



**HAL**  
open science

# Psychophysical characterization of auditory temporal and frequency streaming capacities for listeners with different levels of musical expertise

Patrick Susini, Nicolas Wenzel, Olivier Houix, Emmanuel Ponsot

► **To cite this version:**

Patrick Susini, Nicolas Wenzel, Olivier Houix, Emmanuel Ponsot. Psychophysical characterization of auditory temporal and frequency streaming capacities for listeners with different levels of musical expertise. *JASA Express Letters*, 2023, 3, 10.1121/10.0020546 . hal-04180667

**HAL Id: hal-04180667**

**<https://hal.science/hal-04180667>**

Submitted on 13 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AUGUST 11 2023

# Psychophysical characterization of auditory temporal and frequency streaming capacities for listeners with different levels of musical expertise

Patrick Susini; Nicolas Wenzel; Olivier Houix; Emmanuel Ponsot



*JASA Express Lett* 3, 084402 (2023)

<https://doi.org/10.1121/10.0020546>



View  
Online



Export  
Citation

CrossMark



ASA

Advance your science and career as a member of the  
**Acoustical Society of America**

LEARN MORE

# Psychophysical characterization of auditory temporal and frequency streaming capacities for listeners with different levels of musical expertise

Patrick Susini,<sup>a)</sup> Nicolas Wenzel, Olivier Houix, and Emmanuel Ponsot

STMS, IRCAM, Sorbonne Université, CNRS, Ministère de la Culture, 75 004 Paris, France

patrick.susini@ircam.fr, n.wenzel01@gmail.com, olivier.houix@ircam.fr, emmanuel.ponsot@ircam.fr

**Abstract:** Temporal and frequency auditory streaming capacities were assessed for non-musician (NM), expert musician (EM), and amateur musician (AM) listeners using a local-global task and an interleaved melody recognition task, respectively. Data replicate differences previously observed between NM and EM, and reveal that while AM exhibits a local-over-global processing change comparable to EM, their performance for segregating a melody embedded in a stream remains as poor as NM. The observed group partitioning along the temporal-frequency auditory streaming capacity map suggests a sequential, two-step development model of musical learning, whose contributing factors are discussed. © 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

[Editor: Qian-Jie Fu]

<https://doi.org/10.1121/10.0020546>

Received: 16 May 2023 Accepted: 20 July 2023 Published Online: 11 August 2023

## 1. Introduction

Musical learning provides listeners with enhanced capacities on many sensory, cognitive, and attentional components of auditory processing (e.g., Herholz and Zatorre, 2012; Talamini *et al.*, 2017), along with specific processing styles (Wenhardt *et al.*, 2019). Interestingly, studies investigating these effects using complex melodic stimuli can be divided into two categories, depending on whether they were concerned with listeners' capacity (i) for organizing a multi-scale auditory stream along its *temporal* dimension, or (ii) for segregating multiple streams along the *frequency* dimension.

- (i) A first set of studies showed that musicians exhibit a more detail-oriented cognitive style that prioritizes local compared to global auditory information along the *temporal dimension*. Bouvet *et al.* (2011) showed that non-musician participants are faster and more accurate in processing global than local temporal information. Ouimet *et al.* (2012) and Black *et al.* (2017) investigated how musical expertise modulates this global advantage and found that compared to non-musicians, musicians classified as having an intermediate or high level of expertise both exhibited a reduced global advantage, which was mediated by an enhanced ability to process local information. Susini *et al.* (2020) used a novel experimental design, the local-global task (LGT), to assess how non-musicians and expert musicians compare a target melody with a probe melody modified either at a local level (musical intervals), a global level (melodic contour), or both. Results revealed an even more prominent reorganization for expert musicians, where the initial global dominance observed for non-musicians was reversed to the benefit of local information. Such effects are hypothesized to result from the long-term “analytic” processing acquisition associated with musical learning (e.g., Bever and Chiarello, 1974), since accurate pitch monitoring requires focused attention at the level of each single note.
- (ii) A second set of studies have focused on the processes involved to split a sequence into two or several streams (van Noorden, 1975), often addressing the question of auditory streaming along the *frequency dimension*—albeit necessarily involving time. In particular, the interleaved melody recognition task (IMRT), an objective and indirect measure of auditory streaming initially proposed by Dowling (1973) and further adapted by Bey and McAdams (2002), was developed to assess auditory streaming formation along the frequency dimension (Moore and Gockel, 2012). The IMRT involves a probe melody compared with a mixed sequence composed of a target melody (identical to the probe melody, or slightly modified) interleaved with a distracting melody with a certain mean frequency separation, which changes the difficulty of the task. Bey and McAdams (2002, 2003) already reported a difference between non-musicians and musicians having only a moderate musical experience. Wenhardt *et al.* (2019) tested listeners with strong musical experience and found that they performed better when they had absolute pitch. Overall, these results show that musicians possess an enhanced ability to process distinct streams out of a sound mixture, consistent with their daily experience in extracting melodies out of

<sup>a)</sup> Author to whom correspondence should be addressed.

an orchestral background—one of the most complex functions of the auditory system (Zendel and Alain, 2009; Johnson *et al.*, 2021)—and suggest that this benefit could even become stronger in musicians with higher levels of expertise.

These two sets of studies suggest that various levels of musical expertise could provide various benefits regarding auditory streaming capacities along the temporal and frequency dimensions. Yet, it is still unclear what is the minimum expertise level required to observe changes along the temporal and frequency dimensions, and we still do not know whether learning yield simultaneous or sequential benefits along both dimensions. In order to further address these questions, the present study considered three groups of listeners characterized by different levels of musical experience: non-musicians, amateur musicians, and expert musicians, which were all tested in LGT (Experiment 1) and IMRT (Experiment 2). This investigation also included a Backward Digit Task (BDT, Experiment 3) to assess working memory.

## 2. Methods

Experimental sessions lasted a total of  $\sim 2.5$  h, including a welcome introduction, instructions, questionnaire completion, and a minimum of 10-min breaks between experiments.

### 2.1 Participants

Eight expert musicians (one female; mean age:  $38.25 \pm 15.8$  years), eight amateur musicians (five females; mean age:  $29.25 \pm 10.4$  years) and ten non-musicians (three females; mean age:  $28.7 \pm 10$  years) were recruited, and all took part in the three experiments. Our inclusion criteria were defined before the study. Non-musicians were participants having no musical training or practice. Expert musicians were individuals considering themselves to be musicians, having everyday practice, more than six years of theoretical and instrumental musical learning in French institutions (e.g., “Conservatoire National à Rayonnement Régional”), and playing with other musicians in bands or orchestral ensembles. Amateur musicians are often rather weakly defined, encompassing many different profiles (Zhang *et al.*, 2020). We here defined amateur musicians as individuals with less than five years of musical training, having occasional solo instrumental practice, who declared not considering themselves as musicians, and importantly, having no collective practice in a band/orchestra experience, which we assumed might be of particular influence when assessing streaming capacities (see discussion). All participants reported to be right-handed, with normal hearing, and without absolute pitch ability (see supplementary material,<sup>1</sup> Appendix A, for details on participants and apparatus).

### 2.2 Experiment 1: LGT

In the LGT, a target melody was followed by a comparison melody that was identical to the target (up to a random frequency transposition) or differed by a modification applied at a local, global, or local and global level. Listeners had to decide whether the two melodies were similar or different, in two distinct attention-directed tasks: a local and a global task. Stimuli and procedures used to measure these effects were constructed according to Susini *et al.* (2020). Global advantage and global-to-local interference effects respectively account for the ability to favor global information over local information, and the ability to process local information independently of global information (see supplementary material,<sup>1</sup> Appendix B, for details).

Results—Raw performance scores correspond to the percentage of correct responses obtained in each task and condition. Because our previous study (Susini *et al.*, 2020) showed it was more appropriate to characterize perceptual reorganization by contrasting differences in performance between specific tasks and conditions rather than focusing on differences in average performance (thus being comparable with previous works on this topic, e.g., Bouvet *et al.*, 2011; Ouimet *et al.*, 2012), the present study specifically focused on the two previously-introduced indexes reflecting global advantage (G Advantage) and global-to-local interference (G-to-L Interference) effects (see Susini *et al.*, 2020 or supplementary material,<sup>1</sup> Appendix B, for details). These indexes were computed following the exact same approach and are plotted in Fig. 1(a) against each other. Mixed analyses of variance (ANOVA) conducted on each index (G Advantage and G-to-L Interference) separately revealed a significant group effect [ $F(2, 23) = 8.91, p < 0.01, \text{partial-}\eta^2 = 0.43$ ] only for G Advantage. Non-musicians (NM) were mainly grouped in the upper right part of the plot and on average exhibited both positive global advantage and global-to-local interference effects, but only the former was significant because of one outlier participant (t-tests: G-Advantage:  $p = 0.028$ , G-to-L Interference:  $p > 0.05$ ). In contrast, expert musicians (EM) exhibited a significantly negative global advantage effect ( $p = 0.003$ ), or in other words, a local advantage, and the global-to-local interference effect was not significant ( $p > 0.05$ ). This difference between NM and EM regarding the global advantage index is consistent with our previous findings obtained following the same experimental procedure (Susini *et al.*, 2020). Regarding the group of amateur musicians (AM), we found a trend similar to that of EM, where the global advantage is canceled out (here, not significantly different from zero,  $p > 0.05$ ) and a non-significant global-to-local interference ( $p > 0.05$ ). *Post hoc* t-tests (Bonferroni-Holm-corrected) conducted to compare the different groups regarding the G-Advantage index revealed significant differences between NM and EM ( $p < 0.01$ ), between NM and AM ( $p < 0.05$ ), but no significant difference between EM and AM ( $p > 0.05$ ).

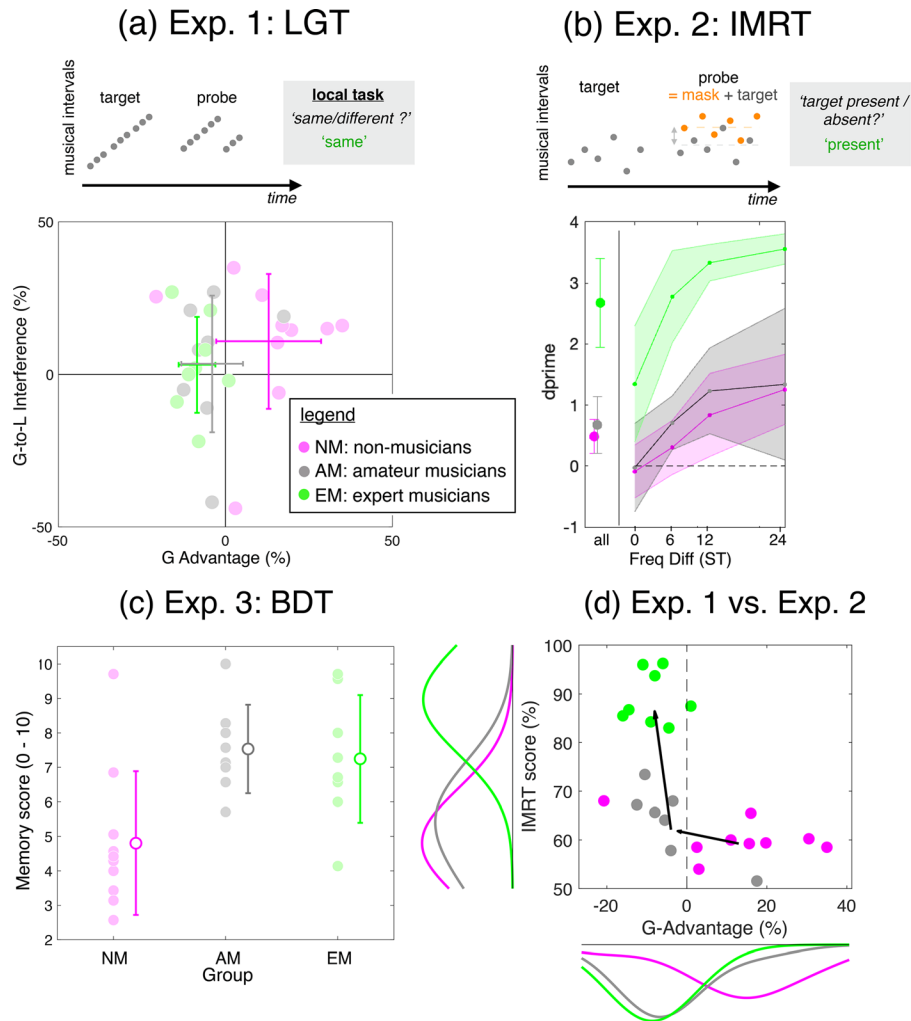


Fig. 1. Results obtained in this study for the three groups of participants (NM, AM, and EM). (a) Exp.1. Individual indexes for global advantage and global–local interference effects obtained in the LGT; (b) Exp. 2. Perceptual sensitivity ( $d'$ ) obtained in the IMRT for different frequency separation conditions (0, 6, 12, and 24 ST), as well as after aggregating all these conditions. (c) Exp. 3. Individual memory scores obtained in the BDT. Error-bars show standard deviation (SD) for each group and condition. (d) Relationship between G-advantage indexes (Experiment 1) and IMRT scores (Experiment 2), with estimated kernel density functions (MATLAB function scatterhist) projected on each dimension. Arrows show differences across group means.

### 2.3 Experiment 2: IMRT

The ability to parse auditory information into streams was investigated using an IMRT. The procedure used was constructed following [Wenhart et al. \(2019\)](#) and [Bey and McAdams \(2002, 2003\)](#). A probe melody was followed by a comparison melody composed of a target melody interleaved with another distractor melodic sequence. The target melody was identical to the probe (up to a random pitch transposition) or further differed by two notes. Listeners had to decide whether the probe melody was present or not in the composite sequence for different mean frequency separation between the target melody and the sequence of distractors (0, 6, 12, and 24 ST); see supplementary material,<sup>1</sup> Appendix C, for details.

Results—Individuals' performance in the IMRT was assessed by computing their perceptual sensitivity ( $d'$ ) for each of the four frequency separations between the target melody and the distractor sequence (0, 6, 12, and 24 ST), allowing us to discard confounds related to response strategies (response biases).  $d'$  and criterion indexes were computed following the implementation described in [Wenhart et al. \(2019\)](#), allowing a direct comparison with their results.

Figure 1(b) shows that sensitivity increased with frequency separation for each of the three groups. Sensitivity for EM was above chance ( $d' \sim 1.5$ ) when the target melody was presented in the same frequency region as the distractor sequence (0 ST), and rapidly reached high values ( $d' > 3$ ), with most individuals showing ceiling effects for the 24-ST condition. Results for EM were comparable to those observed in [Wenhart et al. \(2019\)](#) for the group of musicians with

relative pitch. Sensitivity for NM and AM was at chance ( $d' \sim 0$ ) for 0 ST, and then increased by a similar trend for larger frequency separations, but it always remained largely below the performance of EM, with a maximum value of  $d' \sim 1.5$  for the maximum frequency separation tested of two octaves (24 ST). Detailed results from the statistical analyses conducted on “response bias” are presented in the (see supplementary material,<sup>1</sup> Appendix C, Fig. C1).

A two-way repeated-measures ANOVA conducted on dprimes with separation (0, 6, 12, and 24 ST) as within-subject factor and group (EM, AM, NM) as between-subject factor revealed a significant effect of group [ $F(2, 23) = 57.85$ ,  $p < 0.0001$ ,  $\text{partial-}\eta^2 = 0.83$ ] and separation [ $F(3, 69) = 42.31$ ,  $p < 0.0001$ ,  $\text{partial-}\eta^2 = 0.65$ ] and no significant interaction between these factors. *Post hoc* t-tests for each ST-conditions revealed that EM performance was significantly higher than that of AM and NM ( $p < 0.001$ ) for each frequency separation, and we found no significant difference between AM and NM (all  $P_s > 0.05$ ). For EM, *post hoc* t-tests (Bonferroni-Holm-corrected) revealed a significant difference between the 0ST-condition and all other separation conditions (6 ST:  $p < 0.001$ , 12 ST:  $p < 0.0001$ , 24 ST:  $p < 0.0001$ ). For AM, *post hoc* t-tests revealed a significant difference between the 0-ST and both the 12-ST ( $p < 0.01$ ) and 24-ST ( $p < 0.01$ ) conditions. For NM, *post hoc* t-tests revealed a significant difference between the 0-ST condition and the 12-ST ( $p < 0.05$ ) and 24-ST ( $p < 0.01$ ) conditions, as well as between the 6-ST and the 24-ST conditions ( $p < 0.05$ ).

#### 2.4 Experiment 3: BDT

A BDT was used to assess the working memory span (Talamini *et al.*, 2017); see supplementary material,<sup>1</sup> Appendix D, for details.

Results—Individual memory scores are presented in Fig. 1(c), for each group of participants. The performance of EM and AM were globally higher than that of NM. Average scores were 4.8 for NM (SD = 0.57), 7.5 for AM (SD = 0.63), and 7.24 for EM (SD = 0.63). A one-way ANOVA revealed a main effect of group [ $F(2, 23) = 6.39$ ,  $p = 0.006$ ,  $\text{partial-}\eta^2 = 0.35$ ]. *Post hoc* t-tests revealed significant differences between NM and EM ( $p < 0.05$ ), between NM and AM ( $p < 0.05$ ), and no significant difference between EM and AM.

#### 2.5 Questionnaire

After performing the three experiments, all participants completed a self-report questionnaire addressing musical experience and abilities, as well as their educational history regarding musical practice. The full questionnaire and average answers obtained by the participants of each group to each of the four parts are presented in the supplementary material,<sup>1</sup> Appendix E.

Results—The most important differences across groups were observed in Parts 1 (professional relationship with music/sound) and 3 (reported hearing abilities); Part 2 (listening habits/pleasure with music) of the questionnaire did not account for the difference between our three groups. Average scores for Part 4 (musical practice), were significantly different between AM (2.8) and EM (7.2) ( $p < 0.001$ ). Interestingly, amateurs did not consider themselves musicians. In addition, unlike musicians, they all answered having a self-taught training without having followed academic training in a musical institution.

### 3. Discussion

*Auditory streaming along the temporal dimension: the local-global task*—In line with Black *et al.* (2017), Ouimet *et al.* (2012), and Bouvet *et al.* (2011), the data of the LGT (Experiment 1) provide further evidence [Fig. 1(a)] that NM exhibits a significant “global advantage” along the temporal dimension of melodic sequences. However, while results from Black *et al.* and Ouimet *et al.* only suggested a decrease in the global advantage effect with musical experience, here we found that EM shifted the global prioritization and exhibited a significant “local advantage.” The present results show that EM and NM are clearly clustered in two separate groups with little overlap, particularly along the global advantage index [Fig. 1(a)], which replicates our previous findings employing the exact same design (Susini *et al.*, 2020). In addition, the present data interestingly reveal that the global advantage and global-to-local interference effects are suppressed in AM: their behavior is comparable to that of EM. In Black *et al.* (2017), accuracy results obtained for three groups with different musical experience (low, medium, high) and for three stimuli configurations (compatible, the global and local levels were always moving in the same direction; incompatible, the global and local levels were always moving in opposite directions, and neutral) revealed: a difference between local and global conditions in participants with low musical experience for the three stimuli configurations; no difference in participants with medium musical experience for the compatible condition; no difference in participants with high musical experience for the compatible and the incompatible conditions. Their results suggested that the difference between global and local processing varies progressively with the musical experience. Although both the paradigms and the metrics between our studies differ, these results overall support the hypothesis that musical training is associated with a perceptual reorganization on the temporal dimension that reshapes the initial global dominance to the benefit of local information and that this re-organization takes place quite early in the musical learning stage, i.e., in the transition from NM to AM.

*Auditory streaming along the frequency dimension: the interleaved melody recognition task*—In line with Bey and McAdams (2002, 2003) and Wenhart *et al.* (2019), performance in the IMRT (Experiment 2) increased with the mean



frequency difference between the target and distracting embedded melodies for all three groups, but the increase rate of these curves appeared to depend on musical expertise. For NM and AM, performance was at chance for 0 ST, followed by a quasi-linear increase with ST, and the overall trend of both groups was extremely similar. Performance for EM was significantly higher for the four mean frequency separations [Fig. 1(b)] and followed a different trend: it was already well above chance for 0 ST, and reached a plateau for 12- and 24-ST conditions due to ceiling effects. These data thus interestingly show that AM do not perform better than NM for segregating auditory streams along the frequency dimension. As in [Wenhart et al. \(2019\)](#), the well-above-chance performance of experts for the 0-ST condition suggests that, in contrast to the other two groups, they can exploit bottom-up processes involving primitive analysis of acoustic cues to create perceptual units such as auditory streams in order to make their comparisons. [Bey and McAdams \(2003\)](#) did not observe a significant difference between musicians and non-musicians, who were at chance in the 0-ST condition, but their criterion for considering participants as musicians (playing a musical instrument for at least 3 years) was actually more comparable to our inclusion criterion for the AM group. Taken together with the present results, it suggests that a high level of musical expertise is required to observe changes in capacities for segregating embedded melodic streams along their frequency dimension.

*Potential factors contributing to the observed temporal-frequency streaming partitioning and limitations of our study*—Overall, we observed a marked discrepancy between the results concerning the group of AM, whose behavior was either similar to EM (Experiment 1) or to NM (Experiment 2). This uncoupling between the results of the two tasks, as represented by the specific performance indexes derived highlighting group differences, is even further evidenced plotting G-Advantage against average IMRT scores [Fig. 1(d)]. This trend was not observed when considering the G-to-L Interference of the LGT (see supplementary material,<sup>1</sup> Fig. C2, panel d), suggesting that this partitioning is specific to the G-advantage index. This puzzling result, showing auditory streaming capacities changes on the temporal but not the frequency dimension in AM, suggests that musical training would yield to sequential, not simultaneous acquisition of auditory capacities. Yet, the present study has several limitations that do not allow us to firmly attest that the observed group effects, although clear-cut, are indeed causally linked to musical training itself.

First of all, it is important to remind that the present study remains observational and that only randomized, longitudinal studies accounting for individual differences across experimental group can firmly attest any causal relationship between the characteristics of musical training and the enhancement of auditory capacities by controlling for predisposition factors (e.g., [Schellenberg, 2020](#)). Aware of this limitation, and because LGT and IMRT both rely on a temporal pairwise comparison procedure, we also assessed working memory capacities of our three groups using a BDT. The results obtained in that task revealed that EM outperformed NM, which is in line with the results from [George and Coch \(2011\)](#). We observed that AM performed as high as EM in the BDT. This result parallels the group partitioning observed on the G-advantage index in the LGT, but not the one observed in the IMRT. If the processing engaged in the LGT actually recruits stronger working memory capacities compared to the IMRT, this would explain the observed uncoupling between LGT and IMRT. However, we believe this is unlikely because the G-advantage index considered in the LGT is a normalized index computed from participants' scores differences for detecting global and local modifications, which should similarly engage working memory. Again, since the present study was not designed as a long-term study, it cannot be excluded that our groups may have differed in working memory before any musical learning. Moreover, it must be acknowledged that EM listeners were on average older than the other two groups, and as such age could be a confounding variable in the observed effects in Experiments 2 and 3, age strongly affecting working memory scores ([Grassi et al., 2017](#)). The age difference between our groups was not significant, but it could be due to the small sample size of our cohorts.

The profiles of AM specifically considered in the present study might also be of critical importance to understand the observed results. Long-term and intensive musical practice is shown to be beneficial to several auditory skills ([Strait et al., 2010](#); [Yoo and Bidelman, 2019](#)), so the observed asymmetrical effects could simply reflect that musical practice and training were not important enough to yield observable effects on AM in the IMRT. We intentionally recruited AM defined as musicians with limited training (<3 years) but more importantly having no collective practice (e.g., playing in a band or an orchestra), which we assumed could be important for the present streaming tasks. Indeed, collective musical practice relies on mechanisms that are analogous to the segregation of auditory information along the frequency dimension: constant sharing of attention and parsing of the different melodic streams produced by other musicians, dividing attention between one's actions and those of others while monitoring the overall ([Keller, 2008](#)). An interesting interpretation, which at this stage remains speculative, is that the lack of differences observed between AM and NM in the IMRT might be specifically related to the absence of collective playing experience of our targeted sample of participants, i.e., that individual practice of a musical instrument would not sufficient to learn how to separate different auditory streams varying within the same frequency range. The outcomes of the neurophysiological study of [Tervaniemi et al. \(2006\)](#) corroborates this idea, suggesting that neural facilitation observed in AM compared to NM might not be advanced enough to yield observable differences in encoding of temporally and spectrally complex sound information.

#### 4. Conclusion and perspectives

The present study revealed a differentiated partitioning of listeners with various levels of musical expertise (NM, AM, EM) along the temporal-frequency auditory streaming capacity map, which could not be observed in previous studies that

assessed the capacity of their cohorts on one dimension of streaming only. This result suggests that musical training would yield to sequential, not simultaneous effects on the processing capacities of complex auditory information, meaning that perceptual changes would continue well beyond the first years of training/practice, i.e., when transitioning from amateur to expert. It thus naturally illustrates the multi-dimensional nature of musical expertise, which is relevant to go beyond an often made too simplistic musician vs non-musician dichotomy.

As discussed previously, the present work does not allow stating that the observed differences truly reflect an effect related to the level of musical expertise of our listeners, since other group differences existing before musical training could have contributed. Further work, including cross-sectional studies with larger sample sizes as well as longitudinal studies (where participants are randomly assigned to the experimental groups), is needed to address the contribution and interaction of (i) musical training and (ii) inheritance/predisposition capacities (memory, perfect pitch or inherited tendencies to focus more on details as in autism; see [Wenhart and Altenmüller, 2019](#)) on the benefits of auditory perceptual abilities. In particular, further empirical studies are required to address whether the development of auditory processing capacities along the frequency dimension might indeed, as we speculate, reflect a peculiar effect of collective musical practice (in a band or orchestral ensemble). Although the degree of playing with others naturally increases from amateur to expert musicians, many amateur musicians do make music in groups (church choirs, amateur orchestras or pop-rock bands, or just by playing with friends/family), so this hypothesis could be easily addressed. Finally, future studies could benefit from a different perspective for characterizing the individual musical profile of both non-musicians and musicians based on their perceptual skills (e.g., through the PROMS; [Law and Zentner, 2012](#)), as well as the influence of the type of instrument played by the musicians recruited, classical vs improvisational, instrumental vs vocal.

## References

<sup>1</sup>See supplementary material at <https://doi.org/10.1121/10.0020546> files containing Appendixes A-E and files containing stimuli examples.

- Bever, T. G., and Chiarello, R. J. (1974). "Cerebral dominance in musicians and nonmusicians," *Science* **185**(4150), 537–539.
- Bey, C., and McAdams, S. (2002). "Schema-based processing in auditory scene analysis," *Percept. Psychophys.* **64**(5), 844–854.
- Bey, C., and McAdams, S. (2003). "Postrecognition of interleaved melodies as an indirect measure of auditory stream formation," *J. Exp. Psychol. Human Percept. Perform.* **29**(2), 267–279.
- Black, E., Stevenson, J. L., and Bish, J. P. (2017). "The role of musical experience in hemispheric lateralization of global and local auditory processing," *Perception* **46**(8), 956–975.
- Bouvet, L., Rousset, S., Valdois, S., and Donnadiou, S. (2011). "Global precedence effect in audition and vision: Evidence for similar cognitive styles across modalities," *Acta Psychol.* **138**(2), 329–335.
- Dowling, W. J. (1973). "The perception of interleaved melodies," *Cogn. Psychol.* **5**(3), 322–337.
- George, E. M., and Coch, D. (2011). "Music training and working memory: An ERP study," *Neuropsychologia* **49**(5), 1083–1094.
- Grassi, M., Meneghetti, C., Toffalini, E., and Borella, E. (2017). "Auditory and cognitive performance in elderly musicians and nonmusicians," *PLoS One* **12**(11), e0187881.
- Johnson, N., Shiju, A. M., Parmar, A., and Prabh, P. (2021). "Evaluation of auditory stream segregation in musicians and nonmusicians," *Int. Arch. Otorhinolaryngol.* **25**, e77–80.
- Herholz, S. C., and Zatorre, R. J. (2012). "Musical training as a framework for brain plasticity: Behavior, function, and structure," *Neuron* **76**(3), 486–502.
- Keller, P. E. (2008). "Joint action in music performance," in *Enacting Intersubjectivity: A Cognitive and Social Perspective to the Study of Interactions*, edited by F. Morganti, A. Carassa, and G. Riva (IOS Press, Amsterdam), pp. 205–221.
- Law, L. N., and Zentner, M. (2012). "Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills," *PLoS One* **7**(12), e52508.
- Moore, B. C., and Gockel, H. E. (2012). "Properties of auditory stream formation," *Philos. Trans. R. Soc. B* **367**(1591), 919–931.
- Ouimet, T., Foster, N. E., and Hyde, K. L. (2012). "Auditory global-local processing: Effects of attention and musical experience," *J. Acoust. Soc. Am.* **132**(4), 2536–2544.
- Schellenberg, E. G. (2020). "Correlation = causation? Music training, psychology, and neuroscience," *Psychol. Aesthetics Creativity Arts* **14**(4), 475–480.
- Strait, D. L., Kraus, N., Parbery-Clark, A., and Ashley, R. (2010). "Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance," *Hear. Res.* **261**(1–2), 22–29.
- Susini, P., Jiaouan, S. J., Brunet, E., Houix, O., and Ponsot, E. (2020). "Auditory local–global temporal processing: Evidence for perceptual reorganization with musical expertise," *Sci. Rep.* **10**(1), 16390.
- Talamini, F., Altòè, G., Carretti, B., and Grassi, M. (2017). "Musicians have better memory than nonmusicians: A meta-analysis," *PLoS One* **12**(10), e0186773.
- Tervaniemi, M., Castaneda, A., Knoll, M., and Uther, M. (2006). "Sound processing in amateur musicians and nonmusicians: Event-related potential and behavioral indices," *Neuroreport* **17**(11), 1225–1228.
- van Noorden, L. P. A. S. (1975). "Temporal coherence in the perception of tone sequences," Doctoral dissertation, Eindhoven University of Technology.
- Wenhart, T., and Altenmüller, E. (2019). "A tendency towards details? Inconsistent results on auditory and visual local-to-global processing in absolute pitch musicians," *Front. Psychol.* **10**, 31.
- Wenhart, T., Hwang, Y. Y., and Altenmüller, E. (2019). "Enhanced auditory disembedding in an interleaved melody recognition test is associated with absolute pitch ability," *Sci. Rep.* **9**(1), 7838.



- Yoo, J., and Bidelman, G. M. (2019). "Linguistic, perceptual, and cognitive factors underlying musicians' benefits in noise-degraded speech perception," *Hear. Res.* **377**, 189–195.
- Zendel, B. R., and Alain, C. (2009). "Concurrent sound segregation is enhanced in musicians," *J. Cognitive Neurosci.* **21**(8), 1488–1498.
- Zhang, J. D., Susino, M., McPherson, G. E., and Schubert, E. (2020). "The definition of a musician in music psychology: A literature review and the six-year rule," *Psychol. Music* **48**(3), 389–409.