Test re-test reliability of virtual acoustic space identification (VASI) test in young adults with normal hearing

Kavassery Venkateswaran Nisha[*], Prabuddha Bhatarai, Kruthika Suresh, Shashish Ghimire, Prashanth Prabhu

Department of Audiology, All India Institute of Speech and Hearing, Manasagangothri, Mysuru, 570 006, India

A R T I C L E   I N F O

A B S T R A C T

Background: Recent developments in virtual acoustic technology has levered promising applications in the field of auditory sciences, especially in spatial perception. While conventional auditory spatial assessment using loudspeakers, interaural differences and/or questionnaires are limited by the availability and cost of instruments, the use of virtual acoustic space identification (VASI) test has widespread applications in spatial test battery as it overcomes these constraints.
Purpose: The lack of test-retest reliability data of VASI test narrows its direct application in auditory spatial assessment, which is explored in the present study.
Methods: Data from 75 normal-hearing young adults (mean age: 25.11 y $\pm$ 4.65 SD) was collected in three sessions: baseline, within 15 min of baseline (intra-session), and one week after baseline session (inter-session). Test-retest reliability was assessed using the intra-class correlation coefficient (ICC), coefficient of variation (CV), and cluster plots.
Results: The results showed excellent reliability for both accuracy and reaction time measures of VASI, with ICC values of 0.93 and 0.87, respectively. The CV values for overall VASI accuracy and reaction time 9.66% and 11.88%, respectively. This was also complemented by the cluster plot analyses, which showed 93.33% and 96.00% of temporal stability in the accuracy and reaction time measures, indicative of high test-retest reliability of VASI test in auditory spatial assessment.
Conclusions: The high temporal stability (test-retest reliability) of VASI test validates its application in spatial hearing test battery.

1. Introduction

Auditory spatial perception is the ability to perceive and tune with the sound source's direction in the three-dimensional (3D) plane (Blauert, 1997). Unlike vision, spatial hearing is the only directional receptor that works in a three-dimensional (3-D) plane operating in a full 360° range. Spatial hearing is often regarded as a guiding system for vision in determining the source location and its position relative to other objects in the space, as its equally effective in darkness as in bright light (Letowski and Letowski, 2012).

Being such a versatile system, spatial hearing often is challenged by physiological and ecological limitations. The extent of preciseness in auditory spatial perception judgments in a given environment depends on the physiological status of the auditory system, including age (Abel and Hay, 1996; Häusler et al., 1983; Newton and Hickson, 2021; Noble et al., 1994, 1997), degree of hearing loss (Abel and Hay, 1996; Häusler et al., 1983; Newton and Hickson, 2021; Noble et al., 1994, 1997), their auditory experience (Roffler and Butler, 1968), knowledge of listening strategies (Dufour et al., 2007; Neelon and Jenison, 2004; Sosa et al., 2010), familiarity with the surrounding environment (Brown & May 2006; King, 1999; Knudsen, 1984; Noble and Byrne, 1990), and psychological status of the listener (motivation, attention, tiredness, etc). It also depends on the type of rehabilitative options, including hearing aids (Leeuw and Dreschler, 1987; Noble and Byrne, 1990) or type and duration of spatial training (Habib and Besson, 2009; Majdak et al., 2010; Polley et al., 2006; Zahorik et al., 2006; Zahorik and

Wightman, 2001).

Assessment of auditory spatial acuity is usually explored using either real and virtual sources (Dorman et al., 2016; Drennan et al., 2001; Lorenzi et al., 1999a; 1999b; Zhong et al., 2016). The use of loudspeakers poses limitations of infrastructure, cost-and space involved. The introduction of a virtual acoustic space identification (VASI) (Nisha and Kumar, 2016, 2017) test simulating the sound locations virtually in the head overcomes these limitations. It produces relatively higher objectively verifiable results than perceptual questionnaires.

VASI test employs user display for presentation and response acquisition corresponding to the virtual reality experiment in 8 virtual locations (located 45° apart) spanning 360° in the horizontal plane at 0° vertical elevation (Nisha & Kumar, 2017). These spatially rendered stimuli were obtained by convolving the direct sound component with non-individualized head-related transfer functions (HRTFs). The HRTFs used in the virtualization techniques for VASI is obtained from the Sound Lab version) version 6.7.3 (NASA, Ames Research Center, USA, 2012), a publicly available tool for generating auditory spatial stimuli. Literature evidence points out that slab 3D utilizes higher-order Ambisonics techniques to convolve the HRTFs with target sound (Wenzel et al., 2000).

One of the earliest other best-known systems that make use of a similar Ambisonic technique (used in slab 3D) is the CAVE Automatic Virtual Environment (CAVE). It is a virtual reality (VR) system developed at the Electronic Visualization Laboratory, the University of Illinois at Chicago (Cruz-Neira et al., 1993). This VR system employs four identical speakers to simulate direction and distance effects. This virtual environment system is possible only in the laboratory system and is not available to use in the general public. Another VR technique similar to the one used in the study was designed by Nguyen et al. (2009). Using non-individualized HRTFs techniques, the authors reported high accuracy of auditory spatial rendering across seven directions in the frontal plane (−30°, −20°, −10°, 0°, 10°, 20°, & 30° azimuths). Based on these results, the authors urged the need for dynamic, interactive virtual environments to enhance the applicability of VR in psychoacoustic experiments. No standardized tool in audiological setups is currently available to screen out auditory spatial deficits, with only limited infrastructure and a high cost-benefit ratio. To this end, the use of VASI test in clinical setups seems to a viable option for spatial screening and early identification. The VASI test can have promising implications in rehabilitative audiology too: to infer the training effects (Nisha & Kumar, 2017) or to check the efficacy of the manufacturers claim on spatial processing algorithms in hearing aids. However, the utility of VASI in clinical setups can be advocated only if the test results are reliable across measurement sessions. A measure's reliability and validity must be developed before it can be used in testing or to record a clinical intervention (Ruscetta et al., 2005).

To establish reliability, test-retest responses are systematically correlated, and the variance in each administration is analyzed. Test-retest reliability allows the experimenter/researcher to verify the temporal stability of a scale. Other factors like the in motivation and focus, headset location, test setting, listener posture, and listener guidance are all likely to affect the stability of test outcomes over time (Ruscetta et al., 2005). Test-retest reliability accounts for all these inherent variabilities and quantifies the consistency with which test results can be interpreted. In other words, test-retest reliability provides a measure of the test's repeatability through measurements, i.e., it quantifies how likely a test is to yield reliable outcomes at one administration compared to those achieved at a subsequent administration. In clinical settings, where Audiologists perform a test several times, this level of consistency is critical (e.g., to document the efficacy of medication on hearing, outcomes of

hearing aids, cochlear implants and/or assistive listening devices, and auditory training). On similar lines, establishing the test-retest reliability of VASI test can document its temporal stability for auditory spatial evaluation across sessions and time. This, in turn, can set the stage for its application in audiological setups for quantification of rehabilitative (hearing aids or training) outcomes. The present study aimed to establish the test-retest reliability of the VASI test (Nisha & Kumar, 2016, 2017) in normal-hearing listeners across different age groups. The specific objectives of the study were to understand the temporal stability of VASI using three parameters: accuracy scores, reaction time, and spatial errors measured across three different timelines.

2. Methods

2.1. Participants

A total of 75 participants (37 females and 38 males) aged 20–40 years (mean age: 25.11 y; ± 4.65 SD) volunteered for the study. The participants were either staffs or students of different departments of the All India Institute of Speech and Hearing, Mysuru, India. None of the participants had any prior knowledge of the test procedure, used in study prior to their inclusion. The participation in the study was voluntary and no financial compensation was provided.

All participants met the following criteria: (a) age range of 20–40 years, (b) normal hearing sensitivity in both ears (pure-tone air and bone conduction thresholds ≤15 dB HL at octave frequencies of 500–4000 Hz (Clark, 1981), (c) no history of otological diseases, prior ear-related surgeries, use of ototoxic drugs, or other systemic disorders such as diabetes, etc. which can affect hearing, (d) no known history of attention, cognition and neurological deficits, and (e) no known musical, abacus training or previous experience in psychoacoustic testing on the auditory spatial domain. All participants attended three testing sessions conducted in a silent room with ambient noise levels <35 dB (A) (as recorded on android and iOS software Decibels X). The room was also free from auditory and visual distracters.

2.1.1. Estimation of sample size

To estimate sample size in the study with test-retest reliability, the target reliability ρ, the minimum acceptable reliability ρ_0, and the number of repetitions of measurement π is considered. The sample size was estimated using G*Power version 3.1.9.4 (Faul et al., 2007). The intraclass correlation coefficient (ICC) is used to define the target level of reliability. The type I and type II errors (ε_1 and ε_2) are assumed to be 0.05 and 0.20, respectively (Walter et al., 1998). Since we had to measure the three sessions (baseline, intrasession, and intersession), with the minimum acceptable reliability r = 0.50 and expected reliability of $\rho = 0.80$ (Morse et al., 2002), the required sample size becomes 59 participants. This provides an actual power of 0.95. To account for any unexpected variability of the data, 75 adults were recruited for the study.

2.1.2. Informed consent and ethical guidelines

Participants were explained the purpose of the research study, their tasks, and their benefits before signing a written informed consent. Ethical guidelines for bio-behavioral research formulated by the institutional board (Venkatesan, 2009) were followed.

2.2. Procedure

At the start of the study, an otological history, otoscopy, and tympanometry were conducted, accompanied by a pure tone audiometric evaluation to ensure that the participants met the inclusionary criteria. Following the preliminary investigations of

normal hearing, all the participants were tested for their spatial acuity in closed-field using the VASI test (Nisha & Kumar, 2017). The test was administered three times, including a baseline test and two re-tests. Re-tests were administered after 15 min of the same session (intra-session), and another re-test was administered one week after the first session (inter-session). Accuracy and reaction time scores were collected during each timeline of measurement.

2.2.1. Virtual acoustic space identification (VASI) test

VASI is a test for spatial acuity implemented using illusionary spatial percepts within the head in a closed field. The test employs virtual auditory technology to create eight spatial percepts within the head: midline front (0° azimuth), midline back (180° azimuth), 45° toward the right ear (R45), 90° toward the right ear (R90), 135° toward the right ear (R135), 45° toward the left ear (L45), 90° toward the left ear (L90), and 135° toward the left (L135). The Sound lab (Slab 3D), version 6.7.3b was implemented to generate the VASI stimuli (Wenzel et al., 2000). The HRTF of Slab 3D corresponding to particular azimuth was convolved with 250 ms white noise to generate VASI stimuli. The HRTF used in SLab 3D is comparable to the head models provided in Center for Image Processing and Integrated Computing (Algazi et al., 2001) database and is shown to produce reliable lateralization responses (Miller et al., 2014). Paradigm software (Perception Research Systems, 2007) was used to control the stimulus delivery and response acquisition.

Each virtual acoustic stimulus is played ten times (total presentation = 8 locations* played ten times = 80 times) randomly, at a presentation level of 65 dB SPL using headphones (Sennheiser HDA 300, Wedemark, Germany). Before administering the test, the output of noise burst at each of the 8 locations was objectively calibrated using the sound level meter (SLM, Bruel and Kjaer 2270, Naerem, Denmark) by a trained acoustical engineer. Each respondent was given a trial run to familiarize themselves with the test stimuli and responses acquisition process before the start of test. The testing started only after the participant stated that they were confident with the test. The user interface used in the test is shown in Fig. 1. Familiarization phase included presenting the stimuli after

mouse click on the user interface, based on which sound corresponding to particular azimuth was emanated. The familiarization phase took approximately 10 min.

In the testing phase, the participants were randomly given sound at each location (Fig. 1), and were instructed to click on one of the eight locations which delivered the sound. The participants' responses were captured by clicking the mouse pointer on the virtual location (Fig. 1). The participants were also asked to maintain a static head position during the presentation of stimuli with their head facing the user interface (i.e., midline front to replicate 0° azimuth). The static positioning of head ensured the accurate replication of virtual locations and helped us to exercise control on head movements, which otherwise would affect spatial processing (Brimijoin and Akeroyd, 2012). Head movements were allowed during the response acquisition i.e., between the time duration of offset of stimuli to the time the participant registered the target location by a mouse click. The test was ended after 80 trials (8 locations * 10 trials per location) had been completed. The output of the test, containing the sequence of the stimuli trails and corresponding responses, was stored in excel form.

2.2.2. Scoring and analysis

Each correctly identified trial corresponding to the virtual location was given a score of '1', while the wrong identification was awarded a score of '0'. The output responses from the excel file were analyzed using a confusion matrix script running (Gnanateja, 2014) on MATLAB version (R2021a) (Mathworks Inc, Natick). The output of the script yielded a stimulus-response grid, which is used to study the pattern of errors and the accuracy scores. The accuracy of VASI scores for each location (location-specific) and overall VASI accuracy scores was derived from the stimulus-response grid. Similarly, reaction time for each correctly identified virtual location and corresponding overall reaction time were also calculated.

2.3. Statistical analysis

The data obtained from all the participants across timelines (baseline, intra-, and inter-session) was subjected to statistical analyses using IBM Statistical package for social sciences (SPSS) version 25.0 (SPSS Inc., Chicago). Descriptive statistics with mean and standard deviations of the accuracy and reaction time scores were obtained. Based on Shapiro Wilk's test of normality, parametric analysis was performed using one-way repeated measure ANOVA. One-way ANOVA was performed to see if any significant differences in VASI accuracy and reaction time scores could be observed between the testing sessions. The test-retest reliability of the evaluations was assessed using multiple approaches: intraclass correlation coefficient (ICC), coefficient of variation (CV), and cluster plot analyses. The details of the test-retest analyses and findings were interpreted using the following guidelines.

1. The intraclass correlation coefficient (ICC) is the test to determine the relative homogeneity across different test sessions with respect to the overall observable heterogeneity within the test sessions. It is measured on a scale of 1—10, where '1' is complete reliability, and '0' indicates no reliability (Koo and Li, 2016).
2. Similarly, the coefficient of variation (CV) was calculated to measure the relative variability. The CV shows the deviation in percent from the mean threshold value below which 68 percent of the variability between sessions is supposed to lie (Brown, 1998).
3. Furthermore, cluster plot analysis was performed to represent the agreement between the overall VASI scores obtained at the baseline and the two re-test sessions (baseline vs. intra-test
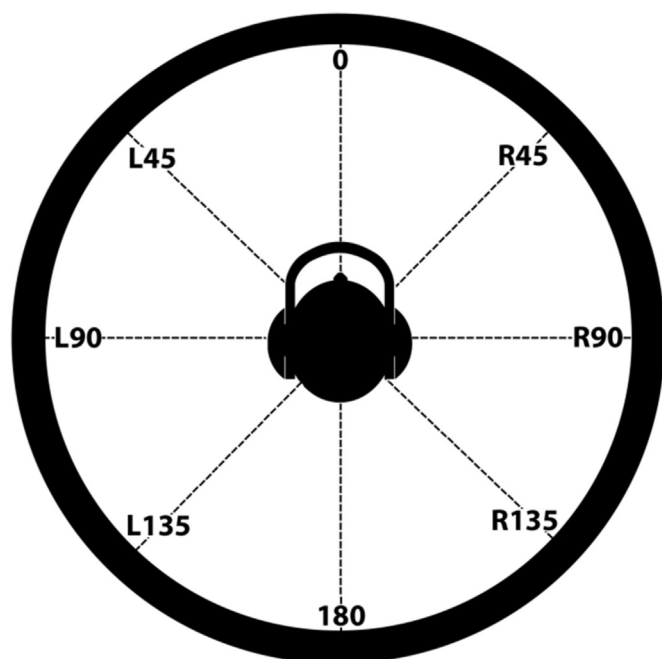

Fig. 1. The user interface of the virtual acoustic space identification (VASI) test.

session; baseline vs. inter-test session). The limits of agreement are shown in a 95% confidence interval where variations between two measurements should fall within the limits for its inclusion in the reliability analysis.

3. Results

The study aimed to examine the temporal stability of the VASI test using test-retest reliability measures for accuracy and reaction time. The Shapiro Wilk test results showed that the overall and location-wise VASI data adhered to normality ($p > 0.05$) at all three timelines of measurement.

3.1. Effect of the timeline of measurement on VASI accuracy scores

The mean accuracy scores (central midline) and the standard deviation (error bars) for each virtual location and overall combined VASI scores across the three measurement phases showed a striking similarity, as observed in Fig. 2. On visual inspection of Fig. 2, it is also evident that although there were slight differences in the location judgments when VASI accuracy is compared across virtual locations (midline 0 and 180 scores being better than judgments in the lateral plane, i.e., right or left planes). The VASI accuracy scores in midline reached the ceiling in most participants, compared to other lateral planes. However, no such differences were apparent amongst the three measurements either within each location (across all 8 locations) or the overall VASI score. Complimentary to this observation, the results of repeated measure ANOVA (three timelines of measurements × nine virtual scores: overall plus eight virtual locations) also statistically verified the lack of significant differences ($p > 0.05$) across the three measurements, as shown in Table 1.

3.2. Effect of the timeline of measurement on reaction time

The mean reaction time (central midline) and the standard

Table 1
Results of one-way repeated measure ANOVA for the effect of measurement phase on VASI accuracy scores.

Virtual auditory space location	Measurement phase effect F (2, 148)	p-value
R45	1.41	0.24
R90	1.28	0.27
R135	0.03	0.97
180	1.31	0.21
L135	0.94	0.39
L90	0.15	0.86
L45	0.77	0.47
0	0.13	0.87
Overall	1.99	0.14

deviation (error bars) for each virtual location and overall VASI scores across the three measurement phases showed a striking similarity in reaction time scores across the locations, as seen in Fig. 3. Unlike the accuracy measures (midline scores were better than scores in lateral planes), the reaction time scores were notably similar across measurements in all virtual locations. Overall scores also showed similar temporal stability. These observations were also statistically verified using repeated measure ANOVA (3 timelines of measurements × nine virtual scores), as shown in Table 2. The lack of significant difference ($p > 0.05$) at all VASI locations (8 virtual locations and one overall score) is indicative of statistically similar response time VASI scores across the measurement phases.

3.3. Effect of the timeline of measurement on the pattern of spatial errors

Table 3 shows the results of confusion matrix analyses documenting the pattern of spatial errors for each virtual location. A close visual inspection of spatial error patterns showed that virtual sounds presented in the left were more confused than in the right plane. Complimentary to this result, variability (SD) within each
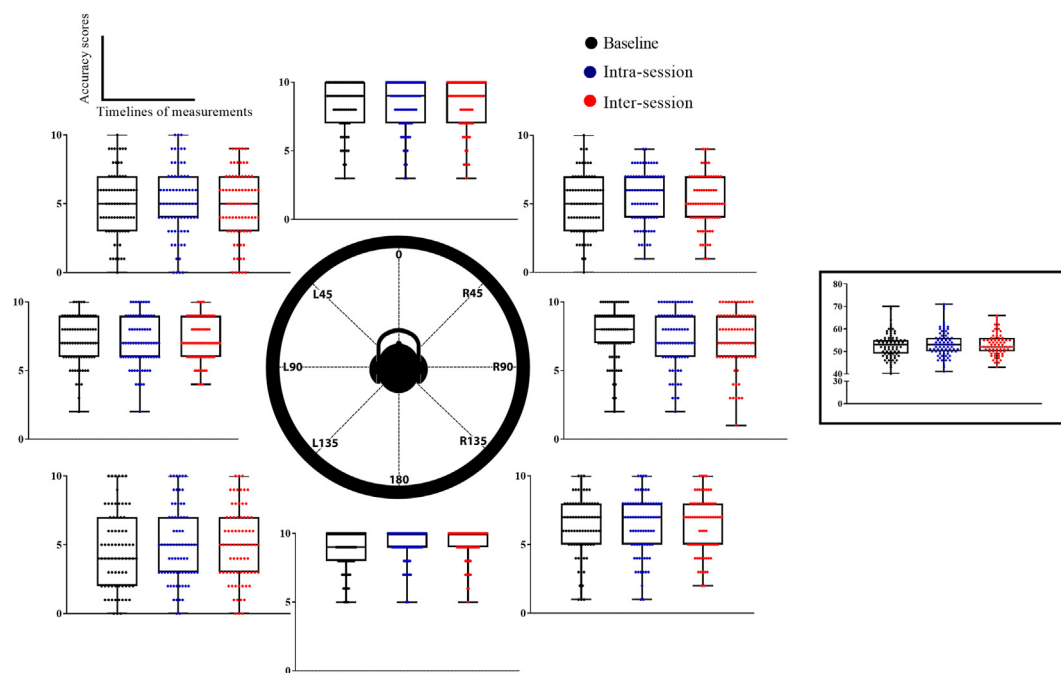

Fig. 2. Boxplot along with individual VASI accuracy data across virtual locations tested. The inner dummy head corresponds to the VASI interface, depicting the eight virtual locations used in the study. The inner panels corresponding to each virtual location denote the corresponding VASI scores, with the centre line of box plot showing the mean and the error bars denoting the one standard deviation (SD). The outer panel corresponds to the overall VASI accuracy scores of each participant along with their mean and SD.
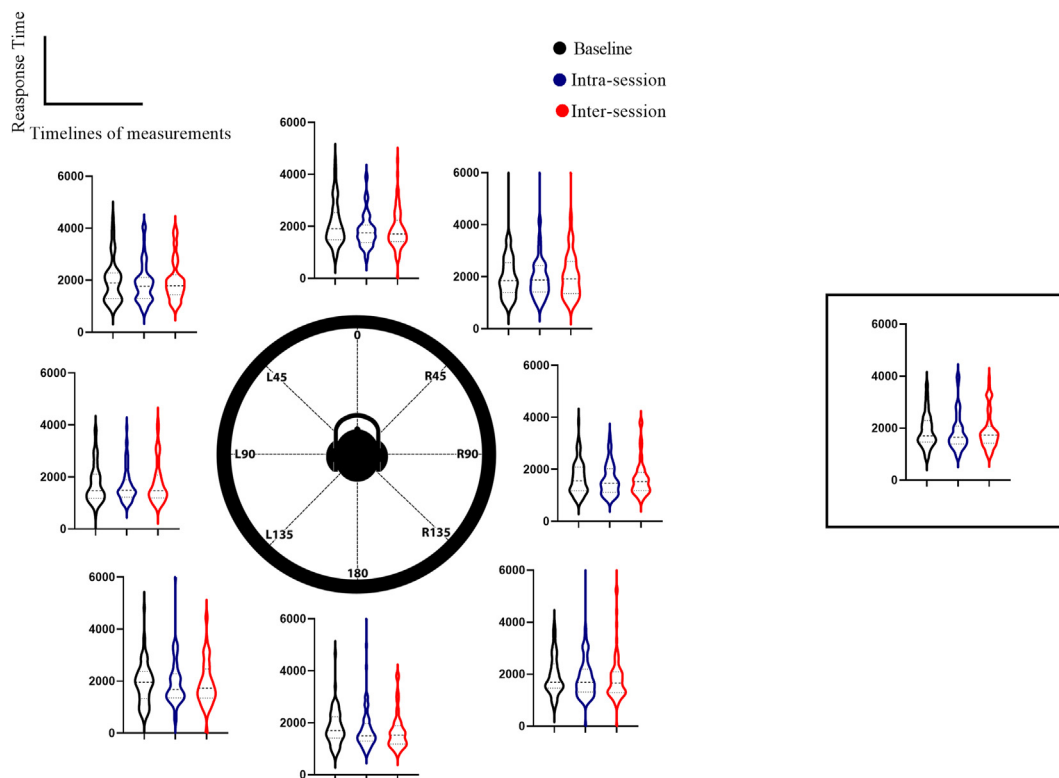
Fig. 3. Violin plot showing response time across different locations. The dotted line depicts the mean while the width of the violin plot denote the spread of the data for each virtual location across all the phases. The outer panel shows the overall reaction time.

Table 2
Results of one-way repeated measure ANOVA for the effect of measurement phase on VASI response time.

Virtual auditory space location	Measurement phase effect $F(2,148)$	p-value
R45	0.50	0.60
R90	0.39	0.67
R135	0.49	0.60
180	0.11	0.89
L135	0.73	0.48
L90	0.59	0.56
L45	1.49	0.23
0	4.31	0.11
Overall	0.32	0.72

virtual location also demonstrated high similarity across the measurement phase, indicative of clear overlap in spatial performance of the participants across the three sessions.

3.4. Test-retest reliability

The results of objective analyses of test-retest reliability using Interclass correlation (ICC) ranged from 0.60 to 0.86 (mean 0.77 ± 0.12 SD) and 0.60 to 0.84 (mean 0.75 ± 0.09 SD) for VASI accuracy and reaction time, respectively, as shown in Table 4. The ICC values are indicative of moderate to high reliability across different virtual locations tested. Similarly, the ICC obtained for overall VASI accuracy (0.93) was also high, suggestive of a high degree of similarity in VASI scores across the three timelines of measurements. The results of coefficient of variation also complimented the ICC scores, as shown in Table 4. The coefficient of variation (CV) for overall accuracy and reaction time scores is 9.66% and 11.99%, suggestive of the measure's very high and moderate

stability, respectively. Similarly, CV for location-wise VASI scores varied across positions, with CV scores not exceeding 12.00%. Similarly, the CV was scattered between 5.43%—16.73% range for reaction time. The CV for both the parameters (accuracy and reaction time) stayed below 20.00%, indicating temporal stability above chance factor in overall and location-wise VASI scores.

The results of test-retest variance data (VASI accuracy, left panels, and reaction time scores-right panels) using cluster plot analyses are depicted in Fig. 4. The cluster plots for overall VASI accuracy scores with baselines on the x-axis and re-test (inter/intra scores) on y-axis show that only 5/75 observations were out of the limits of variance (95% confidence intervals, blue shaded area in top panels in Fig. 4), indicative of only 6.67% error in accuracy scores and very high (93.33%) test-retest reliability. A similar trend was observed for reaction time, with 3/75 observations emerging outside the variance limits, suggestive of test-retest reliability of 96.00% and a minimal error (4.00%).

4. Discussion

The findings of the study revealed high test-retest reliability of both VASI accuracy and reaction time scores, measured at baseline and two-retest sessions (Intra and inter) on three measures: Interclass-correlation (ICC), coefficient of variation (CV), and cluster plot analyses. The descriptive data (Figs. 2 and 3) revealed similarities in VASI accuracy and the participants' response time across all the locations tested in the study. This was verified through results of One-way repeated measure ANOVA (Table 1 & Table 2), which showed no significant differences ($p > 0.05$) in both VASI accuracy and reaction time measures across the three test sessions conducted. These findings highlight the reliability of the test, accounting for no change in the performance of participants on the

Table 3

Pattern of spatial errors across virtual locations in baseline and retest (intra and inter) sessions. Confusion matrix denoting the mean overall response provided for each VAS location for baseline (top white panels), intra-session (grey middle panels) and inter session (black panels). The diagonal (bold) represents accurate mean spatial judgments scores for each VAS location. For VAS location and number relation refer Fig. 1.

Response VAS location		Target VAS location							
	Session	R45	R90	R135	180	L135	L90	L45	0
R45	Baseline	**4.84**	1.69	1.31	0.22	0	0	0.02	0.07
	Intra	**5.29**	1.88	1.39	0.12	0.02	0	0	0.02
	Inter	**4.84**	1.78	1.56	0.12	0.02	0	0.02	0.03
R90	Baseline	3.20	**7.43**	2.09	0.06	0.02	0	0	0.14
	Intra	2.93	**7.07**	1.95	0.02	0.03	0	0	0.10
	Inter	3.07	**6.97**	1.78	0	0.07	0	0	0.08
R135	Baseline	1.73	0.79	**6.40**	0.29	0.02	0	0.02	0.02
	Intra	1.65	1.01	**6.53**	0.34	0.02	0	0.03	0.06
	Inter	1.87	1.22	**6.46**	0.10	0	0	0.02	0.08
180	Baseline	0.12	0	0.15	**8.68**	0.18	0.05	0.12	0.57
	Intra	0.05	0.02	0.13	**9.14**	0.08	0.05	0.24	0.99
	Inter	0.05	0	0.19	**9.26**	0.07	0	0.18	0.98
L135	Baseline	0	0.02	0.03	0.16	**4.46**	1.29	2.66	0.12
	Intra	0	0	0	0.02	**4.61**	0.95	2.43	0.23
	Inter	0.02	0	0	0.08	**4.8**	1.14	2.65	0.06
L90	Baseline	0	0.02	0	0.02	3.72	**6.68**	2.82	0.02
	Intra	0	0	0	0.02	3.12	**7.05**	2.68	0.02
	Inter	0	0	0	0	2.9	**7.17**	2.82	0.02
L45	Baseline	0	0.02	0	0.05	1.33	1.96	**4.32**	0.16
	Intra	0	0	0	0.02	1.64	1.86	**4.48**	0.14
	Inter	0	0.02	0	0.05	1.83	1.65	**4.10**	0.13
0	Baseline	0.11	0.02	0.02	0.52	0.27	0.02	0.04	**8.74**
	Intra	0.07	0.02	0	0.28	0.49	0.08	0.12	**8.35**
	Inter	0.15	0.10	0	0.38	0.32	0.03	0.22	**8.61**

Table 4

Interclass correlation and co-efficient of variation (CV) of VASI accuracy and reaction time across the different timeline of measurement.

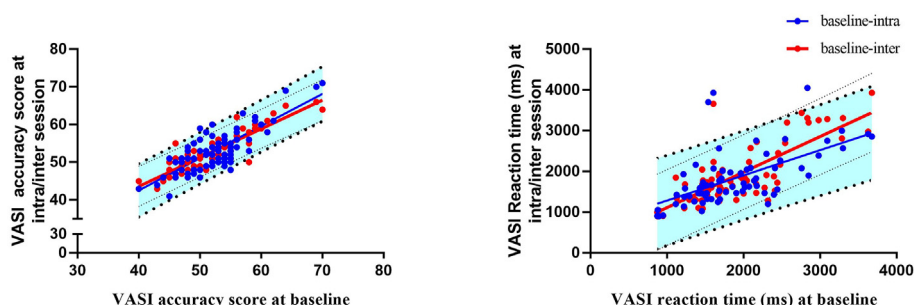| Test-retest reliability	Measure	Virtual location								
		R45	R90	R135	180	L45	L90	L135	0	Overall
ICC	Accuracy	0.60	0.66	0.79	0.83	0.86	0.61	0.86	0.77	0.93
	Reaction time	0.77	0.84	0.63	0.72	0.83	0.80	0.60	0.70	0.87
CV (in %)	Accuracy	8.53	11.33	7.67	9.87	6.13	8.54	7.57	8.86	9.66
	Reaction Time	12.00	5.99	6.33	5.43	16.73	5.94	8.95	12.54	11.99

Fig. 4. Cluster plots (top panels) of overall VASI accuracy scores (left) and reaction time (right). Symbols of different colors shows the agreement between overall baseline and re-test scores. The dotted line indicates limits of variance (95% confidence interval), while the blue shaded area reflects the maximum permissible region of variance.

VASI parameters (accuracy and reaction time) over time. The ICC analyses was performed to measure the relative homogeneity within sessions related to total observed variations between sessions. ICC values above 0.75 are considered to be of good reliability (Koo and Li, 2016). In the current study, the ICC for overall VASI scores was '0.93' (Table 4), indicated excellent test-retest reliability. Similarly, each location-wise ICC scores suggest a moderate to a high degree of reliability across sessions (Table 4). The results of the cluster plot (Fig. 4) also clearly demonstrated very good agreement for accuracy and reaction time measures which suggests good reliability.

VASI test has several applications as different conditions can impair the spatial localization of an individual. The types and degree of hearing loss degrade spatial localization tasks. There have been no gender-wise differences in the spatial localization tasks in the literature (Newton and Hickson, 1981; Nilsson et al., 1973). Hence, we did all the analysis by pooling both males' and females' data as a single component. Similarly, the localization errors (where localization occurs opposite to the actual sound source, also known as reversal errors) are more for sound sources with narrow-band

sounds and sounds spectrally limited to less than 8 kHz (Nakabayashi, 1974). Since our test uses the burst of white noise, it resolves the possible errors causing due to the spectrum of noise. Generally, the reversal error rate, as shown in different studies, ranged from 2% to 12% (Bronkhorst, 1995; Carlile et al., 1997). Compared to localization in the natural environment, the errors and confusion are more in the virtual environment and ranges from 12 to 20% for individualized HRTFs and 15—35% for non-individualized HRTFs (Besing and Koehnke, 1995; Bronkhorst, 1995). The use of non-individualised HRTFs to synthesise spatial percepts in VASI test makes the test portable and increases the utility of VASI test in evaluating spatial perception at home environments with minimal equipments. The high reliability found in our study helps in interpreting the results of the VASI test more confidently. Also, the excellent reliability scores suggest that this tool can be an easy way to train young adults who have poor spatial performance and document the changes based on auditory training in clinical setups. Since this test can be administered with only a computer and a headphone as an accessory, this tool can be easily incorporated into any clinical setting and has excellent prospects for spatial hearing assessments in adults.

4.1. Limitations and future directions

The reliability assessment was restricted only to young adults. Future studies can assess the reliability in children and older adults too. The study was carried out only on those with normal hearing sensitivity. It will be interesting to know the reliability of VASI in the clinical population, such as hearing impairment, central auditory processing disorder, auditory neuropathy, etc.

5. Conclusions

The study attempted to evaluate the test-retest reliability of VASI, which is a portable low-cost tool that can assess spatial perception in variety of settings. The results of the study demonstrated excellent test-test reliability of VASI, demonstrating high temporal stability of VASI test in terms of both accuracy and reaction times measures. The high test-retest reliability of VASI marks its effectiveness in auditory spatial evaluations and can be used to evaluate the training related outcomes in clinical population.

Declaration of competing interest

There is no conflict of interest to disclose. This is a non-funded research.

References

Abel, S.M., Hay, V.H., 1996. Sound localization the interaction of aging , hearing loss and hearing protection. Scand. Audiol. 25, 3—12. https://doi.org/10.3109/01050399609047549.
Algazi, V.R., Duda, R.O., Thompson, D.M., Avendano, C., 2001. The CIPIC HRTF database. In: Proceedings of the IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575). IEEE, pp. 99—102. https://doi.org/10.1109/ASPAA.2001.969552.
Besing, J.M., Koehnke, J., 1995. A test of virtual auditory localization. Ear Hear. 16, 220—229. https://doi.org/10.1097/00003446-199504000-00009.
Blauert, J., 1997. Spatial Hearing: the Psychophysics of Human Sound Localization.

MIT Press, Cambridge (Massachusetts). https://doi.org/10.1037/025238.
Brimijoin, W.O., Akeroyd, M.A., 2012. The role of head movements and signal spectrum in an auditory front/back illusion. Iperception 3, 179—181. https://doi.org/10.1068/i7173sas.
Bronkhorst, A.W., 1995. Localization of real and virtual sound sources. J. Acoust. Soc. Am. 98, 2542—2553. https://doi.org/10.1121/1.413219.
Brown, C.E., 1998. Coefficient of variation. In: Applied Multivariate Statistics in Geohydrology and Related Sciences. Springer Berlin Heidelberg, pp. 155—157. https://doi.org/10.1007/978-3-642-80328-4_13.
Brown, C.H., May, B.J., 2006. Comparative mammalian sound localization. In: Sound Source Localization. Springer-Verlag, pp. 124—178. https://doi.org/10.1007/0-387-28863-5_5.
Carlile, S., Leong, P., Hyams, S., 1997. The nature and distribution of errors in sound localization by human listeners. Hear. Res. 114, 179—196. https://doi.org/10.1016/S0378-5955(97)00161-5.
Clark, J.G., 1981. Uses and abuses of hearing loss classification. ASHA 23, 493—500.
Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., 1993. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. Proc. 20th Annu. Conf. Comput. Graph. Interact. Tech. SIGGRAPH 1993, 135—142. https://doi.org/10.1145/166117.166134.
Dorman, M.F., Loiselle, L.H., Cook, S.J., Yost, W.A., Gifford, R.H., 2016. Sound source localization by normal-hearing listeners, hearing-impaired listeners and cochlear implant listeners. Audiol. Neurotol. 21, 127—131. https://doi.org/10.1159/000444740.
Drennan, W., Gatehouse, S., Howell, P., VanTasell, D., Lund, S., 2001. Localization and speech-identification ability of hearing-impaired listeners using phase-preserving amplification. J. Acoust. Soc. Am. 109,. https://doi.org/10.1121/1.4744373, 2377—2377.
Dufour, A., Touzalin, P., Candas, V., 2007. Rightward shift of the auditory subjective straight ahead in right- and left-handed subjects. Neuropsychologia 45, 447—453. https://doi.org/10.1016/j.neuropsychologia.2006.05.027.
Faul, F., Erdfelder, E., Lang, A., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39, 175—191. https://doi.org/10.3758/BF03193146.
Gnanateja, N., 2014. Consonant confusion matrix. Available at: https://in.mathworks.com/matlabcentral/fileexchange/46461-consonant-confusion-matrix. (Accessed 21 December 2021).
Habib, M., Besson, M., 2009. What do music training and musical experience teach us about brain plasticity?. In: Music Perception, pp. 279—285. https://doi.org/10.1525/mp.2009.26.3.279.
Häusler, R., Colburn, S., Marr, E., 1983. Sound localization in subjects with impaired hearing: spatial-discrimination and interaural-discrimination tests. Acta Otolaryngol. Suppl. 400, 1—62. https://doi.org/10.3109/00016488309105590.
King, A.J., 1999. Auditory perception: does practice make perfect? Curr. Biol. 9, 143—146. https://doi.org/10.1016/S0960-9822(99)80084-0.
Knudsen, E.I., 1984. The role of auditory experience in the development and maintenance of sound localization. Trends Neurosci. https://doi.org/10.1016/S0166-2236(84)80081-8.
Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropr. Med. 15, 155—163. https://doi.org/10.1016/j.jcm.2016.02.012.
Leeuw, A.R., Dreschler, W.A., 1987. Speech understanding and directional hearing for hearing-impaired subjects with in-the-ear and behind-the-ear hearing aids. Acad. Med. Centre 16, 16. https://doi.org/10.3109/01050398709042152.
Letowski, T.R., Letowski, S.T., 2012. Auditory spatial perception : Auditory localization. Ames Res lab. https://doi.org/10.21236/ada562292.
Lorenzi, C., Gatehouse, S., Lever, C., 1999a. Sound localization in noise in normal-hearing listeners. J. Acoust. Soc. Am. 105, 1810—1820. https://doi.org/10.1121/1.426719.
Lorenzi, C., Gatehouse, S., Lever, C., 1999b. Sound localization in noise in hearing-impaired listeners. J. Acoust. Soc. Am. 105, 3454—3463. https://doi.org/10.1121/1.424672.
Majdak, P., Goupell, M.J., Laback, B., 2010. 3-D localization of virtual sound sources: effects of visual environment, pointing method, and training. Attention, perception. Psychophys 72, 454—469. https://doi.org/10.3758/APP.72.2.454.3-D.
Miller, J.D., Godfroy-Cooper, M., Wenzel, E.M., 2014. Using published HRTFS with Slab3D: metric-based database selection and phenomena observed. In: 20th International Conference on Auditory Display. ICAD2014).
Morse, J.M., Barrett, M., Mayan, M., Olson, K., Spiers, J., 2002. Verification strategies for establishing reliability and validity in qualitative research. Int. J. Qual. Methods 1, 13—22. https://doi.org/10.1177/160940690200100202.
Nakabayashi, K., 1974. Sound localization on the horizontal plane. J. Acoust. Soc. Japan 30, 151—160.
Neelon, M.F., Jenison, R.L., 2004. The temporal growth and decay of the auditory motion aftereffect. J. Acoust. Soc. Am. 115, 3112—3123. https://doi.org/10.1121/1.1687834.
Newton, V.E., Hickson, F.S., 2021. Sound localization Part II: a clinical procedure. J. Laryngol. Otol. 95, 41—48. https://doi.org/10.1017/S0022215100090381.
Newton, V.E., Hickson, F.S., 1981. Part II: a clinical procedure. J. Laryngol. Otol. 95, 41—48. https://doi.org/10.1017/S0022215100090381.
Nguyen, K., Suied, C., Viaud-Delmon, I., 2009. Spatial audition in a static virtual environment: the role of auditory-visual interaction. J. Virtual Real. Broadcast. 6, 1—11.
Nilsson, R., Lidén, G., Rosén, M., Zöller, M., 1973. Directional hearing, three different test-methods. Scand. Audiol. 2, 125—131. https://doi.org/10.3109/

01050397309044945.

Nisha, K.V., Kumar, A.U., 2017. Virtual auditory space training-induced changes of auditory spatial processing in listeners with normal hearing. J. Int. Adv. Otol. 13, 118–127. https://doi.org/10.5152/iao.2017.3477.

Nisha, K.V., Kumar, A.U., 2016. Effect of localization training in horizontal plane on auditory spatial processing skills in listeners with normal hearing. J. Indian Speech Lang. Hear. Assoc. 30, 28–39. https://doi.org/10.4103/jisha.JISHA_2_17.

Noble, W., Byrne, D., 1990. A comparison of different binaural hearing aid systems for sound localization in the horizontal and vertical planes. Br. J. Audiol. 24, 335–346. https://doi.org/10.3109/03005369009076574.

Noble, W., Byrne, D., Lepage, B., 1994. Effects on sound localization of configuration and type of hearing impairment. J. Acoust. Soc. Am. 95, 992–1005. https://doi.org/10.1121/1.408404.

Noble, W., Byrne, D., Ter-Horst, K., 1997. Auditory localization, detection of spatial separateness, and speech hearing in noise by hearing impaired listeners. J. Acoust. Soc. Am. 102, 2343–2352. https://doi.org/10.1121/1.419618.

Perception Research Systems, 2007. Paradigm Stimulus Presentation [WWW Document]. URL. http://www.paradigmexperiments.com.

Polley, D.B., Steinberg, E.E., Merzenich, M.M., 2006. Perceptual learning directs auditory cortical map reorganization through top-down influences. J. Neurosci. 26, 4970–4982. https://doi.org/10.1523/JNEUROSCI.3771-05.2006.

Roffler, S.K., Butler, R.A., 1968. Factors that influence the localization of sound in the vertical plane. J. Acoust. Soc. Am. 43, 1255–1259.

Ruscetta, M.N., Palmer, C.V., Durrant, J.D., Grayhack, J., Ryan, C., 2005. Validity, internal consistency, and test/retest reliability of a localization disabilities and handicaps questionnaire. J. Am. Acad. Audiol. 16, 585–595. https://doi.org/10.3766/jaaa.16.8.7.

Sosa, Y., Teder-Sälejärvi, W.A., McCourt, M.E., 2010. Biases of spatial attention in vision and audition. Brain Cognit. 73, 229–235. https://doi.org/10.1016/j.bandc.2010.05.007.

Venkatesan, S., 2009. Ethical Guidelines for Bio Behavioral Research Involving Human Subjects. AIISH.

Walter, S.D., Eliasziw, M., Donner, A., 1998. Sample size and optimal designs for reliability studies. Stat. Med. 17, 101–110. https://doi.org/10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E.

Wenzel, E.M., Miller, J.D., Abel, J.S., 2000. Sound Lab: a real-time, software-based system for the study of spatial hearing. In: 108th AES Convention. The Audio Engineering Society, Paris, France, pp. 1–27.

Zahorik, P., Sundareswaran, V., Wang, K., Tam, C., 2006. Perceptual recalibration in human sound localization: learning to remediate front-back reversals. J. Acoust. Soc. Am. 120, 343–359. https://doi.org/10.1121/1.2208429.

Zahorik, P., Wightman, F.L., 2001. Loudness constancy with varying sound source distance. Nat. Neurosci. 4, 78–83. https://doi.org/10.1038/82931.

Zhong, X., Sun, L., Yost, W., 2016. Active binaural localization of multiple sound sources. Robot. Autonom. Syst. 85, 83–92. https://doi.org/10.1016/j.robot.2016.07.008.