scientific reports

OPEN

Individual differences in temporal order judgment

Leah Fostick[1]✉, Moti Zwilling[2] & Harvey Babkoff[3]

Large individual differences can be observed in studies reporting spectral TOJ. In the present study, we aimed to explore these individual differences and explain them by employing Warren and Ackroff (1976) framework of direct identification of components and their order (direct ICO) and holistic pattern recognition (HPR). In Experiment 1, results from 177 participants replicated the large variance in participants' performance and suggested three response patterns, validated using the K-Means clustering algorithm. In Experiment 2, the introduction of three tone-pairs to 90 participants, as opposed to a single pair, markedly decreased the propensity of participants to utilize HPR. Experiment 3 assessed 85 participants and demonstrated that diotic presentations significantly reduced the prevalence of HPR utilization. These results confirmed that direct ICO is used mainly when stimuli are difficult to group, and when grouping is possible, HPR is used. The results also show that these strategies are flexible and change according to the experimental manipulation. Spectral TOJ provides an opportunity to observe individual differences in auditory perception and facilitates the investigation of the diverse psychoacoustic cues underlying its performance.

Keywords Temporal order judgment, Individual differences, Direct identification, Holistic pattern recognition

Temporal order judgment (TOJ) is a task designed to assess the individual's temporal resolution, the ability to track time-based changes in auditory signals[1], by measuring the minimal time interval between two sounds required to report their order. This interval can be measured as the inter-stimulus interval (ISI), the silent gap between the offset of the first sound to the onset of the second sound, or as stimulus-onset-asynchrony (SOA), the time separating the onset of the first stimulus from the onset of the second. The study of TOJ began with the seminal work of Hirsh and Sherrick[2], who tested TOJ within and across modalities. Within the auditory modality, a variety of TOJ paradigms have been tested to date, including manipulating the parameters by which each of the two sounds may be identified. One of these paradigms uses monaural stimulation and two sounds of different frequencies (referred to as Spectral TOJ or Monaural TOJ). With this paradigm, participants report the temporal order by the pitch of the sounds and respond either High-Low or Low–High. Another widely used paradigm consists of stimulating both ears with the same frequency tones. With this paradigm, the participants respond either Right-Left or Left–Right (referred to as Dichotic, Binaural, or Spatial TOJ). These paradigms were used over the years to measure temporal resolution in normal populations and a variety of at-risk populations, such as dyslexic readers[3–9], aging adults[10–13], sleep-deprived young adults[14,15], participants with attention deficit hyperactivity disorder[16], and aphasic patients[17,18].

Sequence perception was extensively studied in a series of studies by Warren et al.[19–23]. They usually had 3–5 sounds in their sequences and found that accuracy rates were much higher in some sequences than others. They suggested that participants use different strategies in sequence perception and theorized that two possible mechanisms exist to report the sequence of auditory stimuli. When the sequence includes stimuli that are different from each other in pitch, timber, or duration, they are segregated into their components, and the sequence is perceived through *direct identification of components and their order* in the sequence (direct ICO); when the characteristics of the stimuli are relatively similar, they are integrated and perceived globally through *holistic pattern recognition* (HPR)[20]. In direct ICO, listeners perceive the characteristics of each sound in the sequence and then decide on the sounds' order. In HPR, the listeners perceive the sequence as one unit and report the order based on differentiating one pattern from the other. Direct ICO versus HPR in TOJ tasks was also suggested by others[17,24–27].

Spectral TOJ is widely reported in the literature [e.g.,[2–18,24–28]]. A close inspection of the reported data shows a large variance in participants' accuracy levels. For example, Fostick and Babkoff[28] showed that 49% of 388 participants (in 13 different experiments) had an accuracy level higher than 75% at very short ISI values

[1]Department of Communication Disorders, Auditory Perception Lab in the Name of Laurent Levy, Ariel University, 40700 Ariel, Israel. [2]Department of Economics and Business Administration, Ariel University, 40700 Ariel, Israel. [3]Department of Psychology, Bar-Ilan University, 5290002 Ramat-Gan, Israel. ✉email: leah.fostick@ariel.ac.il

(0–8 ms), while 18% had an accuracy level less than 75% even at very long ISI values (150–300 ms). Semal and Demany[29] and Mathias[30] reported that some participants could perceive the order of two tones that differ in their pitch, while others performed poorly, even after training. Other studies have also reported groups of participants with very high accuracy levels at very short ISIs[5,10,17,24,31] while other participants perform very poorly even when the two sounds are separated by relatively very long ISI's. Reports on participants who had difficulties performing TOJ with long ISIs are scarce. This information is presented explicitly in Fostick and Babkoff[28] and in Fostick and Babkoff[24], where 26% of participants had an accuracy level below 75% even at ISIs longer than 200 ms. Implicitly, the difficulties in performing TOJ with long ISIs are presented in Szymaszek et al.[10], who stated: *"If the participant could not report the order of stimuli with ISI of 160 ms, recalibration was done, and the session started with an ISI of 320 ms"* (p. 139).

This large variability in accuracy can be explained by Warren's suggestion of two mechanisms for sequence perception, if, at least for Spectral TOJ, they are employed by different participants performing the same sequence. Accordingly, in the present study, we aimed to test whether manipulating the ability to use direct ICO or HPR will affect participants' performance, such that increasing the ability to segregate the sounds will increase the number of participants using direct ICO while increasing the ability to integrate the sounds will increase the number of participants using HPR. To accomplish this goal, we first replicated the findings of large individual differences in spectral TOJ (Experiment 1) and then manipulated the ability to integrate the sound into a pattern (Experiments 2 and 3).

Experiment 1: Replication

Large variance in spectral TOJ was reported explicitly in one study[28] and only implicitly in others[5,10,17,24,31]. Therefore, the aim of Experiment 1 was to test spectral TOJ on a large sample and replicate the individual differences suggested in previous studies.

Results and discussion

A schematic illustration of the spectral TOJ task is presented in Fig. 1. Participants' accuracy by ISI is presented in Fig. 2, showing large variability in performance as suggested in previous studies[5,10,17,24,31]. As can also be seen in Table 1, the accuracy range was wide for all ISIs, extending from a range of 62% (ISI of 90 and 120 ms) to 75% (ISI of 15 ms). As can also be seen in Table 1, the Coefficient of Variance (CoV), reflecting the variance relative to the mean performance in each ISI (calculated as the standard deviation divided by the mean), is 20% and larger in almost all ISIs, indicating high variability in performance[32]. The accuracy range in each ISI showed that some participants scored 100% in all ISIs, while others barely reached 50%. When manipulating the time interval between sounds (ISI), accuracy is expected to increase with time. Such an increase is evident in the mean, albeit small: an increase of 10% for an increase of 238 ms in ISI. Interestingly, the minimum accuracy level, which was expected to increase with ISI, along with the accuracy range and SD, stayed similar across all ISI. Also, some participants scored 100% correct even at ISI of 2 and 5 ms, while others continued to score below 50% even at a long ISI of 240 ms. This extensive range of performance in each ISI refutes a possible argument about a ceiling/floor effect in the performance. It suggests that some participants find spectral TOJ very easy and some very difficult.

The data of Experiment 1 replicate previous studies and show large individual differences in performing spectral TOJ. Next, to understand what underlies these differences, we attempt to explore whether the data is scattered randomly or whether it can be grouped. TOJ threshold is the ISI required for an accuracy of 75%. In this sample of 177 participants, 96 participants (54%) had scores higher than 75% in the shortest and longest ISIs (Fig. 3a), and 44 (25%) had scores lower than 75% in the shortest and longest ISI (Fig. 3b). The former group can be defined as High-Level Performers (HLPs) and the latter as Low-Level Performers (LLPs). Since these two groups did not cross 75%, thresholds could not be computed. Only 37 participants (21%) showed a psychophysical function of response accuracy by ISI, with accuracy lower than 75% in the shortest ISI and higher than 75% in the longest ISI (Fig. 3c). This group can be defined as Mid-Level Performers (MLPs), and its mean TOJ threshold was 65.01 ms (SD = 52.19 ms).

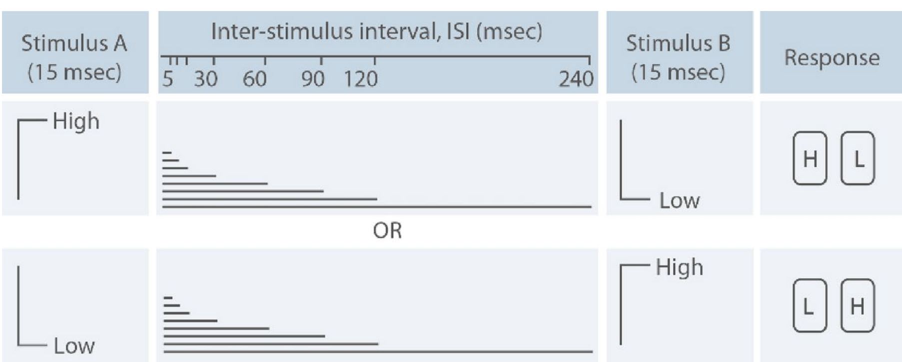

Fig. 1. A schematic illustration of spectral TOJ task presenting the two possible combinations of high and low tones, with different ISI values with each tone pair.

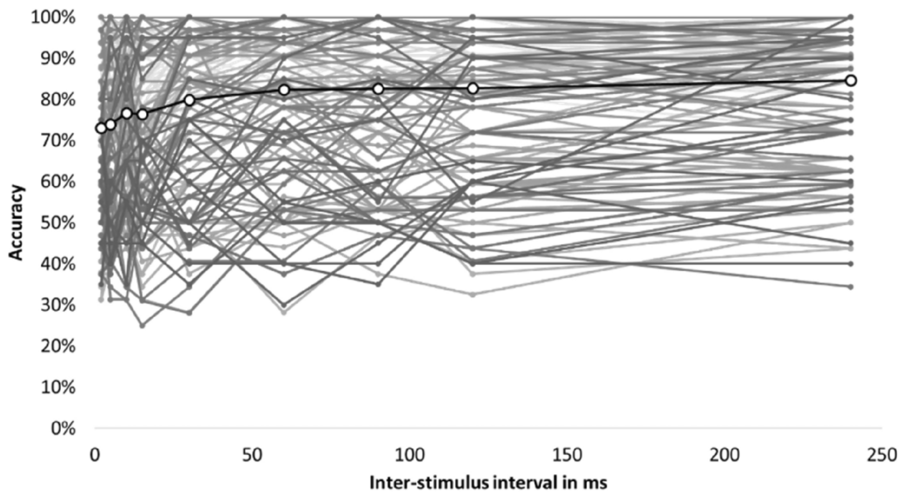

Fig. 2. Spectral TOJ accuracy by inter-stimulus interval (ISI) for 177 participants. The bold black line represents the participants' average.

ISI in ms	Accuracy range	Median (%)	Mean (%)	SD (%)	CoV (%)
2	31–100%	78	75	20	26
5	31–100%	75	75	20	26
10	31–100%	81	78	19	24
15	25–100%	84	78	20	25
30	28–100%	88	81	19	23
60	28–100%	91	83	18	21
90	38–100%	91	84	18	20
120	38–100%	91	84	18	20
240	34–100%	91	85	16	18

Table 1. Spectral TOJ results of each ISI: Accuracy range, Median, Mean, SD. ISI = inter-stimulus interval, SD = standard deviation, CoV = coefficient of variation.

To validate our categorization of participants' performance in the spectral TOJ task, we employed the K-Means clustering algorithm[33]. This unsupervised machine learning method is recognized for its efficiency in identifying naturally occurring groupings within datasets, particularly regarding participant accuracy across varying ISIs. Each participant's accuracy rates across the nine different ISIs served as multidimensional input variables for the clustering algorithm. The algorithm aimed to investigate whether participant data would naturally divide into clusters corresponding to the previously identified groups of HLP, MLP, and LLP. Figure 4 illustrates the application of the Within-Cluster Sum of Squares (WCSS) and the Elbow Method alongside Average Silhouette Scores to determine the optimal cluster count. The analysis suggested that *a three-cluster solution* provided a balance between minimal intra-cluster variation and maximal inter-cluster differentiation, consistent with our predefined categorizations.

Figure 5 depicts the confusion matrix comparing the K-Means derived classifications with our initial categorizations based on accuracy at the shortest and longest ISIs. A Chi-square test indicated no significant difference between the two classifications ($\chi^2_{(2)} = 4.970$, $p = 0.083$). For HLPs, precision (true positives divided by the number of true positives plus the number of false positives) was 99%, and sensitivity (the number of true positives divided by the number of true positives plus the number of false negatives) was 92%, resulting in an F1-score (2 * (precision * sensitivity) / (precision + sensitivity)) of 95%. For LLP, precision was 85%, sensitivity was 75%, and F1-score was 80%. For MLP, precision was 61%, sensitivity was 81%, and F1-score was 70%. The overall accuracy of the K-means classification compared to your classification was 85%. The precision rate for MLP was rather low, reflecting the heterogeneity of this group. Although all participants in this group had a low accuracy rate at ISI = 2 ms and a high accuracy rate at ISI = 240 ms, the rate at which participants' accuracy increased with ISI varied.

It is expected that when participants rely on the perception of each sound in a sequence for judging their order, order judgment will be more difficult and, therefore, with lower accuracy when ISI is short. When ISI is long, order judgment will be easier, and, therefore, with higher accuracy. Accordingly, it could be argued that MLPs, who show low accuracy at short ISIs and an increase in performance with the increase in ISIs, use direct ICO. This group comprised a minority of participants (21% of the 177 participants) and had calculable

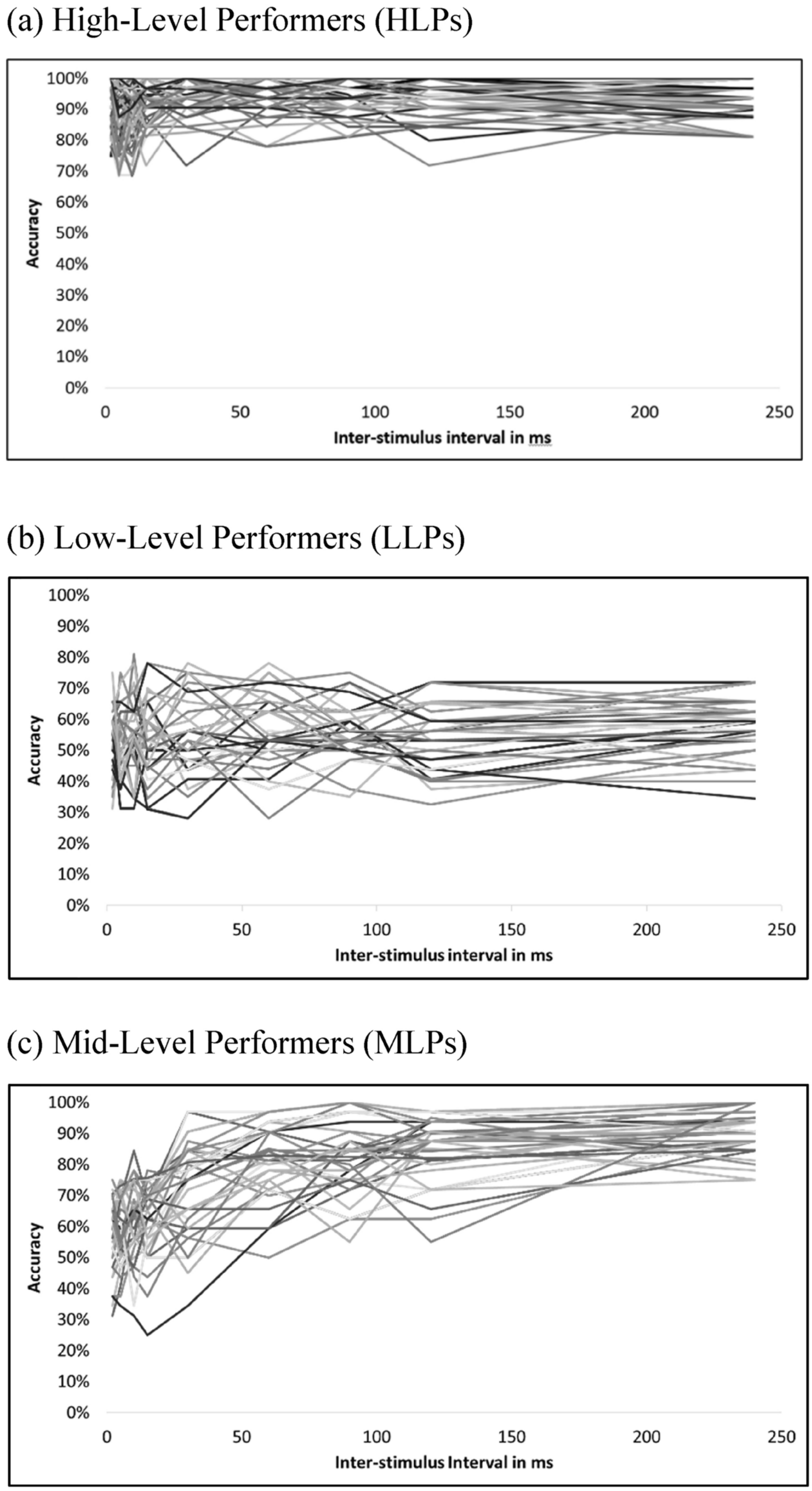

Fig. 3. The study sample (n = 177) divided into participants whose accuracy level were (**a**) above 75% in the shortest ISI (High-Level Performers, HLPs); (**b**) below 75% at the longest ISI (Low-Level Performers, LLPs); and (**c**) below 75% at the shortest ISI and above it in the longest ISI (Mid-Level Performers, MLPs).

thresholds. Fostick and Babkoff[28] presented thresholds of 13 spectral TOJ experiments. Their arithmetic mean weighted with the number of participants is 58 ms, corresponding to the thresholds found in the present study for the same subset of participants. In HPR, since the judgment does not require the perception of each sound separately, the performance is not expected to change with ISI. Such is the performance of the HLPs and LLPs. Accordingly, it could be argued that these groups use HPR but with different levels of efficiency.

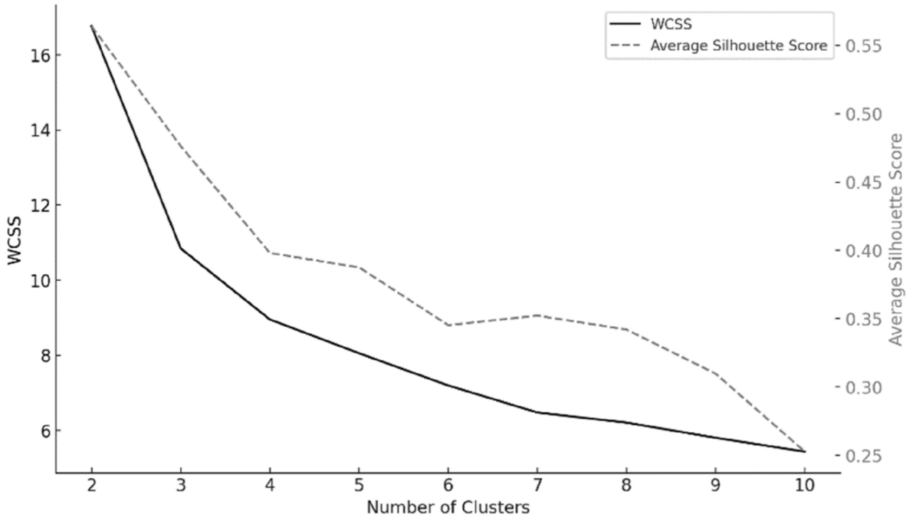

Fig. 4. Evaluation of Clustering: WCSS and Silhouette Scores Across Different Cluster Counts.

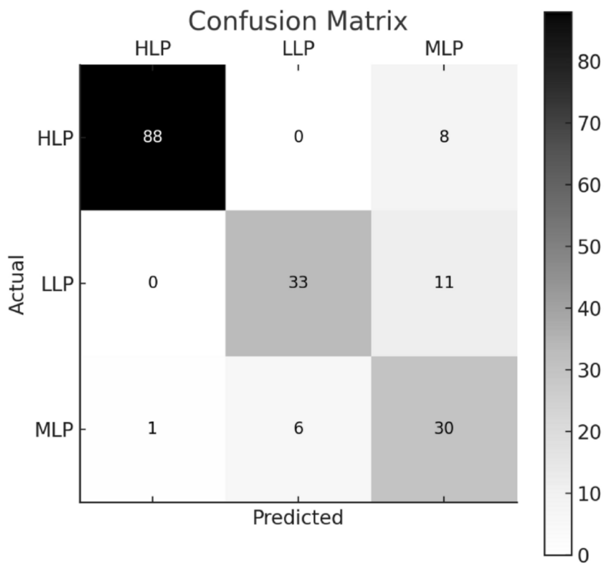

Fig. 5. Confusion matrix for classifying participants' performance by accuracy in the shortest and longest ISIs (Actual) versus classification by K-Means (Predicted).

Experiment 2: Roving frequencies

Participants who use HPR, learn the pattern of each of the two sequences and remember them. Then, in each trial, they identify which of the two patterns was presented. Assuming this, manipulating participants' ability to learn and remember the pattern of the two sequences will affect their ability to use HPR. To test this hypothesis, in Experiment 2, we used three roving tone pairs. Thus, to use HPR, participants will need to remember six different patterns, with an expected occurrence of 16%, rather than two patterns with an expected occurrence of 50%. Accordingly, we hypothesize that roving frequencies will decrease HLP and increase MLP, whose judgment relies on the perception of each tone in a pair, rather than remembering their pattern. Alternatively, if roving frequencies will only affect the difficulty level of the task, without affecting the use of HPR and direct ICO, then the prevalence of both HLP and MLP will decrease, and the prevalence of LLP will increase compared to the condition of one fixed tone pair. In addition, if roving frequencies increase the tendency to use MLP on top of HLP, compared to a condition of two fixed frequencies, because of the difficulty of remembering six patterns, then both conditions will result in similar thresholds. Alternatively, if the difference between the roving and the fixed condition was the result of increasing the difficulty level of the roving condition, then its threshold will be higher than that of the fixed condition.

Results and discussion

HLP, MLP, and LLP, were defined according to their accuracy rate at the shortest and longest ISIs, in the same manner as in Experiment 1 (Fig. 6). Figure 7 presents the prevalence of HLP, MLP, and LLP in the experimental conditions of Roving Frequencies and Fixed Frequencies. To make sure the results of Fixed Frequencies are not related to the frequencies used, the results of the Roving Frequencies condition for 1 and 2 kHz are presented separately. The Roving Frequencies condition showed a lower prevalence of HLP and a higher prevalence of MLP, than the fixed condition ($\chi^2_{(1,2)} = 24.89$, $p < 0.001$). This supports the hypothesis that roving frequencies will make it challenging to learn and remember the patterns of the high-low and low–high orders, thus limiting HPR (HLPs) and increase the use of direct ICO (MLPs). There was also a difference between the Fixed condition and the Roving condition when analyzing only the 1 and 2 kHz data ($\chi^2_{(1,2)} = 16.12$, $p < 0.001$) and no difference between the Roving conditions with all frequencies and that with only 1 and 2 kHz ($\chi^2_{(1,2)} = 3.188$, $p = 0.20$). This confirms that the differences observed between the Roving and Fixed Frequencies conditions are not due to the frequencies tested, but due to the inclusion of multiple tone pairs.

As detailed in Experiment 1, the defining criterion for different response groups was 75% accuracy, which is required for calculating TOJ threshold. According to this definition, thresholds can be calculated for participants in the MLP group in different conditions. MLP thresholds for the Fixed condition (Mean = 82.11, SD = 54.53) were similar to that of the Roving condition with all frequencies (Mean = 104.17, SD = 61.62, t(54) = − 1.16, $p = 0.251$) and to the 1 and 2 kHz only (Mean = 74.29, SD = 51.77, t(45) = 0.46, $p = 0.650$). A Chi-square Goodness-of-Fit test on the LLP group between the conditions showed that it is not different from an expected equal prevalence ($\chi^2_{(2)} = 1.29$, $p = 0.52$). These results suggest that the difference in response patterns was not due

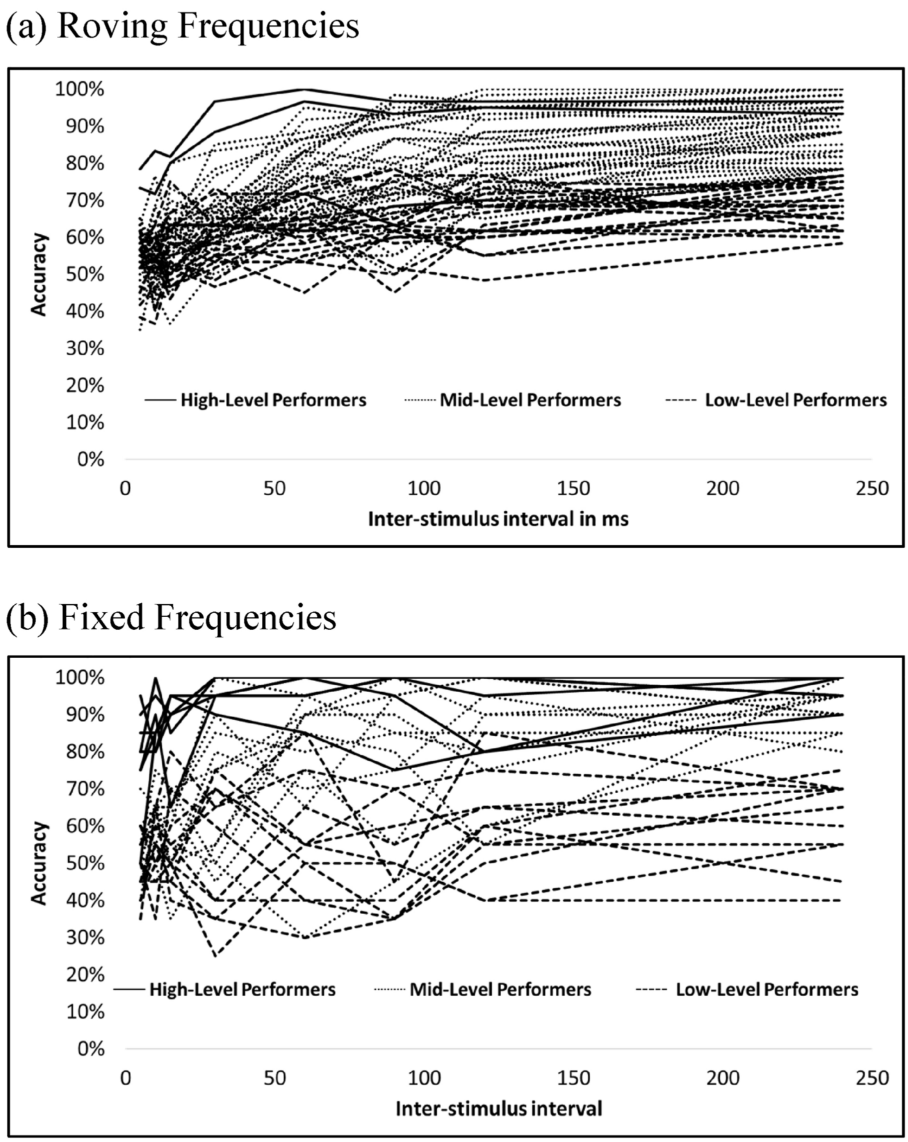

Fig. 6. Accuracy data for individual participants in (**a**) Roving Frequencies and (**b**) Fixed Frequencies conditions.

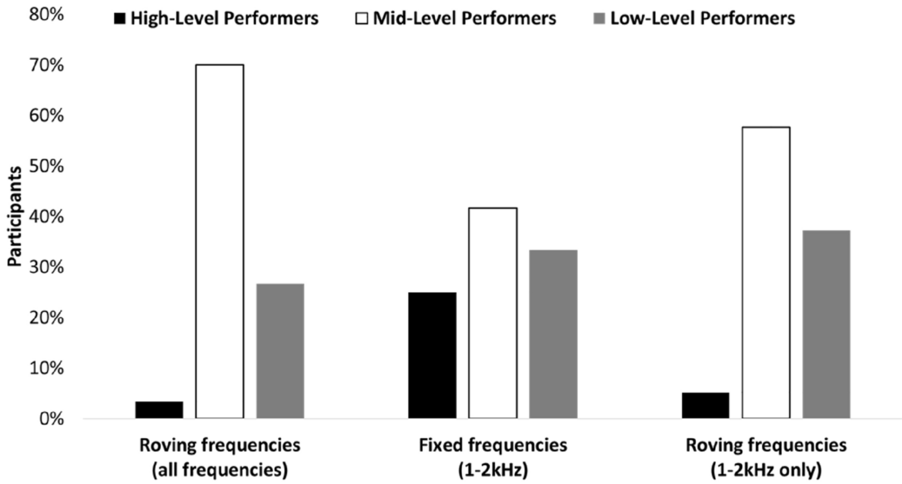

Fig. 7. Prevalence of HLP, MLP, and LLP in the Roving Frequencies condition (including all three tone pairs: 0.8–1.6 kHz, 1–2 kHz, and 1.2–2.4 kHz), Fixed Frequencies condition (included one tone pair of 1–2 kHz), and the Roving Frequencies condition again, but presenting only the results for the 1–2 kHz pair.

to a general increase in the task's difficulty level or fatigue due to larger number of trials in the Roving condition. Therefore, we conclude that is it is due to a change in the perceptual strategies participants used to perform it. The results showed that when HPR is difficult, the prevalence of HLP decreases along with an increase in MLP (and not in LLP), pointing to a possible change in perceptual strategies from HPR to direct ICO.

The Fixed and Roving condition were compared using a between-subjects' design. This can limit the conclusion that the experimental manipulation caused the difference observed. Indeed, the distribution observed in the Fixed condition is different from that observed in Experiment 1 that was also done with fixed frequencies. Although a within-subjects' design is preferrable, it is problematic for spectral TOJ that is subjected to a significant learning effect[34]. Nevertheless, the distribution of performance in most TOJ tasks is similar[28]. As can be seen from Experiments 1 and 2, the difference between experiments of the same condition (i.e., Experiment 1 and Experiment 2—Fixed frequencies condition), is smaller than the difference between experiments of different conditions (i.e., Experiment 2—Fixed and Roving conditions).

Experiment 3: Presentation

The principles of stream segregation pose that sounds presented from the same spatial origin will be fused into the same stream more than those presented from different locations, which will be segregated. Moreover, sounds delivered to the same ear excite the same cochlea and can be processed both peripherally and centrally. In contrast, sounds delivered to different ears excite different cochleae and can only be processed centrally. In TOJ, the two sounds in each pair are delivered one after the other and can be presented in three ways: monaurally (i.e., both tones presented to the same single ear), diotically (i.e., each tone presented to both ears synchronously at the same time), or dichotically (i.e., each tone presented to a different ear). In monaural and diotic presentations, the same cochlea is excited and processes the sounds. In these conditions, grouping is more trivial than in the dichotic presentation when different cochleae are excited, and order processing is done centrally.

If the way of presentation affects the ability to group sounds, it can also affect the tendency to use HPR and direct ICO. Accordingly, we hypothesize that monaural and diotic presentation will facilitate HPR, resulting in a high prevalence of HLP. Alternatively, dichotic presentation will limit the use of HPR, and increase the use of direct ICO and increasing the prevalence of MLP.

Results and discussion

HLP, MLP, and LLP, were defined in the same way as in Experiments 1 and 2 (Fig. 8). Figure 9 presents the prevalence of HLP, MLP, and LLP according to the way of presentation. HLP, MLP, and LLP distribution for Monaural and Diotic presentations are similar ($\chi^2 = 0.43$, $p = 0.81$), and different from the distribution in the Dichotic presentation condition (Dichotic vs. Monaural: $\chi^2 = 34.35$, $p < 0.001$; Dichotic vs. Diotic: $\chi^2 = 30.41$, $p < 0.001$). This suggests that, as hypothesized, sounds presented on the same cochlea facilitate HPR, and result in more HLP in Monaural and Diotic presentations. Sounds presented in different cochleae result in more MLP in Dichotic presentations due to the facilitation of direct ICO. Contrary to our expectations, the prevalence of LLPs also increased in Dichotic presentation.

The present experiment was conducted using a between-subjects design. As in Experiment 2, the use of a between-subjects rather than the preferred within-subjects design was done to avoid a learning effect, which was shown to be large in spectral TOJ[34]. Yet, the performance distribution of the diotic conditions was similar to the distribution observed in Experiment 1, both performed in the same conditions.

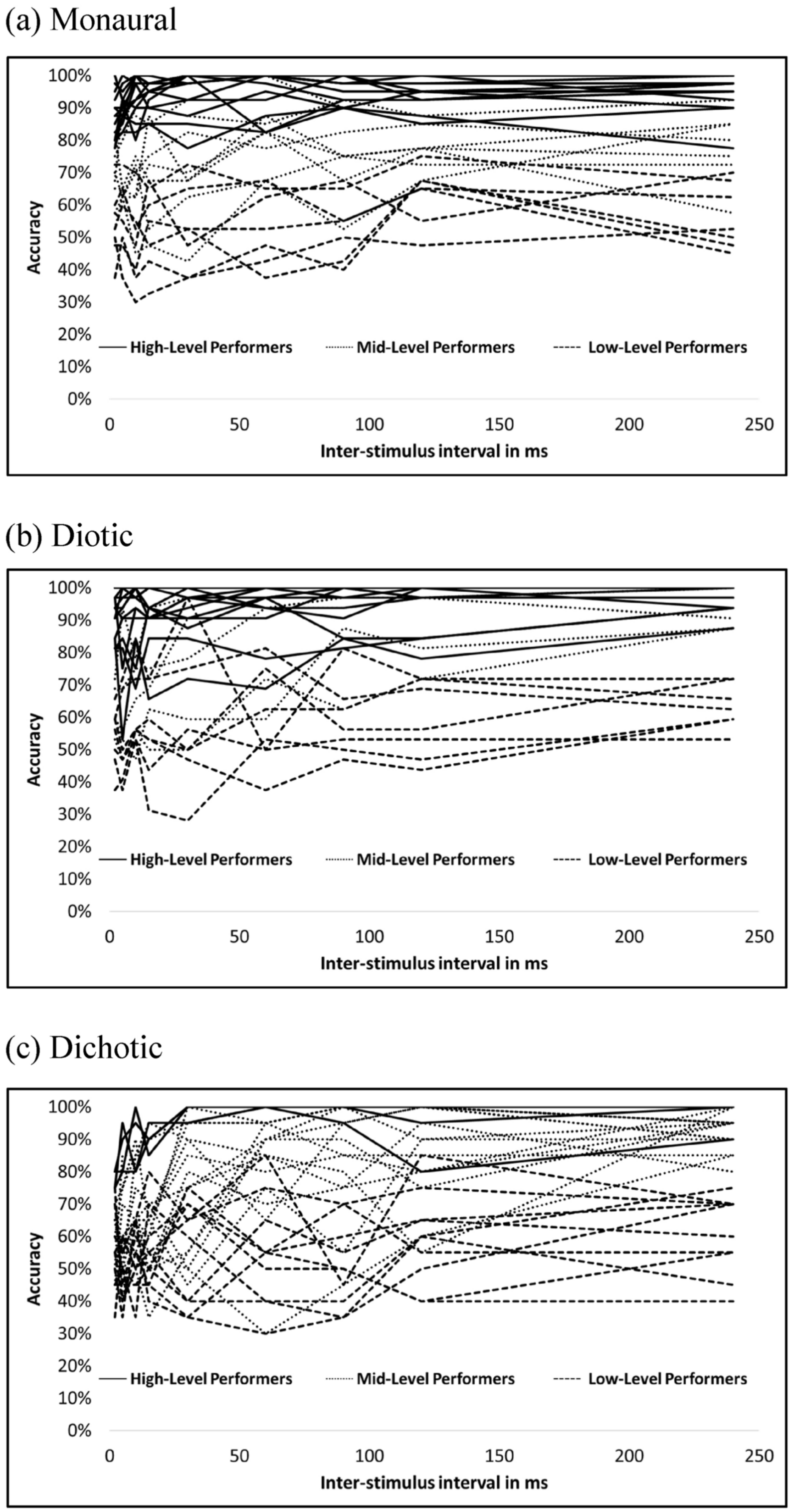

Fig. 8. Accuracy data for individual participants in (**a**) Monaural, (**b**) Diotic, and (**c**) Dichotic presentation.

General discussion

The aim of the present study was to test whether the variance observed in spectral TOJ is related to differences in perceptual strategies related to direct ICO and HPR. Experiment 1 confirmed the large inter-individual differences in spectral TOJ reported in the literature. This variance consists of participants whose performance is high even at the shortest ISI (HLPs), participants whose performance is low even at the longest ISI (LLPs), and participants whose performance changes with ISI (MLPs). The unsupervised K-Means algorithm validated a classification

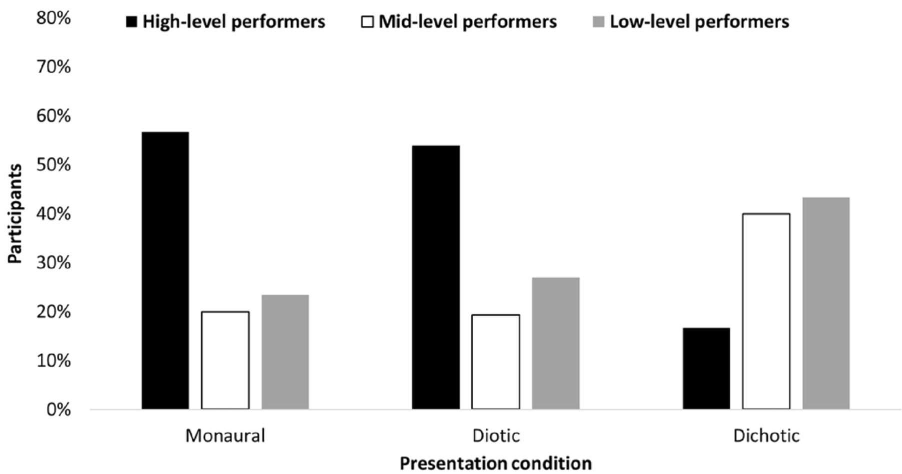

Fig. 9. Prevalence of HLP, MLP, and LLP according to presentation conditions: monaural (both tones presented to the same single ear), diotic (each tone presented to both ears synchronously at the same time), and dichotic (each tone presented to a different ear).

into three groups. Experiments 2 and 3 showed that the prevalence of HLP is larger under conditions of pattern identification and sound grouping, while the prevalence of MLP was larger under conditions that limit them.

The conditions that increase/decrease the number of HLP and MLP correspond to Warren and Ackroff[21] suggestion that direct ICO is used mainly when stimuli are difficult to group, and when grouping is possible, HPR is used. The data of the present study showed that both the frequency and spatial location manipulations affect HLP prevalence for conditions related to HPR and MLP for direct ICO. Also, in accordance with Warren and Ackroff[21], the data of Experiments 2 and 3 show flexibility in using direct ICO and HPR. When the experimental manipulations limit judgments based on either HPR or direct ICO, the number of participants, mainly in the groups of HLPs and MLPs, changed accordingly.

However, if according to Warren[19], direct ICO or HPR is decided based on the characteristics of the sequence, all (or most) participants should have employed the same strategy for a given sequence, with different perceptual strategies for different sequences. The existence of different types of performers (LLP, MLP, HLP) suggests that neither of these is correct. The results of the present study showed that for the same sequence, different participants employ different perceptual strategies. These results highlight the ability to use temporal resolution with additional psychoacoustic information than time-based changes in the presented sounds.

LLPs pose an intriguing observation. Their relatively high prevalence (30%) cannot exclude them as outliers or explain their performance with amusia. LLPs' performance is unaffected by the increase in ISI, suggesting they do not use direct ICO. Manipulating the ability to remember the High-Low and Low–High patterns did not affect their performances, suggesting they may already had difficulty learning and remembering the High-Low and Low–High patterns even in Fixed Frequencies. In dichotic presentation, the prevalence of LLPs increased, although we hypothesize that they do not use direct ICO. This increased number of LLPs may reflect an increase in the task's difficulty level but should be further explored.

Spectral TOJ, a task designed for studying temporal processing, provides a unique opportunity to learn about the different cues available for sequence perception and to explore people's tendency to choose one perceptual strategy over another [see also[23–30]]. The data of the current study showed that the tendency to choose either strategy is not necessarily predetermined and is based partially on the cues provided by the sequence presented and partially on a personal tendency of the participants. This leads to a key question of the factors underlying the use of each perceptual strategy. Do peripheral factors affect the tendency to perceive some cues better than others? We reject this hypothesis since previous studies[29,35] showed that musicians and non-musicians with similar peripheral processing differ in their spectral TOJ response patterns.

Since spectral TOJ involves rapidly presented short tones with different frequencies, the tendency to perceive psychoacoustic information may underline the observed differences in perceptual strategies. The masking pattern of one frequency by the other and the temporal fine structure (TFS) at the output of an auditory filter centered between the two tones, differ between low–high and high-low sequences. Therefore, capturing their information can facilitate HLP, while disregarding them can facilitate MLP[36]. The performance of LLP can be affected by difficulty attending to the proper psychoacoustic information, such as the TFS cues. It could be also they have difficulty in pitch perception that is not met by the training and testing conducted prior to the experiment, as also suggested previously[37]. Other explanation can be misunderstanding of the task, or mislabeling the stimuli (e.g., labeling low–high as high-low). Therefore, future studies should focus on manipulating masking and TFS information, as well as testing cognitive abilities such as learning, working memory, and attention control. Moreover, the ISI setting may be unsuitable for spectral TOJ, as many of the participants do not have psychophysical function of response accuracy by ISI. Future studies should explore whether shortening sound duration, allowing overlap between sounds, or other modification, will increase the number of participants with psychophysical function, or will decrease them due to additional perceptual cues that it will provide. Future studies can also focus on the effect of musical training on perceptual strategy. As the current study's inclusion criteria disregarded the participants' musical education, it is not clear whether it can explain the performance observed in the current study. Although it is unlikely that all participants in the HLP or MLP groups are musicians, this line of research is interesting to pursue.

Method

Participants

Different participants were recruited for each experiment. Experiment 1 included 177 undergraduate students (104 females, 73 males) aged 20–35 years (mean = 26.3 sd = 1.7); Experiment 2 included 90 undergraduate students (55 females, 35 males) aged 20–35 years (mean = 26.4, sd = 2.6). One group performed the classical TOJ paradigm with Fixed Frequencies (n = 30), similar to previous experiments using this design [e.g.,[10,17,18,25,26,28–31]], and the other performed spectral TOJ with Roving Frequencies (n = 60). Since the roving frequencies condition can be confusing, we anticipated it could result in a larger SD than the fixed condition. To address the potential impact of unequal variances on statistical comparisons, we a priori recruited a larger sample size for the roving condition, ensuring greater statistical power for the challenging condition; Experiment 3 included 85 undergraduate students (48 females, 42 males), aged 20–35 years (mean = 25.7 sd = 2.1). Participants were randomly divided into three conditions: Monaural (n = 30), Diotic (n = 26), and Dichotic (n = 30). The participants in all experiments had hearing thresholds of at least 20 dB HL for 0.5, 1, 2, and 4 kHz. Participants with learning disabilities or ADHD were excluded.

Task and stimuli

All experiments were carried out in the method of constant stimuli. In Experiments 1 and 3, Spectral TOJ was carried out with stimuli and design similar to those reported previously[6,13,24,28,35,38]. In all experiments, the tones were 15 ms pure tones with rise/fall times of 1 ms and were presented at 40 dB SL. In Experiments 1 and 3, each tone pair included 1 and 1.8 kHz. The use of two pure tones which are around one-octave apart is common in these TOJ tasks [e.g.,[4,25,28,39]], although some used closer frequencies[17,27,31] and other more distant[10,11,25,27]. The low tone was set to 1 kHz to make sure the tonality is perceived properly with duration of 15 ms[40]. In Experiment 2, the Fixed Frequencies condition included tone pairs of 1 and 2 kHz tones and in the Roving Frequencies condition, the frequencies of the tone pairs were 0.8 and 1.6 kHz, 1 and 2 kHz, and 1.2 and 2.4 kHz, and they were presented randomly along the task. All tone pairs were one-octave apart so as to keep the perceptual difference between them similar. Participants were asked to report the order in which they heard the tones (either high-low or low–high), by pressing keyboard keys marked as Low and High, while the silent gap separating them was manipulated. Since stimulus duration was constant, this silent gap was defined as an ISI, the time from the offset of the first stimuli to the onset of the second stimuli. Nine ISI values of 2, 5, 10, 15, 30, 60, 90, 120, and 240 ms were randomly used. Each ISI value was repeated 16 times, producing 288 trials (9 ISIs × 2 presentation orders × 16 repetitions). The Roving Frequencies condition of Experiment 2 included 864 trials (9 ISIs × 6 presentation orders × 16 repetitions). After every 32 trials, participants received a short recess. In Experiments 1 and 2, the tones were presented diotically, i.e., stimuli were presented to both ears synchronously. In Experiment 3, the tones were presented diotically (as in Experiments 1 and 2), monaurally, and dichotically. In the monaural condition, stimuli were presented to the right ear only. In the dichotic condition, stimuli were presented to both ears asynchronously, one after the other.

The experiments were preceded by a training session with tones of the same duration as in the experiment. The first part of the training was designed to present participants with the low and high sounds. The second part associated the sounds with the appropriate keyboard keys (high or low). The third part tested whether they correctly reported the pitch presented with the sound (with a criterion of at least 20 correct answers out of 24 trials). All participants passed this test. The fourth part of the training presented participants with the tone pairs. The tone pairs in this part of the training had the longest ISI (240 ms) and participants were asked to report the order of the two tones presented. Participants received feedback for their responses during the second and fourth parts of the training, but no feedback was presented during the experiment.

Apparatus

The hearing screening test was performed using a Danplex DA64 audiometer. The experiment was performed on a Dell laptop computer, and the sounds were delivered through TDH-49 headphones.

Procedure

The study was approved by the Ariel University Institutional Review Board and was carried out according to Good Clinical Practice guidelines[41]. Participants provided written informed consent at the beginning of the experiment, prior to the hearing screening. In addition, participants' absolute threshold for the stimuli used in each of the experiments was measured prior to the experimental tasks, in order to calibrate the hearing level in the experiment to 40 dB SL. The experiments, including the screening and training procedures, took 45 – 60 min.

Data availability

The datasets analyzed during the current study are available in the Kaggle repository: https://www.kaggle.com/datasets/leahfostick/spectral-toj-for-177-participants/data.

Received: 16 May 2024; Accepted: 19 December 2024
Published online: 07 January 2025

References

1. Plack, C. J. (2005). Auditory perception. *Handbook of cognition*, 7198.
2. Hirsh, I. J. & Sherrick, C. E. Jr. Perceived order in different sense modalities. *J. Exp. Psychol.* **62**(5), 423 (1961).
3. Reed, M. A. Speech perception and the discrimination of brief auditory cues in reading disabled children. *J. Exp. Child Psychol.* **48**(2), 270–292 (1989).
4. Tallal, P. Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang.* **9**(2), 182–198 (1980).
5. Ben-Artzi, E., Fostick, L. & Babkoff, H. Deficits in temporal-order judgments in dyslexia: Evidence from diotic stimuli differing spectrally and from dichotic stimuli differing only by perceived location. *Neuropsychologia* **43**(5), 714–723 (2005).
6. Fostick, L., Bar-El, S. & Ram-Tsur, R. Auditory temporal processing as a specific deficit among dyslexic readers. *Online Submission* **2**(2), 77–88 (2012).
7. Fostick, L., Bar-El, S. & Ram-Tsur, R. Auditory temporal processing and working memory: Two independent deficits for dyslexia. *Psychol. Res.* **2**(5), 308–318 (2012).
8. Fostick, L., Eshcoly, R., Shtibelman, H., Nehemia, R. & Levi, H. Efficacy of temporal processing training to improve phonological awareness among dyslexic and normal reading students. *J. Exp. Psychol.: Hum. Percept. Performance* **40**(5), 1799 (2014).
9. Fostick, L. & Revah, H. Dyslexia as a multi-deficit disorder: Working memory and auditory temporal processing. *Acta Psychol.* **183**, 19–28 (2018).
10. Szymaszek, A., Szelag, E. & Sliwowska, M. Auditory perception of temporal order in humans: The effect of age, gender, listener practice and stimulus presentation mode. *Neurosci. Lett.* **403**(1–2), 190–194 (2006).
11. Szymaszek, A., Sereda, M., Pöppel, E. & Szelag, E. Individual differences in the perception of temporal order: The effect of age and cognition. *Cognit. Neuropsychol.* **26**(2), 135–147 (2009).
12. Ben-Artzi, E., Babkoff, H. & Fostick, L. Auditory temporal processes in the elderly. *Audiol. Res.* **1**(1), e6 (2011).
13. Fostick, L. & Babkoff, H. Temporal and non-temporal processes in the elderly. *J. Basic Clin. Physiol. Pharmacol.* **24**(3), 191–199 (2013).
14. Babkoff, H., Zukerman, G. I. L., Fostick, L. & Ben-Artzi, E. L. I. S. H. E. V. A. Effect of the diurnal rhythm and 24 h of sleep deprivation on dichotic temporal order judgment. *J. Sleep Res.* **14**(1), 7–15 (2005).
15. Fostick, L., Babkoff, H. & Zukerman, G. Effect of 24 hours of sleep deprivation on auditory and linguistic perception: A comparison among young controls, sleep-deprived participants, dyslexic readers, and aging adults. *J. Speech Lang. Hear. Res.* **57**(3), 1078–1088 (2014).
16. Fostick, L. The effect of attention-deficit/hyperactivity disorder and methylphenidate treatment on the adult auditory temporal order judgment threshold. *J. Speech Lang. Hear. Res.* **60**(7), 2124–2128 (2017).
17. Fink, M., Churan, J. & Wittmann, M. Temporal processing and context dependency of phoneme discrimination in patients with aphasia. *Brain Lang.* **98**(1), 1–11 (2006).
18. von Steinbüchel, N., Wittmann, M., Strasburger, H. & Szelag, E. Auditory temporal-order judgement is impaired in patients with cortical lesions in posterior regions of the left hemisphere. *Neurosci. Lett.* **264**(1–3), 168–171 (1999).
19. Warren, R. M. Auditory pattern recognition by untrained listeners. *Percept. Psychophys.* **15**(3), 495–500 (1974).
20. Warren, R. M. & Byrnes, D. L. Temporal discrimination of recycled tonal sequences: Pattern matching and naming of order by untrained listeners. *Percept. Psychophys.* **18**(4), 273–280 (1975).
21. Warren, B. M. & Ackroff, J. M. Two types of auditory sequence perception. *Percept. Psychophys.* **20**(5), 387–394 (1976).
22. Warren, R. M. Perception of acoustic sequences: Global integration versus temporal resolution. In *Thinking in sound: The cognitive psychology of human audition* (eds McAdams, S. & Bigand, E.) 37–68 (Clarendon Press/Oxford University Press, 1993).
23. Warren, R. M. & Bashford, J. A. When acoustic sequences are not perceptual sequences: The global perception of auditory patterns. *Percept. Psychophys.* **54**(1), 121–126 (1993).
24. Fostick, L. & Babkoff, H. Different response patterns between auditory spectral and spatial temporal order judgment (TOJ). *Exp. Psychol.* **60**(6), 432–443 (2013).
25. Bao, Y. et al. Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. *Cognition* **129**(3), 579–585 (2013).
26. Bao, Y. et al. Auditory perception of temporal order: A comparison between tonal language speakers with and without non-tonal language experience. *Acta Neurobiol. Exp* **74**, 98–103 (2014).
27. Zhang, Z., Lin, X. & Bao, Y. Holistic temporal order judgment of tones requires top-down disentanglement. *PsyCh J.* **12**(4), 491–499 (2023).
28. Fostick, L. & Babkoff, H. Auditory spectral versus spatial temporal order judgment: Threshold distribution analysis. *J. Exp. Psychol. Hum. Percept. Perform.* **43**(5), 1002–1012 (2017).
29. Semal, C. & Demany, L. Individual differences in the sensitivity to pitch direction. *J. Acoust. Soc. Am.* **120**(6), 3907–3915 (2006).
30. Mathias, S. R., Micheyl, C. & Bailey, P. J. Stimulus uncertainty and insensitivity to pitch-change direction. *J. Acoust. Soc. Am.* **127**(5), 3026–3037 (2010).
31. Fink, M., Churan, J. & Wittmann, M. Assessment of auditory temporal-order thresholds: A comparison of different measurement procedures and the influences of age and gender. *Restorative Neurol. Neurosci.* **23**(5–6), 281–296 (2005).
32. Shechtman, O. (2013). The coefficient of variation as an index of measurement reliability. In *Methods of clinical epidemiology* (pp. 39–49). Berlin, Heidelberg: Springer.
33. Celebi, M. E., Kingravi, H. A. & Vela, P. A. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst. Appl.* **40**(1), 200–210 (2013).
34. Fostick, L., Wechsler, S. & Peretz, E. Short-term learning effect in different psychoacoustic measures. *J. Basic Clin. Physiol. Pharmacol.* **25**(3), 307–312 (2014).
35. Fostick, L. & Babkoff, H. The effects of psychophysical methods on spectral and spatial TOJ thresholds. *Sensors* **22**(13), 4830 (2022).
36. Moore, B. C. J. *Auditory Processing of Temporal Fine Structure: Effects of Age and Hearing Loss* (World Scientific, 2014).
37. Demany, L., Lucas, T. & Semal, C. Pitch priming in sequences of two sounds. *J. Acoust. Soc. Am.* **140**(3), 2056–2063 (2016).
38. Fostick, L. Card playing enhances speech perception among aging adults: Comparison with aging musicians. *Eur. J. Ageing* **16**(4), 481–489 (2019).
39. Fink, M., Ulbrich, P., Churan, J. & Wittmann, M. Stimulus-dependent processing of temporal order. *Behav. Process.* **71**(2–3), 344–352 (2006).
40. Gelfand, S. A. (2017). *Hearing: An introduction to psychological and physiological acoustics*. CRC Press.
41. Guidelines, I. C. H. Integrated addendum to ICH E6 (R1): guideline for good clinical practice E6 (R2). *Current Step* **2**, 1–60 (2015).

Author contributions

LF supervised the data collection; LF and MZ analyzed the data; All authors wrote and reviewed the main manuscript text.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.F.