined for Air conduction (AC) test frequencies including 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz. (2) Bone-conduction (BC) test frequencies included 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. (3) the Timer/Interviewer recorded the test time. The timing started when the first tone was presented at the first AC test frequency and ended when the tester stored the threshold for the last BC test frequency. The timing was paused during the change from supra-aural headphones to bone oscillator transducers. (4) Data, including ATs in decibels and test time in minutes and seconds, were recorded in the PDS. When ATs indicated a need for masking, the test time was paused, and masking was administered. Although masked thresholds were noted, it was not included in data collection and data analysis. Only unmasked thresholds were analyzed. After the test was finished, a fifteen-minute break was given during which refreshments were provided. Upon their return, participants switched procedures and testers, such that one participant was administered the other procedure by the other tester. For example, participant 01 was administered Procedure A by Tester X and participant 02 was administered Procedure B by Tester Z. They switched after the break so that participant 01 was administered Procedure B by Tester Z and participant 02 was administered Procedure A by Tester X. This took place for approximately 1 hour for each pair of participants, including pauses and breaks in between.

Testing continued until all recruited participants had completed the procedures and ATs and test time had been recorded. An audiogram of the test result for Procedure A (MHW), which indicated masked thresholds, was given to the recruited participants via email two weeks after completing the procedure.

The Rev MHW flowchart is shown in Figure 3.

Organizing and Storing Data

PDS were stored and organized accordingly. A soft copy was made and stored on an external hard drive. Data were recorded in a Microsoft Excel sheet and were provided to the biostatistician for data analysis.

Data Processing and Analysis

Statistical Analysis

Agreement between MHW and the Rev MHW was evaluated using Bland-Altman analysis, a method-comparison analysis designed to assess agreement between two clinical measurement procedures rather than their correlation.¹² Bland-Altman analysis was selected because the objective was to evaluate agreement between two clinical measurement procedures rather than correlation or association. For each paired measurement, the difference between procedures (Rev MHW – MHW) was plotted against the mean of the two procedures. The mean difference (bias) was determined to assess systematic differences between methods, while

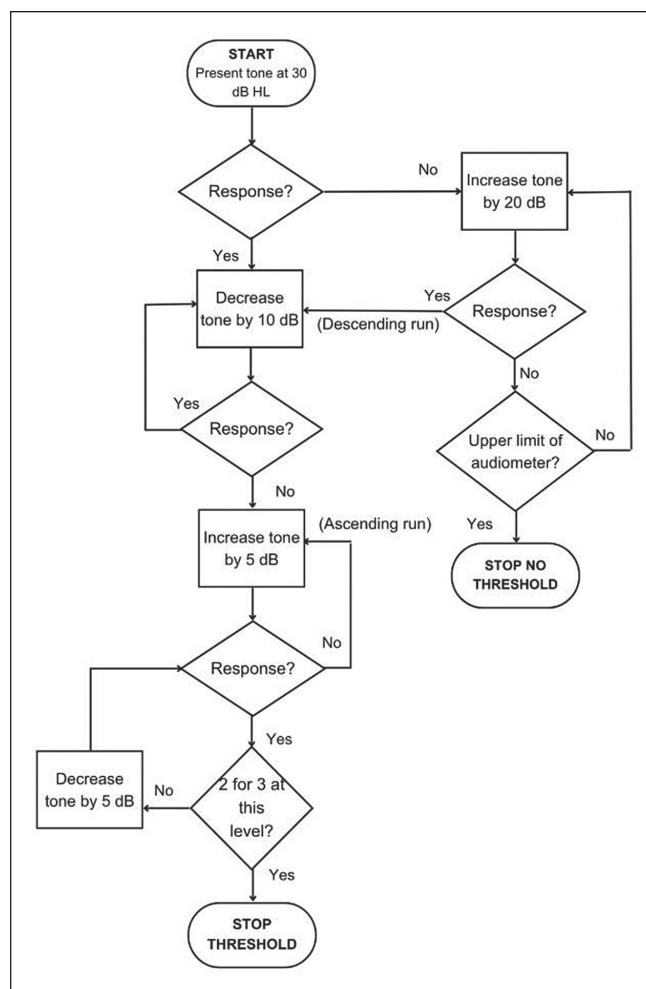


Figure 3. Rev MHW Down-up procedure flowchart.⁴

the standard deviation (SD) of the differences was used to establish the 95% limits of agreement (LOA), calculated as the mean difference $\pm 1.96 \times \text{SD}$. These limits represent the range within which 95% of the differences between the two procedures are expected to lie.

Bland–Altman plots were generated to visually assess agreement and identify any evidence of proportional bias or trends in measurement differences across the range of hearing-threshold values. The analysis was performed using all paired hearing-threshold measurements obtained from both air-conduction and bone-conduction testing.¹³

A paired t-test was performed to determine whether a significant difference existed in test duration between the two procedures.¹⁴ Mean test times, SDs, mean differences, and t-statistics were calculated, and statistical significance was assessed at $\alpha = 0.05$. No missing data occurred, and no sensitivity analyses were performed.

RESULTS

Demographic Characteristics

Table 2 shows the participants' demographics.

Data Presentation and Interpretation

A total of 600 paired measurements were included in the Bland–Altman analysis, encompassing both AC and BC thresholds across all tested frequencies and ears.

The mean difference between procedures (Rev MHW – MHW) was 0.55 dB (SD = 6.23 dB), indicating minimal systematic bias. The 95% limits of agreement ranged from –11.65 to 12.75 dB, with 569 of 600 paired measurements (94.8%) falling within these limits. Additionally, 496 of the 600 paired hearing-threshold measurements (82.7%)

Table 2. Demographic and Audiometric Characteristics of the Participants (N = 30)

Variable	Total Number of Participants (N = 30)
Age (years)	
Mean	35.20
Standard Deviation	10.60
Sex, n (%)	
Female	19 (63%)
Male	11 (37%)
Age brackets (years), N (%)	
Early adulthood (18–30)	13 (43%)
Early to mid-adulthood (31–43)	11 (37%)
Middle adulthood (44–60)	6 (20%)
AT^a based on AC PTC^b of MHW (PTA^c = 500Hz, 1000Hz, and 2000Hz), N (%)	
0–25 dB AC PTA (Normal hearing)	25 (83%)
>25 dB AC PTA (Hearing loss)	5 (17%)

^aAuditory threshold; ^bAir conduction pure-tone average; ^cPure-tone average

differed by no more than ± 5 dB, indicating that most of the thresholds obtained using the revised procedure were within one routine audiometric step of those obtained using the conventional procedure. Visual inspection of the Bland–Altman plot showed that the differences were generally distributed around the mean bias line without substantial evidence of proportional bias, suggesting that the level of disagreement did not increase systematically with higher or lower threshold values (Figure 4).

To evaluate differences in administration time between the two procedures, a paired t-test was performed using the test durations obtained from the same participants under

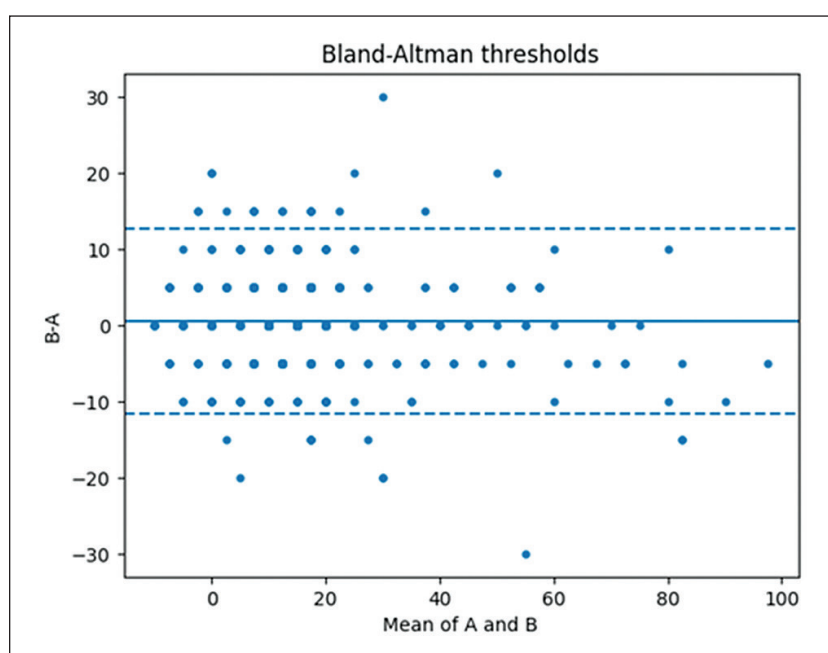


Figure 4. Bland–Altman plot of MHW and Rev MHW.

Table 3. Comparison of Test Duration between the MHW and Rev MHW Procedures

Variable	Mean $\pm$ SD		Mean Difference (A – B)	95% CI for the Mean Difference	t(df = 29)	p-value
	MHW (Procedure A)	Rev MHW (Procedure B)				
Test Duration (minutes)	6.88 $\pm$ 0.99	5.39 $\pm$ 0.85	1.50 min	1.12 to 1.87 min	8.20	<0.001

both conditions. Descriptive statistics and the results of the paired t-test are presented in Table 3. The Rev MHW reduced test duration by an average of 1.50 minutes (95% CI, 1.12–1.87 minutes), and this reduction was statistically significant ($t(29) = 8.20$, $p < 0.001$).

Overall, the analyses demonstrated a high level of agreement between MHW and the Rev MHW for hearing-threshold measurements while showing that Rev MHW required less time to administer.

DISCUSSION

An audiometric test that is rapid and accurate is desired by every audiologist, especially in institutions with high patient loads. The standard MHW required a repeated 10dB descent and two (2) 5dB incremental ascents. By doing away with the repeated 10 dB descent and going straight to a 5dB descent, a significant amount of time is saved without a significant effect on diagnostic accuracy.

Agreement between MHW and Rev MHW

Using Bland–Altman analysis (which evaluates systematic bias and limits of agreement between two measurement methods) to assess the agreement between MHW and Rev MHW, the results revealed a mean difference of only 0.55 dB between the two procedures, indicating negligible systematic bias. Given that audiometric thresholds are typically measured in 5 dB increments and that small variations in threshold estimation are expected during repeated testing, this degree of bias is clinically insignificant. Therefore, Rev MHW appeared capable of producing hearing-threshold measurements that are closely agreed to those obtained using MHW.

The 95% LOA were calculated to be –11.65 dB and 12.75 dB. Although some variability was observed between procedures, the mean bias was small (0.55 dB), indicating minimal systematic differences in threshold estimation. Given that pure-tone audiometric thresholds are routinely measured in 5 dB increments, and that test–retest variability of approximately ± 5 to ± 10 dB has been reported in clinical audiometry, the observed limits of agreement suggested that differences between procedures were generally within a clinically reasonable range.¹⁵ Visual inspection of the Bland–Altman plot showed that the differences were distributed around the mean bias line without substantial evidence of proportional bias, suggesting that the magnitude of disagreement did not systematically increase across the range of threshold values.

Table 4. Comparison of AT Classifications: MHW and Rev MHW

Procedure	Total Number of Participants (N = 30)	
	MHW	Rev MHW
AT^a Classification (based on AC PTA^b)		
Normal (0–25 dB)	25	23
Mild (26–40 dB)	2	4
Mild to moderate (41–55 dB)	2	1
Moderately severe (56–70 dB)	0	1
Severe (71–90 dB)	1	1

^aAuditory Thresholds; ^bAir Conduction Puretone Average

Looking at the obtained thresholds from MHW, it showed that five participants had hearing loss based on their pure-tone average. However, the thresholds obtained from the Rev MHW revealed that seven participants had hearing loss, which showed that the AT classification of 2 participants had changed. Table 4 shows the AT classification of the participants from MHW when compared to Rev MHW.

The following changes in AT classification from MHW to Rev MHW were observed: (1) Participants with mild hearing loss increased slightly from 2 to 4 participants. (2) Participants with mild to moderate hearing loss decreased slightly from 2 to 1 participant. (3) Participants with moderately severe hearing loss increased slightly from 0 to 1 participant. The 2 participants, previously classified as normal hearing, changed their classification to mild hearing loss. While 1 participant, previously classified as mild to moderate hearing loss, changed classification to moderately severe hearing loss. Despite these classification changes, the Bland–Altman analysis demonstrated minimal systematic bias and acceptable agreement between procedures.^{12,13} None of the previously classified hearing loss participants were reclassified to normal hearing. Therefore, no patient with a classified hearing loss was missed with the Rev MHW. Furthermore, when applied in a clinic setting, the Rev MHW successfully identified individuals with hearing loss and thereby enabled affected individuals to receive the hearing care service that they require.

In addition, there is ± 5 dB test/retest and even inter-rater difference that is acceptable according to Stuart et al.¹⁵

Effects of test duration

An important finding of this study was the 22% reduction in test time (Table 3) associated with Rev MHW. This supported the research hypothesis that reducing the step size from 10 dB to 5 dB in the second ascending run

leads to a faster test time. The change to a 5 dB step size in the second ascending run resulted in a reduced action step, meaning that there is a smaller decibel step that needs to be crossed to obtain a threshold. A reduction of approximately 1.5 minutes represents a meaningful improvement in workflow when large numbers of patients are evaluated daily.

When choosing a method for conducting audiologic procedures, performing the test accurately in the least amount of time in a busy clinic is imperative. With a statistically significant difference in total test time, with a reduction of 10 dB to 5 dB steps in the second descending run of administering the manual method in PTA, clinicians may consider the Rev MHW a practical alternative. This reduction represents a meaningful improvement in efficiency, particularly in settings where large numbers of individuals are screened or evaluated. Shorter testing times may enhance patient comfort, reduce examiner workload, and improve overall clinical productivity while maintaining measurement accuracy.⁶ For example, a faster method may prevent fatigue in patients, particularly in older adults, which could affect consistency in their response criteria during testing. Furthermore, a faster method will provide the clinician with time that they could choose to devote to other important clinical tasks, such as documentation, patient and professional correspondence, or even a rest break.

Taken together, these findings suggested that Rev MHW provides hearing-threshold measurements that closely agree with those obtained through MHW while offering the additional advantage of faster administration. Future studies involving larger and more diverse populations may further establish the generalizability and clinical applicability of these findings.

To the authors' knowledge, no previous studies have evaluated the agreement between the standard MHW and a revised version incorporating a reduced down step during the second ascending run, while also examining its effect on test duration. Jervall and Arlinger investigated the effect of reducing the step size from 5 dB to 2 dB in both the ascending and descending phases of the manual audiometric procedure. Their findings demonstrated good test-retest reliability for the 2 dB step size; however, the smaller step size resulted in longer test durations. The increased testing time was attributed to the greater number of threshold crossings required to establish an AT. In contrast, the present study demonstrated that reducing the second down step from 10 dB to 5 dB maintained comparable hearing-threshold measurements while significantly reducing test duration.

Limitations of the Study

One of the limitations of this study was the total test time measured in the Methods section. Particularly, this study only included threshold determination, which is Step 3 in the Data Collection phase in Figure 1. It does not include the indirect tasks that go with successfully completing the test procedure (e.g., Step 1 and 2 of the Data Collection

phase) such as explaining the procedure, providing reminders in between tasks, and the actual changing of transducers from supra-aural headphones to BOT. These indirect tasks contributed to a successful and efficient administration of the test, which may have influenced the measured test duration; however, it was unlikely to change the study conclusion.

The Bland–Altman analysis was performed using all paired threshold measurements; however, multiple observations were obtained from each participant across frequencies and ears, which may limit the statistical independence of the observations. Observations were clustered within participants and were not statistically independent. Consequently, Bland–Altman estimates and their confidence intervals may underestimate uncertainty. Mixed-effects agreement methods may be considered in future studies.

Masked thresholds were not included. The accuracy of ATs would increase if masked thresholds were obtained and analyzed. However, it may be inferred that the results of the study will still apply with masked thresholds since the procedure for obtaining ATs remains the same for obtaining masked thresholds.

Because participants were recruited from a single academic medical center and consisted primarily of adults with normal hearing, the findings may not be generalizable to pediatric populations, individuals with profound hearing loss, or patients who are unable to provide reliable behavioral responses during pure-tone audiometry. Additional studies involving larger and more diverse populations, including individuals with varying degrees and types of hearing loss, are needed to establish the broader clinical applicability of the Rev MHW procedure.

CONCLUSION

The Rev MHW maintained clinically similar threshold estimates with the standard MHW in determining ATs. In addition, Rev MHW significantly reduced test duration compared with MHW. Overall, the results support the potential use of Rev MHW as a more efficient alternative procedure for determining ATs in adults while maintaining broadly comparable threshold measurements.

Acknowledgments

The authors are grateful to the faculty of the UPM Master in Clinical Audiology program and to the members and staff of the UP PGH Ear Unit for their patience, help, advice, teachings, and insights.

They would also like to thank the participants who have given their time so that this research can come to fruition.

Statement of Authorship

Both authors certified fulfillment of ICMJE authorship criteria.

Author Disclosure

Both authors declared no conflicts of interest.

Funding Source

The study was funded by the authors.

REFERENCES

1. Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. *J Speech Hear Disord*. 1959 Nov;24(4):330-45. doi: 10.1044/jshd.2404.330. PMID: 14405102
2. Jerger J. The evolution of the audiometric pure-tone technique. *Hearing Review*. 2018;25(9):12-18.
3. American Speech-Language-Hearing Association. Guidelines for manual pure-tone threshold audiometry. *ASHA*. 1978;20(4):297-301.
4. Vencovsky V, Rund F. Pure tone audiometer. In: *Proceedings of the 20th Annual Conference on Technical Computing*; 2012; Prague, Czech Republic [Internet]. [cited 2025 Jul 21]. [about 5 p.]. Available from: http://dsp.vscht.cz/konference_matlab/MATLAB12/full_paper/088_Vencovsky.pdf
5. Jerlvall L, Arlinger S. A comparison of 2-dB and 5-dB step size in pure-tone audiometry. *Scand Audiol*. 1986;15(1):51-6. doi: 10.3109/01050398609045954. PMID: 3704542.
6. Poling GL, Kunnel TJ, Dhar S. Comparing the accuracy and speed of manual and tracking methods of measuring hearing thresholds. *Ear Hear*. 2016;37(5):e336-e340. doi: 10.1097/AUD.0000000000000317. PMID: 27232075; PMCID: PMC4996736.
7. Beck RM, Ramos BF, Grasel SS, Ramos HF, Moraes MF, Almeida ER, et al. Comparative study between pure tone audiometry and auditory steady-state responses in normal hearing subjects. *Braz J Otorhinolaryngol*. 2014 Jan-Feb;80(1):35-40. English, Portuguese. doi: 10.5935/1808-8694.20140009. PMID: 24626890; PMCID: PMC9443973.
8. Margolis RH, Morgan DE. Automated pure-tone audiometry: an analysis of capacity, need, and benefit. *Am J Audiol*. 2008;17(2):109-113. doi: 10.1044/1059-0889(2008/07-0047). PMID: 18840703.
9. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med*. 2007 Oct 16;4(10):e296. doi: 10.1371/journal.pmed.0040296. PMID: 17941714; PMCID: PMC2020495.
10. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011 Jan;64(1):96-106. doi: 10.1016/j.jclinepi.2010.03.002. PMID: 21130355.
11. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol*. 2012;8(1):23-34. doi: 10.20982/tqmp.08.1.p023. PMID: 22833776; PMCID: PMC3402032.
12. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986 Feb 8;1(8476):307-310. PMID: 2868172.
13. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999 Jun;8(2):135-160. doi: 10.1177/096228029900800204. PMID: 10501650.
14. Pagano M, Gauvreau K. *Principles of Biostatistics*. 2nd ed. New York, NY: Chapman & Hall/CRC; 2018. doi: 10.1201/9780429489624.
15. Stuart A, Stenstrom R, Tompkins C, Vandenhoff S. Test-retest variability in audiometric threshold with supraaural and insert earphones among children and adults. *Audiology*. 1991;30(2):82-90. doi: 10.3109/00206099109072873. PMID: 1877901.