Discrimination thresholds for interaural-time and interaural-level differences in naïve listeners: Sex differences and learning

Beverly A. Wright [a,b]

[a]Department of Communication Sciences and Disorders, Knowles Hearing Center,
Northwestern Institute for Neuroscience,
2240 Campus Drive,
Northwestern University, Evanston, IL 60208
b-wright@northwestern.edu

Huanping Dai [c]

[c]Department of Speech, Language, and Hearing Sciences,
1131 E. 2nd Street
University of Arizona, Tucson, AZ 85721
huanpingdai@gmail.com

[b] Corresponding author

ABSTRACT

The two primary cues to sound-source location on the horizontal plane are interaural time differences (ITDs) and interaural level differences (ILDs). Here we asked whether the ability to discriminate small changes in each of these interaural cues differs between the sexes. We tested one group of males (n = 43) and females (n = 94) on ITD discrimination at 0.5 kHz and a separate group of males (n = 80) and females (n = 166) on ILD discrimination at 4 kHz. None of the participants had any prior experience with psychoacoustic tasks. Testing of each participant was completed in a single testing session of 4-5 blocks of 60 trials. For ILD discrimination, the overall mean threshold, as well as the mean threshold for each block, was statistically significantly lower for males than for females. Despite that, males and females learned at an equal rate over the course of testing. For ITD discrimination, in contrast, thresholds did not differ significantly between the sexes for the overall mean or for any block. There also was no statistically significant learning across blocks for either sex. For both tasks and both sexes, the individual thresholds spanned a wide range. The presence of a statistically significant sex difference and learning for ILD but not for ITD discrimination, along with a larger effect size for ILD than for ITD discrimination, suggests that the factors responsible for these outcomes acted upon an ILD-specific neural pathway, and not upon an ITD-specific pathway, nor any pathway common to the two cues. Because the ILD and ITD specific pathways are most separable initially, the factors associated with sex and learning may have acted upon the ILD-specific pathway at an early stage.

Key words: sound localization; gender difference; within-session learning; ILD and ITD pathways; sound lateralization; human psychophysics; male advantage

HIGHLIGHTS

- Interaural differences in time (ITDs) and level (ILDs) are sound-localization cues.
- There is a sex difference for ILD, but not for ITD, discrimination.
- The sex difference for ILD discrimination favors males.
- There is within-session learning for ILD, but not for ITD, discrimination.
- The learning on ILD discrimination does not differ between the sexes.
- The factors responsible for these results likely act on an ILD-specific pathway.

1. INTRODUCTION

Processing information pertaining to the position and orientation of objects in space is a critical sensory and cognitive function. One factor that appears to influence spatial-processing ability in the visual domain is the sex of the observer. There are numerous reports that, on average, males perform statistically significantly better than females on visual spatial tasks including mental-rotation tasks, spatial-perception tasks, and, in some cases, spatial visualization tasks (for meta analyses, see: Linn and Petersen, 1985; Voyer, Voyer, and Bryden, 1995; Maeda and Yoon, 2013; Lauer, Yhang, and Lourenco, 2019). There are also a few initial indications that the male advantage in spatial abilities extends to the auditory domain for tasks that involve multiple sound sources, non-dominant sound-localization cues, or moving sounds. For example, on average, males performed statistically significantly better than females on: (1) horizontal localization of a single sound source either amidst a multi-source, cocktail-party-like, environment (Zündorf et al., 2011) or in the presence of a single distractor sound source at a different location (Lewald and Hausmann, 2013); (2) vertical localization of a sound source monaurally (in the right ear) (Lewald, 2004); and (3) the judgment of the terminal distance, the time of arrival, and the duration of looming sound sources (Schiff and Oldak, 1990; Neuhoff et al., 2009; Grassi, 2010). However, it is unclear to what extent sex differences are also present at the fundamental level of the initial processing of the individual cues used for sound-source localization. Here we begin to address this question by focusing on possible sex differences in the processing of the two primary cues for sound-source localization on the horizontal plane: interaural time differences (ITDs) and interaural level differences (ILDs). ITDs and ILDs arise because a sound arrives earlier and often at a higher sound level at the ear that is closer to the sound source. These interaural differences vary systematically in magnitude as the horizontal position of the sound source changes. ITDs are the predominant sound-localization cue at frequencies below ~1600 Hz, while ILDs are the predominant cue at higher frequencies (Strutt, 1907; for an overview, see Middlebrooks, 2015). We selected ITDs and ILDs for examination because localization on the horizontal plane is the primary function of spatial hearing in humans, and because there is an extensive literature on the processing of these cues.

While both ITDs and ILDs contribute to the perception of sound-source location, the role that each cue plays individually is difficult to discern when sounds are in the free field, as in everyday life. However, in the laboratory, the sensitivity to each cue can be assessed separately using stimuli presented over headphones. In that case, the two cues can be manipulated independently, allowing the magnitude of one cue to vary while the magnitude of the other cue is fixed. Under those circumstances, a sound image is perceived within the head, and its lateral position depends on the magnitude of the manipulated interaural difference. The smallest discernable change in the magnitude of the interaural difference in time (ITD) and in level (ILD) can then be evaluated individually. Hence, it is possible to assess sex differences for ITD discrimination and ILD discrimination separately.

We are aware of only two previous reports of sex differences in ITD and ILD discrimination. First, in a meeting abstract, Langford (1994; see McFadden, 1998, Fig.

1, for data and additional description) noted that, on average, performance was better for males than for females for both ITD discrimination and ILD discrimination with bandpass noise. The mean ITD-discrimination threshold with a 0.6-0.8 kHz noise band was 86 μs for males, but was 113 μs for females. Similarly, the mean ILD-discrimination threshold with a 2-4 kHz noise band was 3.1 dB for males, but was 4.0 dB for females. However, while these results are indicative of sex differences, the sample sizes were relatively small (~25 per group), and it is not clear from either the meeting abstract or later description whether the sex differences were statistically significant.

Second, Saberi et al. (2004) reported that, on average, performance was better for males than for females for interaural-difference discrimination with click stimuli. In that investigation, the data were pooled across ITD- and ILD-discrimination tasks by converting individual discrimination thresholds to z scores. The mean z scores were approximately one-quarter of a unit lower (better) for males than for females, both for isolated single clicks and for the second of two clicks presented in rapid succession. In each case, the sample size was large (n= 77 M and 225 F) and the sex difference was statistically significant ($p < 0.05$). However, because the data were pooled across the two interaural cues it is not possible to determine from the results provided whether those sex differences arose from differences in sensitivity between males and females for ITDs only, ILDs only, or for both cues.

In sum, though previous reports point toward possible sex differences--favoring males--in the ability to discriminate small changes in interaural cues, the data in those reports were either from a relatively small number of participants, or were not confirmed with statistical tests, or were not analyzed separately by cue. Therefore, we measured ITD- and ILD-discrimination thresholds in large groups of young adults, and, for each discrimination task (ITD or ILD), we compared thresholds between males and females. We examined ITD discrimination at 0.5 kHz (overall n=137) and ILD discrimination at 4 kHz (overall n=246), thereby testing each cue within its own natural, dominant, frequency region of operation. None of the participants had any previous experience with psychoacoustic tasks, and we tested the two tasks in entirely different groups. That is, the participants who performed one task, did not perform the other. Thus, for each task, the reported performance reflects naïve performance. Participants completed 4-5 blocks of testing trials so we evaluated sex differences based on both the mean threshold across blocks and the threshold for each individual block. We also evaluated whether the thresholds improved across the individual blocks and whether any such learning differed between the sexes. We report a statistically significant sex difference, favoring males, for ILD discrimination, but no statistically significant sex difference for ITD discrimination. We also report statistically significant learning for both sexes for ILD discrimination, but no statistically significant learning for either sex for ITD discrimination.

2. METHODS

2.1. Participants

A total of 383 young adults participated. Of those, 246 participants were tested on ILD discrimination (166 females and 80 males; mean age: 21.2 years, sd = 3.6, range: 18 to 43 years, median 20 years) and 137 participants were tested on ITD discrimination (94 females and 43 males; mean age: 21.3 years, sd = 4.1, range: 18 to 39 years, median 20 years). The greater number of female than male participants simply reflects the sex distribution of the population that responded to the recruitment fliers posted on the campus. The unequal numbers of males and females were taken into account in all of the analyses. All participants self-reported normal hearing. None of the participants had any previous experience on psychoacoustic tasks. All participants were paid for their participation. The procedures were approved by the Northwestern University Office for the Protection of Research Subjects.

2.2. Tasks and Stimuli

The tasks were interaural-time-difference (ITD) discrimination and interaural-level-difference (ILD) discrimination. For both tasks, on each trial, a standard stimulus and a comparison stimulus were presented. Each stimulus consisted of two pure tones of identical frequency presented, one to each ear, over headphones (Sennheiser HD265). The resulting percept was of a sound image within the head. The lateral position of the image was determined by the relative levels and phases of the two tones. Both tones were 300 ms in duration including 10-ms rise/fall cosine ramps, started and ended simultaneously, and had a nominal stimulus level of 70 dB SPL. The phase of the tone presented to the right ear was randomized across trials.

For the ITD-discrimination task, the two tones had a frequency of 0.5 kHz and were always presented at 70 dB SPL to the two ears (an ILD of 0 dB). For the standard stimulus, there was no phase difference between the tones presented to the two ears (an ITD of 0 µs), yielding a sound image on the midline. For the comparison stimulus, there was a phase difference between the tones presented to the two ears (comparison ITD; an ITD that was > 0 µs); the phase of the tone to the left ear was delayed relative to the phase of the tone to the right ear by the comparison ITD, yielding a sound image that was closer to the right ear. The ITD was in the ongoing phase. The envelope ITD was fixed at 0 µs.

For the ILD-discrimination task, the two tones had a frequency of 4 kHz and were always presented in phase to the two ears (an ITD of 0 µs). For the standard stimulus, there was no level difference between the tones presented to the two ears (an ILD of 0 dB); the two tones were presented at 70 dB SPL to both ears, yielding a sound image on the midline, midway between the two ears. For the comparison stimulus, there was a level difference between the tones presented to the two ears (comparison ILD; an ILD that was > 0 dB); the tone to the left ear was presented at 70 dB SPL minus 0.5 times the comparison ILD and the tone to the right ear was presented at 70 dB SPL plus 0.5 times the comparison ILD (an ILD that was > 0 dB), yielding a sound image that was closer to the right ear. We divided the ILD magnitude between the two ears to reduce the size of the level change within a single ear with the aim of discouraging the use of a monaural absolute-level cue. We chose not to use an alternative strategy of roving the overall stimulus level across observation intervals for three reasons. First, level roving can be distracting, particularly for naïve participants like here, which can increase the difficulty of the assigned task, and thereby potentially have a large unwanted impact on the outcome. Second, roving is intended to prevent participants from using an irrelevant cue, and the assumption in using roving is that participants actually ignore the irrelevant (roved) cue. However, when participants are having trouble solving the assigned task, they sometimes systematically base their decisions on the irrelevant (roved) cue (Wright and McFadden, 1990; Wright and Saberi, 1999). Third, level roving is not used consistently in the literature, and was not used in the previous published report of sex differences in ILD/ITD discrimination (Saberi et al., 2004; there is not enough information in the Langford, 1994, abstract to determine whether level roving was used). Therefore, not using level roving aids comparisons across investigations.

Both tasks used a two-alternative forced-choice paradigm. On each trial, the standard stimulus and the comparison stimulus were presented in random order, with an inter-stimulus interval of 650 ms. Participants were asked to select which stimulus was farther to the right (the comparison stimulus) by pressing a key on a computer keyboard. Visual feedback ("Correct!!" or "Wrong") was provided after each response. Before testing, participants were given examples of the standard stimulus and of a clearly discriminable comparison stimulus. All participants were tested in a sound-attenuated booth.

2.3. Threshold Estimation

For both the ITD- and ILD-discrimination tasks, the discrimination threshold was estimated using a 3-down/1-up adaptive tracking procedure. In each block of 60 trials, the comparison ITD or ILD value decreased after every three consecutive correct responses, and increased after each incorrect response. When the comparison value switched from increasing to decreasing or vice versa, the value at which the switch occurred was labeled a reversal. The first three or four reversals were discarded. The largest remaining even number of reversal values, denoted the ‘usable’ reversals, were then averaged to obtain an estimate of the threshold, which corresponded to a percent correct of PC=79.4% on the psychometric function (Levitt 1971). Any block with fewer than seven total reversals was excluded from subsequent analyses. For ITD discrimination, the starting value of the comparison ITD was 1 µs, and the step sizes were multiplications/divisions by $10^{0.2}$ until the third reversal and by $10^{0.05}$ for the remaining reversals (Saberi, 1995). The minimum allowed value was 1 µs. The maximum allowed value was 650 µs, based on the typical largest naturally occurring time delay between the two ears in humans (e.g., Feddersen, Sandel, Teas and Jeffress, 1957). For ILD discrimination, the starting value of the comparison ILD was 6 dB and the step sizes were 0.5 dB until the third reversal and 0.25 dB for the remaining reversals. The minimum allowed value was 0 dB. There was no constraint on the maximum value. The starting values and step sizes were chosen to be consistent with previous investigations (Zhang and Wright, 2007; Zhang and Wright 2009; Ortiz and Wright, 2010). A previous analysis of adaptive tracks for ITD and ILD discrimination obtained with these parameters revealed that, for both tasks, the final threshold estimates were based on approximately equal numbers of usable reversals, obtained over approximately equal numbers of trials, and that the usable reversal values fell within a small range of the final threshold estimates. (Zhang and Wright, 2007). Thus, it appears that the different initial trajectories of the tracks for the two tasks did not strongly influence the reliability of threshold estimation.

The vast majority of participants completed 4 or 5 blocks of trials (ILD: n=78 M, 160 F; ITD: n=40 M, 89 F). A few participants completed only 2 or 3 blocks of trials (ILD: n=2 M, 6 F; ITD: n=3 M, 5 F). The data of all of the participants were included in the analyses of the grand mean and individual thresholds. Only the data of participants who completed 4 or 5 blocks of trials were included in the analyses of learning.

3. RESULTS

On average, males performed better than females for both ILD and ITD discrimination, but the sex difference was statistically significant only for ILD discrimination. For ILD discrimination, the mean threshold was statistically significantly lower (better) for males (n=80) at 4.6 dB (SE: 0.2 dB) than for females (n=166) at 5.3 dB (SE: 0.2 dB) (t-test for independent groups: T(149.92)=2.24; p= 0.0267; Cohen's d=0.36; mean difference=0.68 dB, 95% confidence interval: 0.08 dB to 1.23 dB) (Fig. 1A). The raw median, first-quartile, and third-quartile threshold values were also lower for males than for females; furthermore, all of the threshold values that were outliers (plus signs), reflecting the worst performance, were from females, and were higher than the worst performance from males (Fig. 1B; box plots). While there was a statistically significant male advantage for ILD discrimination, the individual thresholds spanned a wide, essentially overlapping, range in both groups: from 1.4 to 10.2 dB in males and from 1.6 to 11.8 dB in females (Fig. 1B; circles). The mean threshold over all individuals regardless of sex (n=246) was 5.0 dB (SE: 0.1 dB). At a more detailed, block-by-block, level, a 2-sex (n=78 M, 160 F, participants who completed at least 4 blocks, see Methods) by 4-block (blocks 1-4) ANOVA, with repeated measures on block, revealed a statistically significant main effect of sex ($F(1,236)=7.26$, $p=0.0076$, $\eta_p^2=0.0298$) and a main effect of block ($F(3,708)=25.26$, $p=0.0000$, $\eta_p^2=0.0967$), but no interaction between sex and block ($F(3,708)=0.55$, $p=0.6473$, $\eta_p^2=0.0023$) (Fig. 1C). Thus, based on the current statistical analyses, males outperformed females on average, and both sexes learned at similar rates across blocks, so the magnitude of the male advantage remained unchanged across blocks.

For ITD discrimination, while males performed better than females on average, unlike for ILD discrimination, the sex difference was not statistically significant. All analyses of the ITD-discrimination thresholds were conducted on ITD thresholds in logarithmic units [$10 \log_{10}$(ITD in µs), where ITD is the threshold value in µs]. Where feasible in reporting, the logarithmic values are converted back to the more familiar units of raw µs. The mean ITD-discrimination threshold was not statistically significantly different between males (n=43) at 64.3 µs (SE: 4.2 µs) and females (n=94) at 69.8 µs (SE: 4.5 µs) (t-test for independent groups: $T(80.68)=0.7331$, $p= 0.4656$; Cohen's $d=0.16$; mean difference=0.36 in $10 \log_{10}$(ITD in µs), 95% confidence interval: -0.61 to 1.32 in $10 \log_{10}$(ITD in µs))(Fig. 2A). Though the mean thresholds did not differ significantly, all three of the raw quartile values were numerically lower for males than for females (Fig. 2B; box plots). Again, individual thresholds varied widely in both groups, ranging from 18.8 to 251.2 µs in males and from 13.3 to 304.1 µs in females (Fig. 2B; circles). The mean threshold over all individuals regardless of sex (n=137) was

68 μs (SE: 3.6 μs). At the block-by-block level, a 2-sex (n=40 M, 89 F, participants who completed at least 4 blocks, see Methods) by 4-block repeated-measures ANOVA revealed no statistically significant main effect of sex ($F(1,127)=0.11$, $p=0.7356$, $\eta_p^2=0.0009$), no main effect of block ($F(3,381)=0.67$, $p=0.5714$, $\eta_p^2=0.0052$), and no interaction between sex and block ($F(3,381)=1.41$, $p=0.2402$, $\eta_p^2=0.0110$) (Fig. 2C). Thus, based on the current analyses, neither the mean difference between the sexes, nor the learning across blocks for either sex, was statistically significant.

4. DISCUSSION

The primary outcome of the current study was that there was an overall statistically significant sex difference, favoring males, for ILD discrimination but no statistically significant sex difference for ITD discrimination. In addition, there was statistically significant learning for ILD discrimination but no statistically significant learning for ITD discrimination, and the learning on ILD discrimination did not differ significantly between the sexes. Here we first discuss performance variability in these results. We then compare these results to previous reports of sex differences, first in ILD and ITD discrimination specifically, and then in auditory and visual spatial abilities in general. Finally, we consider potential implications of the distinctions in sex differences and learning between ILD and ITD discrimination.

4.1. Performance variability

While the present results document a statistically significant sex difference for ILD discrimination, and not for ITD discrimination, it is important to consider these results in relation to the overall spread of the individual thresholds and the confidence interval for the sex difference for each task. For each task, the thresholds were spread across wide, essentially overlapping, ranges for both sexes: spanning approximately 1 to 12 dB for ILD discrimination (Fig. 1B) and approximately 13 to 300 μs for ITD discrimination (Figs. 2B). Thus, for each task, any actual mean threshold difference between the sexes was small relative to the range and variability within each sex. This large variability and the large overlap should be kept in mind when interpreting the mean sex differences on these tasks in the current investigation. For example, for ILD discrimination, although the sex difference was statistically significant, one cannot rule out the possibility that the actual effect could be negligible. After all, the lower bound of the 95% confidence interval for the sex difference was only 0.08 dB, which would be negligible. Similarly, for ITD discrimination, although the sex difference was not statistically significant, one cannot rule out the possibility that the actual effect could be nontrivial. The upper bound of the 95% confidence interval was 1.32 in 10 $\log_{10}$(ITD in μs). That translates to ITD thresholds for females being 1.35 times those for males, which would be a nontrivial sex difference favoring males. At the same time, the lower bound of the confidence interval was -0.61 in 10 $\log_{10}$(ITD in μs). That translates to ITD thresholds for females being 0.87 times those for males, which actually would be a sex difference favoring females, rather than males. Nevertheless, for ITD discrimination, the confidence interval for the sex difference clearly straddled zero, leading to no statistically significant sex difference, and the effect size (Cohen's d = 0.16) was minimal. For ILD discrimination, in contrast, the confidence interval did not straddle zero, leading to a statistically significant sex difference, and the effect size (Cohen's d =

0.36) was between small (Cohen’s d = 0.2) and medium (Cohen’s d = 0.5) (Cohen, 1988). Moreover, the effect size for ILD discrimination was more than twice that for ITD discrimination. Taken together, there appears to be sufficient evidence in the current data of a measurable mean sex difference for ILD discrimination, but not for ITD discrimination.

4.2. Comparison to previous reports of sex differences in ILD and ITD discrimination

For ILD discrimination, the sex difference observed here appears to be consistent with that previously reported in the meeting abstract by Langford (1994; also see McFadden, 1998, Fig. 1). In both cases, the sex difference had the same valence (M>F) and was similar in magnitude (0.7 here vs. 0.9 dB for Langford). The similarities in the outcomes between the two reports are noteworthy given the many differences between the two investigations, such as in the number of participants (80 M and 166 F here vs. ~25 per group for Langford), the stimulus parameters (300-ms, 4-kHz tone here vs. 60-ms, 2-4 kHz narrowband noise for Langford), and the measurement procedures (two-interval forced-choice here vs. three-interval oddity for Langford). Thus, the sex difference in ILD discrimination seems to be robust.

For ITD discrimination, the current data appears to differ from Langford’s (1994; also see McFadden, 1998, Fig. 1) report of a male advantage on that task. While the mean ITD-discrimination threshold was lower in males than females in both reports, the magnitude of the difference was larger for Langford (27 μs) than here (5.5 μs), suggesting that there is a discrepancy between the outcomes. The apparent discrepancy between the two outcomes could be attributable to any of the variety of differences between the two investigations. For example, while we used a 0.5-kHz pure tone, Langford used a 0.6-0.8 kHz bandpass noise. Therefore, Langford's participants could have used a cue based on the interaural delay of the amplitude envelope rather than a cue based on stimulus phase, a distinction that could lead to a sex difference if males also happen to be better at encoding amplitude envelopes than females. Likewise, while our participants were truly naïve to psychoacoustic tasks, it is uncertain whether Langford's participants were. At a minimum, Langford's participants also completed testing on ILD discrimination and their other experience with psychoacoustic tasks is unspecified. Therefore, Langford's participants could have had an opportunity to learn on ITD discrimination between sessions (Ortiz and Wright, 2010) (if they had previous experience with ITD discrimination) or to generalize learning from ILD discrimination to ITD discrimination (Ortiz and Wright, 2010) (if they had previous experience with ILD discrimination). Those potential opportunities could lead to a statistically significant sex difference if males also happen to learn or generalize at a faster rate than females between sessions (even though the learning rate does not appear to differ between the sexes within sessions (Fig. 2C). The discrepancy could also come from other differences, such as in the stimulus duration (300-ms here vs. 60-ms for Langford) or in the measurement procedures (two-interval forced-choice here vs. three-interval oddity for Langford).

Finally, the present results are potentially consistent with the report of Saberi et al. (2004) of a male advantage in interaural discrimination for click stimuli. The two reports are consistent in the sense that, when we combined all of the data across the ILD- and ITD-discrimination tasks by converting all of the threshold values to z scores, as Saberi et al. did, there was a statistically significant male advantage (T(232.68)=2.24, p=0.0260; n=123 M, 260 F; Cohens d=0.27; mean difference=0.25 in z scores; 95% confidence interval: 0.030 to 0.46 in z scores), as Saberi et al. reported. However, because Saberi et al. only reported the results for the combined data, it is not possible to determine whether the sex difference they observed was carried by the ILD cue, like here, or instead was carried by the ITD cue, by both cues equally, or by neither cue alone.

One factor that we did not monitor in the present investigation is the potential influence of the menstrual cycle on the discrimination performance of the females. There are several indications that performance on tasks involving interaural timing cues fluctuates over the course of the menstrual cycle. For example, in one report, sensitivity to binaural beats, a phenomenon associated with the tracking of interaural phase differences, was better overall in males (n = 20) than in females (n = 20), but it fluctuated over the course of the menstrual cycle in females, being worst during preovulation and best (reaching male performance) during premenstruation (Tobias, 1965). The fluctuation was initially noticed in a cross-sectional evaluation of the 20 female participants and later evaluated longitudinally in 3 additional female participants. Similarly, in another report, ITD discrimination for pulse trains fluctuated over the course of the menstrual cycle being worst during the midluteal phase and best during menstruation (n = 9 females, tested longitudinally) (Haggard and Gaston, 1978). In that report, no sex difference was documented, because only females were tested, but the authors appear to have assumed that when the performance of females was best it was most male-like. Thus, it is possible that the sensitivity to ITDs, and, by extension, the sensitivity to ILDs, in the present group of females fluctuated during the cycle. If so, that influence is obscured here because we did not track the cycle in these females. We note however that regardless of how much performance fluctuated with the menstrual cycle, there was no overall sex difference on ITD discrimination, and there was an overall sex difference on ILD discrimination.

4.3. Comparison to previous reports of sex differences in auditory and visual spatial abilities

The male advantage in the sex difference observed here is in the same direction as the sex differences related to spatial abilities that have been reported previously for audition and vision. As mentioned in the introduction, for audition, a male advantage has been reported for horizontal localization of a sound source in the presence of distractor sound sources at different locations (Zündorf et al., 2011; Lewald and Hausmann, 2013), monaural localization on the vertical plane (Lewald, 2004), the perception of looming sound sources (Schiff and Oldak, 1990; Neuhoff et al., 2009; Grassi, 2010), and the lateralization of click stimuli (Saberi et al., 2004). The same direction of sex difference, favoring males, is also evident for both ILD and ITD discrimination in reports without statistical analyses (Langford, 1994; also see McFadden, 1998, Fig. 1), as well as in the current data for ITD discrimination, for which the sex difference was not statistically significant, but for which the raw mean, and all quartile values (median, 25th and 75th percentiles) were numerically lower (better) for males than females (see Fig. 2B). Thus, compared with females, males appear to have generally superior spatial hearing. This male advantage is also in the same direction as the sex differences for spatial abilities in vision (as also mentioned in the Introduction: Linn and Petersen, 1985; Voyer, Voyer, and Bryden, 1995; Maeda and Yoon, 2013; Lauer, Yhang, and Lourenco, 2019), supporting the idea that the sex difference in spatial abilities may be more global, rather than modality specific.

Males have also been reported to outperform females, on average, on other non-spatial auditory tasks, including, for example, detection of a tone of fixed frequency presented in a multi-tone masker with randomized frequency components (Neff, Kessler, and Dethlefs, 1996; McFadden, Pasanen, Maloney, Leshikar, Pho, 2018), frequency discrimination (Rammsayer and Troche, 2012), loudness discrimination (Rammsayer and Troche, 2012), and--during adolescence only--forward masking and temporal-interval discrimination (Huyck and Wright, 2018) (for review, see McFadden, 1998). However, this male advantage is not universal across sensory abilities. For example, even in audition, it does not extend to absolute threshold, for which females on average outperform males (e.g., Flamme, Deiters, and Needham, 2011). The male advantage also does not extend uniformly to other sensory modalities. For example, based on objective measures of task performance, females are more sensitive than males to tastes (Fikentscher, Roseburg, Spinar, and Bruchmüller, 1977; Yoshinaka, Ikebe, Uota, Ogawa, Okada, Inomata, Takeshita, Mihara, Gondo, Masui, Kamide, Arai, Takahashi, and Maeda, 2016) and particularly to odors (for reviews see Doty and

Cameron, 2009; Sorokowski, et al., 2019). Interestingly, females also outperform males on the ability to remember the location of objects in relation to one another (e.g., Silverman, Choi, and Peters, 2007). This female advantage for object location memory, along with the male advantage for spatial abilities, is consistent with the proposal that evolutionary pressure favors females in tasks closely related to gathering, and males in tasks closely related to hunting (e.g., Silverman et al. 2007).

As for the effect sizes for the sex differences in spatial abilities, three points seem noteworthy. First, the effect size for the sex difference for ILD and ITD discrimination combined is consistent across the current investigation and that of Saberi et al. (2004). As noted above, when the current data were analyzed the same way as in Saberi et al., Cohen's d=0.27. That effect size is similar to Saberi et al.'s description that, overall, the thresholds for males were better than for females by a quarter of one standard deviation. Thus, the effect size for these discrimination tasks appears to be large enough to be replicable. Second, when evaluated separately between ILD and ITD discrimination, the magnitude of the effect size for ILD discrimination can be at least as large as Cohen's d=0.36. Third, the effect size for discrimination is only about half the magnitude of the effect sizes previously reported for more complex spatial-hearing tasks including sound localization in the presence of distractors (Cohen's d=0.61, Zündorf et al., 2011) and the perception of looming sound sources (d = 0.76 and 0.78, Neuhoff et al., 2009). Thus, for these spatial-hearing tasks, it appears that the effect size for sex differences differs across different classes of task, and may depend on complexity, such that tasks with greater complexity may yield larger effects. A taskdependency of the effect size for sex differences is well established for visual spatial tasks. For example, based on meta-analyses, effect sizes in adults range from 0.56 to 0.73 for mental rotation, from 0.48 to 0.64 for spatial perception, and from 0.13 to 0.23 for spatial visualization (Linn and Petersen, 1985; Voyer, Voyer, and Bryden, 1995; Maeda and Yoon, 2013). These effect sizes span approximately the same range as for spatial-hearing tasks.

4.4. Behavioral distinctions between ILD and ITD discrimination: sex difference and learning

Perhaps the most striking aspect of the present results is that they revealed two behavioral distinctions between ILD and ITD discrimination: one in sex differences and the other in learning. This outcome is all the more remarkable in that, as far as the participant was concerned, the particular cue was irrelevant; only the behavioral task was relevant, which was to judge whether the sound image was farther to the right. Yet, the cue was relevant as far as the sex difference and learning are concerned.

For learning, there was statistically significant within-session improvement across blocks for ILD discrimination (at an equal rate for both sexes), but performance did not improve significantly for ITD discrimination. The present distinction in within-session learning between ILD and ITD discrimination is clearer than in previous investigations in which performance over the initial blocks of testing was documented, though not analyzed statistically over the same time frame as here (Ortiz and Wright, 2010; Sand and Nilsson, 2014). Statistically significant improvement for ILD but not for ITD discrimination has also been documented for across-session learning. In those reports, training extended across multiple sessions, and the amount of learning was assessed relative to a control group who completed pre- and post-training tests with no training in between (*statistically significant learning for ILD*: Wright and Fitzgerald, 2001; Zhang and Wright, 2009; Kumpik, Ting, Campbell, Schnupp, and King, 2009; Gao, Yan, Huang, Li, and Zhang, 2020; *no statistically significant learning for ITD*: Wright and Fitzgerald, 2001; Zhang and Wright, 2007; Gao, Yan, Huang, Li, and Zhang, 2020; but see Rowan and Lutman, 2006, 2007 for exceptions showing learning for ITD). A similar pattern was also noted in blind people, whose discrimination thresholds were lower than for sighted people for both interaural cues, but more so for ILDs than for ITDs (Nilsson and Schenkman, 2016). In contrast, statistically significant learning during the gap between the first two training sessions has been documented for both ILD and ITD discrimination following training on each task alone (ILD or ITD only) (ILD: Sand and Nilsson, 2014; ITD: Ortiz and Wright, 2009, 2010; Sand and Nilsson, 2014). Taken together, it appears that statistically significant learning on ILD discrimination occurs within the first session, between the first two sessions, and over multiple sessions, while statistically significant learning on ITD discrimination tends to occur only between the first two sessions. Thus, ILD discrimination seems to be more malleable overall than ITD discrimination.

For sex differences, a distinction between ILD and ITD discrimination has not been documented until now. As mentioned above, in previous studies of sex differences involving ILD and ITD discrimination, either the sample sizes were relatively small and no statistical analyses were reported (Langford, 1994), or the data for the ILD and ITD cues were not analyzed separately (Saberi et al., 2004). The present results illustrate that sex differences can be associated with a specific interaural cue. Thus, these results extend the specificity of sex differences from the previously known more general level of the perceptual task to the more detailed level of the perceptual cue. These results also indicate that ILD discrimination is more susceptible than ITD discrimination to the effect of whatever factor(s), such as hormonal influences, are responsible for the observed sex differences.

The apparently greater influences of sex and learning on ILD discrimination than ITD discrimination accentuate the separation in the processing of the two cues. While it is not known exactly what factors are responsible for the unequal influences of sex and learning on the two cues, we can speculate about where along the processing pathway those factors may have had their effect. It is well established that ILDs and ITDs are initially processed in separate pathways (for an overview, see Middlebrooks, 2015), and more recent work suggests that, alongside cue-specific representations (Wood, Town, Bizley, 2019), the two cues are merged into a common, cue-independent, representation of sound-source location in auditory cortex (e.g., Salminen, Takanen, Santala, Lamminsalo, Altoè, and Pulkki, 2015; Higgins, McLaughlin, Rinne, and Stecker, 2017; Wood, Town, Bizley, 2019). Given this framework, we can infer that the factors that led to the observed sex difference and learning acted upon the ILD-specific pathway, rather than upon the ITD-specific pathway, or any pathway common to the two cues. Because the ILD and ITD pathways are most separable initially, the factors associated with sex and learning may have acted upon the ILD-specific pathway at an early stage.

5. ACKNOWLEDGMENTS

We thank Robert Baudo, Alex Clain, and Ruijing Ning for data collection, and Ruijing Ning for data concatenation.

6. REFERENCES

Cohen J., 1988. Statistical power analysis for the behavioral sciences (2nd). Hillsdale: Erlbaum.

Doty R.L., Cameron E.L., 2009. Sex differences and reproductive hormone influences on human odor perception. Physiol. Behav. 97, 213-228. https://doi.org/10.1016/j.physbeh.2009.02.032. Erratum in: Physiol. Behav. (2009). 98, 517.

Feddersen W.E., Sandel, T.T., Teas, D.C., Jeffress, L.A., 1957. Localization of high-frequency tones. J. Acoust. Soc. Am., 29, 988-991. https://doi.org/10.1121/1.1909356

Fikentscher R., Roseburg B., Spinar H., Bruchmüller W., 1977. Loss of taste in the elderly: sex differences. Clin. Otolaryngol. Allied. Sci. 2, 183-189. https//doi.org/10.1111/j.1365-2273.1977.tb01356.x.

Flamme G.A., Deiters K., Needham T. 2011. Distributions of pure-tone hearing threshold levels among adolescents and adults in the United States by gender, ethnicity, and age: Results from the US National Health and Nutrition Examination Survey. Int. J. Audiol., 50, S11-20. https://doi.org/10.3109/14992027.2010.540582.

Gao X., Yan T., Huang T., Li X., Zhang Y.X., 2020. Speech in noise perception improved by training fine auditory discrimination: far and applicable transfer of perceptual learning. Sci. Rep. 10, 9320. https://doi.org/10.1038/s41598-020-76295-9.

Grassi M., 2010. Sex difference in subjective duration of looming and receding sounds. Perception. 39, 1424-1426. https://doi.org/10.1068/p6810.

Haggard M., Gaston J.B., 1978. Changes in auditory perception in the menstrual cycle. Br. J. Audiol. 12, 105-18. https://doi.org/10.3109/03005367809078862.

Higgins N.C., McLaughlin S.A., Rinne T., Stecker G.C., 2017. Evidence for cue-independent spatial representation in the human auditory cortex during active listening. Proc. Natl. Acad. Sci. U S A. 114, E7602-E7611. https://doi.org/10.1073/pnas.1707522114.

Huyck J.J., Wright B.A., 2018. Transient sex differences during adolescence on auditory perceptual tasks. Dev Sci. 21, e12574. https://doi.org/10.1111/desc.12574.

Kumpik D., Ting J., Campbell R.A., Schnupp J.W., King A.J., 2009. Specificity of binaural perceptual learning for amplitude modulated tones: a comparison of two training methods. J. Acoust. Soc. Am. 125, 2221-32. https://doi.org/10.1121/1.3082102.

Lauer J.E., Yhang E., Lourenco S.F., 2019. The development of gender differences in spatial reasoning: A meta-analytic review. Psychol Bull. 145, 537-565. https://doi.org/10.1037/bul0000191.

Langford T.L., 1994. Individual differences in sensitivity to interaural disparities of time and of level. J. Acoust. Soc. Am. 96, 3256-3257. https://doi.org/10.1121/1.411025.

Levitt H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am., 49, 467-477. https://doi.org/10.1121/1.1912375

Lewald J., 2004. Gender-specific hemispheric asymmetry in auditory space perception. Brain Res. Cogn. Brain Res. 19, 92-99. https://doi.org/10.1016/j.cogbrainres.2003.11.005.

Lewald J., Hausmann M., 2013. Effects of sex and age on auditory spatial scene analysis. Hear Res. 299, 46-52. https://doi.org/10.1016/j.heares.2013.02.005.

Linn M.C., Petersen A.C., 1985. Emergence and characterization of sex differences in spatial ability: a meta-analysis. Child Dev. 56, 1479-98.

Maeda, Y. and Yoon, S.Y., 2013. A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: Visualization of rotations (PSVT: R). Educ. Psychol. Rev., 25, 69-94. https://doi.org/10.1007/s10648-012-9215-x.

McFadden D., 1998. Sex differences in the auditory system. Develop. Neuropsych. 14, 261–298. https://doi.org/10.1080/87565649809540712.

McFadden D., Pasanen E.G., Maloney M.M., Leshikar E.M., Pho M.H., 2018. Differences in common psychoacoustical tasks by sex, menstrual cycle, and race. J. Acoust. Soc. Am. 143, 2338. https://doi.org/10.1121/1.5030998.

Middlebrooks J.C., 2015. Sound localization. Handb. Clin. Neurol. 129, 99-116. https://doi.org/ 10.1016/B978-0-444-62630-1.00006-8.

Neff D.L., Kessler C.J., Dethlefs T.M. 1996. Sex differences in simultaneous masking with random-frequency maskers. J. Acoust. Soc. Am. 100, 2547-50. https://doi.org/10.1121/1.417364.

Neuhoff J.G., Planisek R., Seifritz E., 2009. Adaptive sex differences in auditory motion perception: looming sounds are special. J. Exp. Psychol. Hum. Percept. Perform. 35, 225-34. https://doi.org/10.1037/a0013159.

Nilsson M.E., Schenkman B.N., 2016. Blind people are more sensitive than sighted people to binaural sound-location cues, particularly inter-aural level differences. Hear Res. 332, 223-232. https://doi.org/10.1016/j.heares.2015.09.012.

Ortiz J.A., Wright B.A., 2009. Contributions of procedure and stimulus learning to early, rapid perceptual improvements. J. Exp. Psychol. Hum. Percept. Perform. 35, 188-194. https://doi.org/ 10.1037/a0013161.

Ortiz J.A., Wright B.A., 2010. Differential rates of consolidation of conceptual and stimulus learning following training on an auditory skill. Exp. Brain. Res. 201, 441-51. https://doi.org/10.1007/s00221-009-2053-5.

Rammsayer T.H., Troche S.J. 2012. On sex-related differences in auditory and visual sensory functioning. Arch. Sex. Behav. 41, 583-90. https://doi.org/10.1007/s10508-011-9880-8.

Rowan D., Lutman M.E., 2006. Learning to discriminate interaural time differences: an exploratory study with amplitude-modulated stimuli. Int. J. Audiol. 45, 513-20. https://doi.org/10.1080/14992020600801434.

Rowan D., Lutman M.E. 2007. Learning to discriminate interaural time differences at low and high frequencies. Int. J. Audiol. 46, 585-94. https://doi.org/10.1080/14992020701524828.

Saberi K, 1995. Some considerations on the use of adaptive methods for estimating interaural-delay thresholds. J. Acoust. Soc. Am. 98, 1803-6. https://doi.org/10.1121/1.413379.

Saberi K., Antonio J.V., Petrosyan A., 2004. A population study of the precedence effect. Hear. Res. 191, 1-13. https://doi.org/10.1016/j.heares.2004.01.003.

Salminen N.H., Takanen M., Santala O., Lamminsalo J., Altoè A., Pulkki V., 2015. Integrated processing of spatial cues in human auditory cortex. Hear. Res. 327, 143-52. https://doi.org/10.1016/j.heares.2015.06.006.

Sand A., Nilsson M.E., 2014. Asymmetric transfer of sound localization learning between indistinguishable interaural cues. Exp Brain Res. 232, 1707-1716. https://doi.org/10.1007/s00221-014-3863-7.

Schiff W., Oldak R., 1990. Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. J. Exp. Psychol. Hum. Percept. Perform. 16, 303–316. https://doi.org/10.1037//0096-1523.16.2.303.

Silverman I., Choi J., Peters M. 2007. The hunter-gatherer theory of sex differences in spatial abilities: data from 40 countries. Arch Sex Behav. 36, 261-268. https://doi.org/10.1007/s10508-006-9168-6.

Sorokowski P., Karwowski M., Misiak M., Marczak M.K., Dziekan M., Hummel T., Sorokowska A. 2019. Sex differences in human olfaction: a meta-analysis. Front Psychol. 10, 242. https://doi.org/10.3389/fpsyg.2019.00242.

Strutt J.W. (1907). On our perception of sound direction. Philos. Mag. 13: 214–232., https://doi.org/10.1080/14786440709463595

Tobias J. V., 1965. Consistency of sex differences in binaural-beat perception. Internat. Audiol. 4, 179-182. https://doi.org/10.3109/05384916509074132.

Voyer D., Voyer S., Bryden M.P. 1995. Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. Psychol Bull. 117, 250-70. https://doi.org/10.1037/0033-2909.117.2.250.

Wood K.C., Town S.M., Bizley J.K., 2019. Neurons in primary auditory cortex represent sound source location in a cue-invariant manner. Nat. Commun. 10, 3019. https://doi.org/10.1038/s41467-019-10868-9.

Wright B.A., Fitzgerald M.B., 2001.Different patterns of human discrimination learning for two interaural cues to sound-source location. Proc Natl Acad Sci U S A. 98, 12307-12312. https://doi.org/10.1073/pnas.211220498.

Wright B.A., McFadden D., 1990. Uncertainty about the correlation among temporal envelopes in two comodulation tasks. J Acoust Soc Am. 88, 1339-1350. https://doi.org/10.1121/1.399711.

Wright B.A., Saberi K. 1999. Strategies used to detect auditory signals in small sets of random maskers. J Acoust Soc Am. 105, 1765-1775. https://doi: 10.1121/1.426714.

Yoshinaka M., Ikebe K., Uota M., Ogawa T., Okada T., Inomata C., Takeshita H., Mihara Y., Gondo Y., Masui Y., Kamide K., Arai Y., Takahashi R., Maeda Y., 2016. Age and sex differences in the taste sensitivity of young adult, young-old and old-old Japanese. Geriatr. Gerontol. Int. 16, 1281-1288. https://doi.org/10.1111/ggi.12638.

Zhang Y., Wright B.A., 2007. Similar patterns of learning and performance variability for human discrimination of interaural time differences at high and low frequencies. J. Acoust. Soc. Am. 121, 2207-2216. https://doi.org/10.1121/1.2434758.

Zhang Y., Wright B.A., 2009. An influence of amplitude modulation on interaural level difference processing suggested by learning patterns of human adults. J. Acoust. Soc. Am. 126, 1349-1358. https://doi.org/10.1121/1.3177267.

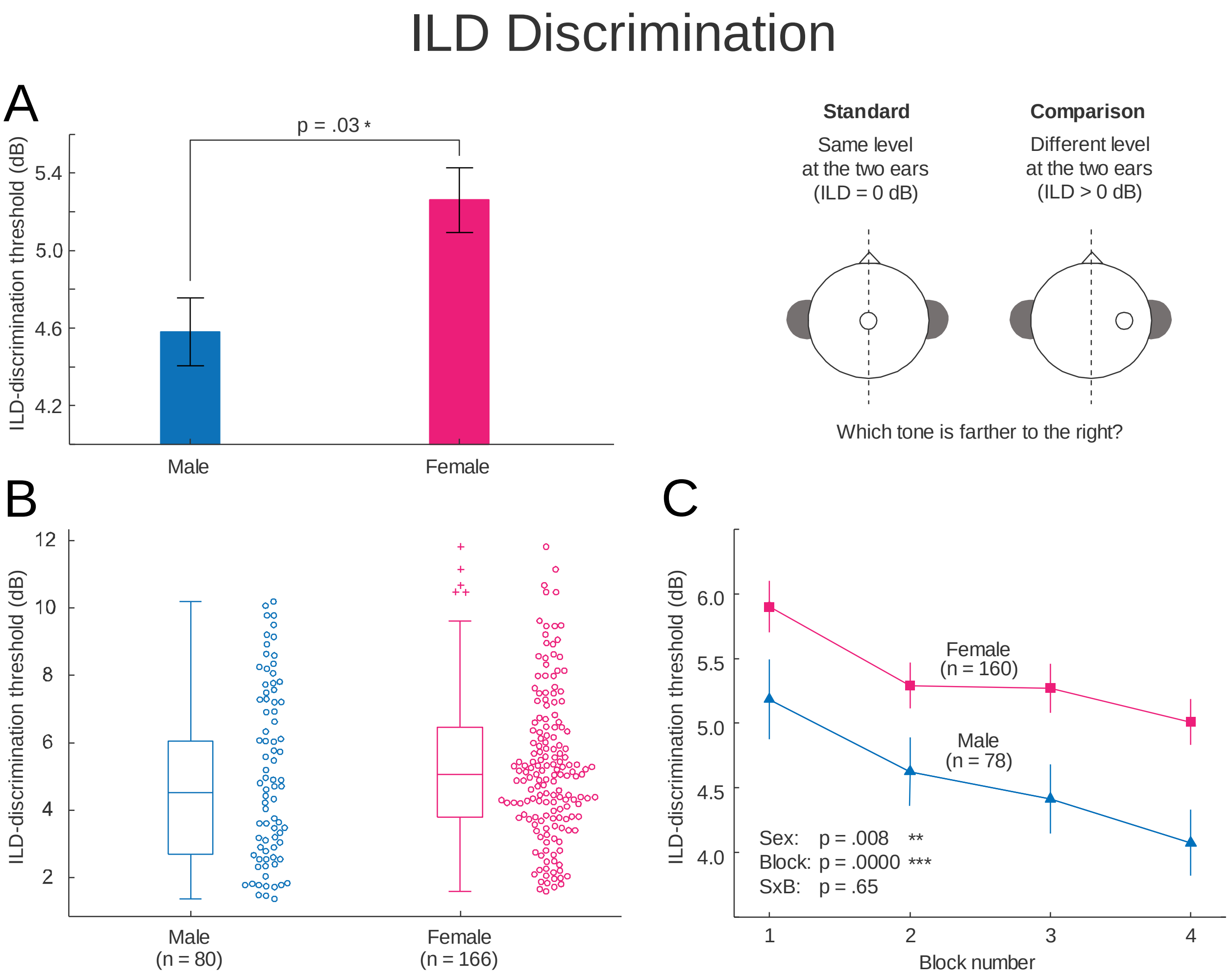
Zündorf I.C., Karnath H.O., Lewald J. 2011. Male advantage in sound localization at cocktail parties. Cortex. 47, 741-749. https://doi.org/10.1016/j.cortex.2010.08.002.

7. FIGURE CAPTIONS

Figure 1: ILD-discrimination thresholds were statistically significantly lower for males than for females. **(A)** Mean ILD-discrimination thresholds (bars) and +/- 1 SEM (error bars); **(B)** Individual thresholds (circles), and box plots indicating the median, the 25th and 75th percentiles, and the 1.5-interquartile ranges of the thresholds; and **(C)** Mean ILD-discrimination thresholds (symbols) and +/- 1 SEM (error bars) for each of the first four blocks. Data are shown separately for males (represented in blue; triangles in C; n = 80 in A and B and n = 78 in C) and females (represented in magenta; squares in C; n = 166 in A and B and n = 160 in C). Schematics illustrate the ILD-discrimination task. * = p < .05; ** = p < .01; *** = p < .001.

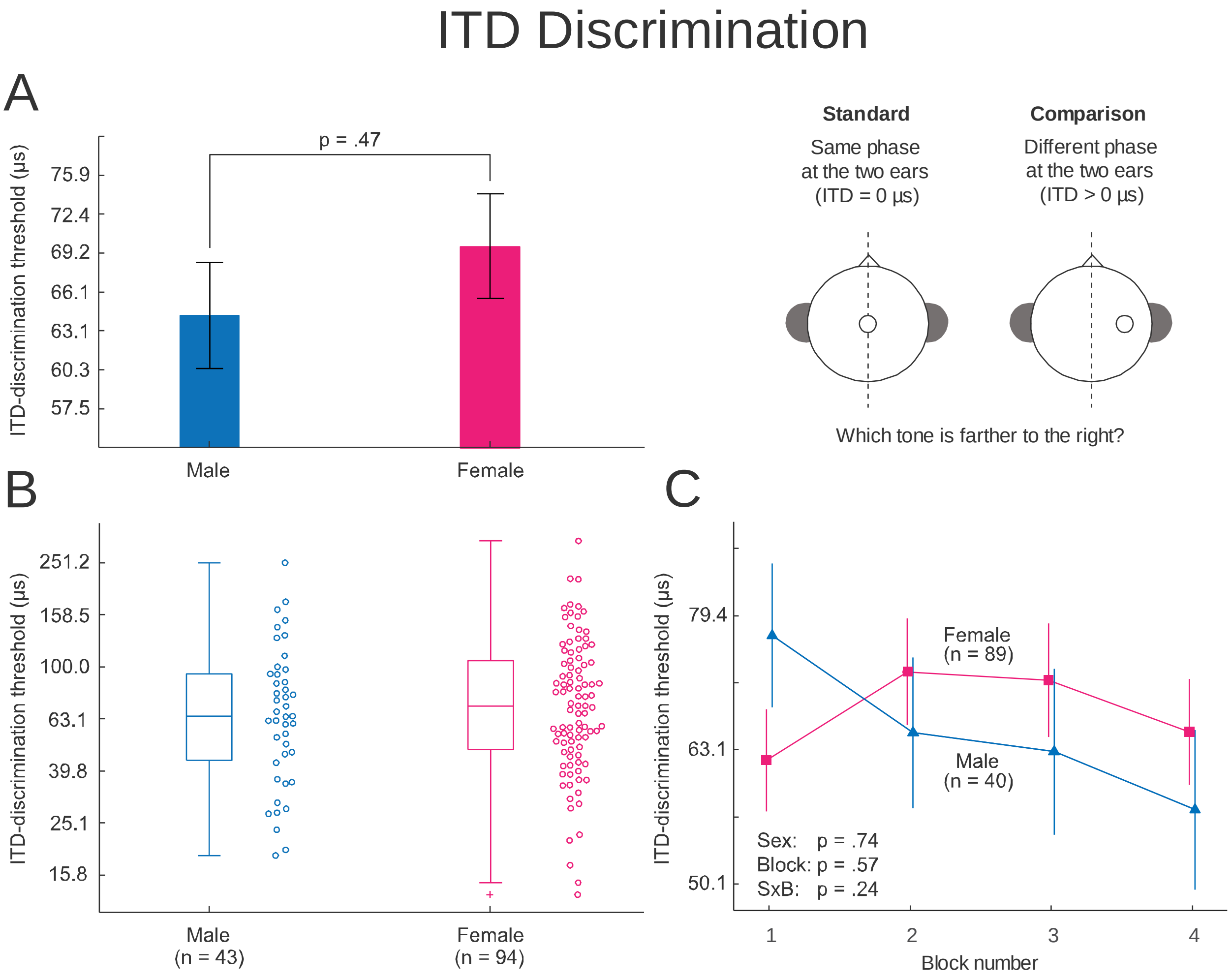
Figure 2: ITD-discrimination thresholds did not differ statistically significantly between males and females. As in Fig. 1, but for ITD discrimination: males (n = 43 in A and B and n = 40 in C) and females (n = 94 in A and B and n = 89 in C).

FIGURE 1



ILD Discrimination
A
p = .03 *
ILD-discrimination threshold (dB)
5.4
5.0
4.6
4.2
Male
Female
Standard
Same level
at the two ears
(ILD = 0 dB)
Comparison
Different level
at the two ears
(ILD > 0 dB)
Which tone is farther to the right?
B
ILD-discrimination threshold (dB)
12
10
8
6
4
2
Male
(n = 80)
Female
(n = 166)
C
ILD-discrimination threshold (dB)
6.0
5.5
5.0
4.5
4.0
Female
(n = 160)
Male
(n = 78)
Sex: p = .008 **
Block: p = .0000 ***
SxB: p = .65
1
2
3
4
Block number

FIGURE 2


ITD Discrimination
A
ITD-discrimination threshold (µs)
75.9
72.4
69.2
66.1
63.1
60.3
57.5
p = .47
Male
Female
Standard
Same phase
at the two ears
(ITD = 0 µs)
Comparison
Different phase
at the two ears
(ITD > 0 µs)
Which tone is farther to the right?
B
ITD-discrimination threshold (µs)
251.2
158.5
100.0
63.1
39.8
25.1
15.8
Male
(n = 43)
Female
(n = 94)
C
ITD-discrimination threshold (µs)
79.4
63.1
50.1
Female
(n = 89)
Male
(n = 40)
Sex: p = .74
Block: p = .57
SxB: p = .24
1
2
3
4
Block number