國立臺灣大學電機資訊學院資訊網路與多媒體研究所
碩士論文
Graduate Institute of Networking and Multimedia
College of Electrical Engineering and Computer Science
National Taiwan University
Master's Thesis

MR-Drum：設計輔助初學者學習打鼓的階段式學習框架與 MR 系統
MR.Drum: Designing Mixed Reality Interfaces to Support Structured Learning Micro-Progression in Drumming

王哲瑋
Che-Wei Wang

指導教授：陳彥仰博士
Advisor: Mike Y. Chen, Ph.D.

中華民國 114 年 1 月
Jan, 2025

doi:10.6342/NTU202500720

國立臺灣大學碩士學位論文
口試委員會審定書
MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

MR-Drum: 設計輔助初學者學習打鼓的階段式學習框架與 MR 系統

MR-Drum: Designing a Micro-Progression Learning Framework and Mixed-Reality System to Support Learning to Drum

本論文係 王哲瑋 （學號 R10944037）在國立臺灣大學資訊網路與多媒體研究所完成之碩士學位論文，於民國 114 年 1 月 6 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Graduate Institute of Networking and Multimedia on 6 January 2025 have examined a Master's Thesis entitled above presented by WANG, CHE-WEI (student ID: R10944037) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

陳彥仰 （指導教授 Advisor）	鄭龍磻	蔡欣叡
	陳炳宇	余能豪

系（所）主管 Director: 鄭卜壬

doi:10.6342/NTU202500720

誌謝

首先，我要感謝我的指導教授陳彥仰教授在撰寫這篇論文的過程中提供了寶貴的建議和討論，並且在論文研究之外也給予我深刻的指引以及幫助我突破我的思考領域的啟發此外，也感謝教授提供了優越的實驗室環境以及打造了正向、互助、積極、多元、幽默有趣的實驗室同儕氛圍。

感謝學長姐以及同儕，碁家丞 (Daniel)、范家安、翁大為 (David)、柯柏丞 (柯博士)、鄭家宇、王舜昱 (Curtis)、張巧如、陳譽、張宇承 (Malik)、呂品君、許育綸 (Eason)、蘇允中 (William)、伍恩輝這幾年提供的啟發、鼓勵、幫助、歡樂、幽默、醫療。

特別感謝呂品君，她對於這個研究在重初期題目發想、中期方向突破、到後期撰寫論文準備投稿，都提供了非常非常非常多的幫助和啟發，可以說沒有呂品君這篇論文就不會存在。

感謝過程中參與實驗的自學鼓手以及鼓手老師們，給予很多有幫助的建議與回饋，以及在實驗結束後仍然提供各種幫助。感謝所有在投稿 CHI 的 reviewer 們，不吝給予本篇論文詳細明確的改進方向。感謝口試委員鄭龍磻教授、余能豪教授和蔡欣叡教授和陳炳宇教授，在學位考試時對報告和論文提出深入而專業的問題，讓我對於自己的研究有更深刻的反思和理解。

此碩士學位論文之大部分研究成果同時發表於人機互動領域頂尖會議—ACM CHI 2025。

最後還是說一句，陳譽加油!!! You are next !!!

doi:10.6342/NTU202500720

摘要

學習打鼓具有挑戰性，因為需要雙手和雙腳同時獨立演奏多種節奏。我們進行了兩項形成性研究，以了解：(1) 專業鼓手教學方法，以及 (2) 鼓手當前的自學方式與痛點。所有受訪教師皆會將複雜的節奏與肢體動作拆解，並透過結構化進程進行教學，然而這種方法至今尚未被 HCI 研究探討。基於這些發現，我們開發了一種全新的微進程學習框架（micro-progression learning framework），專為初學者設計，將理解進程（擊鼓順序與節奏）與肢體協調進程拆分並結構化為 16 個學習階段。此外，我們設計了 MR-Drum，這是一款混合實境（MR）系統，透過第一人稱視角顯示虛擬四肢，以示範節奏、肢體動作及鼓面動態，並提供可調節的節奏速度與自動錯誤檢測功能。一項與教學影片的比較性研究顯示，MR-Drum 顯著提升了錯誤率與節奏準確度，並在理解能力、技能發展與使用者體驗方面獲得顯著較高的偏好，所有受試者皆更偏好 MR-Drum 作為學習工具。

關鍵字：以人為本的運算；互動系統與工具；應用計算；使用者介面設計；教育/學習；使用者體驗設計；虛擬/擴增實境；原型設計與實作

doi:10.6342/NTU202500720

Abstract

Learning drumming is challenging because multiple rhythms must be performed independently and simultaneously using both hands and feet. We conducted two formative studies to understand: 1) professional drumming instructors' teaching methods, and 2) drummers' current self-learning practices and pain points. All instructors deconstructed complex rhythms and limb movements and then used structured progression to teach drumming, which has not been explored by HCI research to date. Based on these findings, we developed a novel micro-progression learning framework for novice drummers that divides and structures *comprehension progression (drum sequence and rhythm)* and *limb coordination progression* into 16 stages. We also designed MR-Drum, a mixed-reality system that provides a first-person view of virtual limbs to demonstrate rhythm, limb, and drum surface dynamics, with adjustable tempo and automatic error detection. A summative user study vs. instructional videos showed that MR-Drum significantly improved error rate and timing accuracy, was significantly preferred for comprehension, skill development, and user experience, and was preferred overall by all participants.

Keywords: Human-centered computing; Interactive systems and tools; Applied computing; User Interface Design; Education/Learning; User Experience Design; Virtual/Augmented Reality; Prototyping/Implementation

doi:10.6342/NTU202500720

Contents

doi:10.6342/NTU202500720

doi:10.6342/NTU202500720

List of Figures

doi:10.6342/NTU202500720

doi:10.6342/NTU202500720

doi:10.6342/NTU202500720

List of Tables

doi:10.6342/NTU202500720

Chapter 1

Introduction

Learning to play musical instruments requires mastering complex musical structures and coordinating physical movements to perform the music. Most popular instruments, such as the piano, guitar, violin, saxophone, and flute, primarily rely on the hands and the fine motor skills of the fingers, whereas instruments like pipe organs and drums demand significant coordination between the hands and feet. Drumming presents a unique challenge in that it requires the coordination of hands and feet, each maintaining independent rhythms, along with frequent interleaving of the right and left hands on the same drum surfaces. This unique combination of rhythmic multitasking and limb coordination makes drumming a particularly challenging instrument to learn, even for players who already have experience in other popular instruments.

A fundamental approach to learning musical instruments is to break down the music into more manageable subdivisions, facilitating learning and practice [24, 21]. In musical notation, these subdivisions are represented by bars—segments of music bounded by vertical lines. A common technique in music education is to focus on one or more bars at a time, enabling learners to develop their skills incrementally before integrating the sections back into the complete piece.

For drumming, drum set patterns can be deconstructed into rhythm progression and limb progression, which are key methods in rhythm and drumming education [33, 28]. These techniques break down the coordination of independent rhythms and limb movements into smaller, manageable steps, making complex music easier to learn.

doi:10.6342/NTU202500720

There has been extensive and exciting research in HCI for learning musical instruments, supporting different phases of musical performance using techniques such as visualization [14, 1, 15], haptic feedback [9, 12], and auto-detection of errors [14, 1, 15]. However, prior work has treated music as a whole, without addressing rhythm or limb progression. This is particularly crucial for drumming novices, as even a single bar of music can be challenging to comprehend and perform due to the multiple independent rhythms and precise timing required for hands and feet.

To understand how professional drumming instructors structure learning progressions for novice drummers, we conducted a formative study with 8 professional instructors. The study focused on: (1) their teaching methods and (2) the common challenges students face, along with the techniques and exercises used to help students overcome these difficulties. The instructors described teaching methods for novice to intermediate players that emphasized comprehension before practicing on the full drum set, gradually progressing from simplified rhythms, drums, and limb movements to full rhythm and limb coordination. Through this process, we collected a total of 82 exercises from the instructors and developed a novel micro-progression framework with two dimensions: *comprehension progression (drum sequence and rhythm)* and *limb progression* encompassesing 16 structured stages, as shown in Figure **??**.A.

To inform system functionality and UI design, we conducted a second formative study to understand the content, tools, and challenges associated with self-learning practices across a diverse range of drumming experiences. All 8 drummers reported a preference for self-learning by watching and following instructional videos on YouTube. However, participants identified several key pain points with this approach: (1) third-person perspectives misalign with users' limb movements and physical drum sets; (2) video playback speed is controlled by percentage (e.g., 75% or 50%) rather than beats per minute (BPM) used by music; (3) videos lack the ability to isolate specific elements, such as focusing solely on the main rhythm without ornaments; and (4) there is no error detection or feedback, requiring users to self-monitor, which is error prone, or to record and review their practice, which consumes valuable practice time.

doi:10.6342/NTU202500720

Based on the micro-progression framework and the UI implications from the formative studies, we designed MR-Drum, a mixed-reality micro-progression learning system that uses virtual limbs to demonstrate rhythm, limb, and drum surface dynamics. It provides a visual metronome for tempo adjustment by BPM, and supports automatic, real-time error detection and visualization.

In a comparative user study (n=12), MR-Drum significantly improved comprehension, skill development, and drumming performance vs. instructional video-based learning methods, and was preferred overall by all participants.

In summary, our key contributions are as follows:

1. We introduce a novel micro-progression framework for learning drumming, drawing on the instructional methods of professional drumming instructors. This framework encompasses 16 structured stages along two dimensions: *rhythm* and *limb* micro-progression.

2. We used a user-centered design process to inform and iteratively develop MR-Drum, a mixed-reality system that applies our micro-progression framework to provide learners with guided, first-person perspective demonstrations of drumming rhythms and limb coordination.

3. We empirically validated the effectiveness of MR-Drum's micro-progression approach, finding significant improvements in learners' comprehension, skill development, and performance accuracy compared to standard video-based instruction.

4. We have open-sourced MR-Drum[1], enabling other researchers and practitioners to experience and build upon our progress to extend to other musical instruments.

[1]Github repo URL anonymized for review.

doi:10.6342/NTU202500720

Chapter 2

Background on Drum Sets/Kits and Notations

A standard drum set (or kit), often referred to as a basic five-piece drum set, combines cymbals, drums, and tom-toms to provide rhythmic foundation and tonal variety. A standard drum set consists of 8 playable components as shown in Figure 2.1(A).

Drummers use both hands and feet simultaneously to produce a variety of tones and rhythms. For tone control, the hands are tasked with striking up to seven drum kit components, necessitating frequent switching and movement between them, as illustrated in Figure 2.1(A) as orange (left hand) and blue (right hand) regions. The right foot controls the bass drum, while the left foot operates the hi-hat pedal, which serves both rhythmic and tonal functions. The hi-hat pedal can adjust the hi-hat’s tone, similar to the tonal control provided by the piano’s left pedal. However, left foot movements are typically not required at the beginner level.

Drum notation uses a five-line staff, similar to traditional music notation, where each line, space, and symbol corresponds to a specific part of the drum kit. Instead of indicating pitch, the notation maps instruments like the hi-hat, snare drum, bass drum, and toms to designated positions on the staff, as shown in Figure 2.1(B). For example, X-shaped note heads typically represent cymbals, while standard round note heads indicate drums. The sequence of left and right hand movements is indicated by “L” and “R” above the notes,

doi:10.6342/NTU202500720

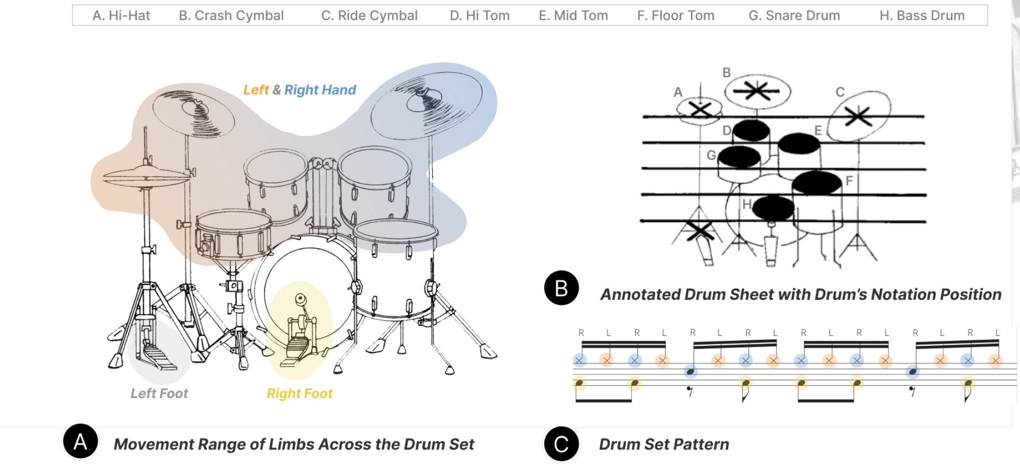

Figure 2.1: Drum set/kit and drum sheet/notation: (A) Components of a standard drum set and the limbs responsible for striking them, color coded as orange (left hand), blue (right hand), grey (left foot), and yellow (right foot); (B) Drum notation uses a five-line staff, similar to traditional music notation, where each line, space, and symbol corresponds to a specific part of the drum set; and (C) drum sheet specifying the sequence and timing of drum components being played, with the sequence of left and right hand movements indicated by "L" and "R" above the notes, which we color coded to show the limbs used for each note.

as shown in Figure 2.1(C).

doi:10.6342/NTU202500720

Chapter 3

Related Work

3.1 Instrument Learning Progression

Taking progressive steps is crucial for mastering a music piece, as attempting to learn the entire composition at once can be overwhelming. Several practice methods share this core principle, including sectioning and integration [21], as well as the isolation of specific elements [24]. In ”Rhythmic Training” [33], a comprehensive guide to rhythm exercises across various levels of complexity, the approach begins with establishing a solid quarter-note pulse and gradually progresses towards the target rhythm through systematic subdivision and the selective removal of notes. This method facilitates the internalization of rhythm by breaking it down into simpler, more manageable exercises, enabling musicians to develop a more intuitive understanding of complex rhythmic patterns. Not only rhythm but also limb coordination development requires a progressive breakdown. Instruments like the pipe organ and drum set demand the coordinated use of both hands and feet, embodying the challenging concept of interdependence. In the classic drumming textbook ”The Art of Bop Drumming” [28], this challenge is addressed through four-way coordination exercises that decompose each bar with a focus on both rhythm and limb micro-progression to build interdependence effectively.

There has been extensive and exciting research in HCI for learning musical instruments in general, that focuses primarily on visualization, haptic feedback, and auto-detection of errors. For instance, haptic guidance methods such as electrical muscle stimulation (EMS)

doi:10.6342/NTU202500720

and pneumatic artificial muscles have been employed to train percussion techniques, emphasizing rhythm and stroke strength [6, 10, 17, 9]. Visual feedback-based systems like P.I.A.N.O. provide live performance evaluations, while other systems deliver real-time instructions via screen displays [14, 11], projection methods [35, 1], or even by integrating haptic feedback with mixed reality (MR) in drum games [15]. Some approaches also aim to internalize rhythm comprehension for percussion instruments through the combined use of haptic, visual, and auditory cues [12]. However, prior work has focused on treating music as a whole, without adopting progressive steps, which is particularly important for drums when even a single bar of music can be challenging to comprehend and perform because of the multiple independent rhythms for hands and feet and limb micro-progression and rhythm micro-progression become essential. Therefore, MR.Drum is the first interactive music system to apply structured learning with micro-progression techniques for drumming.

3.2 MR Instrument Learning

Learning a musical instrument involves developing complex motor skills, particularly in coordinating limb movements, which MR (Mixed Reality) significantly enhances by providing a first-person perspective. MR allows learners to closely observe and imitate professional movements[18], offering a unique advantage over traditional methods that rely on external screens. By immersing learners in a virtual environment, MR helps map movements directly to real instruments, facilitating more effective development of limb coordination and muscle memory.

In guitar learning, for example, the focus is on chord shapes. Players can learn correct hand postures by aligning their hands with static virtual hands displayed on their head-mounted displays (HMDs) [31]. In contrast, interactive piano learning systems emphasize fine motor skills, employing dynamic virtual hands to support group learning. Students can observe the teacher's hand movements through their HMDs [3], learn from expert performances from a first-person perspective [16], or train with visual feedback on discrepancies in their hand postures [18]. For violin players, who require both fine motor

doi:10.6342/NTU202500720

skills and broader arm motions, a 3D avatar in an AR environment has been proposed to enhance the sense of presence and performance quality [4].

Drumming requires the simultaneous coordination of up to three limbs, relying heavily on gross motor skills, making MR particularly suitable for imitation learning. Systems such as ”Halvatar” [26] and Teach Me Drums [22] have applied VR to drum learning, but these approaches do not allow learners to see the real drum set, making it difficult to accurately translate virtual movements to the physical drums. Additionally, these systems do not incorporate foot movements, which are crucial in full drum set practice. In contrast, MR.Drum integrates a 3D first-person view with real-time visualization of both the real drums and virtual limbs, supporting 4-way coordination, including foot movements, to offer a more complete and effective learning experience.

doi:10.6342/NTU202500720

Chapter 4

Formative Study #1: Professional Instructors

The first of two formative studies aimed to understand the teaching methods, techniques, and best practices of professional instructors, in order to help design a self-learning progression framework for novice drummers. It focused on two aspects: 1) teaching methods and tools used and 2) common challenges students encounter and the corresponding teaching techniques and exercises used to help students overcome them.

4.1 Study Design, Procedure, and Participants

4.1.1 Study Design and Procedure

The research method we used was a live teaching demonstration followed by semi-structured interviews, with participants using the think-aloud method to explain their reasoning at each step. Instructors demonstrated their teaching methods using sample drum set patterns we provided as well as their teaching materials. We asked instructors to list common obstacles that their students encounter and then demonstrate the sequence of techniques that instructors would use to help students overcome each of those obstacles. The interview transcripts and observation notes taken by researchers were analyzed using thematic analysis [2] to systematically identify and interpret patterns (themes) in the data.

doi:10.6342/NTU202500720

Each session began with a study overview and demographic information, followed by a teaching demonstration and interviews. After completing the session, the participants received nominal compensation for their participation. Each session lasted approximately 60 minutes.

4.1.2 Participants

We recruited 8 professional drumming instructors (T1-T8), with ages ranging from 22 to 38 (M = 28.5, SD = 5.7), comprised of four males and four females, via posting on social media and drummer forums. Participants had an estimated drumming experience of 2500 to 7500 hours (M = 4600, SD = 1801), 3 to 15 years of teaching experience (M = 5.4, SD = 4.1), and have taught 10 to 150 students (M = 47.5, SD = 50.6).

4.2 Findings

All instructors' teaching methods followed structured progression, *"breaking down"* (T6) and *"starting with simple rhythms"* (T4) then *"progressively added complexity"* (T7) towards the full drum set patterns. Although the specific set of exercises is tailored and adapted to each student based on their needs and obstacles at the moment, the instructors' overarching teaching process consisted of the following two phases: 1) a *comprehension phase* using a single drum, followed by 2) a *practice phase* on the full drum set that focused on rhythm and limb coordination.

4.2.1 Comprehension Phase

All instructors started by verbalizing the rhythm by mimicking the sounds of drums (i.e., the vocal phrasing of drum patterns [28]), to help students understand the rhythm and its connection to the three key drums. The tempo was slowed down and drum patterns were simplified to only three key drums using the following syllable sounds: 1) "*dong*": for the bass drum, 2) "*da*" : for the snare drum, and 3) "*ts*": for the hi-hat cymbal. When students' timing was incorrect for this exercise, instructors would first use metronomes to

doi:10.6342/NTU202500720

help correct students' internal timing and use clapping to further help guide the timing.

To help students understand limb coordination, instructors simplified the full drum patterns to focus only on upper limb patterns using a single drum (i.e., no foot patterns and no drum transitions). Instructors would first demonstrate the rhythm using both hands and drum sticks on a practice pad, then the student would replicate it on a snare drum. The objectives of this exercise were to demonstrate proper stroke technique and ensure the correct left and right hand sequencing.

4.2.2 Practice Phase

After students had achieved sufficient comprehension, instructors then had them use the full drum set to practice rhythm and limb coordination. Instructors would ask students to play the drum patterns to their best ability, assess and prioritize the problems to address, and then tailor exercises correspondingly.

Tempo progression All instructors slowed down the tempo as needed to *"ensure that each beat was executed correctly"* (T2). Some *"started by slowing temp to 50 BPM from 100 BPM"* (T1) while others *"gradually reduced tempo until students could consistently play correctly, then gradually increased in small increments (5-10 BPM)"* (T4).

Rhythm progression All instructors used a progression of beat resolution from simple to complex, *"starting with the 1, 2, 3, 4 main beats first, then add subdivision 1/2 beats and 1/4 beats"* (T6, T7), using the vocalization of 1/2 beats (eighth note) and 1/4 beats (sixteenth note) that were *"1& 2& 3& 4&, pronounced 'one and, two and, three and, four and', and 1e&a 2e&a 3e&a 4e&a"* (T4), respectively.

When students had rhythm errors, instructors *"focused on correcting mistakes on the main downbeats first, ignoring minor offbeat errors initially"* (T1), and *"simplified rhythm by removing 1/4 beats (sixteenth notes) first, then 1/2 beats, leaving the 4 main beats as the foundation for practice"* (T8).

doi:10.6342/NTU202500720

Limb coordination progression Instructors emphasized *"starting with simple, repetitive rhythms, and then slowly introduce more difficult ones"* (T2, T4, T5) by *"overlaying more variable and challenging rhythms on top"* (T7) to gradually increase complexity. All instructors used the right hand, which plays the hi-hat that typically has consistent and repetitive rhythms, to establish such stable foundation to help students with their internal timing. *"Hi-hat is typically used as the foundational element in practice and is rarely omitted, as it is key to maintaining stability"* (T5).

Breaking complex limb coordination involving both hands and feet into simpler, incremental steps, instructors had students *"starting with the right hand on the hi-hat due to its consistent and minimal variation"* (T4). Then *"limb combinations begin with simpler patterns, with the right hand playing the hi-hat as it is usually the easiest"* (T6), *"pairing the right hand with the foot to build coordination incrementally"* (T3), emphasizing that *"practicing with pairs of limbs together improved synchronization"* (T4).

Instructors also identified the feet as the most challenging limbs to control, as *"students are unused to coordinating hands and feet, causing their feet to follow their hands unintentionally"* (T4). *"If a student struggled with consecutive double kicks on the bass drum, I go back to fundamental foot exercises"* (T2). *"When coordination between hands and feet is difficult, I isolate the hands first and then gradually integrate the feet"* (T3).

4.3 Learning Progression Framework Design

Based on the progression patterns used by the eight professional instructors for novice drum players, we designed a complete learning framework and placed the specific exercises used by each instructor in this 4x4 framework, as shown in Figure 4.1.

The resulting learning framework consists of two progression axes:

- **Comprehension progression (drum sequence and rhythm)**: shown as the vertical axis, consisting of 4 levels: 1) drum sequence, followed by rhythm progression from 2) main beats to 3) 1/2 beats and 4) 1/4 beats.
- **Limb coordination progression**: shown as the horizontal axis, consisting of 4 lev-

doi:10.6342/NTU202500720

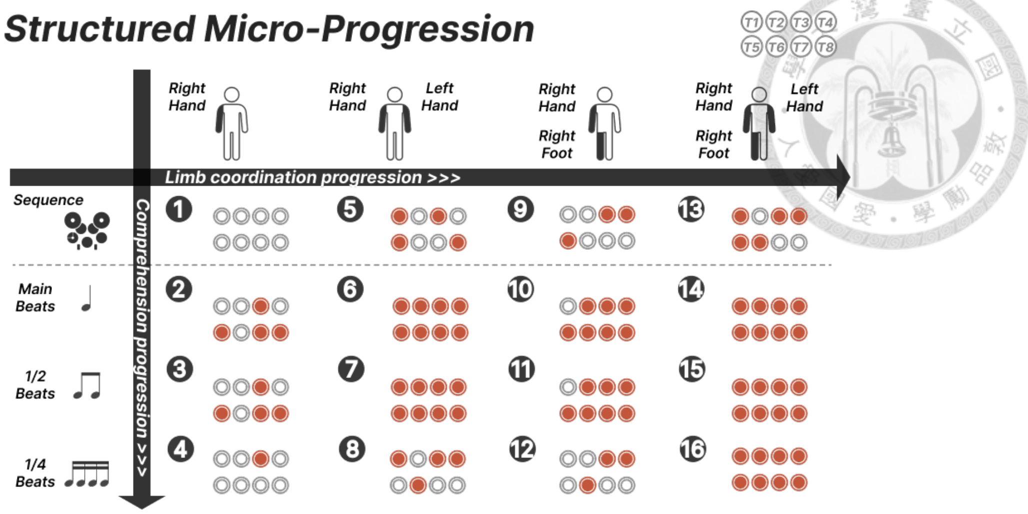

Figure 4.1: Mapping of all 8 instructors' 82 exercises to our micro-progression learning framework, where the horizontal axis shows limb coordination progression and the vertical axis shows comprehension progression. Each circle represents a professional drumming instructor, with orange indicating that an exercise matching the stage was used by the instructor.

els: 1) right-hand only, 2) both hands, 3) right hand + foot, and 4) all limbs.

Combining the two dimensions of progression, this 4x4 framework has a total of 16 micro-progression stages for self-learning novice players.

Figure 4.2 shows an example of a drum set pattern and its rhythm progression with all limbs (i.e., micro-progression stages 13-16). The progression starts with drum sequence comprehension that does not require correct timing, then progressively adds more detailed patterns from the 4 main beats to 1/2 and 1/4 beats.

Figure 4.3 show the same drum set pattern and its limb coordination progression with all notes (i.e., micro-progression stages 4, 8, 12, 16). The progression starts with right-hand only, then progressively add left hand, feet, then all limbs.

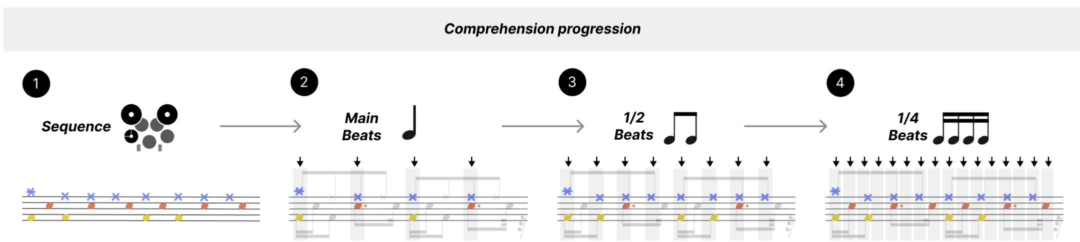

Figure 4.2: *Comprehension progression* consists of 4 levels, shown using an example drum sheet and arrows highlighting the beats: (1) drum sequence level (play the correct drum in the correct sequence, timing not required).; (2) main beats level (play only the 4 main beats with correct timing);(3) 1/2 beats level (adding 1/2 beats); and (4) 1/4 beats level (adding 1/4 beats).

doi:10.6342/NTU202500720

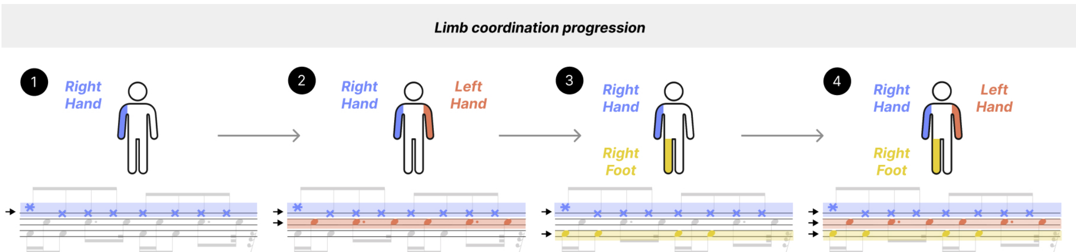

Figure 4.3: *Limb coordination progression* consists of 4 levels, shown using an example drum sheet, color coded to show limb mapping: (1) right-hand level (play only the notes assigned to the right hand), (2) right-hand + left-hand level; (3) right-hand + right-foot level, and finally the (4) right-hand + left-hand + right-foot level.

doi:10.6342/NTU202500720

Chapter 5

Formative Study #2: Self-learning Practices

To help inform system functionality and UI design, the second formative study aimed to understand the content, tools, and pain points of self-learning practices across a wide range of drumming experiences.

5.1 Study Design, Procedure, and Participants

5.1.1 Study Design and Procedure

The research method we used was a semi-structured interview, also analyzed by thematic analysis [2]. We asked participants to describe their sources of self-learning materials, tools they used, their positive and negative aspects, challenges encountered, and how they would overcome them.

The session began with an introduction to the study and the collection of demographics, including the Goldsmith Musical Sophistication Index (Goldsmith-MSI) questionnaire [23], a comprehensive self-report inventory designed to assess individual differences in musical engagement, skills, and behaviors across the general population.

doi:10.6342/NTU202500720

	Active Engagement	Perceptual Abilities	Musical Training
Mean	39.5 out of 63	46.8 out of 63	27.9 out of 49
Std	9.9	7.5	7.6
Range	25-52	40-61	15-36

Table 5.1: Goldsmith Music Sophistication Index of participants of self-learners formative study.

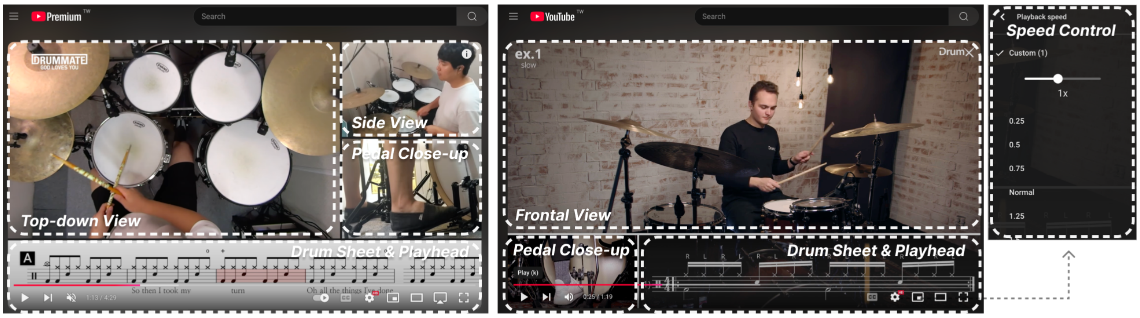

Figure 5.1: Examples of popular instructional videos for learning drumming from Youtube[1]. These videos provide synchronized notation and multiple camera angles showing drum performance, such as top-down, side, pedal close-up, and frontal views. The playback speed are controlled using the video players' speed control, adjusted by percentage.

.

5.1.2 Participants

We recruited 8 drummers (S1-S8) through social media posts and drummer forums, targeting individuals with varying levels of experience. The participants ranged in age from 20 to 31 years (M = 24.2, SD = 4.0), and comprised six males and two females. Among them 3 had less than 50 hours of drumming practice, 2 had 50-500 hours, and 3 had more than 500 hours. Participants' Goldsmith Musical Sophistication Index [23]'s item of Active Musical Engagement(AE), Self-reported Perceptual Abilities(PA), and Musical Training(MT) is shown in Table 5.1.

5.2 Findings: Content and Tools of Self-learning Practices

All participants reported that they selected music and drum set patterns to learn based on personal interests. All currently use video-based learning, by searching for instruction

[1]video credit: DRUMMATE (https://www.youtube.com/Drummate) and DrumX (https://www.youtube.com/@drumx4421)

doi:10.6342/NTU202500720

videos on Youtube (all) and by joining Facebook Groups (S5). When instruction videos could not be found for the music of interest, participants would resort to *"finding and reading books"* (S1) and *"listening to audio-only clips"* (S4). Among all self-learning methods, all participants preferred video-based demonstrations the most, as they can practice by *"using the phone to playback instructional videos and drumming along with them"* (S5) and *"mimicking the demonstrator's movements"* (S3).

Figure 5.1 shows two example instructional videos used by study participants, including both paid and free content. The instructional videos provide 1) third-person views from the front, side, or top-down that show the drummers' upper body movement and the drum sets, 2) a close-up view of the foot and pedal, 3) a drum sheet with a synchronized playhead that visually indicates the current position in the notation, and 4) an audio-only metronome. All participants commented on the cognitive efforts required to re-orient and re-map the drums and limbs, as the third-person views required *"extra effort to transform the orientations in my brain"* (S3) and led to *"forgetting what the demonstrated actions were because the drum set in the video isn' t the same as mine"* (S4).

To help comprehend rhythm and limb movement, all participants used the playback speed control built-in to the video players to first slow and then speed up the tempo as needed. For example, *"When I can' t keep up with the original speed, I adjust to 0.5x speed and then gradually increase by 10 BPM each time."* (S6). However, all reported that the playback speed controls could not match their desired BPM to practice at, as both Youtube and Facebook's video players only supported fixed percentage increments, such as 0.25x, 0.5x, and 0.75x playback speeds. *"I usually start practicing by slowing the song down to 60 BPM, then gradually increasing by 10 BPM each time. However, it's not possible to adjust the video speed to match the BPM I want. For example, if the original demonstration was at 120 BPM and I wanted to practice at 60, 70, and 80 BPM, I simply could not align the video speed to match."* (S1) Furthermore, *"at slower playback, the sound becomes distorted"* [S1].

When practicing on the drum set, all participants used the approach: *"continuously repeat [play, detect error]"* (S5) cycles. To detect errors, all participants used their ears

doi:10.6342/NTU202500720

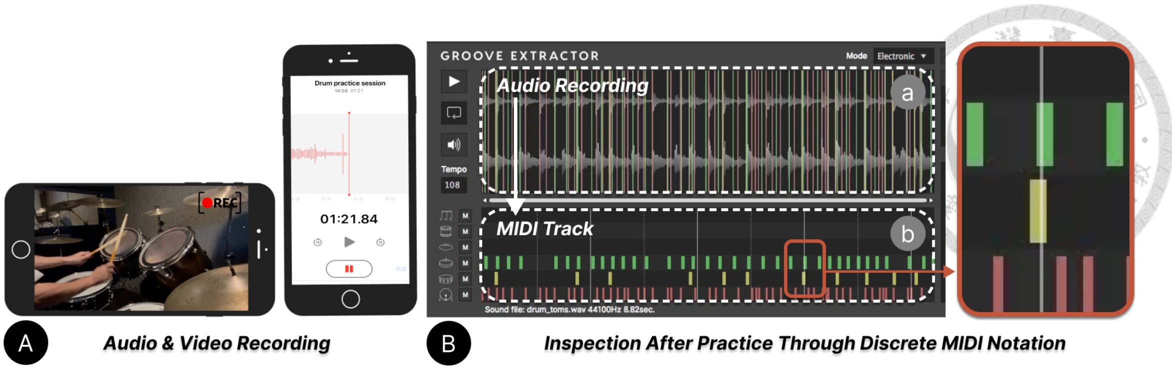

Figure 5.2: Tools used for tracking errors by participants, including: (A) recording audio or video using phones, and (B) using music editing software Logic Pro to convert audio recording (a) into discrete MIDI notes then (b) visually identify rhythm inaccuracies vs. timing gird lines.

to monitor in real-time as they played and compared to the metronome sound to determine whether their timing was accurate. However, all commented that simultaneous self-monitoring while playing was challenging, e.g., *"when my attention is focused on controlling the movements of my hands and feet, I can' t simultaneously focus on listening to identify what I got wrong."* (S1)

To improve error detection, some participants used additional tools, shown in Figure 5.2. Four participants recorded audio and video for review. One participant (S1) went a step further by importing practice recordings into the music editing software Logic Pro, which features a "Drummer Track" function that converts continuous analog audio recordings into discrete MIDI (Musical Instrument Digital Interface) notes. These MIDI notes are displayed visually, enabling the detection of drum and timing errors. However, these non-real-time error tracking tools interrupted practice and took up valuable practice time. *"Inspecting recorded practices interrupts my drumming practice, so I usually wait until the end of a practice session to do so. But sometimes, I would find out that I was playing incorrectly throughout the entire session."* (S2)

5.3 UI and System Design Implications

We summarize users' feedback, including pain points, needs, and also positive aspects of drumming instructional videos, into the following key design implications:

doi:10.6342/NTU202500720

1. **First-person view of demonstration that matches users' drum set** to reduce cognitive efforts required to re-orient and re-map the drums and limbs from third-person views of instructional videos.

2. **Close-up side view of the foot and pedal** as *"foot movement and timing are unclear when viewed from the top or front angles. Only a side-angle close-up camera on the pedal can clearly show the height, angle, and exact motion of the foot."* (S2)

3. **Tempo control based on BPM**: to exactly match drummers' practice methods, which involves adjusting in increments of 10 BPM, rather than using percentage-based video playback speeds that lack precise tempo alignment.

4. **Only show the limb movements for the subset of drum patterns users are focusing on** to reduce information overload when users want to focus only on a specific subset of drum patterns, such as the main rhythm vs. ornaments, which are embellishments or decorative techniques. e.g. *"Complex drum set pattern includes both the main rhythm and ornaments. When I start practicing a new piece, I prefer to focus on the main rhythm first. However, the ornaments in the demonstration video can distract me from observing how the main rhythm is played."* (S2)

5. **Automatic real-time error detection and visualization** to improve error tracking accuracy, reduce manual and cognitive effort, and maximize practice time. *"When the drum set pattern is too complex, I need to listen to the recording multiple times to identify my mistakes, and sometimes I can't figure them out at all."* (S4)

doi:10.6342/NTU202500720

Chapter 6

UI Design and System Implementation

6.1 UI and System Design

We used an iterative approach to refine the UI designs, incorporating feedback from four users (2 males, 2 females) aged 19 to 28 (M = 24.3, SD = 3.3). Among them, three had UI design experience, and two had drumming experience. For each of the above design implications, we explored several UI design options, implemented as prototypes, and refined the designs based on feedback.

The design overview and the comparison between before-user-feedback and after-user-feedback are shown in Figure 6.2.

1. **First-person view of demonstration that matches users' drum set.** We rendered AR virtual limbs that extend from the user to demonstrate drumming in first-person perspective on the user's own drums in mixed-reality mode. Users reported that the virtual hands sometimes obstructed their view of their real hands during use. To address this issue, we implemented four transparency levels (25%, 50%, 75%, 100%) and four visualization options: Stick-only, Stick + Hand, Stick + Hand + Forerm, and Stick + Hand + Full Arm, as shown in Figure 6.1. Users reported that a transparency level of 50% combined with the Stick + Hand + Full Arm option was the most effective, providing clear demonstrations while allowing them to see their own hands.

doi:10.6342/NTU202500720

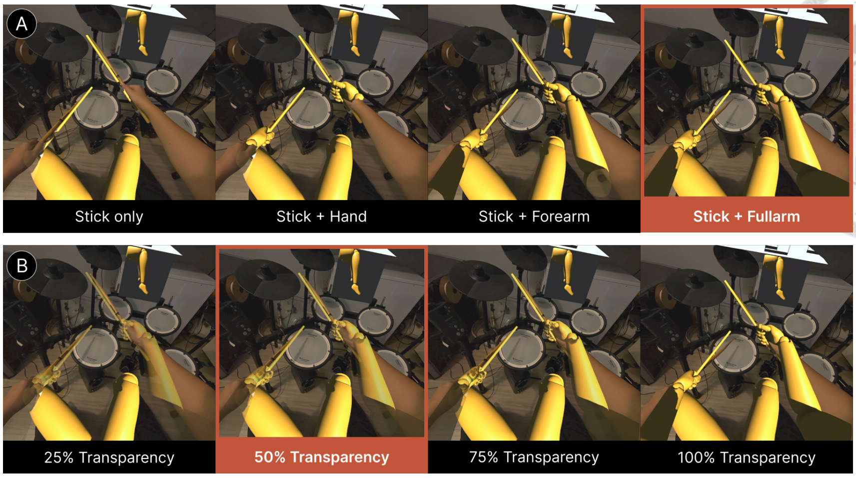

Figure 6.1: Design exploration of virtual limb visualization: (A) Display options for the arm: stick only, stick + hand, stick + hand + forearm, stick + full arm, progressing from partial to full representation; (B) Transparency levels: 25%, 50%, 75%, 100%. The most preferred design was: stick + full arm at 50% transparency.

Users noted that using the same color for all virtual limbs made it difficult to distinguish them. To improve clarity, we adopted the color-coding approach used in popular drumming VR games [27], assigning distinct colors to each limb: orange for the left hand, blue for the right hand, and yellow for the legs. These colors were selected from colorblind-friendly palettes to ensure accessibility.

Users also found it challenging to identify the exact timing when the virtual drumstick struck a drum surface. To address this issue, we adopted an approach inspired by popular drumming VR games [25], where the drum surface is momentarily highlighted upon being struck. This visual feedback improved timing recognition during practice.

2. **Close-up side view of the foot and pedal.** We added an AR mirror to display foot movements clearly, as shown in Figure 6.2(d), after testing several positions for the mirror layout.

3. **Tempo control based on BPM.** Our initial design allowed users to adjust both the AR virtual limb demonstration and the audio-only metronome to their desired BPM speed, with the BPM displayed on an AR panel.

   After testing the prototype, users noted that while the metronome cycles in four beats, even a brief distraction could make it unclear which beat they were currently

doi:10.6342/NTU202500720

on within the cycle. To address this, we incorporated a visual metronome inspired by the popular app Tempo [8], providing a clear visual representation of the beat progression to complement the auditory feedback, as shown in Figure 6.2(b).

4. **Only show the limb movements for the subset of drum patterns users are focusing on.** We implemented AR virtual limbs to support both complete demonstration and that matches each micro-progression stage.

5. **Automatic real-time error detection and visualization.** We designed an AR drum sheet featuring a live playhead to indicate the current playing position. Additionally, the AR drum sheet detects and visually highlights timing errors in real-time, as shown in Figure 6.2(c). Users can use these visual cues to pinpoint timing errors and compare their movements with the virtual limb demonstrations. This feedback allows users to identify mistakes as they occur and adjust their technique to correct them.

6. **Other.** During prototype testing, users often inquired about their current practice progress. To address this, we introduced an AR progression status panel, enabling users to easily monitor their progress within the micro-progression framework at any time.

6.2 System Implementation

We developed the MR interface using Unity (2022.3.34f1) with Minis[34], a Unity plugin that adds MIDI input device support to Unity's new Input System to receive input MIDI signal from electronic drum sets. MR-Drum runs on a PC connected to a Quest 3 passthrough AR HMD and an electronic drum set. The electronic drum set sends the user's drumming signals in MIDI format to the MR-Drum system running on the PC, which analyzes the user's practice performance.

For the drum set, we chose the Roland TD-1DMK, as Roland is the top electronic and acoustic drum brand in the US and top 3 globally [30], and the TD-1DMK electronic drum

doi:10.6342/NTU202500720

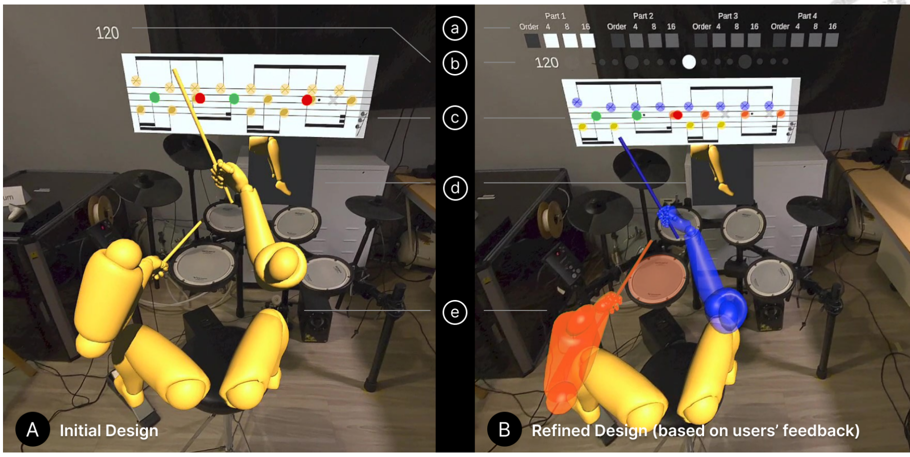

Figure 6.2: Before vs. after user interface designs: (A) initial UI design, and (B) refined UI design based on user feedback. The initial design included (b) a BPM display, (c) a synchronized drum set pattern with playhead and error feedback, (d) an AR mirror for virtual foot movement, and (e) a first-person virtual limb demo. The refined design added the following enhancements: (a) micro-progression status panel, (b) visual metronome with adjustable BPM, (c) error feedback color-coded to match virtual limb colors, and (e) optimized limb transparency with color-coded limbs.

set uses the standard, five-piece drum set layout.

For limb demonstration, we hired a professional drummer with over 10 years of experience and 5,000 practice hours to record demonstrations using Quest 3 controllers mounted drumsticks. The recorded performance is processed per limb and by beat to generate drumming demonstrations for each of the 16 stages of the micro-progression learning framework.

doi:10.6342/NTU202500720

Chapter 7

Summative User Experience Study

To understand how MR-Drum supports self-learning of drumming, we conducted a comparative assessment vs. the most popular and the most preferred self-learning method, instructional videos. We collected objective performance measurements, subjective preference ratings, and qualitative feedback. The two performance metrics were based on what professional instructors used to evaluate students' practice performance: 1) *error rate*, i.e., "whether the correct notes are played" (T6, T7, T8), and 2) *timing error*, i.e., "correct timing" (T5, T6, T8).

For subjective preference ratings, we focused on the outcome of the two learning phases: 1) comprehension as the outcome of the comprehension phase, and 2) skill development as the outcome of the practice phase, and additionally, 3) user experience.

7.1 Study Design

The study used a within-subjects design with one independent variable of two conditions: practicing with MR-Drum vs. instructional videos, in counter-balanced ordering. To control for the two conditions, we provided MR-Drum's AR visual metronome and music sheet with real-time playhead and error visual cues to the instructional video interface, as shown in Figure 7.1. For both conditions, participants verbally asked the experiment operator to play/stop the virtual limbs and videos for them.

To avoid learning effect, we selected two drum set patterns of similar difficulty and

doi:10.6342/NTU202500720

counter-balanced across the two conditions. The dependent variables of this study were: 1) objective performance metrics measurement, and 2) subjective preference rating.

The performance metrics are calculated with the following steps:

- Error Rate

  1. Each note in the drum set pattern is matched with the closest participant-played note within a time difference of less than 0.25 seconds.
  2. extra notes = notes played by a participant that are not matched with any drum set pattern' s note
  3. missing notes = drum set pattern' s notes that are not matched with any notes played by a participant
  4. error rate = (# extra notes + # missing notes) / (# drum set pattern' s notes)

- Timing Error Per Note

  1. Each note in the drum set pattern is matched with the closest participant-played note within a time difference of less than 0.25 seconds.
  2. extra notes = notes played by a participant that are not matched with any drum set pattern' s note
  3. missing note = drum set pattern' s notes that are not matched with any notes played by a participant
  4. penalty = 0.25 seconds
  5. time error per note = [sum(time difference between matched pair of notes) + (# extra notes + # missing notes) * penalty] / (# drum set pattern' s notes).

The subjective preference rating was collected using a 10-point strength-of-preference [7, 5, 20]. Participants first selected their preferred condition and then rated the degree of difference on a 5-point scale. This method, which is more sensitive to utility differences than independent ratings, was applied to assess comprehension, skill development, user experience, and overall preference.

doi:10.6342/NTU202500720

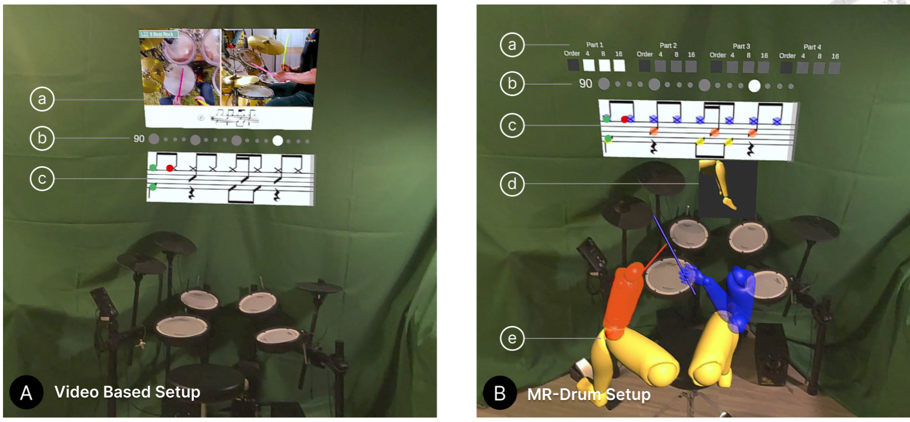

Figure 7.1: Mixed-reality user interfaces viewed via a Quest 3 headset: (A) Instructional videos and (B) MR-Drum: The video-based setup features (a) a video player, while the MR-Drum interface includes (a) a 16-stage micro-progression status panel, (d) an AR mirror for virtual foot movement, and (e) virtual limbs. The (b) visual metronome and (c) drum set pattern with real-time error feedback are the same to control for the differences between the two conditions.

7.2 Study Procedure

After introducing the experiment procedure and addressing participants' questions, participants engaged in two 15-minute drumming practice sessions under two different conditions, with a 10-minute break between the two sessions.

Before each practice session, participants completed a pre-practice test. After each session, we collected their fast-sickness scores [13] to ensure they were not experiencing VR motion sickness, followed by a post-practice test. During the pre- and post-practice tests, participants performed five trials at each of three tempos: 60, 90, and 120 BPM. For each tempo, the average of the two best-performing trials was recorded as their performance.

Upon completing both sessions, participants reported their subjective preference ratings and answered qualitative questions. The study duration was about 60 minutes.

7.3 Apparatus

Pariticpants wore a Quest 3 headset to view the mixed-reality interface, and practiced drumming on a Roland TD-1DMK electronic drum kit. The TD-1DMK features a standard five-piece drum set layout that closely mirrors that of traditional acoustic drum kits,

doi:10.6342/NTU202500720

	Active Engagement	Perceptual Abilities	Musical Training
Mean	32.3 out of 63	41.9 out of 63	21.9 out of 49
Std	8.7	8.4	9.9
Range	23-45	26-56	9-37

Table 7.1: Goldsmith Music Sophistication Index (Gold-MSI) of participants in the summative study.

typically comprising a bass drum, snare drum, one or more toms, hi-hat, and cymbals. This standard setup facilitates a seamless transition for drummers adapting to different kits, enhancing their adaptability across various drum set configurations.

7.4 Participants

We recruited 12 participants (6 females, 6 males) between 20 and 29 years old (M = 24, SD = 2.39). 8 participants indicated that they had little or no prior experience playing the drum, among which 4 had experience with drumming for 70-80 hours, 5 has experience with other instruments. Participants' Goldsmith Music Sophistication Index (Gold-MSI) [23] on Active Musical Engagement (AE), Self-reported Perceptual Abilities (PA), and Musical Training (MT) are shown in Table 7.1. Regarding AR/VR usage frequency, 11 participants reported less than once per year, and 1 reported once every 3 months. Participants received a nominal compensation for their participation.

7.5 Results and Discussion

7.5.1 Objective Metrics: Performance and Improvement After Practice

Participants' performance and improvement is shown in Figure 7.2. Wilcoxon Signed Rank test was used for pair-wise statistical significance. The effect size of each pairwise comparison is calculated as r=Z/$\sqrt{n}$ and interpreted using guidelines 0.1-0.3 (small), 0.3-0.5 (moderate), and $\geq$ 0.5 (large effect) [29].

doi:10.6342/NTU202500720

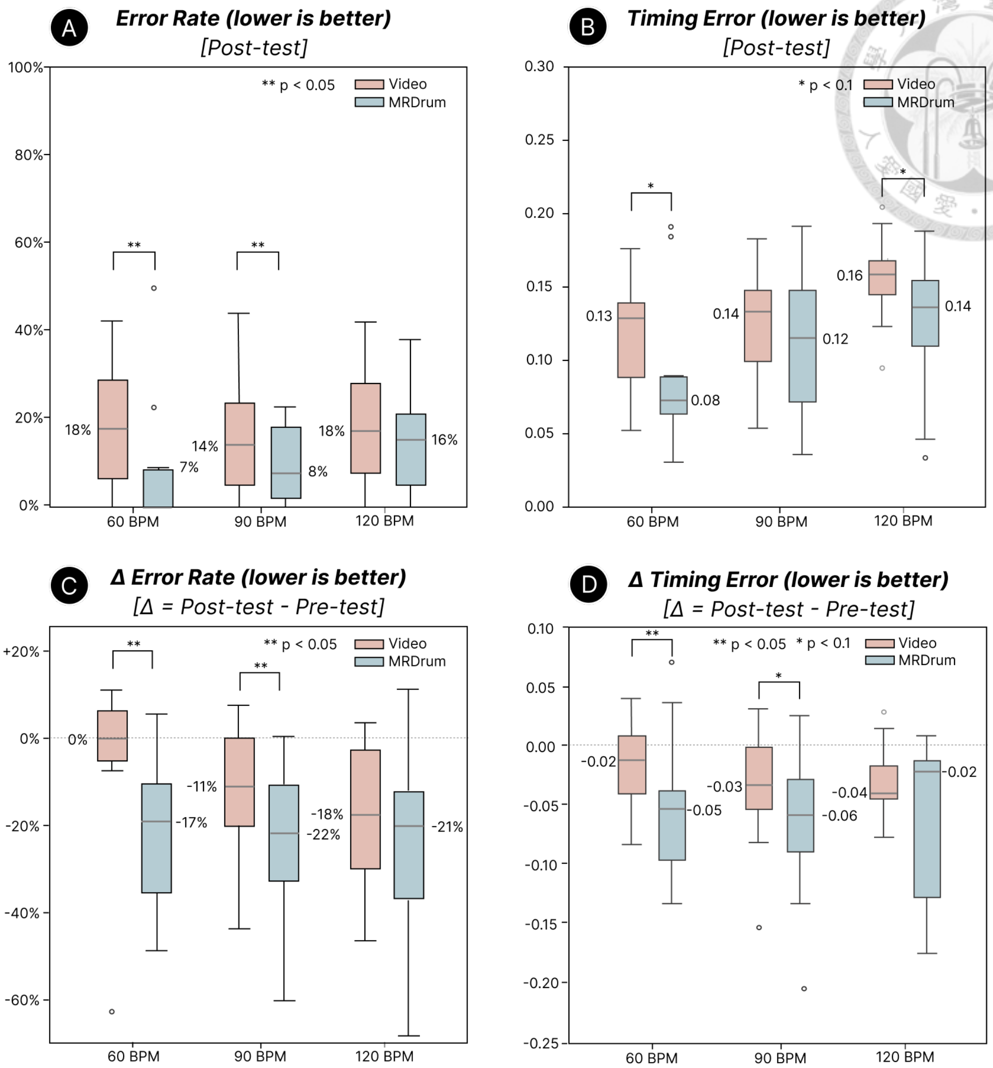

Figure 7.2: Participants' performance (A, B) and improvement (C, D) after practicing with MR.Drum, represented by participants' error rate (A), change in error rate (Δ error rate) (B), timing error (C), and change in timing error (Δ timing error) (D) per note at 60 BPM, 90 BPM, and 120 BPM, respectively.

Figure 7.2(A)(B) shows participants' error rate and time error per note after practicing with each setup. Participants' error rate are significantly lower after practicing with MR.Drum than with Video setup (p < 0.05, $r > 0.5$) at 60BPM and 90BPM. Lower at 120 BPM but not significant. Participants' time error per note are significantly lower after practicing with MR.Drum than with Video setup (p < 0.1, r = 0.3-0.5) at 60BPM and 120BPM. Lower at 90 BPM but not significant. Participants without drumming experience had higher error rate in pre-test (0.4 vs 0.17), but not in post-test (0.16 vs 0.13), and higher time error per note in pre-test (0.18s vs 0.15s), but not in post-test (0.13s vs 0.13s)

Figure 7.2(C)(D) shows participants' error rate improvement and time error per note

doi:10.6342/NTU202500720

improvement after practicing. Participants' error rate decrease significantly more after practicing with MR-Drum than with Video-based setup ($p < 0.05$, $r > 0.5$) at 60BPM and 90BPM. error rate decrease more with MR.Drum at 120 BPM but not significant. Participants' time error per note decrease significantly more after practicing with MR-Drum than with Video-based setup at 60 BPM ($p < 0.05$, $r > 0.5$) and 90 BPM ($p < 0.1$, $r = 0.3\text{-}0.5$). time error per note does not decrease more with MR-Drum at 120 BPM.

All participants reported low sickness level (range 0-1 out of 10) after every practice session.

We found that after practice, MRDrum's Error Rate, Delta Error Rate, and Time Error are significantly better at slower BPMs (60 BPM, 90 BPM) but not at faster BPM (120 BPM). The reason is that at higher BPMs, in addition to a lack of understanding of the drum set pattern, hand motor skills (whether participants have sufficient muscle strength to perform high-speed drumming) also significantly influenced their performance. *"If the student can play the drum set pattern correctly at a slow speed, it indicates that they have understood it, then gradually increasing the speed helps their hands adapt to the faster movements."* (T4) Participants cannot achieve noticeable muscle strength improvement within a 15-minute practice session. *"I know how to play it, but my hands can't move that fast, and my feet can't keep up either—it gets tiring."* (P6)

7.5.2 Subjective Preference Rating: Comprehension

Figure 7.3 shows the 10-point strength-of-preference ratings. Two-tailed, Wilcoxon signed-rank test was used for statistical significance. The effect size of each pairwise comparison is calculated as $r=Z/\sqrt{n}$ and interpreted using guidelines 0.1-0.3 (small), 0.3-0.5 (moderate), and ≥ 0.5 (large effect) [29].

Participants rated MR.Drum significantly more helpful for rhythm comprehension (91.6%, $p < .01$, $r > .5$), limb comprehension (100%, $p < .01$, $r > .5$), drum transition comprehension (100%, $p < .01$, $r > .5$), and 100% of participants rated MR.Drum overall more helpful for comprehension($p < .01$, $r > .5$).

doi:10.6342/NTU202500720

Rhythm Comprehension. More than half of the participants identified *"micro-progression of rhythm subdivisions"* (P2 – P7, P9, P12) as key to understanding rhythm. Participants noted that these subdivisions provide step-by-step guidance, especially for beginners, making complex rhythms more approachable. *"The subdivision breakdown helps me build internal rhythm by gradually reconstructing the parts"* (P12). However, some participants mentioned challenges with overly detailed stages. *"Some stages feel unnecessarily broken down; I wish I could highlight only the most useful ones for me"* (P11).

Limb Comprehension. All participants found the *"first-person perspective of the virtual tutor"* (P1–P12) crucial for understanding limb movements. *"The first-person perspective allows me to directly learn the movements and clearly see which hand to use"* (P8). Many highlighted the value of virtual limbs demonstrating directly on their drum sets, removing the need for *"mental mapping"* (P2, P4, P5, P12). Simplified subdivisions also supported limb coordination, making practice more effective. *"Breaking down subdivisions allows for more focused practice and better coordination at higher speeds"* (P2). Without this breakdown, participants noted, *"Watching everything at once is overwhelming, and it' s easy to miss some limb movements, like the left hand"* (P7).

7.5.3 Subjective Preference Rating: Skill Development

Figure 7.3 shows the 10-point strength-of-preference ratings. Participants rated MR.Drum significantly more helpful for rhythm skill development (91.6%, $p < .01$, $r > .5$), limb skill development (100%, $p < .01$, $r > .5$), and 100% of participants rated MR-Drum overall more helpful for skill development($p < .01$, $r > .5$).

Rhythm Skill Development. Half of the participants stated that *"rhythm subdivisions"* (P1, P3, P9–P12) improved their sense of timing and internal beat. *"The rhythm subdivisions help me understand the structure better and build on unfamiliar patterns step by step"* (P11). This gradual layering of unfamiliar rhythms onto a familiar base was noted as particularly effective. However, some participants mentioned that the virtual tutor lacked the dynamic body rhythm of a real drummer. *"The virtual tutor focuses on arms, but a*

doi:10.6342/NTU202500720

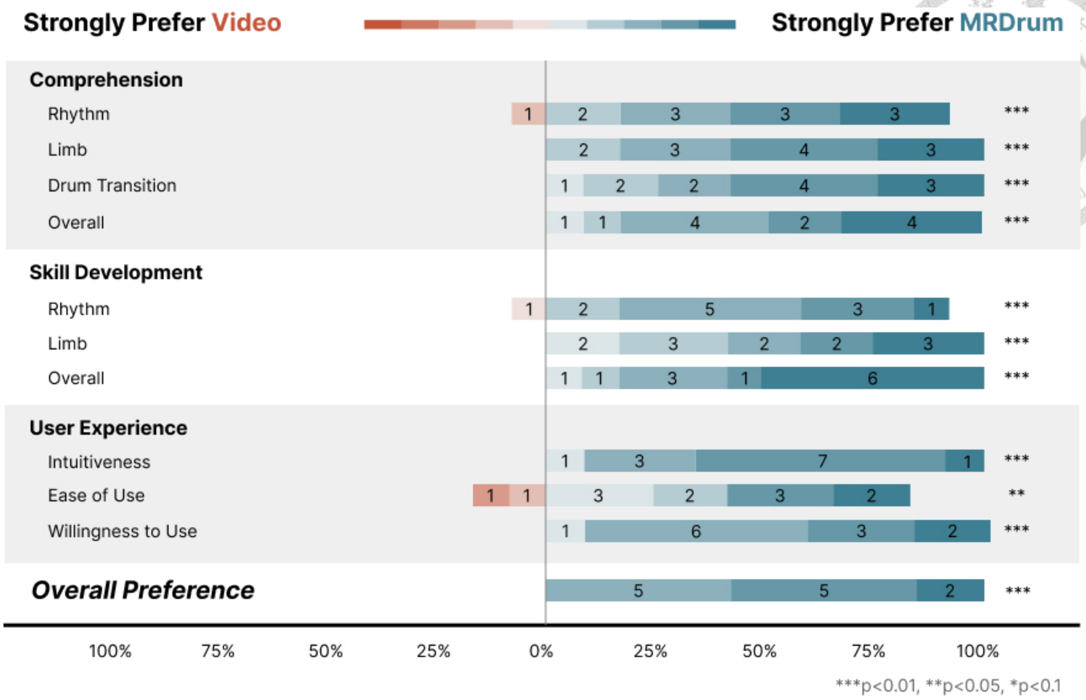

Figure 7.3: 10-point strength-of-preference subjective ratings comparing the video-based practice method and the MR-Drum practice method across three categories: Comprehension, Skill Development, and User Experience. Each shaded block shows the number of participants (out of twelve) who were assigned the corresponding score.

real drummer's whole-body rhythm conveys a better sense of flow" (P4).

Limb Skill Development. Over half of the participants indicated that *"direct imitation of the virtual tutor"* (P1, P2, P4, P5, P8–P11) aided limb skill development. "The virtual hands overlap with my movements, making it very easy to follow" (P4). Immediate feedback during errors also helped participants refine their limb coordination. "Mimicking the virtual limbs helps establish muscle memory, and the instant feedback corrects mistakes" (P11). Many participants also found limb micro-progression stages effective (P3, P5, P6, P7, P11, P12). "Training different limbs through individual stages helps locate errors and maintain control in later stages" (P11). Starting with simple stages was noted as crucial: "Building a solid foundation in the early stages makes limb coordination in later stages more stable" (P7). However, some participants raised concerns about the naturalness of virtual limb movements. "The flexibility of the virtual hands matters because if it looks unnatural, it's harder to imitate" (P4).

doi:10.6342/NTU202500720

7.5.4 Subjective Preference Rating: User Experience

Figure 7.3 shows the 10-point strength-of-preference ratings. Participants rate MR-Drum significantly better in intuitiveness (100%, $p < .01$, $r > .5$), ease of use (83%, $p < .05$, $r > .5$), and 100% of participants prefer MR-Drum for overall experience($p < .01$, $r > .5$).

Intuitiveness. The MRDrum system was highly intuitive in aligning drum positions, limb movements, and rhythms (P2, P4 – P9, P11). Participants emphasized the clarity of the virtual tutor's demonstrations, with one noting, *"The virtual limb matches the drum directly, making it easy to know where to hit and how it sounds"* (P9). The first-person perspective further streamlined learning: "The first-person view lets me mimic exactly where the tutor strikes" (P8). Micro-progression stages also simplified practice by breaking down tasks and *"focusing on key beats"* (P1,P10), *"making complex rhythms approachable"* (P3, P12).

Ease of Use. The step-by-step progression from simple to complex tasks made MR-Drum accessible and effective (P3, P8, P9, P11). Participants appreciated customizable features, such as adjusting BPM and toggling virtual cues: *"Customizing features like rhythm speed felt tailored to my needs"* (P8). One participant described the system as comparable to a real teacher: *"It feels like having a real instructor"* (P11). However, hardware limitations were noted, including the headset's weight and slight pass-through delays, which could impact comfort and usability (P10).

7.5.5 Discussion

Virtual Limbs Demonstrations: All participants preferred the virtual hands demonstration over video demonstrations because they eliminated the need to change perspective or remap movements. *"Watching the virtual teacher once is as effective as watching the video three times, as it doesn't require perspective switching or remapping movements."* (P4) The virtual hands provided clearer directional cues for limb movements. (P3,9,11) *"It's easier to understand which hand to use and where to strike."* (P9)

doi:10.6342/NTU202500720

The movements demonstrated by the virtual hands were easier to remember because they provide more immersive 3D spatial information.(P1,2,7,8,11) *"Videos appear two-dimensional and lack 3D spatial information, making it easy to forget the movements."* (P8) Participants also noted that virtual limbs demonstration is easier to *"imitate and follow along."* (P1,2,5,7,8) and *"Drumming along with virtual teacher"* (P7) helps them memorize the movement and *"comparing movements with the virtual hands"* (P11) made it easier to identify mistakes. Unexpectedly, P12 mentioned, *"It feels like someone is practicing with me, and the sense of companionship makes the practice less lonely."*

Micro Progression Learning Framework: In terms of micro-progression, nearly all participants found step-by-step decomposition helpful in tackling complex rhythms (P1-5, P7-12). *"Breaking the practice into smaller pieces helps me understand each part before moving on"* (P3). Many felt that gradually increasing difficulty built confidence and a sense of accomplishment, rather than facing the entire complexity at once (P1-3, P7, P9-10). *"Starting with simpler patterns before adding complexity boosts my confidence and motivates me to keep going"* (P7). Identifying weak points through these incremental stages allowed for targeted improvement and more efficient practice (P1, P3-4, P8-9, P11-12). *"Focusing on one limb or a single segment of the rhythm makes it easier to correct mistakes and eventually combine everything smoothly"* (P11).

doi:10.6342/NTU202500720

Chapter 8

Discussion, Limitations, and Future Work

8.1 Virtual Limbs and Hands for Learning Musical Instruments

While our study found that first-person view of virtual limbs to be effective for drumming practice, virtual limbs for other musical instruments have produced mixed results. For example, guitARhero [32] evaluated virtual hands for learning guitar using both first-person and mirrored perspectives, but found them both to be less effective than highlighted guitar frets. Whereas PianoSyncAR [19] found virtual hands to be effective for piano, although the virtual hands needed to be slightly shifted horizontally to improve player observation and imitation.

Several factors potentially influence the effectiveness of virtual limb demonstrations. First, the type of motor skill involved, fine (small) vs. gross (large), affects the visibility of movements. For example, the left hand on string instruments like guitar and violin performs small, precise movements that are harder to observe, compared to the larger movements required for piano, and even more so for drumming, which involves the largest limb movements. In drumming, these large movements require earlier preparation to position the limb toward the target drum surface effectively, which benefit the most from virtual

doi:10.6342/NTU202500720

limbs as *"the movement of the virtual hands provides an advanced cue indicating which drum will be played next, giving me more time to move from one drum to another"* (P5).

Second, the degree of spatial separation vs. occlusion between virtual limbs/fingers also impacts visibility. For instance, on a violin, the left hand (fingering strings) experiences significant occlusion between fingers, whereas the right hand (bowing) faces no such occlusion issues, remaining fully visible. For guitar, authors of guitARhero [32] noted that "*the augmentation and the participant's own hand needed to be in the same place, introducing visually cluttered and making it difficult for participants to determine whether they were pressing on the correct frets.*" [32]

Third, whether the mapping from limbs/fingers to a specific playing part of the musical instrument (e.g. a drum, a piano ksey, a saxophone key) is 1-to-1 vs. N-to-1. For example, instruments such as drums, pianos, and guitars are N-to-1, as multiple limbs/fingers can strike and press the same drum/key/fret. These would benefit more from virtual limb/finger demonstration, whereas highlighted keys would be more visible and be sufficient for instruments with keys and valves that are 1-to-1, such as most woodwind and brass instruments, e.g., saxophones, flutes, and trumpets.

While virtual limbs are best suited for musical instruments that use large movements, have low occlusion issues, and played N-to-1, there are several design strategies worth exploring, such as offsetting the position and angle of the virtual limbs to reduce occlusion, visualization techniques to highlight small movements, and using a combination of highlighted frets/key with virtual limbs/fingers.

8.2 Micro-Progression Frameworks for Other Instruments

Our work demonstrated an approach to identifying and developing a micro-progression framework based on the teaching methods and exercises of professional instructors and showed that a mixed-reality implementation of this framework effectively supports self-learning for drumming. The micro-progression concept likely extends other musical instruments, as the progression in rhythm comprehension is broadly applicable. However, the progression of motor skill development is more instrument-specific, varying by cat-

doi:10.6342/NTU202500720

egory. For example, piano and keyboard require precise finger independence, string instruments emphasize fine motor control and fretting, bowing, and plucking techniques, while wind instruments focus on breath control, embouchure, and pressing keys/valves, each with distinct demands on limb coordination.

For example, when playing guitar, the fretting (left) hand and picking/strumming (right) hand can be practiced independently. During left-hand chording practice, the right hand is often simplified to basic strumming (a single sweeping motion across all strings), avoiding more complex picking techniques. Similarly, rhythm is typically simplified in the initial stages of guitar practice, progressing in complexity over time. A common approach is the one-strum practice, where the player strums only on the first beat of each measure, allowing them to focus on chord changes and timing.

These instrument-specific breakdown practices can form the foundation for designing micro-progression frameworks tailored to each instrument. However, developing a comprehensive framework requires in-depth interviews with instructors to understand the nuanced progression strategies and exercises specific to each instrument.

8.3 Designing for Live Instruction

While MR-Drum significantly improves upon instructional videos, professional instructors remain more effective in several key aspects, including diagnosing students' weaknesses and recommending tailored exercises, identifying and correcting improper drumming posture, and providing haptic guidance for rhythm and posture adjustments.

We are exploring ways to leverage mixed-reality and micro-progression to complement live instruction. For example, live capturing of instructors' movements could allow students to view demonstrations from a first-person perspective, enhancing understanding and engagement. Additional utilities, such as error detection and visualization, could further support both instructors and students in refining techniques and improving learning outcomes.

doi:10.6342/NTU202500720

8.4 Study Limitations and Future Directions

Current MR headsets has limitations including weight, frame rate, and limited field of view (FOV) that could impact the overall drumming experience, affecting comfort and visual precision. Interestingly, despite these constraints, users reported that the passthrough drumming experience of the Quest 3 headsets was noticeable but had minimal effects on their practice sessions. The auditory feedback and physical response from striking real drums helped maintain rhythm and timing, compensating for the slight visual lag. *"Even though there was some visual lag, the sound and feel were enough to keep me on track"* (P5). However, advancements in MR hardware, such as lighter headsets with wider FOVs and higher frame rates, and lower-latency headsets such as Apple Vision Pro, would improve the comfort and experience, especially important for extended practice sessions.

The comparative summative study was conducted in-lab with 12 users over relatively short practice sessions, and has other limitations such as novelty effect, where the introduction of mixed-reality (AR) temporarily boosts performance or interest. To mitigate this effect, we used a counter-balanced experimental design. To facilitate further research, we have open-sourced MR-Drum, enabling interested researchers to explore micro-progression and mixed-reality learning for drumming with larger sample sizes, longitudinal studies, and potential extensions to other musical instruments.

8.5 System Limitations and Future Work

MR-Drum's currently requires users to complete all levels. Some participants expressed a preference for greater flexibility, with one noting that *"some levels are too finely divided"* (P1) and that *"it would be beneficial to have more control over level selection"* (P8). Although users can easily pass levels they are already proficient in to confirm that they performed correctly, without spending much time, we are exploring ways to incorporate diagnostic features to better personalize the progression stages.

MR-Drum's micro-progression framework was specifically designed for novices and therefore did not include left foot movements, as these are typically not required at the

doi:10.6342/NTU202500720

beginner level. To support more advanced drumming, we plan to extend the prototype by adding an additional limb progression stage for the left foot, resulting in a total of 4x5 stages.

Furthermore, the current MR-Drum virtual hand and foot demonstrations focus exclusively on the hands and feet, omitting the torso and overall body movement. This creates a gap in representing the full rhythmic flow of a drummer’s body during performance. As one participant observed, *“The virtual tutor focuses on arms, but a real drummer’s whole-body rhythm conveys a better sense of flow”* (P4) To address this, we are exploring techniques to capture and integrate full-body movements into virtual demonstrations, aiming to enhance the sense of rhythm and flow conveyed to learners.

Currently, the virtual limb demonstrations in MR-Drum are manually performed and recorded, allowing for precise capture of movements. However, incorporating AI-generated virtual limb demonstrations could greatly enhance scalability and the availability of music for learning. As one participant noted, *“the way the system breaks down the movements is really useful, but it would be even better if it could have demonstrations tailored to music I like”* (P7).

doi:10.6342/NTU202500720

Chapter 9

Conclusion

Through a series of formative and summative studies, we have developed a novel micro-progression learning framework based on teaching methods of professional instructors by dividing complex musical structures and limb coordination of drumming along *comprehension progression* and *limb coordination progression* into 16 micro-progression levels. Based on drummer's self-learning practices and pain points, we designed and implemented a mixed-reality system, MR-Drum, to support the micro-progression learning framework, providing a first-person, virtual-limb-based interface, We conducted a comparative study vs. instructional videos which showed that MR-Drum significantly improved drumming performance in terms of error rate and timing, was significantly preferred for comprehension, skill development, and usability, and was preferred overall by all participants.

Our approach to analyzing and designing the micro-progression learning framework can be adapted for other musical instruments, enabling more effective learning compared to traditional instructional video methods. Additionally, the design exploration and insights gained from using virtual limbs for drumming demonstrations can benefit researchers and practitioners working on mixed-reality demonstrations in various domains.

To further advance these efforts, we have open-sourced MRDrum, facilitating others to experience and explore micro-progression and virtual limb techniques, and to extend these concepts to other musical instruments and applications.

doi:10.6342/NTU202500720

Bibliography

[1] J. Bae, B. Lee, S. Cho, Y. Heo, and H. Bang. Drum on: interactive personal instrument learning system. In *ACM SIGGRAPH 2012 Posters*, SIGGRAPH '12, New York, NY, USA, 2012. Association for Computing Machinery.

[2] V. Braun and V. Clarke. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4):589–597, 2019.

[3] M. Cai, M. A. Amrizal, T. Abe, and T. Suganuma. Design of an ar-based system for group piano learning. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 20–21, 1, Oct 2019. 1.

[4] A. Campo, A. Michałko, B. Van Kerrebroeck, B. Stajic, M. Pokric, and M. Leman. The assessment of presence and performance in an ar environment for motor imitation learning: A case-study on violinists. *Computers in Human Behavior*, 146:107810, 2023.

[5] P.-Y. Chen, C.-Y. Tsai, W.-H. Wang, C.-J. Lai, C.-A. Fan, S. C. Lin, C.-C. Chi, and M. Y. Chen. Aircharge: Amplifying ungrounded impact force by accumulating air propulsion momentum. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, New York, NY, USA, 2023. Association for Computing Machinery.

[6] A. Ebisu, S. Hashizume, K. Suzuki, A. Ishii, M. Sakashita, and Y. Ochiai. Stimulated percussions: method to control human for learning music by using electrical muscle stimulation. In *Proceedings of the 8th Augmented Human International Conference*, AH '17, New York, NY, USA, 2017. Association for Computing Machinery.

doi:10.6342/NTU202500720

[7] I. Evangelidis. Task sensitivity and noise: How mechanical properties of preference elicitation tasks account for differences in preferences across tasks., may 2023.

[8] Frozen Ape. Tempo: The top paid metronome app on ios. http://www.frozenape.com/tempo-metronome.html. Accessed: December 10, 2024.

[9] T. Goto, S. Das, K. Wolf, P. Lopes, Y. Kurita, and K. Kunze. Accelerating skill acquisition of two-handed drumming using pneumatic artificial muscles. In *Proceedings of the Augmented Humans International Conference*, AHs '20, New York, NY, USA, 2020. Association for Computing Machinery.

[10] S. Holland, A. J. Bouwer, M. Dalgelish, and T. M. Hurtig. Feeling the beat where it counts: fostering multi-limb rhythm skills with the haptic drum kit. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '10, page 21–28, New York, NY, USA, 2010. Association for Computing Machinery.

[11] M. Hosoya, M. Morise, S. Nakamura, and K. Yoshii. A real-time drum-wise volume visualization system for learning volume-balanced drum performance. In J. Baalsrud Hauge, J. C. S. Cardoso, L. Roque, and P. A. Gonzalez-Calero, editors, *Entertainment Computing – ICEC 2021*, pages 154–166, Cham, 2021. Springer International Publishing.

[12] H. Kanke, T. Terada, and M. Tsukamoto. A percussion learning system by rhythm internalization using haptic indication. In *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology*, ACE '15, New York, NY, USA, 2015. Association for Computing Machinery.

[13] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. *Hum Factors*, 53(4):415–426, Aug 2011.

[14] K. Kikuchi, M. Niibori, and M. Kamada. An ar system to practice drums. In L. Barolli, H.-C. Chen, and T. Enokido, editors, *Advances in Networked-Based Information Systems*, pages 367–373, Cham, 2022. Springer International Publishing.

doi:10.6342/NTU202500720

[15] Y. Kweon, S. Kim, B. Yoon, T. Jo, and C. Park. Implementation of educational drum contents using mixed reality and virtual reality. In C. Stephanidis, editor, *HCI International 2018 – Posters' Extended Abstracts*, pages 296–303, Cham, 2018. Springer International Publishing.

[16] K. Labrou, C. H. Zaman, A. Turkyasar, and R. Davis. Following the master' s hands: Capturing piano performances for mixed reality piano learning applications. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA '23, New York, NY, USA, 2023. Association for Computing Machinery.

[17] I. Lee and S. Choi. Vibrotactile guidance for drumming learning: Method and perceptual assessment. In *2014 IEEE Haptics Symposium (HAPTICS)*, pages 147–152, 1, 2014. 1.

[18] R. Liu, E. Wu, C.-C. Liao, H. Nishioka, S. Furuya, and H. Koike. Pianosyncar: Enhancing piano learning through visualizing synchronized hand pose discrepancies in augmented reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 859–868, 1, 2023. 1.

[19] R. Liu, E. Wu, C.-C. Liao, H. Nishioka, S. Furuya, and H. Koike. Pianosyncar: Enhancing piano learning through visualizing synchronized hand pose discrepancies in augmented reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 859–868, Sydney, Australia, 2023. IEEE.

[20] P.-C. Lu, C.-W. Wang, Y. L. Hsu, A. Lopez, C.-Y. Tsai, C.-J. Chang, W. T. M. Tan, L.-C. Lu, and M. Y. Chen. Veer: Exploring the feasibility of deliberately designing vr motion that diverges from mundane, everyday physical motion to create more entertaining vr experiences. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, 2024. Association for Computing Machinery.

doi:10.6342/NTU202500720

[21] G. E. McPherson. *The Child as Musician: A handbook of musical development*. Oxford University Press, 1, 09 2015.

[22] M. Moth-Poulsen, T. Bednarz, V. Kuchelmeister, and S. Serafin. Teach me drums: Learning rhythms through the embodiment of a drumming teacher in virtual reality. In *16th Sound and Music Computing Conference, SMC 2019*, pages 269–273, 1, 2019. 1.

[23] D. Müllensiefen, B. Gingras, J. Musil, and L. Stewart. The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLOS ONE*, 9(2):1–23, 02 2014.

[24] S. Oore. Learning advanced skills on new instruments. *1*, 1(1):60–64, 05 2005.

[25] Paradiddle. Play drums without the limitations of the real world. https://paradiddleapp.com/, 2024. Accessed: December 10, 2024.

[26] J. Pinkl and M. Cohen. Vr drumming pedagogy: Action observation, virtual co-embodiment, and development of drumming. *Electronics*, 12(17):1, 2023.

[27] PotamWorks SAS. Smash drums: Epic vr drumming game. https://www.smashdrums.com/. Accessed: December 10, 2024.

[28] J. Riley. *The art of bop drumming*. Alfred Music Publishing, 1, 1994.

[29] L. H. Robert Rosenthal, Harris Cooper. Parametric measures of effect size. In *The handbook of research synthesis 621*, page 231–244, New York, 1994. Russell Sage Foundation.

[30] Roland Corporation. Company overview: Products, market share, etc. https://ir.roland.com/en/ir/investors/strength/strength-7062976471010175889.html. Accessed: December 10, 2024.

[31] L. R. Skreinig, D. Kalkofen, A. Stanescu, P. Mohr, F. Heyen, S. Mori, M. Sedlmair, D. Schmalstieg, and A. Plopski. guitarhero: Interactive augmented reality guitar

doi:10.6342/NTU202500720

tutorials. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4676–4685, 2023.

[32] L. R. Skreinig, D. Kalkofen, A. Stanescu, P. Mohr, F. Heyen, S. Mori, M. Sedlmair, D. Schmalstieg, and A. Plopski. guitarhero: Interactive augmented reality guitar tutorials. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4676–4685, 2023.

[33] R. Starer. *Rhythmic training*. Mca Music, 1, 1969.

[34] K. Takahashi. Minis: Midi input for new input system. GitHub repository, 2021. Accessed: December 10, 2024.

[35] T. Yamabe, H. Asuma, S. Kiyono, and T. Nakajima. Feedback design in augmented musical instruments: A case study with an ar drum kit. In *2011 IEEE 17th International Conference on Embedded and Real-Time Computing Systems and Applications*, volume 2, pages 126–129, 1, 2011. 1.

s

doi:10.6342/NTU202500720