Measuring the Just Noticeable Difference for Audio Latency

Andreas Schmid
University of Regensburg
Regensburg, Germany
andreas.schmid@ur.de

Maria Ambros
University of Regensburg
Regensburg, Germany
maria.ambros@stud.uni-regensburg.de

Johanna Bogon
University of Regensburg
Regensburg, Germany
johanna.bogon@ur.de

Raphael Wimmer
University of Regensburg
Regensburg, Germany
raphael.wimmer@ur.de

ABSTRACT

All parts of an audio processing chain introduce latency. Previous studies have shown that high audio latency may negatively impact human performance in different scenarios, e.g., when performing live music or when interacting with real-time human-computer systems. However, is not yet known where the human perception threshold for audio latency lies, i.e., what the lowest amount of latency is that musicians might notice. Therefore, we conducted a user study (n=37) using the PEST method to estimate the just noticeable difference (JND) for audio latency under different base latency settings. Our results suggest that base latency influences the perception threshold in a non-linear manner: Participants achieved a mean JND of 49 ms for a base latency of 0 ms, 27 ms for a base latency of 64 ms, and 77 ms for a base latency of 512 ms. Furthermore, the JND was lower for participants with high musical sophistication.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*.

KEYWORDS

latency, time perception, audio latency, just noticeable difference

ACM Reference Format:
Andreas Schmid, Maria Ambros, Johanna Bogon, and Raphael Wimmer. 2024. Measuring the Just Noticeable Difference for Audio Latency. In *Audio Mostly 2024 - Explorations in Sonic Cultures (AM '24), September 18–20, 2024, Milan, Italy*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3678299.3678331

1 INTRODUCTION

Latency is omnipresent in audio applications. Inertia and mechanical play cause inherent delays until an instrument produces sound. Each meter that sound travels through air adds approximately three milliseconds of latency. Digital signal processing such as sampling, buffering, and filtering adds latency. Finally, hearing and processing audio in the brain adds further latency. High action-to-sound latency – the delay between performing an action on an instrument and the generated sound – can affect musicians. Several studies have shown that ensemble performance becomes significantly more difficult as latency increases – for example when playing piano [2, 4], in an orchestra [3], or percussion [18, 19]. Musicians rate live monitoring worse with higher latency [23] and DJ's require more time to synchronize tracks when latency is present [32]. But audio latency can also affect non-musical multimedia applications, such as video games [14]. Therefore, latency is an important factor to consider when designing and working with auditory media.

However, just measuring the impact of latency on musicians' performance ignores their user experience. People can compensate for latency while still being annoyed by it. Therefore, it is necessary to better understand the threshold at which audio latency is actually noticed by musicians and non-musicians. One method to determine the perception threshold for differences between stimuli is the *Just Noticeable Difference* (JND) [13]. The JND is the smallest difference for a parameter between two stimuli that can be reliably distinguished. JND is an essential measure in psychophysics, dating back to its origins in the 19th century [9]. Ever since, researchers have determined the JND for several parameters of auditory stimuli. For example, humans can perceive a difference in amplitude of about 1 dB [24] and a difference in frequency of about 1–3 Hz, depending on the absolute pitch and the tone's complexity [16]. JND has also been used to determine the perception threshold for latency in human-computer interaction. Latency JND is highly dependent on task and medium. For example, humans can perceive a latency difference of less than 20 ms when pushing a touch-sensitive button [21]. When dragging a virtual object over a touch display, the perception threshold can even be as low as two milliseconds [20].

We conducted a user study ($n = 37$) to determine the JND for action-to-sound latency with respect to different base latencies: 0, 64, and 512 milliseconds. Additionally, we took into account the participants' musical sophistication determined via the Gold-MSI questionnaire [10, 26]. We optimized the study apparatus to minimize latency and latency variation introduced by our setup. Our results show that base latency has a statistically significant effect on the JND for audio latency. Interestingly, we found that the mean JND for audio latency was lowest in the condition with a medium amount of base latency. Additionally, we found a statistically significant, moderate negative correlation between participants' musical sophistication and their JND for audio latency: participants with high musical sophistication were able to perceive smaller latency differences.

AM '24, September 18–20, 2024, Milan, Italy

ACM ISBN 979-8-4007-0968-5/24/09
https://doi.org/10.1145/3678299.3678331

2 RELATED WORK

Whenever users interact with a system in real time, latency can deteriorate performance and experience. This includes human-computer systems, video games, and audio applications. The negative effect of latency on performance is especially apparent in collaborative situations.

2.1 Influence of Latency on Ensemble Performance

Several studies investigated the effect of latency on collaborative music performance. The methodology used in those studies is oftentimes similar: performers are located in separate rooms and can hear each other via headphones. Audio signals transmitted between the rooms can be delayed by the experimenter so performance under different latency conditions can be observed. Using this method, Schuett [31] found that upwards of 30 ms of latency, participants started slowing down and were less synchronized when clapping in interlocking rhythms. In a follow-up study, Chafe et al. [3] could replicate Schuett's findings. They found a sweet spot for latency at 11.5 ms. Higher latency led to participants slowing down in tempo as they started waiting for their partner's signal. With lower latency, participants started speeding up. In a study by Chew et al. [4], two professional pianists were asked to play three pieces with different tempos. Different amounts of latency were added to the system (0 – 150 ms). Results suggest that a piece's tempo has influence on the performers' ability to adapt to latency: Fast pieces are possible to perform with up to 50 ms, slow pieces with up to 75 ms of latency. In a similar study, Bartlette et al. [2] found that high latency, especially above 100 ms, strongly affected pace, timing, coordination, and subjective musicality of the performance.

Even though those studies give insights into the practical effects of latency in ensemble performance, they do not measure the perception threshold for audio latency. It is likely that latency is noticed before it starts to affect performance significantly. Additionally, all these studies used very small sample sizes and/or homogeneous participant samples. Therefore, one has to be cautious when generalizing their findings to a broader population.

2.2 Influence of Latency on Audio Applications

Hämäläinen and Mäki-Patola [17] studied the perception threshold for audio latency when playing a theremin. Their participants could reliably perceive differences of 30 ms and above. Lester and Boley [23] investigated how different amounts of latency change performers' subjective perception in live monitoring scenarios. They compared different instruments in two monitoring scenarios: wedge and in-ear monitors. Results suggest that performers accept higher latency in wedge monitoring setups. Furthermore, the accepted amount of latency varies widely across instruments and situations, ranging from 1.4 ms to 42 ms. Jack et al. [18, 19] studied how delayed auditory feedback impacts the perceived quality of digital instruments. They built a custom digital percussion instrument to minimize latency introduced by their apparatus. Two groups of participants, professional percussionists and non-percussionist amateur musicians, rated the quality of the instruments in four delay conditions. The conditions with 0 and 10 ms latency were rated similarly and both were rated significantly better than 10±3 ms and 20 ms. Professional percussionists were more consistent under higher latency and were more aware of the added delay.

Mäki-Patola [25] claims that tight latency thresholds, as usually suggested in instrument design, may not be applicable in all situations. In their review of related literature, they present numerous studies that have found higher acceptance thresholds. Some studies found that musicians can perceive very small differences in latency – as low as ten milliseconds in some cases [11, 12]. However, other studies show that in practical music performance, musicians' timing varies in much higher margins. For example, Rasch et al. [28] found that asynchronies of up to 50 ms are common, Dahl [6] found that professional percussionists jitter by 10 to 40 ms, and Lago and Kon [22] found that latencies of up to 30 ms are common and acceptable with traditional instruments. Sawchuck et al. [29] found that latency tolerance is highly dependent on instrumentation and piece. During their study, 100 milliseconds of latency were considered acceptable for collaborative playing with a piano, whereas only 20 ms were acceptable for accordion players.

2.3 Influence of Latency in Human-Computer Interaction

Halbhuber et al. [14] found that high audio latency has an effect on playing first-person shooters. While added audio latency of 40 ms did not affect players' performance and game experience, participants performed significantly worse during conditions with higher latency. However, latency conditions of this study were spread out very far so the next higher latency condition was at 270 ms of added latency. Therefore, we can not draw conclusions on the JND from this study. Kaaresoja et al. [21] investigated the temporal perception of different feedback modalities for virtual buttons. They found that the point of subjective simultaneity (PSS) for haptic and acoustic feedback was at 19 ms. Jota et al. [20] investigated perception thresholds for latency for low-level HCI tasks: dragging targets and tapping on targets. To display targets, they implemented a special apparatus based on a high-speed projector controlled by an FPGA to minimize latency of the test setup. Results suggest that users are able to perceive latencies of 10 ms in certain situations. However, they argue that performance benefits for latencies lower than 25 to 50 ms (depending on the task) are negligible. Similarly, Deber et al. [7] found that the just noticeable difference for latency perception in interactive applications depends on the input technique. Their results indicate that users perceive lower amounts of latency for direct input compared to indirect input. Furthermore, they confirm previous findings by showing that the the perception threshold for latency is lower for dragging tasks than for tapping.

2.4 Summary

Latency has a measurable effect of human performance in a multitude of task types. Even though several studies show that audio latency affects musical performance [2–4, 31], especially in ensembles, there is a debate to which degree those effects play a role in practice [25] – especially concerning inherent tempo changes of musical performances [6, 22, 28]. Perception thresholds for latency have been measured predominantly for visual tasks, such as pointing [7, 20]. Findings from those studies suggest that latency

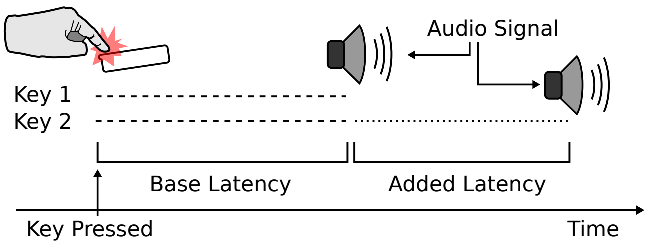

Figure 1: The apparatus for our study. When a participant presses one of the two keys, the apparatus waits for a certain amount of time before playing a sound. One of the keys only has the *base latency* of 0, 64, or 512 ms whereas the other has the same base latency plus a variable *added latency*. The key with the higher latency is randomly chosen for each trial.

perception depends on the task, as well as input and output modalities of a system. Most studies on audio latency focus on its effects, for example on task performance. Even though perception thresholds for other aspects of audio signals, such as pitch [16], rhythm [27, 34], and amplitude [24] have been determined, no such studies exist for audio latency. They are sometimes reported as additional findings [17], but have not been systematically investigated.

3 METHOD

To determine the perception threshold of audio latency, we conducted a within-subject user study. Our methodology was based on *Parameter Estimation by Sequential Testing* (PEST) [33], which is a common method to determine perception thresholds in psychophysics [13, p.55]. Participants were asked to successively press two keys on a keyboard. Both keys triggered an audio signal after a certain delay (Fig. 1). Then, participants had to decide which of the two keys had the lower latency (two-alternative forced choice). Over the course of the study, the latency margin between the two keys was adjusted until the participant's just noticeable difference for audio latency was reached.

Perception thresholds are oftentimes defined as the margin between two stimuli that can be distinguished by participants with a probability of over 75% [13, p.56]. To ensure participants were able to distinguish a given latency difference with at least 75% certainty, we repeated each trial ten times. If eight or more repetitions were completed successfully, the latency margin between the two keys was decreased by 50%. Otherwise, we increased the latency margin by 50%. This way, we could approximate a participant's JND for audio latency with a binary search until the temporal resolution of our apparatus was reached (Fig. 3).

Perception thresholds are oftentimes relative to the magnitude of a stimulus, as modeled by Weber's Law [9, 13, p.134]. Therefore, we investigated the JND for audio latency for different conditions of base latency which was added to both keys. Furthermore, we considered participants' musical sophistication as an additional independent variable.

3.1 Apparatus

The minimum latency of our apparatus, as well as its latency variation, determine the resolution and accuracy of our user study's

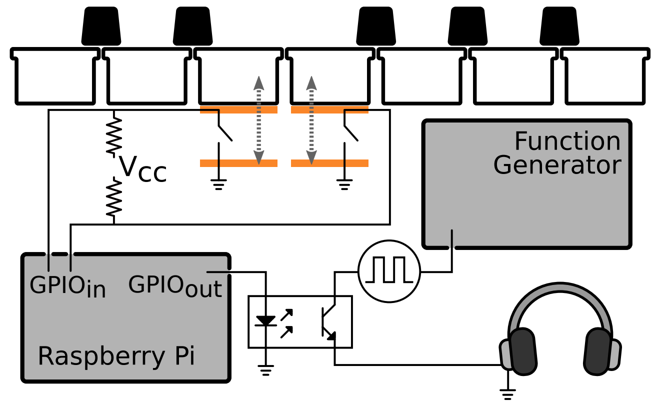

Figure 2: Circuit of the apparatus. Once one of the instrumented keys is pressed, copper contacts touch and an interrupt is triggered on the Raspberry Pi. After a given latency, an optocoupler is closed and a square wave signal can pass from a function generator to the participant's headphones.

results. Related work has shown that humans can visually perceive latency differences as low as two milliseconds in some applications [20]. Therefore, we wanted to design an apparatus with a resolution of about one millisecond and as little variation as possible.

Our first approach was to use a MIDI keyboard and a synthesizer. We conducted latency measurements for the MIDI output of a keyboard[1] and a stage piano[2] using a Raspberry Pi and a two channel oscilloscope.

To determine the moment a key on the instrument was pressed, we connected the first oscilloscope channel to the contacts of one key. Therefore, the oscilloscope measured a voltage change once the key was pressed and its circuit was closed. To determine when a MIDI event was transmitted, the instrument under test was connected to the Raspberry Pi via USB. A Python program on the Raspberry Pi used the *Mido* library[3] to receive MIDI events. The second oscilloscope channel was connected to a GPIO pin on the Raspberry Pi. Once a MIDI event arrived, the GPIO pin was switched from low to high. This change in voltage was registered by the oscilloscope. By pressing the key on the instrument and measuring the time difference between the two rising flanks with the oscilloscope, latency between the key being physically closed and the MIDI event arriving at the computer could be determined. For both devices, latency varied by up to eight milliseconds between individual key presses. We assume that this variation in latency is an artifact of a USB polling rate of 125 Hz [35].

As these variations in latency would significantly reduce the accuracy of our user study's results, we opted for a faster way to trigger an audio signal. The final apparatus is depicted in Fig. 2. We modified the *Nektar Impact GX 61* MIDI keyboard by attaching copper contacts on the bottom of two keys[4], as well as underneath those keys. When one of these keys is pressed, the copper contacts close a circuit, which is in turn registered by a Raspberry Pi's GPIO

[1] Nektar Impact GX 61, https://nektartech.com/impact-gx49-61/
[2] Numa Compact 2x, https://www.studiologic-music.com/products/numa_compact2x/
[3] https://mido.readthedocs.io/
[4] the E and F keys above the middle C

pin. After a set amount of latency, the Raspberry Pi triggers an optocoupler and a signal from a waveform generator[5] can pass to a pair of headphones worn by participants during the study. The generated sound remains audible for as long as the key is pressed. When the key is released, the sound stops after the key's current latency.

We configured the waveform generator to generate a square wave at 500 Hz. Due to the signal's frequency and the resulting semi-period of 1 ms, our apparatus achieves a temporal resolution of about one millisecond. We did consider using a higher frequency to further increase resolution, however initial experiments showed that the resulting audio signal was considered very unpleasant by participants.

To validate our setup's internal latency, we used an oscilloscope to measure the time from the Raspberry Pi registering an interrupt to the first flank of the audio signal passing through the optocoupler. For 25 individual measurements, delay times ranged from 56 to 952 microseconds (mean: 318.7 μs, std: 335.7 μs). These results confirm our assumption that the apparatus' latency variation is mainly caused by the signal's frequency and below one millisecond overall.

3.2 Procedure

We determined participants' JND for audio latency for a given *base latency* with a binary search. Participants were asked to press two keys on a MIDI keyboard. Each of those keys played a sound through participants' headphones – either after a constant *base latency*, or after the *base latency* plus *added latency* (Fig. 1). Then, participants were asked to specify which of the keys had less latency. There was no time limit and they were allowed to press both keys as often as they wanted. Each trial was repeated ten times and for each repetition, we randomized which of the two keys to use for *base latency* and *base latency* plus *added latency*. If participants could correctly identify the key with less latency in at least eight out of ten repetitions, we decreased the *added latency* by 50%. Otherwise, we increased the *added latency* by 50%. At the beginning of the procedure, *added latency* (and therefore margin to the *base latency*) was set to 256 ms. A condition ended after nine rounds when the final amount of *added latency* was one millisecond higher or lower than in the eighth round. An example of the procedure is depicted in Figure 3. Using this procedure, we determined each participant's JND for audio latency for three *base latency* conditions: 0 ms, 64 ms, and 512 ms. The order of *base latency* conditions was fully counterbalanced among participants. Headphone volume could be adjusted by participants to their liking.

3.3 Participants

We recruited 37 participants (12 women, 24 men, one non-binary) for our user study. They were aged between 19 and 29 years (mean: 23.3). Most of our participants were students (35) and among those, 16 studied computer science or a related field. The remaining two participants were working in an engineering-related field. We used *Goldsmiths Musical Sophistication Index* (Gold-MSI) to determine participants' musical sophistication [10, 26]. The Gold-MSI is a validated self-report questionnaire designed to assess participants' ability to engage with music without considering musical education,

[5] UNI-T UTG9005C

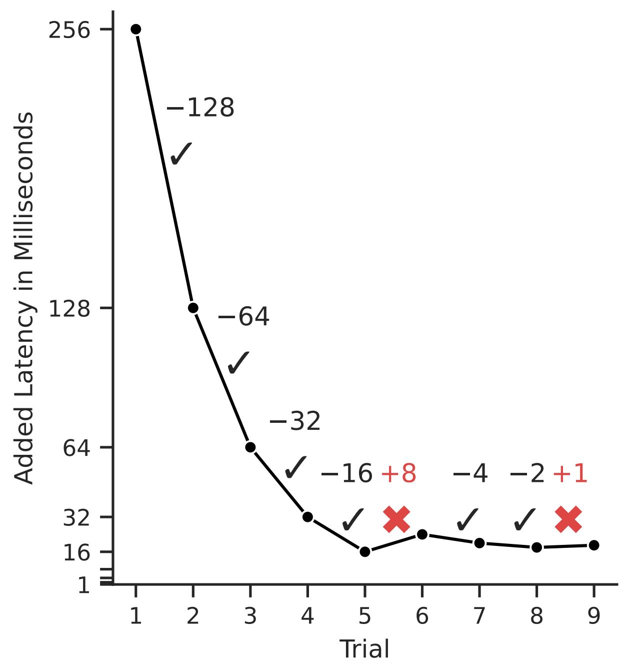

Figure 3: Exemplary visualization of latency differences over the course of nine trials. JND for audio latency was estimated by starting with a very high latency margin of 256 ms. If a participant could determine the key with lower latency in at least eight of ten repetitions, the latency margin was halved. Otherwise, the margin was increased by 50%.

such as the ability to sight-read or to play an instrument. Musical sophistication is measured on a scale of 18 to 126, with 126 being the highest musical sophistication. Total MSI scores across our participants ranged from 42 to 113.

A preliminary analysis of our data has shown that two participants oftentimes pressed only one key on the keyboard, making it impossible to compare latencies between both keys. For another participant, problems with the apparatus occurred during the study and wrong latencies were assigned to the keys. Therefore, we excluded those three participant from further analysis.

4 RESULTS

To evaluate the effect of independent variables – base latency and participants' musical sophistication – on the just noticeable difference for audio latency, we performed a two-part analysis.

With no added base latency, participants achieved JNDs between 10 and 129 ms (mean: 48.8 ms, SD: 33.0 ms). Adding a large base latency (512 ms) clearly increased the JND (mean: 77.1 ms, SD: 57.6 ms, range: 16 – 257 ms). Interestingly, participants performed best in the conditions with a moderate base latency (64 ms), achieving JNDs between 5 and 73 ms (mean: 27.1 ms, SD: 17.0 ms) (Fig. 4).

Shapiro-Wilk tests show that the normal distribution of JND was violated for all conditions of base latency ($p < 0.01$ for all conditions). We therefore used the non-parametric Friedman Test to test

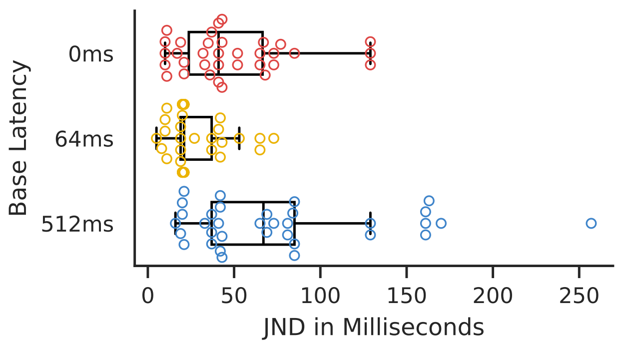

Figure 4: Just Noticeable Difference for audio latency for the three different base latency conditions. Post-hoc tests show significant differences between all three conditions.

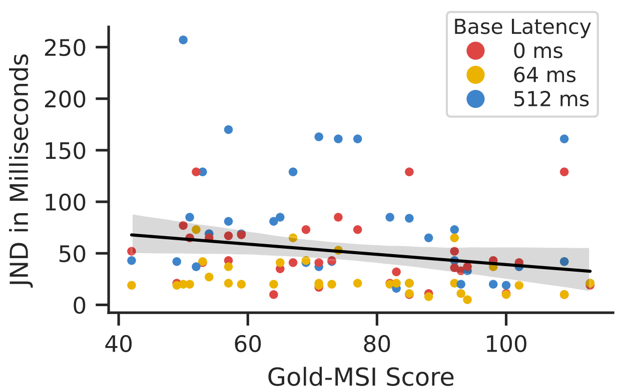

Figure 5: There is a moderate negative correlation between participants' self-assessed musical sophistication via the Gold-MSI score and the JND for audio latency.

for effects of base latency on JND. We could find a significant effect of base latency on JND ($\chi^2(2) = 24.582, p < 0.001$). Bonferroni-corrected post-hoc pairwise Wilcoxon tests show significant effects of base latency on JND between all conditions. We found a significant difference (p=0.001) with *large* effect ($d = 0.825$) between base latencies of 0 ms and 64 ms, a *medium* effect ($d = 0.693$) between base latencies 0 ms and 512 ms ($p = 0.012$), and a *very large* effect ($d = 1.177$) between base latencies of 64 ms and 512 ms ($p < 0.001$). Effect sizes were determined according to Cohen [5].

To analyze the relationship between participants' musical sophistication and their just noticeable difference for audio latency, we tested for a correlation between both variables. For each participant, we aggregated JNDs from all three conditions to a mean JND. A Pearson test shows a statistically significant, moderate negative correlation between this mean JND and participants' score in the Gold-MSI questionnaire ($r(32) = -0.345, p = 0.046$) (Fig. 5).

During our study, participants could take as long as they wanted to decide which key had lower latency. We also did not limit the

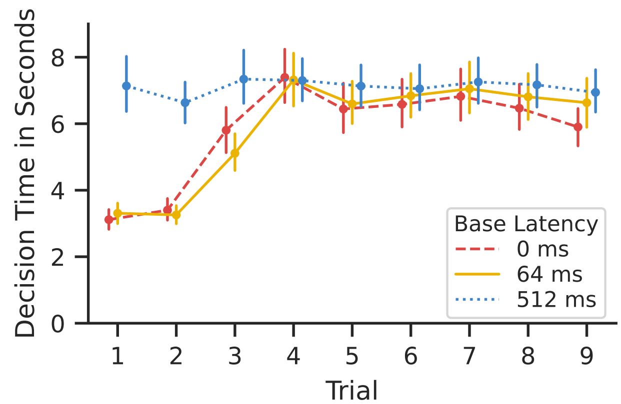

Figure 6: Participants' decision time in seconds for different trials and latency condition. Decision time is defined as the time from the first key press until the final decision.

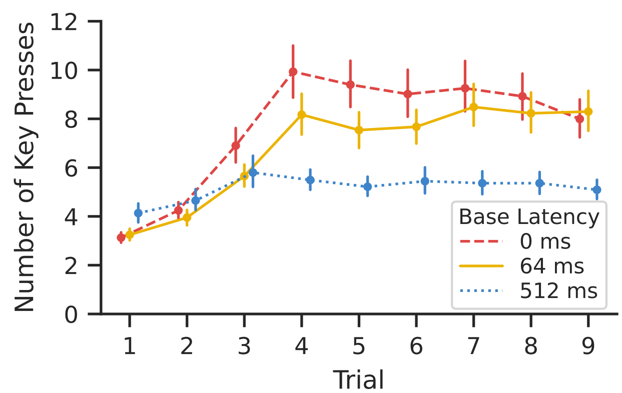

Figure 7: Number of key presses until participants made their choice of which key has lower latency.

amount of key presses before they had to decide. Therefore, decision time and the amount of key presses can be an indicator for participants' confidence in their decision. Even though an investigation of participants' behavior during the study is beyond our study's objectives, an exploratory analysis regarding decision time and number of key presses might be interesting for further research.

Mean *decision time* – the time from the first key press until a decision is made – per trial is depicted in Fig. 6. Interestingly, decision time was constant across all trials for high base latency, whereas for low and medium base latency, decision time was lower for the first three trials. The number of key presses is fairly constant for high base latency (around 4 – 6 times across all nine trials) (Fig. 7). For low and medium base latency, participants performed clearly more key presses starting at the fourth trial (around 8 – 10 key presses). For both, decision time and key presses, participants behaved similarly for low and medium base latency.

5 DISCUSSION, LIMITATIONS, FUTURE WORK

We conducted a user study to determine the just noticeable difference for action-to-sound audio latency in different base latency settings (0, 64, and 512 ms) and found a significant effect of base latency on JND. Interestingly, JND was the lowest with a medium base latency of 64 ms. This suggests a non-linear relationship between base latency and the JND for audio latency – a contradiction to Weber's Law [9] but potentially related to the observations by Chafe et al. [3]. As our study's objective was to investigate *whether* base latency has an effect on JND for audio latency, we can only speculate on reasons for this non-linear relationship. A pragmatic explanation could be humans' familiarity with audio latency, as in practice, audio signals rarely reach listeners' ears without delay. Another possible reason for this effect could be the interplay between the keys' haptic feedback and the timing of the audio signal.

Furthermore, we found a moderate negative correlation between participants' self-assessed musical sophistication and their JND for audio latency. This finding goes in line with earlier research suggesting that musical background can have an effect on the perception of auditory signals. For example, Ehrlé and Samson [8] found that percussionists achieve a lower JND for musical timing than non-musicians or classically trained musicians.

Our study is one of the first to systematically investigate the effect of base latency on the JND for audio latency. Therefore, we chose large margins of base latency to increase the probability to find an effect if it exists. However, this approach limits the possibility to analyze, explain, and model the effect of base latency on the JND in detail. Therefore, future studies should consider more and smaller base latency steps.

Furthermore, as we optimized our apparatus for minimum base latency, we did not investigate effects of latency jitter in our study. Findings regarding such varying latency are inconclusive: while sudden latency spikes affect performance and user experience [15], latency jitter has shown no effects in a similar study [30]. However, both of those studies focus on the context of human-computer interaction. As latency jitter is omnipresent in digital audio processing [1], future research should investigate the effect of latency jitter on the perception of action-to-sound audio latency. We also did not investigate effects of other factors on latency perception, e.g., audio volume or waveform shape. Lastly, even though we recruited participants with a wide range of musical sophistication, none of them were professional musicians. As our results suggest that people with higher musical sophistication can identify smaller latency margins, the perception threshold might be even lower for experienced musicians.

In this paper, we presented results of a user study to determine the just noticeable difference for audio latency. We used a custom apparatus optimized for low latency, achieving a resolution of one millisecond. Our findings indicate that base latency has an influence on JND with a non-linear relationship. Furthermore, participants' musical sophistication was negatively correlated with JND for audio latency. Future studies should further investigate the non-linear relationship between base latency and JND, as well as the influence of the input modality on the JND for audio latency.

REFERENCES

[1] Andrew P. McPherson, Robert H. Jack, and Giulio Moro. 2016. Action-Sound Latency: Are Our Tools Fast Enough?. In *Proceedings of the International Conference on New Interfaces for Musical Expression.* http://www.nime.org/proceedings/2016/nime2016_paper0005.pdf

[2] Christopher Bartlette, Dave Headlam, Mark Bocko, and Gordana Velikic. 2006. Effect of Network Latency on Interactive Musical Performance. *Music Perception* 24, 1 (Sept. 2006), 49–62. https://doi.org/10.1525/mp.2006.24.1.49

[3] Chris Chafe, Michael Gurevich, Grace Leslie, and Sean Tyan. 2004. Effect of Time Delay on Ensemble Accuracy. (April 2004), 4.

[4] E. Chew, R. Zimmermann, A. Sawchuk, C. Kyriakakis, C. Papadopoulos, A. François, G. Kim, A. Rizzo, and A. Volk. 2004. Musical Interaction at a Distance : Distributed Immersive Performance. https://www.semanticscholar.org/paper/Musical-Interaction-at-a-Distance-%3A-Distributed-Chew-Zimmermann/375b168675cdaebe8394f7d0cb06bb504338a874

[5] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). Routledge, New York. https://doi.org/10.4324/9780203771587

[6] Sofia Dahl. 2000. The Playing of an Accent - Preliminary Observations from Temporal and Kinematic Analysis of Percussionists. *Journal of New Music Research* 29, 3 (Sept. 2000), 225. https://doi.org/10.1076/jnmr.29.3.225.3090

[7] Jonathan Deber, Ricardo Jota, Clifton Forlines, and Daniel Wigdor. 2015. How Much Faster is Fast Enough? User Perception of Latency & Latency Improvements in Direct and Indirect Touch. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15).* Association for Computing Machinery, New York, NY, USA, 1827–1836. https://doi.org/10.1145/2702123.2702300

[8] Nathalie Ehrlé and Séverine Samson. 2005. Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners. *Brain and Cognition* 58, 1 (June 2005), 133–147. https://doi.org/10.1016/j.bandc.2004.09.014

[9] Gustav Theodor Fechner. 1860. *Elemente der Psychophysik.* Vol. 2. Breitkopf und Härtel.

[10] Daniel Fiedler and Daniel Müllensiefen. 2015. Validierung des Gold-MSI-Fragebogens zur Messung Musikalischer Erfahrenheit von Schülerinnen und Schülern der Sekundarstufen an allgemeinbildenden Schulen. In *Theoretische Rahmung und Theoriebildung in der musikpädagogischen Forschung*, Anne Niessen and Jens Knigge (Eds.). Number 36 in Musikpädagogische Forschung. Waxmann, Münster; New York, 199–219. https://www.pedocs.de/frontdoor.php?source_opus=12622

[11] Steven A. Finney. 1997. Auditory Feedback and Musical Keyboard Performance. *Music Perception: An Interdisciplinary Journal* 15, 2 (1997), 153–174. https://doi.org/10.2307/40285747

[12] Adrian Freed, Amar Chaudhary, and Brian Davila. 1997. Operating Systems Latency Measurement and Analysis for Sound Synthesis and Processing Applications. International Computer Music Association, Thessaloniki, Greece, 479–481.

[13] Simon Grondin. 2008. *Psychology of Time.* Emerald Group Publishing. Google-Books-ID: oWf2OqH3aJYC.

[14] David Halbhuber, Annika Köhler, Markus Schmidbauer, Jannik Wiese, and Niels Henze. 2022. The Effects of Auditory Latency on Experienced First-Person Shooter Players. In *Proceedings of Mensch und Computer 2022 (MuC '22).* Association for Computing Machinery, New York, NY, USA, 286–296. https://doi.org/10.1145/3543758.3543760

[15] David Halbhuber, Valentin Schwind, and Niels Henze. 2022. Don't Break my Flow: Effects of Switching Latency in Shooting Video Games. *Proceedings of the ACM on Human-Computer Interaction* 6, CHI PLAY (2022), 229:1–229:20. https://doi.org/10.1145/3549492

[16] A. J. Houtsma and J. Smurzynski. 1990. Pitch identification and discrimination for complex tones with many harmonics. *Journal of the Acoustical Society of America* 87, 1 (1990), 304–310. https://doi.org/10.1121/1.399297

[17] Perttu Hämäläinen and Teemu Mäki-Patola. 2004. Latency Tolerance for Gesture Controlled Continuous Sound Instrument without Tactile Feedback. In *Proceedings of the International Computer Music Conference (ICMC2004).* The International Computer Music Association, Coral Gables, Florida, USA, 8.

[18] Robert H. Jack, Adib Mehrabi, Tony Stockman, and Andrew McPherson. 2018. Action-sound Latency and the Perceived Quality of Digital Musical Instruments - Comparing Professional Percussionists and Amateur Musicians. *Music Perception* 36, 1 (Sept. 2018), 109–128. https://doi.org/10.1525/mp.2018.36.1.109

[19] Robert H. Jack, Tony Stockman, and Andrew McPherson. 2016. Effect of latency on performer interaction and subjective quality assessment of a digital musical instrument. In *Proceedings of the Audio Mostly 2016.* ACM, Norrköping Sweden, 116–123. https://doi.org/10.1145/2986416.2986428

[20] Ricardo Jota, Albert Ng, Paul Dietz, and Daniel Wigdor. 2013. How fast is fast enough? a study of the effects of latency in direct-touch pointing tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13).* Association for Computing Machinery, New York, NY, USA, 2291–2300. https://doi.org/10.1145/2470654.2481317

[21] Topi Kaaresoja, Stephen Brewster, and Vuokko Lantz. 2014. Towards the Temporally Perfect Virtual Button: Touch-Feedback Simultaneity and Perceived Quality in Mobile Touchscreen Press Interactions. *ACM Transactions on Applied Perception* 11, 2 (July 2014), 1–25. https://doi.org/10.1145/2611387

[22] Nelson Posse Lago and Fabio Kon. 2004. The Quest for Low Latency: 30th International Computer Music Conference, ICMC 2004. *International Computer Music Conference, ICMC Proceedings* (2004). http://www.scopus.com/inward/record.url?scp=85105778604&partnerID=8YFLogxK

[23] Michael Lester and Jon Boley. 2007. The Effects of Latency on Live Sound Monitoring. Audio Engineering Society. https://www.aes.org/e-lib/browse.cfm?elib=14256

[24] A. W. Mills. 1960. Lateralization of High-Frequency Tones. *The Journal of the Acoustical Society of America* 32, 1 (Jan. 1960), 132–134. https://doi.org/10.1121/1.1907864

[25] Teemu Mäki-Patola. 2005. Musical Effects of Latency. *Musical Effects of Latency* Suomen musiikintutkijoiden symposiumin satoa (2005), 82–85.

[26] Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. 2014. The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLOS ONE* 9, 2 (Feb. 2014), e89642. https://doi.org/10.1371/journal.pone.0089642 Publisher: Public Library of Science.

[27] Ruijing Ning, Samuel J. Trosman, Andrew T. Sabin, and Beverly A. Wright. 2019. Perceptual-learning evidence for inter-onset-interval- and frequency-specific processing of fast rhythms. *Attention, Perception, & Psychophysics* 81, 2 (Feb. 2019), 533–542. https://doi.org/10.3758/s13414-018-1631-7

[28] R. A. Rasch. 1979. Synchronization in Performed Ensemble Music. *Acta Acustica united with Acustica* 43, 2 (Sept. 1979), 121–131.

[29] A. A. Sawchuk, E. Chew, R. Zimmermann, C. Papadopoulos, and C. Kyriakakis. 2003. From remote media immersion to Distributed Immersive Performance. In *Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence (ETP '03)*. Association for Computing Machinery, New York, NY, USA, 110–120. https://doi.org/10.1145/982484.982506

[30] Andreas Schmid, David Halbhuber, Thomas Fischer, Raphael Wimmer, and Niels Henze. 2023. Small Latency Variations Do Not Affect Player Performance in First-Person Shooters. *Proceedings of the ACM on Human-Computer Interaction* 7, CHI PLAY (Oct. 2023), 381:197–381:216. https://doi.org/10.1145/3611027

[31] Nathan Schuett. 2002. The Effects of Latency on Ensemble Performance. (May 2002), 56.

[32] Laurent S. Simon, Arthur Vimond, and Emmanuel Vincent. 2013. Effects of audio latency in a disc jockey interface. Montreal, Canada, 035037. https://doi.org/10.1121/1.4800187

[33] Martin Taylor and Douglas Creelman. 1967. PEST: Efficient Estimates on Probability Functions. *The Journal of the Acoustical Society of America* 41 (April 1967), 782–787. https://doi.org/10.1121/1.1910407

[34] Pekcan Ungan and Suha Yagcioglu. 2014. Significant variations in Weber fraction for changes in inter-onset interval of a click train over the range of intervals between 5 and 300 ms. *Frontiers in Psychology* 5 (Dec. 2014). https://doi.org/10.3389/fpsyg.2014.01453

[35] Raphael Wimmer, Andreas Schmid, and Florian Bockes. 2019. On the Latency of USB-Connected Input Devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300650