Small Latency Variations Do Not Affect Player Performance in First-Person Shooters

ANDREAS SCHMID*, University of Regensburg, Germany
DAVID HALBHUBER*, University of Regensburg, Germany
THOMAS FISCHER, University of Regensburg, Germany
RAPHAEL WIMMER, University of Regensburg, Germany
NIELS HENZE, University of Regensburg, Germany

In interactive systems high latency affects user performance and experience. This is especially problematic in video games. A large number of studies on this topic investigated the effects of constant, high latency. However, in practice, latency is never constant but varies by up to 100 ms due to variations in processing time and delays added by polling between system components. In a large majority of studies, these variations in latency are neither controlled for nor reported. Thus, it is unclear to which degree small, continuous variations in latency affect user performance. If these unreported variations had a significant impact, this might cast into doubt the findings of some studies. To investigate how latency variation affects player performance and experience in games, we conducted an experiment with 28 participants playing a first-person shooter. Participants played with two levels of base latency (50 ms vs. 150 ms) and variation ($\pm 0\,ms$ vs. $\pm 50\,ms$). As expected, high base latency significantly reduces player performance and experience. However, we found strong evidence that small variations in latency in the order of $\pm 50\,ms$, do not affect player performance significantly. Thus, our findings mitigate concerns that previous latency studies might have systematically ignored a confounding effect.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; • **Applied computing → Computer games**.

Additional Key Words and Phrases: latency, gaming

ACM Reference Format:
Andreas Schmid, David Halbhuber, Thomas Fischer, Raphael Wimmer, and Niels Henze. 2023. Small Latency Variations Do Not Affect Player Performance in First-Person Shooters. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 381 (November 2023), 20 pages. https://doi.org/10.1145/3611027

1 INTRODUCTION

System latency, the time between user input and system response, is a inherent property of interactive systems. A system's end-to-end latency is comprised of delays added by hardware, polling rates, processing time, and the time needed to transfer data over a network. As interactions between humans and computers are effectively feedback loops [6], where users constantly react to the system's response, high latency affects user experience and performance. This effect is

*Both authors contributed equally.

Authors' addresses: Andreas Schmid, University of Regensburg, 93053, Regensburg, Germany, andreas.schmid@ur.de; David Halbhuber, University of Regensburg, 93053, Regensburg, Germany, david.halbhuber@ur.de; Thomas Fischer, University of Regensburg, 93053, Regensburg, Germany, thomas1.fischer@stud.uni-regensburg.de; Raphael Wimmer, University of Regensburg, 93053, Regensburg, Germany, raphael.wimmer@ur.de; Niels Henze, University of Regensburg, 93053, Regensburg, Germany, niels.henze@ur.de.

2573-0142/2023/11-ART381
https://doi.org/10.1145/3611027

especially problematic in real-time applications, such as video games [9, 38, 63], or in psychological studies [53–55], where participants' reaction time is measured. With e-sports having become a multi-million dollar business and the replication crisis in psychology [50], understanding and counteracting latency and its effects is more important than ever. Consequently, manufacturers of gaming hardware advertise their products as "low latency" and even include latency-measuring technology directly into the hardware, e.g., *Nvidia Reflex* [59] in gaming monitors.

Research focusing on latency has a long history, starting with MacKenzie and Ware's seminal work in 1993 [49], in which the authors showed that adding latency increases movement time and error rate in pointing tasks. Since then, scientific community and gaming enthusiasts developed a number of methods for measuring the latency of different systems (for example, desktop PCs [7, 28, 33, 57], smartphones [15, 37], virtual reality systems [16], or input devices [71]). In practice, the base latency of a system is not constant. It varies e.g., depending on the polling rate of the USB connection [71], game loops, processing times, and vertical display synchronization. However, in most studies which investigate the effect of latency on user performance, a constant amount of latency is added to the system response for each condition, and only these constant values are reported. Rarely do authors measure or report how large the latency variations of their setup are.

This might be a problem. In case latency variation has a noticeable effect on users' performance in real-time applications, consequences would be significant for several sub-disciplines of human-computer interaction and psychology, including games research. As it is present in all interactions between humans and computers, latency variation could confound results of user studies investigating time-critical phenomenons. It could therefore, if not controlled for by thoroughly measuring the apparatus' latency distribution, invalidate findings in the worst case. Therefore, effects of latency variation on users could contribute to the replication crisis [54, 55], leading to non-existent effects being measured, reported, and published [62].

There is evidence that lower system latency with higher variation affects users more than higher latency with lower variation [13, 26, 69], as users can compensate for known constant latency by adjusting their behavior. Clicking slightly before anticipated events or aiming slightly ahead of targets are examples of such compensation strategies. However, studies investigating varying latency frequently use blocks of constant latency with sudden changes or large latency ranges [13, 26, 69]. In practice, latency variation is much more subtle with standard deviations clearly below 100 milliseconds around the mean latency [30]. Currently, it is unknown how small variations in local latency, as they occur outside of a laboratory setting, affect user experience and performance.

In this work, we investigate how small variations of local latency (latency jitter) influence user experience and performance in gaming. *Latency jitter* is the variation of local latency that occurs with each input, as opposed to *network jitter*, which is caused by buffering of packages in network communication [74]. We conducted a within-subjects study (n = 28) to investigate how latency jitter influences game experience and performance when playing a fast-paced first-person shooter (FPS). We utilized a FPS since previous work showed that they are particularly negatively affected by latency [9]. To operationalize latency jitter, we varied the level of mean base latency (*low* = 50 ms vs. *high* 150 ms) and the level of variation (*low* = $\pm 0\,ms$ vs. *high* = $\pm 50\,ms$). To maximize internal validity, we used a system optimized for extremely low and constant end-to-end latency for our study apparatus. Our work aims to answer the following research questions:

- RQ_1**:** *"How does base latency influence performance and game experience in first-person shooter games?"*
- RQ_2**:** *"How does latency variation influence performance and game experience in first-person shooter games?"*

RQ_3**:** *"How does the interaction between base latency and latency variation modulate the effects on performance and game experience?"*

The results of our data analysis consolidate previous findings and show that a high latency affects game performance and experience (RQ_1). However, we found no effect of latency variation on neither the players' performance nor their experience (RQ_2). Furthermore, investigating the interaction between base latency and its variation, we found no effect on most of our measures. However, we found that players derived a greater sense of meaning when playing the game with low base latency and high variation compared to a high latency with low variation (RQ_3).

Our findings are crucial to latency and video games research since we show that local *latency jitter* does, generally, not affect performance and experience, thus, validating previous work defining latency as a constant.

2 RELATED WORK

A expanding amount of work addresses latency and its impact on users and video game players. In this section, we first provide an overview of how latency arises in interactive systems and the problems users experience due to latency. We then focus on the impact of latency in video games and its effects on player performance and experience. We highlight previous work investigating latency variability and how volatile latency adversely affects users and players. Lastly, we conclude this section with a summary in which we spotlight why latency inherent variability needs to be accounted for when investigating its effects.

2.1 Sources of Latency

A system's end-to-end latency, the time between user input and system response, is comprised of several partial latencies [7, 8]. When an input event is triggered, for example by physically closing the contacts of a mouse button, an event is transferred from the input device to the computer via USB. However, the input device itself contributes to end-to-end latency as it takes time to scan and de-bounce buttons and since USB polling rates are limited [71]. The input event is registered by the operating system's kernel and passed on to the user space, where input callbacks of application toolkits are triggered. Task scheduling, high system load caused by background applications, as well as input handling of the application toolkit can delay this process [7]. An application, e.g., a video game, then reacts to this input.

In network applications, such as multiplayer games, events also have to be transferred to a server that sends back a response. Depending on the type of connection, bandwidth, and physical distance to the server, network round-trip times can have a significant impact on latency.

Most applications update their state in a loop and re-paint regions if necessary. The latency added by re-painting depends on the used graphics toolkit or game engine, as well as the complexity of the rendered content [58]. This step is highly resource-intensive for modern video games because of complex physics simulations and high-resolution models and textures. Once an image is rendered to a frame buffer, it is sent to a monitor and displayed to the user. In addition to the time it takes to transfer an image to the monitor (display response time [17, 64]), the monitor's refresh rate also contributes to end-to-end latency. A monitor continuously updates its content at a fixed rate. For a 60 Hz monitor, a new image is gradually displayed every 16.67 milliseconds. If vertical synchronization (VSYNC) is enabled, the system's frame buffer is synchronized to the monitor's refresh rate, so images are always drawn from top to bottom. Even though this prevents screen tearing, images might get delayed by up to one monitor refresh cycle in the worst case.

2.2 Effect of Latency in Interactive Systems

As latency delays the feedback loop of users reacting to a system's state and the system responding to users' input, it has a direct influence on task difficulty and task completion time. In an early study on latency in remote control tasks, Sheridan and Ferrel [61] show that task time and the number of open loop movements increase linearly with added delay. MacKenzie and Ware [49] show that latency increases task time and difficulty in pointing tasks. They also incorporate latency as a factor for the *index of difficulty* into the *Fitts' Law* model [22]. Ever since, task time and throughput have become established measures for the effect of latency on atomic actions, such as pointing. For example, Teather et al. [65] compared the effects of latency and spatial jitter in pointing tasks with two different input methods. They found that latency deteriorates throughput and increases movement time significantly more than spatial jitter. Friston et al. [23] investigated the effects of small latencies on *Fitts' Law* and *Steering Law* [2] tasks. Their results are in line with findings in previous studies [49, 52, 65]. Furthermore, Friston et al. found that latency in pointing tasks predominantly increases the time users spend correcting their movements.

The perceivable threshold for latency is strongly dependent on input modality and task. Jota et al. [32] found that the just noticeable difference for latency for touch events is 64 milliseconds on average. Ng et al. [51] found that users can perceive a latency difference of as low as one millisecond when dragging. Kaaresoja et al. [34] used empirical data to establish latency guidelines for different feedback modalities of virtual buttons on capacitive touch screens. They recommend latencies of 5 – 50 ms for tactile feedback, 20 – 70 ms for audio feedback, and 30 – 85 ms for visual feedback. In the studies mentioned above, a custom-built apparatus with extremely low latency was used to measure perception thresholds as accurately as possible.

2.3 Latency in Video Games

As video games are real-time applications, oftentimes requiring quick reactions from players, they are especially prone to the effects of latency. Accordingly, there is a large body of research on how latency influences players, how much latency is tolerable in different genres, and how to counteract latency with predictive models either on the gaming system or in the game [24, 29, 40, 60].

Eg et al. [18] used a simple 2D game to investigate how latency affects players' performance and quality of experience (QoE). In their game, players had to follow a moving circle with the mouse cursor and click it as quickly as possible. In a within-groups study, participants played this game with different amounts of latency added to the the mouse. Higher latency lead to higher task time and lower QoE among participants. Beigbeder et al. [5] investigated the effects of network latency and packet loss on player performance in the first person shooter *Unreal Tournament 2003*. They controlled latency and packet loss for individual players with a network emulator. High latency significantly affected players' precision and kill/death-ratios.

Claypool and Claypool [10] analyzed game genres in terms of susceptibility to latency and found that different latency threshold exist for different genres. They categorize first person shooters as one of the genres that is most susceptible to latency. This is in line with findings from Armitage [4] who determined the latency threshold for players of the first person shooter *Quake 3* to be around 150 milliseconds. However, this study is quite old and from today's standpoint, much lower latency thresholds should be applied. For example, Liu et al. [47, 48] conducted two user studies in which participants playing the first person shooter *Counter Strike: Global Offense* under different latency conditions. They found that latency linearly degrades game experience and players' scores starting at 25 milliseconds and that local latency affects players more than network latency. In similar work, Liu et al. [44] showed that latency negatively influences players' navigation capabilities in

FPS games. Players playing with a higher latency required more time to move their avatar in a game-winning position, than players playing with a lower latency.

Due to the low perception threshold for latency on the one hand, and physical limits for data transmission rates on the other hand [4, 38], new approaches for compensating latency with predictive models are emerging. For example, Halbhuber et al. [27] trained a CNN to compensate 50 milliseconds of latency by predicting the mouse position in a real time strategy game. Their system could significantly improve game experience among participants.

2.4 Latency Variability

If a system's latency is constant, users can compensate by pressing buttons early or clicking in front of a moving target. However, if there are variations in latency, the system's behavior becomes unpredictable, and compensation strategies no longer work. Therefore, a high but constant base latency might be better than a lower latency with high variability. Some studies have investigated the effect of varying latency in different usage scenarios.

Weber et al. [69], for example, investigated the effects of varying system response time (SRT) on user performance, task load, and user experience. In a within-group study, participants were asked to classify e-mails with a GUI dialog system with two levels of SRT variability (low vs. high). In the condition with low SRT variability, system response time after each user input was either 750 or 3000 milliseconds. In the condition with high SRT variability, system response time was randomly selected between 300 and 3000 milliseconds in 450 ms steps. Despite the higher total time on task with low SRT variability due to the higher mean SRT, task execution time was significantly lower in this condition. Weber et al. explain this effect with temporal expectancy, as users can better predict the system's behavior when SRT variability is lower.

Davis et al. [13] investigated the effects of fixed and variable latency on driving performance and mental load in a driving simulator study. They first gathered baseline data in a pre-study where participants drove without added latency. For their main study, Davis et al. compared high constant latency (700 ms) to varying latency (400 − 1100 ms, mean: 700 ms). Latency variation was generated with a sinusoidal function. In both latency conditions, participants performed significantly worse than the baseline regarding lane offset, average velocity, and task load. With varying latency, lane offset was significantly higher than with constant latency. However, the latency condition had no main effect for average velocity, task load, and motion sickness.

It is worth noting that both studies, Weber et al.'s [69] and Davis et al.'s [13], used extremly high values for mean latency and latency variation. Even though they could find that varying latency impairs users' performance stronger than constant latency, this does not mean that their findings apply to typical gaming systems with much lower latency ranges [30].

Halbhuber et al. [26] investigated the effects of regularly switching latency on game experience and player performance in a browser-based 2D shoot-'em-up game. During a gaming session, latency was either constant or switched between zero and 33 or 66 milliseconds of artificially added latency in different frequencies. They found that overall, the negative effect of latency on game experience and player performance was stronger with switching latency than with a constant high latency.

2.5 Summary

Numerous studies have shown that latency directly affects task difficulty and task time – and in turn deteriorates users' performance [23, 49, 65]. This effect can be modeled accurately for atomic tasks, such as touch-based pointing [32] or target selection with a computer mouse [43]. There is strong empirical evidence that this effect also applies to more complex real-time applications such as video games [4, 10, 24, 45, 47]. However, the impact of latency in video games strongly depends on the game's genre, with fast-paced, dexterity-based games, such as FPS games or racing games, being affected more severely than strategy games [10].

In most studies, latency is assumed to be constant and constant amounts of delay are applied to users' actions to simulate higher system latency. Only a few studies have investigated the effects of varying latency. In those studies, an effect of latency variation on user performance could be found [13, 26, 69]. However, extremely high base latencies and latency ranges were used, so those results are not directly applicable to real-life gaming systems with much lower latency variation [30]. In conclusion, previous work does not answer how low-value latency variation with a high ecological foundation affects performance and experience.

3 APPARATUS: GAME MODIFICATION AND LOCAL LATENCY

In our work, we used the open-source game *Cube 2: Sauerbraten*[1], a fast-paced first-person arena shooter developed in 2004. In *Sauerbraten*, players control an avatar equipped with a virtual weapon. The game's goal is to navigate the avatar through different levels and shoot other entities, such as other players or AI-controlled bots, to survive and gain points. Despite its age, *Sauerbraten* enjoys a lively and consistent community, which recently even launched an official *Steam* fork of *Sauerbraten* called *Tomatenquark*[2]. The rationale to use *Sauerbraten* in our work is threefold: Firstly, FPS games, such as *Sauerbraten*, are susceptible to latency as demonstrated by previous work [9, 47, 48]. In FPSs, latency leads to players being less accurate, scoring fewer points, and having a reduced gaming experience. Secondly, *Sauerbraten* is an open-source project which allows us to modify and control every aspect of the gaming session, such as what weapons players are allowed to use, how AI-controlled bots behave, and which maps are played. Furthermore, in contrast to proprietary video games, we have direct access to the game's source code which makes low-level logging of game events straightforward. Thirdly, preliminary tests have shown that *Sauerbraten* is highly performant and has a very low impact on the system's end-to-end latency. High game performance is crucial in our work since we aim to investigate low amounts of latency variation. Hence, every fluctuation, for example, induced by a game with high demand on system resources, may potentially bias our work.

In the following section, we first highlight our modification to the game to make it fit to be used in a study with high internal validity. Then, we elaborate how we measured the local latency of our setup since the local latency needs to be factored in all future investigations.

3.1 Modification of Cube 2: Sauerbraten

Sauerbraten is typically played against other humans or multiple bots. However, to prevent differing player skills and play styles to confound our work, we modified the game so one player only faces one bot at a time. We used *Sauerbraten's* built-in level 75 bots[3] which corresponds to a medium to hard difficulty. We set the difficulty level of the bot in a way that players are neither under- nor over-challenged by it.

As the map and the player's weapon fundamentally change how the game is played, we restricted both for our study to prevent confounding effects. We restricted rounds to the in-game map *Teahupoo*. Furthermore, we disabled all virtual weapons except the standard pistol which we modified to have unlimited ammunition, and prevented players from changing or picking up new weapons during the gaming session.

[1] http://sauerbraten.org/
[2] https://store.steampowered.com/app/1274540/Tomatenquark/
[3] bot difficulty ranges from 50 to 101

In the next step we removed all keyboard shortcuts which are not essential for our work, such as using a medikit or opening menus such as the map overlay. Lastly, we added custom logging functions to track different game events. We logged the number of shots, hits, misses, players deaths, and bot deaths. Figure 1 shows an aerial view of the game map *Teahupoo* (right) and the in-game view of the player (left).

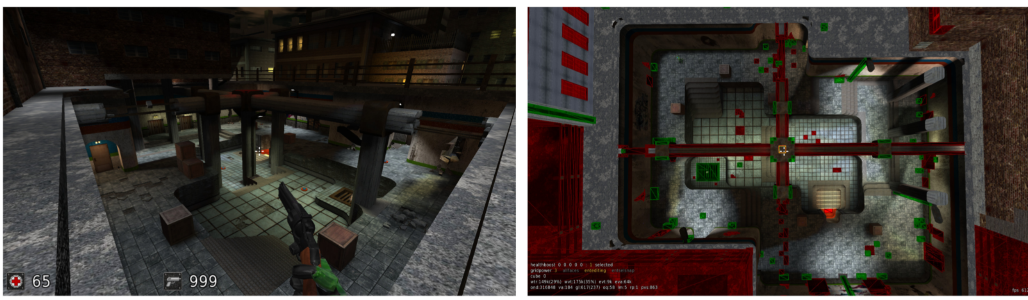

Fig. 1. Shows two screenshots from the first-person shooter game *Cube 2: Sauerbraten*. The left depicts the player's viewport while playing. The screenshot shows the player's weapon, health and ammunition. The right shows an aearial view of the in-game map *Teahupoo* which was used for all gaming rounds.

3.2 Local Latency

To investigate the effects of latency and latency variation on players of our modified FPS game, we needed *(a)* a way to reliably add latency to the test system, and *(b)* a system with very low and constant end-to-end latency.

To add latency to the system, we used a *C* program using the *evdev*[4] library to capture and block input events from physical input devices, similar to Liu and Claypool's *EvLag* [42]. For each input event from a mouse or keyboard, the program creates a thread which waits for a specified amount of time before invoking the input event with a virtual input device provided by *evdev*. The program allows for adding constant or varying delays with uniform or normal distribution. The process is illustrated as pseudo code in listing 1. Additionally, the program can be controlled via inter-process communication using a FIFO queue so added latency can be adjusted between conditions without needing to restart the program.

Listing 1. Pseudo code for the delayed input events.

```
func delayed_event(delay_time, keycode, value):
    wait(delay_time)
    emit(virtual_input_fd, keycode, value)

loop:
    keycode, value = read(input_fd)
    delay_time = random_uniform(min_delay, max_delay)
    t = thread(delayed_event, delay_time, keycode, value)
    t.start()
```

[4]https://linux.die.net/man/4/evdev

For our study, we used a *HP Pavillon Gaming 790* desktop PC[5] running *Debian Buster 5.10* with proprietary *Nvidia* graphics drivers (version 470.103.01). In terms of periphery, we used an *ASUS ROG Strix XG248Q* at 1920 × 1080 pixels with 240 Hz, a *Logitech G15* gaming mouse[6], and a *Logitech G213* gaming keyboard[7].

We measured the end-to-end latency of our system with Schmid and Wimmer's *Yet Another Latency Measuring Device* (YALMD) [57], an Arduino-based device which electrically triggers a button click on an input device and measures the time until a brightness change on the computer's monitor is detected with a photo sensor. We used YALMD to trigger a mouse click leading to a gun shot in *Cube 2: Sauerbraten* and attached the photo diode at the center of the screen so a visible muzzle flash would stop the latency measurement. This way, we could also validate that our method for adding latency to the system is accurate and does not introduce unwanted additional latency. The measured end-to-end latency of our system running *Cube2: Sauerbraten* is between 6.2 and 15.5 ms (M = 9.11 ms, *SD* = 1.4 ms). Detailed measurement results are depicted in Fig. 2. All latency values reported in the remainder of the paper incorporate this local latency without explicitly mentioning it.

4 INVESTIGATING THE EFFECTS OF VARYING LATENCY IN FIRST-PERSON SHOOTER GAMES

To investigate how varying latency affects game experience and player performance in video games, we conducted a within-subjects study with 28 participants playing our modified version of *Cube 2: Sauerbraten*. Participants played with two levels of mean base latency and two levels of latency variation.

The negative effect of high latency on game experience and player performance has been shown in numerous studies [4, 5, 10, 18, 44, 47, 48]. Therefore, we expect to measure the same effect in our study. Furthermore, previous findings [13, 26, 69] suggest that high latency variation affects game experience and player performance more than constant latency with the same mean, as players can adapt their behavior to compensate constant latency. Concluding, we hypothesize that both, high base latency and high latency variation affect game experience and player performance.

4.1 Study Design

In our study we utilized two independent variables (IV) in a 2 × 2 within-design to control for mean base latency and latency variation: (1) Base refers to the mean baseline latency participants played with. Base has two levels: (I) *low* which refers to playing with 50 ms, and (II) *high* which refers to playing with 150 ms of base mean latency. The second IV is Variation and defines how much the actual latency varied around the mean base latency Base. Variation has also two levels: The first level (I) *low* refers to no variation. The second level (II) refers to a variation of ±50 ms. This entails that the actual latency, for example, in a *high* Base / *high* Variation round varied from 100 ms to 200 ms (150 ms ± 50 ms). Latency was applied to each input of the computer mouse (movement and clicks) and keyboard following a uniform distribution. Therefore, during conditions without Variation, all input events are delayed by a constant amount. During conditions with Variation, random delays are added to each event, resulting in jittering mouse movement. In those conditions, it is also possible that the order of rapid consecutive input events changes if high latency is applied to the first event and low latency is applied to the second event. The levels of Base are in line with previous work, which shows that latency in the wild reaches values up to 150 ms [30, 31].

[5] Intel i7-8700 (3.2 GHz), Nvidia GTX 1080, 16 GB DDR4 RAM

[6] input device latency: 2.17 $ms (SD : 0.3\,ms)$ as reported by Wimmer et al. [71]

[7] input device latency: 2.55 $ms (SD : 0.34\,ms)$

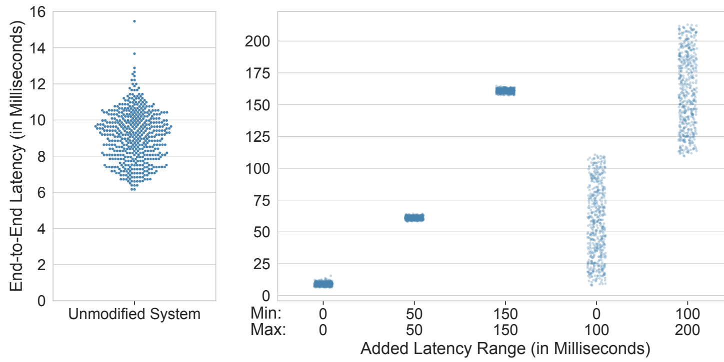

Fig. 2. Results of end-to-end latency measurements for different latency conditions. The plot on the left hand side depicts the system's end-to-end latency running "Cube 2: Sauerbraten" without any added latency. The plot on the right hand side shows the system's end-to-end latency for the different conditions used in our study. Each measurement series consists of 500 individual measurements with random delays in between. The measuring probe was attached at the top left corner of the monitor and VSYNC was disabled.

However, the levels of Variation are constrained by the chosen levels of Base since we are not able to decrease latency below 0 ms. Thus, the lower bound of Base defines the upper bound of Variation.

As our method for measuring the system's end-to-end latency requires measuring probes to be attached to the input devices' circuit boards, as well as bright flashing regions on the screen, we did not measure the system's latency during the study. Therefore, we validated all latency conditions beforehand with the method described in section 3.2. Results can be seen in Fig. 2.

To measure the players' performance and game experience, we utilized a range of dependent variables (DVs). In line with previous work, player performance is measured in three DVs: (1) *Hitrate* [24] – which quantifies the ratio of total shots to successful hits, (2) *KD-Ratio* [25] – which refers to the ratio of player kills and deaths, and lastly (3) *TotalKills* [46] – which corresponds to the total amount of enemy kills per round.

To quantitatively evaluate participants' game experience, we used the 30-item *Player Experience Inventory* (PXI) [1, 66]. We used the PXI, since the instrument was rigorously validated and tested in in multiple studies [66]. Given its multi-dimensionality the instrument allows for an in-depth analysis of player experience, contrary to other work, which, for example, uses single-item questionnaires to assess game experience. The PXI is divided into two categories: (1) functional consequences and (2) psycho-social consequences. The functional consequences dimension is built by five subscales *Ease of Control*, *Progress Feedback*, *Audiovisual Appeal*, *Clarity of Goals*, and *Challenge* and encompasses essential game aspects such as gameplay mechanics, controls, and audio-visual elements. Latency substantially impacts these functional components. For instance, a delay in player actions being registered due to high latency can result in a diminished sense of control [72], a reduced responsiveness [5] or a feeling of an inappropriate challenge [26], ultimately leading to a less satisfying gameplay experience.

The psycho-social consequence dimension of the PXI delves into the social and psychological ramifications of gaming. The dimension describes second-order emotional experiences derived from playing and it also contains five subscales: *Mastery*, *Curiosity*, *Immersion*, *Autonomy*, and *Meaning*. Potentially, latency affects each of the psycho-social subscales individually and differently as each of dimension shapes one crucial aspect of the overall gaming experience.

4.2 Procedure

Participants were met and greeted at the laboratory by an experimenter. Participants were not informed about the exact details of the study (to investigate the effects of varying latency), to prevent a bias induced by the participants' expectations [25]. Hence, participants were just told to test a game. Subsequently, participants gave informed consent to our data collection and were briefed on the further course of the study. After we explained the controls and the objective of the game, each participant played six rounds of *Cube 2: Sauerbraten*. Each round lasted for five minutes. The first and last rounds of the study were always played without artificially added latency (Base) or variation (Variation) to control for exhaustion induced performance degradation. In the remaining four rounds, we altered Base and Variation. Each of the four rounds represents one of the unique combination of Base and Variation. The nested rounds with changing Base and Variation were counterbalanced using a balanced Latin Square design to prevent sequencing effects. After each round, participants filled out the PXI on a separate device and had a short break, which allowed us to alter the game for the next round. Upon finishing all six rounds, participants filled out a demographic questionnaire and the study was concluded. In a short debriefing, we informed participants about the exact purpose of the study. We estimated a total duration of one hour for participation.The study was designed, conducted, and analyzed following the research ethics policy issued by our institution and received clearance per the policy[8].

4.3 Apparatus and Task

As apparatus, we utilized the low-latency hardware setup described in section 3.2. Our modified version of *Sauerbraten* was executed in full-screen mode.

In each of the six rounds, the participants controlled an avatar equipped with a virtual weapon. In the game world, participants were free to roam and fought against one AI-controlled bot. The participants' objective in the game was to shoot the adverse bot as often as possible without getting shot by the bot. After shooting a bot three times, the bot died and respawned at a random location in the game world. If the bot hit the players' character three times, the player character died as well, and also respawned at a random location in the game world. Players obtained points for successfully killing an enemy bot. However, they did not lose points if they did not hit the bot or got killed themselves. This overall procedure was repeated six times (four times for each unique combination of Base and Variation and two times with an unaltered game version).

4.4 Participants

Since previous work showed that the effects of latency are reliably detectable with a relatively small number of participants (Halbhuber et al. [26]: 24 participants per condition, Liu et al. [45]: 25 participants), we recruited 28 participants (24 male, four female) using our institution's mailing list and advertisement in a local gaming club. The participants' age ranged from 20 years to 33 years with a mean age of 24.6 years (SD = 3.4 years). Participants' prior experience with FPS games ranged from 10 hours to 18 000 hours, with 2081 hours (SD = 525 hours) on average. Their self-reported skill level on a 10-point Likert-scale ranged from 2 points to 9 points (M = 5.2 points, SD = 2.1 points). Students participating in our study were eligible to obtain one credit point for their course of study as compensation for their participation.

5 RESULTS

In this section, we report the results of our data analysis. We structure this section by IVs instead of DVs for better readability. Additionally, we only report p-values in body text. However, full

[8]https://link-removed-for-review

inferential statistical data can be found in Table 1. We used an alpha level of .05 for all statistical tests (significance assumed if $p < .05$). The collected data were screened for normal distribution using Shapiro-Wilk tests (Gaussian distribution assumed if $p > 0.05$). All measures, except the *Autonomy* subscale of the PXI ($p = 0.632$), showed a violation of normal distribution (all $p < 0.05$). For inferential assessment of non-parametric data for Base (*low* vs. *high*) and Variation (*low* vs. *high*), we applied a rank-aligned 2 × 2 ART-ANOVA [73] with repeated measures on both factors. Analogously, we used a conventional 2 × 2 ANOVA for the analysis of parametric data. The participant's ID was entered as error term in both ANOVAs to account for random variation induced by individual participants. Effect sizes (η_p^2) are interpreted following the recommendation by Field [21].

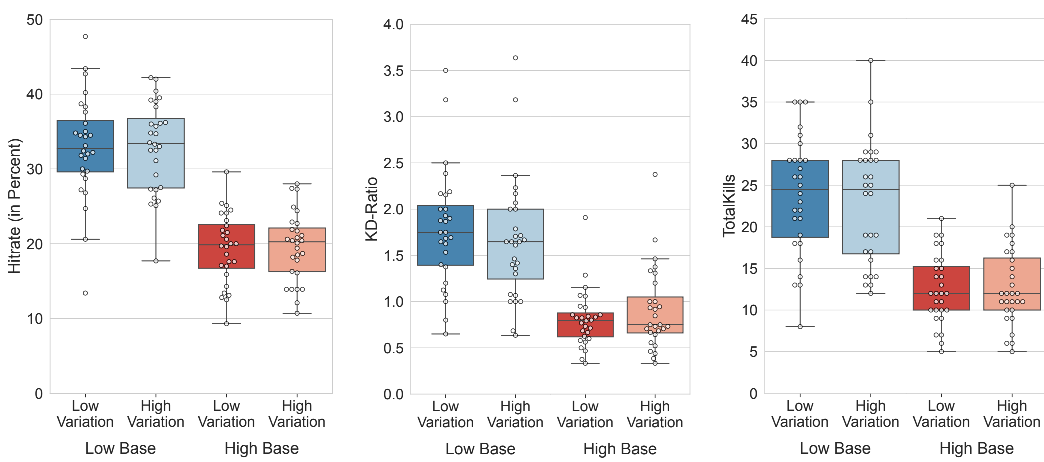

Fig. 3. Depicts boxplots of *Hitrate* (left), *KD-Ratio* (center), and *TotalKills* (right) for each combination of Base and Variation. Participants performed significantly better when playing with *low* Latency than with *high* Latency. They were more accurate, had a better kill-to-death ratio, and killed more bots overall. There was no effect of latency variation on any of the dependent variables in both Base latency conditions.

5.1 Null Hypothesis Testing

5.1.1 Base. ART-ANOVA revealed significant main effects of Base on *Hitrate*, *KD-Ratio*, and *TotalKills* (all $p < 0.024$). Participants performed significantly better when playing with *low* Latency than with *high* Latency. They were more accurate, had a better kill-to-death ratio, and killed more bots overall. Figure 3 depicts *Hitrate* (left), *KD-Ratio* (center), and *TotalKills* (right) grouped by unique combinations of Base and Variation.

ART-ANOVA and ANOVA showed significant main effects of Base on all subscales of the PXI (all $p < 0.015$). Overall, participants had a significantly better game experience when playing with *low* than *high* Base. Participants rated the game as easier to control, were more satisfied with the progress feedback provided by the game, found the game to be more appealing on an audiovisual level, had an easier time grasping the game's goal, and found the challenge provided by the game to be more appropriate when playing with the lower level of Base. Furthermore, players derived a greater extent of mastery, meaning, autonomy, and immersion in the *low* Base conditions.

5.1.2 Variation. We analysed the impact of Variation with the same systematic as Base. ART-ANOVA revealed no significant main effect of Variation on *Hitrate*, *KD-Ratio* and *TotalKills* (all $p >$

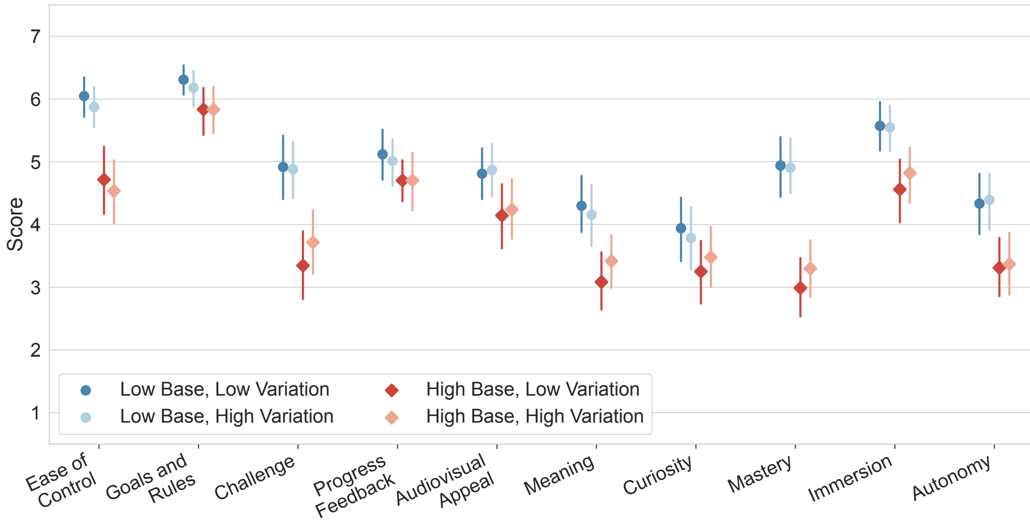

Fig. 4. Depicts the results of *Ease of Control*, *Goals and Rules*, *Challenge*, *Progress Feedback*, *Audiovisual Appeal*, *Meaning*, *Curiosity*, *Mastery*, *Immersion*, and *Autonomy* for each combination of Base and Variation. In general, participants rated the game as easier to control, were more satisfied with the progress feedback provided by the game, found the game to be more appealing on an audiovisual level, had an easier time grasping the game's goal, and found the challenge provided by the game to be more appropriate when playing with the lower level of Base. Furthermore, players derived a greater extent of mastery, autonomy, meaning, autonomy, and immersion in the low Base conditions.

0.365). Furthermore, ART-ANOVA and ANOVA showed no significant main effects of Base on any subscale of the PXI (all $p > 0.131$). We found no significant effect of Variation. Hence, Variation as an isolated factor did not alter players' performance or game experience.

5.1.3 Base × Variation. Lastly, we analysed the interaction Base × Variation and its effects on our measures. ART-ANOVA showed no interaction effect on *Hitrate*, *KD-Ratio*, and *TotalKills* (all $p > 0.157$). Similarly, ART-ANOVA and ANOVA revealed no significant interaction effect on all subscales of the PXI (all $p > 0.174$) except *Meaning*.

ART-ANOVA found a significant interaction effect (Base × Variation) on *Meaning* ($F(1, 27) = 5.585$, $p = 0.025$, $\eta_p^2 = 0.17$ / large). To entangle the interaction effect between Base and Variation, we used paired alpha-corrected (Bonferroni) t-tests. T-tests revealed a significant difference between playing with *low* Latency / *high* Variation and *high* Latency / *low* Variation (t(27 = 4.117, p < 0.001, CI95[0.538, 1.605], $d_{cohens} = 0.62$ / medium [12]), but no effect between *low* Latency / *low* Variation and *low* Latency / *high* Variation (t(27 = 1.317, p = 0.198, CI95[-0.099, 0.465], $d_{cohens} = 0.06$] / small) or between *high* Latency / *low* Variation and *high* Latency / *high* Variation (t(27 = 2.352, p = 0.158, CI95[-0.598, -0.067], $d_{cohens} = 0.02$] / small).

Participants playing with *low* Latency / *high* Variation derived a significantly greater level of *meaning* from playing the game ($M = 4.154$, $SD = 1.306$) compared to participants playing with *high* Latency / *low* Variation ($M = 3.083$, $SD = 1.239$) (Fig. 4).

Table 1. Results of the Base and Variation ART-ANOVA and ANOVA (*) analysis. Each row represents one dependent variable and its analysis for either main effects of Base and Variation or the interaction effect Base × Variation. We found significant negative main effects of Base on all measures and no effect of Variation. Investigating the interaction, we found that players derived a significant greater level of meaning from playing with low Base / high Variation than playing with high Base / low Variation.

Performance-Measures	Base			Variation			Base × Variation		
	$F(1, 27)$	p	η_p^2	$F(1, 27)$	p	η_p^2	$F(1, 27)$	p	η_p^2
Hitrate	257.091	< **0.001**	0.90	0.085	0.772	< 0.01	0.001	0.951	< 0.01
KD-Ratio	5.671	**0.024**	0.17	0.846	0.365	0.03	0.632	0.433	0.02
TotalKills	211.101	< **0.001**	0.88	0.842	0.371	0.02	2.115	0.157	0.07
PXI Subscale									
Ease of Control	40.280	< **0.001**	0.59	2.431	0.131	0.08	0.026	0.872	0.01
Prog. Feedback	6.712	**0.015**	0.19	< 0.001	0.991	< 0.01	0.839	0.367	0.01
A.V. Appeal	25.470	< **0.001**	0.48	0.662	0.442	0.02	0.211	0.649	0.01
Goals & Rules	17.602	< **0.001**	0.39	0.126	0.725	< 0.01	0.026	0.872	0.01
Challenge	39.463	< **0.001**	0.59	1.068	0.311	0.03	2.087	0.160	0.07
Mastery	69.694	< **0.001**	0.72	0.221	0.641	< 0.01	1.946	0.174	0.06
Curiosity	8.326	**0.007**	0.23	0.017	0.896	< 0.01	0.958	0.336	0.03
Immersion	31.721	< **0.001**	0.54	0.675	0.418	0.02	0.775	0.386	0.02
Meaning	17.214	< **0.001**	0.39	1.991	0.169	0.06	5.585	**0.025**	0.17
Autonomy*	23.471	< **0.001**	0.93	0.099	0.698	< 0.01	< 0.001	> 0.999	0.01

5.2 Bayesian Inference

To further examine the effects of latency and its variation, we performed multiple Bayesian 2 × 2 RM-ANOVAs with Base (*low* vs. *high*) and Variation (*low* vs. *high*) as factors. Null hypothesis significance testing (NHST) detects differences between distributions, thus, accepting or rejecting a null hypothesis. While useful for detecting statistical differences in data, NHST cannot determine if an insignificant difference indicates similarity between the studied data. To explore similarity in our data, we utilized a Bayesian analysis, which estimates the probability that the null hypothesis (i.e., no differences in distribution) is true, instead of rejecting it, as NHST does [11, 68]. Unlike NHST, Bayesian inference calculates probabilities for both H_0 and H_1.

We used *JASP* [67] and followed the default prior probability distribution recommended by Wagenmakers et al. [68] for Bayesian inference. For post-hoc testing, we used Bayesian t-tests, and corrected the posterior odds for multiplicity using Westfall's approach [14, 70]. To interpret Bayes factors [35, 39], which indicate the strength of evidence for H_0 over H_1, we followed the guideline of Lee and Wagenmakers [41].

5.2.1 Base. A Bayesian 2 x 2 RM-ANOVA found extreme evidence ($0 < BF_{01} < 0.01$, $error = 0.647\,\%$) for a model that supports a true effect of Base on *Hitrate*, on *KD-Ratio* ($0 < BF_{01} < 0.01$, $error = 0.771\,\%$) and *TotalKills* ($0 < BF_{01} < 0.01$, $error = 1.184\,\%$), which indicates that the gathered performance data is at least a hundred times more likely in support of a distribution in which Base alters *Hitrate*, *KD-Ratio*, and *TotalKills*.

Similarly, we found extreme evidence for accepting a hypothesis that postulates a true effect of Base on the subscales of the PXI, *Ease of Control* ($0 < BF_{01} < 0.01$, $error = 0.754\,\%$), *Audiovisual Appeal* ($0 < BF_{01} < 0.01$, $error = 1.317\,\%$), *Goals and Rules* ($0 < BF_{01} < 0.01$, $error = 1.272\,\%$), *Challenge* ($0 < BF_{01} < 0.01$, $error = 2.068\,\%$), *Mastery* ($0 < BF_{01} < 0.01$, $error = 13.441\,\%$), *Immersion* ($0 < BF_{01} < 0.01$, $error = 1.137\,\%$), *Autonomy* ($0 < BF_{01} < 0.01$, $error = 2.772\,\%$), and *Meaning* ($0 < BF_{01} < 0.01$, *error* < 0.01 %) and very strong evidence on the subscales *Curiosity* (BF_{01} = 0.029, *error* = 0.961 %) and *Progress Feedback* (BF_{01} = 0.061, *error* = 0.842 %).

In summary the Bayesian inference of the effects of LATENCY consolidate our previous findings, which demonstrated that the mean latency – BASE – fundamentally altered player performance and gaming experience.

5.2.2 Variation. A Bayesian RM-ANOVA found moderate evidence for a model that implies no true effect of VARIATION on *Hitrate* (BF_{01} = 5.143, *error* = 0.781 %), on *KD-Ratio* (BF_{01} = 5.061, *error* = 1.195 %) and *TotalKills* (BF_{01} = 5.061, *error* = 1.021 %). Post-hoc investigations using Bayesian t-tests corrected for multiple testing within VARIATION (*low* vs. *high*) found moderate evidence that the achieved *Hitrate* (BF_{01} = 6.758, *error* < 0.001 %) and *KD-Ratio* (BF_{01} = 6.556, *error* < 0.001 %), and the amount of total kills *TotalKills* (BF_{01} = 5.059, *error* = 0.843 %) is not altered by the levels of VARIATION.

Similarly, investigating the effects of VARIATION on gaming experience, we found weak to moderate evidence against a model that supports a true effect of VARIATION on nine of the ten PXI subscales (1 < BF_{01} < 5, 1.094 < *error* < 5.152 %), only the investigation of *Meaning* differed and yielded weak evidence in favor of H_1 (BF_{01} = 0.447, *error* = 1.389 %). Post-hoc investigations using Bayesian t-tests corrected for multiple testing within VARIATION (*low* vs. *high*) revealed weak to moderate evidence that the given scores in the subscales – *Ease of Control* (BF_{01} = 1.828, *error* = 0.012 %), *Progress Feedback* (BF_{01} = 6.049, *error* < 0.001 %), *Audiovisual Appeal* (BF_{01} = 5.403, *error* < 0.001 %), *Goals and Rules* (BF_{01} = 5.050, *error* = 0.052 %), *Challenge* (BF_{01} = 3.760, *error* = 0.013 %), *Mastery* (BF_{01} = 4.705, *error* = 0.040 %), *Curiosity* (BF_{01} = 6.653, *error* < 0.001 %), *Immersion* (BF_{01} = 4.226, *error* = 0.025 %), *Autonomy* (BF_{01} = 6.248, *error* < 0.001 %), and *Meaning* (BF_{01} = 4.202, *error* = 0.024 %) - are not influenced by the levels of VARIATION.

5.2.3 Base × Variation. Lastly, we investigated the interaction between BASE and VARIATION. We found weak to moderate evidence that the interaction does not affect *Hitrate* (BF_{01} = 4.476, *error* = 3.299 %) and *TotalKills* (BF_{01} = 2.114, *error* = 5.099 %). However, we found weak evidence that *KD-Ratio* is influenced by the interaction effect (BF_{01} = 0.371, *error* = 3.840 %).

Investigating the effects of BASE × VARIATION on game experience, we found weak to moderate evidence against accepting a hypothesis that postulates a true effect of BASE × VARIATION on nine out of ten subscales of the PXI: *Ease of Control* (BF_{01} 2.334, *error* = 7.086 %), *Progress Feedback* (BF_{01} = 5.006, *error* = 3.187 %), *Audiovisual Appeal* (BF_{01} = 4.079, *error* = 7.095 %), *Goals and Rules* (BF_{01} = 2.935, *error* = 3.384 %), *Challenge* (BF_{01} = 1.521, *error* = 1.742 %), *Mastery* (BF_{01} = 1.586, *error* = 5.785 %), *Curiosity* (BF_{01} = 2.059, *error* = 2.311 %), *Immersion* (BF_{01} = 1.701, *error* = 5.108 %), and *Autonomy* (BF_{01} = 4.338, *error* = 4.42 %). However, investigating the interaction effect on *Meaning* we found moderate evidence for an interaction effect (BF_{01} = 0.11, *error* = 1.911 %).

6 DISCUSSION

The results of our NHST showed that a high base latency significantly affects player performance and gaming experience. However, we found no significant main effect of latency variation on either the player performance or the experience. While most of our measures were unaffected by the interaction between latency and its variation, inferential analysis showed that players rated playing the game with low base latency and high variation better on the PXI's *meaning* subscale than playing with high latency and low variation. The *meaning* subscale, in general, refers to the level of meaning derived from playing the game. It quantifies how well players connected with the game, and how well they were able to resonate with what is important while playing the game on a psycho-social level [1].

We consolidated our findings in regards to base latency (RQ_1) using a Bayesian analysis, which showed that our data is highly in favor of a model that acknowledges latency as a true effector on player performance and game experience (all $0 < BF_{01} <= 0.061$). Furthermore, the analysis revealed moderate evidence that latency variation (RQ_2) does not alter player performance ($5.062 < BF_{01} < 6.758$) and weak to moderate evidence in support of a model that postulates no effect of latency variation on the gaming experience ($1 < BF_{01} < 5$ for nine out of ten PXI subscales). Investigating the interaction between latency and its variation (RQ_3) revealed that our data is in favor of a model that supports no true effect of the interaction on all measures with weak to moderate evidence (1.5 $BF_{01} < 5.1$), except on the *meaning* subscale of the PXI ($BF_{01} = 0.11$).

In this section, we first discuss and contextualize our findings about the effects of high base latency and latency variation on player performance and game experience, shed light on how base latency and its variation interact. We conclude by discussing our study's limitations and possible future work.

6.1 Effects of Base Latency, Latency Variation and the Interaction

Regarding RQ_1, the influence of constant latency on players, our findings brace previous work, demonstrating that high latency leads to a decrease in player performance and game experience. Liu et al. [47], for example, showed that linearly increasing latency from 25 ms to 125 ms in the FPS game *Counter-Strike: Global Offensive* decreases players' accuracy and overall score. Other work, for instance, by Sabet et al. [56], illustrated that high latency also has negative effects on the subjective gaming experience. Our work is in line with both findings regarding performance and experience. In our work, players had a significantly reduced accuracy, total amount of bot kills, and a worse kill-to-death ratio, while simultaneously deriving a significantly lower quality of game experience when playing with high base latency. In line with previous work, we argue that the lack of responsiveness of the game induces the degradation of performance and experience. Playing with a high base latency led to a discrepancy between in- and output and, thus, to a decreased performance and experience.

To answer RQ_2, our analysis of the influence of latency variation yielded no significant results. Using a Bayesian approach, we found up weak to moderate evidence that latency variation does not alter performance and experience. Our findings, thus, are opposite to previous work investigating the effects of network jitter on player performance and game experience. For example, Amin et al. [3] found that network jitter of 100 ms – which is comparable to our variation of $\pm 50\,ms$, significantly increases task completion time in video games, compared to playing without jitter. Similarly, the authors found that network jitter also significantly decreases the overall gaming experience. Since, we were not able to replicate those findings using local latency jittering, our findings indicate that local and network-based jitter manifest their effect on performance and experience fundamentally differently. This is in line with previous work by Liu et al [48], who showed that local and network latency, generally, influence player performance differently.

Regarding RQ_3, the interaction between base latency and its variation, an inferential analysis suggests that all measures except one subscale of the PXI are unaffected by the interaction. A Bayesian analysis consolidates this finding with weak to moderate evidence. The significant interaction effect for the PXI's *meaning* subscale is supported by the Bayesian analysis with moderate evidence. As the evidence we found regarding the interaction effect is rather weak, additional work is required to further investigate this effect. However, the interaction between base latency and latency variation is likely to have very little relevance in most practical use cases.

6.2 Limitations and Future Work

While we found that latency variation does not affect video game players, our study still has some limitations. Firstly, our sample of participants only partially represents the population of interest. Besides the strong gender imbalance among our participants, most of them were computer science students and, thus, not representing a high level of diversity. Future work, should aim to further generalize our findings by investigating short-term latency variation with a more diversified participant pool including players from different ages, educational levels and cultural backgrounds.

On the same note, as latency is especially interesting in the context of e-sports, future research should rigorously control for the participants' skill levels. While we asked participants in our study to self-rate their skill level in playing FPS games, we did not quantitatively asses their actual skill. Self-rated assessments are heavily biased and typically not highly reliable. The participants' skills are of particular interest, since previous work indicates that more experienced players are more likely to perceive small difference in latency [46]. Hence, it is possible that latency variation, as investigated in this paper, does affect expert players, but not players with a lower level of gaming skill. Furthermore, it is also possible that the player's individual skill level not only alters their performance, but also the felt gaming experience. Previous work indicates that players that perform better in a video game experience a higher level of enjoyment, and thus a higher level of gaming experience, compared to players performing worse [36]. Hence, future research should either control player skill more strictly or measure a reliable metric to use it in the work's statistical analysis.

In addition, it is important to recognize our study's limitation presented by the sample size. With a sample of only 28 participants, there are limits to statistical power and precision. Although previous research suggests that latency effects can be detected even with a smaller sample size, the limited number of participants in our study may make it difficult to identify small effects. Therefore, future studies should investigate local latency variation using a larger sample to improve generalizability, reliability, and the probability to cleary detect effects. Hence, future work should perform an a prior power analysis to ensure a certain power level (e.g., $1 - \beta > .90$ [19, 20]) to detect a certain effect (e.g., $\eta_p^2 > 0.2$).

Moreover, we only tested the effects of varying latency for a FPS game. However, latency's effect strongly depends on the game's genre [10]. For example, in fighting games like *Street Fighter*, frame-perfect input is necessary to perform certain actions, such as blocking opponents' attacks, or successfully performing a combo. With a time window of only 16.67 milliseconds (assuming a 60 Hz game loop), latency jitter might affect such actions much more than fighting a bot in a shooting game. Therefore, our findings can not yet be generalized to the broad landscape of video game genres as further studies are needed. Hence, our study could be replicated with other games to learn about the influence of latency jitter in different genres.

7 CONCLUSION

In this paper, we presented the results of a study ($N = 28$) investigating the effects of local latency jitter on player performance and game experience in a FPS game. Participants played with two levels of mean base latency and two levels of latency variation.

Our work contributes to the extending body of work showing that a high latency reduces player performance and game experience. Furthermore, we highlight that latency's variation as an standalone factor does not significantly influence video gaming session. A Bayesian analysis found that latency variation, as well as the interaction between base latency and latency variation, do not alter performance and experience.

Overall, our findings can be seen as a sigh of relief for all past and future latency research, as the often overlooked factor of latency variance seems to have little to no practical relevance on the outcome of latency studies. Thus, our work illustrates that previous findings may not be confounded by not factoring in latency variation. Nevertheless, we recommend that researchers accurately measure and report the latency and latency variation of their apparatûs for better replicability. Additionally, further research is required to investigate under which circumstances and to which degree latency variation can affect users in different scenarios.

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020). https://doi.org/10.1016/j.ijhcs.2019.102370

[2] Johnny Accot and Shumin Zhai. 1997. Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '97)*. Association for Computing Machinery, New York, NY, USA, 295–302. https://doi.org/10.1145/258549.258760

[3] Rahul Amin, France Jackson, Juan E Gilbert, Jim Martin, and Terry Shaw. 2013. Assessing the impact of latency and jitter on the perceived quality of call of duty modern warfare 2. In *International Conference on Human-Computer Interaction*. Springer, 97–106. https://doi.org/10.1007/978-3-642-39265-8_11

[4] Grenville Armitage. 2003. An experimental estimation of latency sensitivity in multiplayer Quake 3. In *The 11th IEEE International Conference on Networks, 2003. ICON2003*. IEEE, 137–141. https://doi.org/10.1109/ICON.2003.1266180

[5] Tom Beigbeder, Rory Coughlan, Corey Lusher, John Plunkett, Emmanuel Agu, and Mark Claypool. 2004. The Effects of Loss and Latency on User Performance in Unreal Tournament 2003®. In *Proceedings of 3rd ACM SIGCOMM Workshop on Network and System Support for Games* (Portland, Oregon, USA) *(NetGames '04)*. Association for Computing Machinery, New York, NY, USA, 144–151. https://doi.org/10.1145/1016540.1016556

[6] Stuart Card, Thomas Moran, and Allen Newell. 1986. The model human processor - An engineering model of human performance. *Handbook of perception and human performance.* 2, 45–1 (1986).

[7] Géry Casiez, Stéphane Conversy, Matthieu Falce, Stéphane Huot, and Nicolas Roussel. 2015. Looking Through the Eye of the Mouse: A Simple Method for Measuring End-to-end Latency Using an Optical Mouse. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 629–636. https://doi.org/10.1145/2807442.2807454

[8] Géry Casiez, Thomas Pietrzak, Damien Marchal, Sébastien Poulmane, Matthieu Falce, and Nicolas Roussel. 2017. Characterizing Latency in Touch and Button-Equipped Interactive Systems. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 29–39. https://doi.org/10.1145/3126594.3126606

[9] Mark Claypool and Kajal Claypool. 2006. Latency and player actions in online games. *Commun. ACM* 49, 11 (2006), 40–45.

[10] Mark Claypool and Kajal Claypool. 2006. Latency and Player Actions in Online Games. *Commun. ACM* 49, 11 (Nov. 2006), 40–45. https://doi.org/10.1145/1167838.1167860

[11] Ton JM Cleophas and Aeilko H Zwinderman. 2018. *Modern bayesian statistics in clinical research.* Springer.

[12] Jacob Cohen. 2013. *Statistical power analysis for the behavioral sciences.* Routledge.

[13] James Davis, Christopher Smyth, and Kaleb McDowell. 2010. The Effects of Time Lag on Driving Performance and a Possible Mitigation. *IEEE Transactions on Robotics* 26, 3 (June 2010), 590–593. https://doi.org/10.1109/TRO.2010.2046695

[14] Tim de Jong. 2019. A bayesian approach to the correction for multiplicity. (2019).

[15] Jonathan Deber, Bruno Araujo, Ricardo Jota, Clifton Forlines, Darren Leigh, Steven Sanders, and Daniel Wigdor. 2016. Hammer Time!: A Low-Cost, High Precision, High Accuracy Tool to Measure the Latency of Touchscreen Devices. ACM Press, 2857–2868. https://doi.org/10.1145/2858036.2858394

[16] Massimiliano Di Luca. 2010. New Method to Measure End-to-End Delay of Virtual Reality. *Presence: Teleoperators and Virtual Environments* 19, 6 (Dec. 2010), 569–584. https://doi.org/10.1162/pres_a_00023

[17] Federico Dossena and Andrea Trentini. 2022. OpenLDAT—A system for the measurement of display latency metrics. *Journal of the Society for Information Display* n/a, n/a (2022). https://doi.org/10.1002/jsid.1104

[18] Ragnhild Eg, Kjetil Raaen, and Mark Claypool. 2018. Playing with delay: With poor timing comes poor performance, and experience follows suit. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 1–6. https://doi.org/10.1109/QoMEX.2018.8463382

[19] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160. https://doi.org/

10.3758/BRM.41.4.1149

[20] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191. https://doi.org/10.3758/BF03193146

[21] Andy Field. 2013. *Discovering statistics using IBM SPSS statistics.* sage.

[22] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381. https://doi.org/10.1037/h0055392

[23] Sebastian Friston, Per Karlstrm, and Anthony Steed. 2016. The Effects of Low Latency on Pointing and Steering Tasks. *IEEE Transactions on Visualization and Computer Graphics* 22, 5 (May 2016), 1605–1615. https://doi.org/10.1109/TVCG.2015.2446467

[24] David Halbhuber, Niels Henze, and Valentin Schwind. 2021. Increasing Player Performance and Game Experience in High Latency Systems. *Proceedings of the ACM on Human-Computer Interaction* 5, CHI PLAY (2021), 1–20. https://doi.org/10.1145/3474710

[25] David Halbhuber, Maximilian Schlenczek, Johanna Bogon, and Niels Henze. 2022. Better be quiet about it! The Effects of Phantom Latency on Experienced First-Person Shooter Players. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*. 172–181. https://doi.org/10.1145/3568444.3568448

[26] 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 (Oct. 2022), 229:1–229:20. https://doi.org/10.1145/3549492

[27] David Halbhuber, Maximilian Seewald, Fabian Schiller, Mathias Götz, Jakob Fehle, and Niels Henze. 2022. Using Artificial Neural Networks to Compensate Negative Effects of Latency in Commercial Real-Time Strategy Games. In *Proceedings of Mensch Und Computer 2022* (Darmstadt, Germany) *(MuC '22)*. Association for Computing Machinery, New York, NY, USA, 182–191. https://doi.org/10.1145/3543758.3543767

[28] Ding He, Fuhu Liu, Dave Pape, Greg Dawe, and Dan Sandin. 2000. Video-Based Measurement of System Latency. *International Immersive Projection Technology Workshop* 111 (July 2000), 6. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.34.1123&rep=rep1&type=pdf

[29] Niels Henze, Markus Funk, and Alireza Sahami Shirazi. 2016. Software-reduced touchscreen latency. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 434–441.

[30] Zenja Ivkovic, Ian Stavness, Carl Gutwin, and Steven Sutcliffe. 2015. Quantifying and Mitigating the Negative Effects of Local Latencies on Aiming in 3D Shooter Games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 135–144. https://doi.org/10.1145/2702123.2702432

[31] Zenja Ivkovic, Ian Stavness, Carl Gutwin, and Steven Sutcliffe. 2015. Quantifying and Mitigating the Negative Effects of Local Latencies on Aiming in 3D Shooter Games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 135–144. https://doi.org/10.1145/2702123.2702432

[32] 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* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2291–2300. https://doi.org/10.1145/2470654.2481317

[33] Topi Kaaresoja and Stephen Brewster. 2010. Feedback is... late: measuring multimodal delays in mobile device touchscreen interaction. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*. 1–8. https://doi.org/10.1145/1891903.1891907

[34] 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 Trans. Appl. Percept.* 11, 2, Article 9 (jun 2014), 25 pages. https://doi.org/10.1145/2611387

[35] Robert E Kass and Adrian E Raftery. 1995. Bayes factors. *Journal of the american statistical association* 90, 430 (1995), 773–795.

[36] Christoph Klimmt, Christopher Blake, Dorothée Hefner, Peter Vorderer, and Christian Roth. 2009. Player performance, satisfaction, and video game enjoyment. In *Entertainment Computing–ICEC 2009: 8th International Conference, Paris, France, September 3-5, 2009. Proceedings 8*. Springer, 1–12.

[37] Teemu Kämäräinen, Matti Siekkinen, Antti Ylä-Jääski, Wenxiao Zhang, and Pan Hui. 2017. Dissecting the End-to-end Latency of Interactive Mobile Video Applications. In *Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications*. ACM, Sonoma CA USA. https://doi.org/10.1145/3032970.3032985

[38] Ulrich Lampe, Qiong Wu, Hans Ronny, André Miede, and Ralf Steinmetz. 2013. To Frag or to Be Fragged - An Empirical Assessment of Latency in Cloud Gaming. SciTePress - Science and and Technology Publications, 5–12. https://doi.org/10.5220/0004345900050012

[39] Michael Lavine and Mark J Schervish. 1999. Bayes factors: What they are and what they are not. *The American Statistician* 53, 2 (1999), 119–122.

[40] Huy Viet Le, Valentin Schwind, Philipp Göttlich, and Niels Henze. 2017. PredicTouch: A System to Reduce Touchscreen Latency Using Neural Networks and Inertial Measurement Units. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. 230–239.

[41] Michael D Lee and Eric-Jan Wagenmakers. 2014. *Bayesian cognitive modeling: A practical course*. Cambridge university press.

[42] Shengmei Liu and Mark Claypool. 2021. EvLag: A Tool for Monitoring and Lagging Linux Input Devices. In *Proceedings of the 12th ACM Multimedia Systems Conference* (Istanbul, Turkey) *(MMSys '21)*. Association for Computing Machinery, New York, NY, USA, 281–286. https://doi.org/10.1145/3458305.3478449

[43] Shengmei Liu and Mark Claypool. 2021. Game Input with Delay–A Model of the Time Distribution for Selecting a Moving Target with a Mouse. In *MultiMedia Modeling: 27th International Conference, MMM 2021, Prague, Czech Republic, June 22–24, 2021, Proceedings, Part I 27*. Springer, 506–518.

[44] Shengmei Liu and Mark Claypool. 2022. The Impact of Latency on Navigation in a First-Person Perspective Game. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–11.

[45] Shengmei Liu, Mark Claypool, Atsuo Kuwahara, James Scovell, and Jamie Sherman. 2021. The effects of network latency on competitive first-person shooter game players. In *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 151–156.

[46] Shengmei Liu, Mark Claypool, Atsuo Kuwahara, James Scovell, and Jamie Sherman. 2021. L33t or N00b? How Player Skill Alters the Effects of Network Latency on First Person Shooter Game Players. (2021).

[47] Shengmei Liu, Mark Claypool, Atsuo Kuwahara, Jamie Sherman, and James J Scovell. 2021. Lower is Better? The Effects of Local Latencies on Competitive First-Person Shooter Game Players. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12. https://doi.org/10.1145/3411764.3445245

[48] Shengmei Liu, Atsuo Kuwahara, James Scovell, Jamie Sherman, and Mark Claypool. 2021. Comparing the Effects of Network Latency versus Local Latency on Competitive First Person Shooter Game Players. (2021). https://doi.org/10.31219/osf.io/bhxcy

[49] I Scott MacKenzie and Colin Ware. 1993. Lag as a determinant of human performance in interactive systems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. 488–493. https://doi.org/10.1145/169059.169431

[50] Scott E Maxwell, Michael Y Lau, and George S Howard. 2015. Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist* 70, 6 (2015), 487. https://doi.org/10.1037/a0039400

[51] Albert Ng, Michelle Annett, Paul Dietz, Anoop Gupta, and Walter F. Bischof. 2014. In the Blink of an Eye: Investigating Latency Perception During Stylus Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1103–1112. https://doi.org/10.1145/2556288.2557037

[52] Andriy Pavlovych and Wolfgang Stuerzlinger. 2009. The tradeoff between spatial jitter and latency in pointing tasks. ACM Press, 187. https://doi.org/10.1145/1570433.1570469

[53] Richard R. Plant. 2016. A reminder on millisecond timing accuracy and potential replication failure in computer-based psychology experiments: An open letter. *Behavior Research Methods* 48, 1 (March 2016), 408–411. https://doi.org/10.3758/s13428-015-0577-0

[54] Richard R. Plant, Nick Hammond, and Tom Whitehouse. 2003. How choice of mouse may affect response timing in psychological studies. *Behavior Research Methods, Instruments, & Computers* 35, 2 (May 2003), 276–284. https://doi.org/10.3758/BF03202553

[55] Richard R. Plant and Philip T. Quinlan. 2013. Could millisecond timing errors in commonly used equipment be a cause of replication failure in some neuroscience studies? *Cognitive, Affective, & Behavioral Neuroscience* 13, 3 (Sept. 2013), 598–614. https://doi.org/10.3758/s13415-013-0166-6

[56] Saeed Shafiee Sabet, Steven Schmidt, Saman Zadtootaghaj, Babak Naderi, Carsten Griwodz, and Sebastian Möller. 2020. A latency compensation technique based on game characteristics to mitigate the influence of delay on cloud gaming quality of experience. In *Proceedings of the 11th ACM Multimedia Systems Conference*. 15–25. https://doi.org/10.1145/3339825.3391855

[57] Andreas Schmid and Raphael Wimmer. 2021. *Yet Another Latency Measuring Device*. preprint. Open Science Framework. https://doi.org/10.31219/osf.io/tkghj

[58] Andreas Schmid and Raphael Wimmer. 2023. Measuring the Latency of Graphics Frameworks on X11-Based Systems. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 204, 7 pages. https://doi.org/10.1145/3544549.3585779

[59] Seth Schneider. 2020. Introducing NVIDIA Reflex: Optimize and Measure Latency in Competitive Games. https://www.nvidia.com/en-us/geforce/news/reflex-low-latency-platform/

[60] Valentin Schwind, David Halbhuber, Jakob Fehle, Jonathan Sasse, Andreas Pfaffelhuber, Christoph Tögel, Julian Dietz, and Niels Henze. 2020. The Effects of Full-Body Avatar Movement Predictions in Virtual Reality Using Neural Networks. In *26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 28, 11 pages. https://doi.org/10.1145/3385956.3418941

[61] T.B. Sheridan and W.R. Ferrell. 1963. Remote Manipulative Control with Transmission Delay. *IEEE Transactions on Human Factors in Electronics* HFE-4, 1 (1963), 25–29. https://doi.org/10.1109/THFE.1963.231283

[62] Joseph P. Simmons, Leif D. Nelson, and Uri Simonsohn. 2011. False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science* 22, 11 (2011), 1359–1366. https://doi.org/10.1177/0956797611417632

[63] Josef Spjut, Ben Boudaoud, Kamran Binaee, Jonghyun Kim, Alexander Majercik, Morgan McGuire, David Luebke, and Joohwan Kim. 2019. Latency of 30 ms Benefits First Person Targeting Tasks More Than Refresh Rate Above 60 Hz. In *SIGGRAPH Asia 2019 Technical Briefs on - SA '19 (SA '19)*. ACM Press, New York, NY, USA, 110–113. https://doi.org/10.1145/3355088.3365170

[64] Patrick Stadler, Andreas Schmid, and Raphael Wimmer. 2020. DispLagBox: simple and replicable high-precision measurements of display latency. In *Proceedings of the Conference on Mensch und Computer (MuC '20)*. Association for Computing Machinery, New York, NY, USA, 105–108. https://doi.org/10.1145/3404983.3410015

[65] Robert J. Teather, Andriy Pavlovych, Wolfgang Stuerzlinger, and I. Scott MacKenzie. 2009. Effects of tracking technology, latency, and spatial jitter on object movement. IEEE, 43–50. https://doi.org/10.1109/3DUI.2009.4811204

[66] Vero Vanden Abeele, Lennart E Nacke, Elisa D Mekler, and Daniel Johnson. 2016. Design and preliminary validation of the player experience inventory. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. 335–341. https://doi.org/10.1145/2968120.2987744

[67] Eric-Jan Wagenmakers. 2022. JASP - A Fresh Way to Do Statistics. https://jasp-stats.org/. Accessed on 2021-01-01.

[68] Eric-Jan Wagenmakers, Jonathon Love, Maarten Marsman, Tahira Jamil, Alexander Ly, Josine Verhagen, Ravi Selker, Quentin F Gronau, Damian Dropmann, Bruno Boutin, et al. 2018. Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic bulletin & review* 25, 1 (2018), 58–76.

[69] Florian Weber, Carola Haering, and Roland Thomaschke. 2013. Improving the human-computer dialogue with increased temporal predictability. *Human Factors* 55, 5 (Oct. 2013), 881–892. https://doi.org/10.1177/0018720813475812

[70] Peter H Westfall, Wesley O Johnson, and Jessica M Utts. 1997. A Bayesian perspective on the Bonferroni adjustment. *Biometrika* 84, 2 (1997), 419–427.

[71] 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*. ACM Press, 1–12. https://doi.org/10.1145/3290605.3300650

[72] Paul Winkler, Philipp Stiens, Nadine Rauh, Thomas Franke, and Josef Krems. 2020. How latency, action modality and display modality influence the sense of agency: a virtual reality study. *Virtual Reality* 24 (2020), 411–422.

[73] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 143–146. https://doi.org/10.1145/1978942.1978963

[74] Li Zheng, Liren Zhang, and Dong Xu. 2001. Characteristics of network delay and delay jitter and its effect on voice over IP (VoIP). In *ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No.01CH37240)*, Vol. 1. 122–126 vol.1. https://doi.org/10.1109/ICC.2001.936286

Received 2023-02-21; accepted 2023-07-07