guitARhero: Interactive Augmented Reality Guitar Tutorials

Lucchas Ribeiro Skreinig, Denis Kalkofen, Ana Stanescu, Peter Mohr, Frank Heyen, Shohei Mori, Michael Sedlmair, Dieter Schmalstieg and Alexander Plopski

Fig. 1: We present a system that can visualize automatically generated guitar tutorials using two alternative AR display modalities: Users can view augmented instructions using a video see-through head-mounted display or a desktop display.

Abstract— This paper presents guitARhero, an Augmented Reality application for interactively teaching guitar playing to beginners through responsive visualizations overlaid on the guitar neck. We support two types of visual guidance, a highlighting of the frets that need to be pressed and a 3D hand overlay, as well as two display scenarios, one using a desktop magic mirror and one using a video see-through head-mounted display. We conducted a user study with 20 participants to evaluate how well users could follow instructions presented with different guidance and display combinations and compare these to a baseline where users had to follow video instructions. Our study highlights the trade-off between the provided information and visual clarity affecting the user's ability to interpret and follow instructions for fine-grained tasks. We show that the perceived usefulness of instruction integration into an HMD view highly depends on the hardware capabilities and instruction details.

Index Terms—Augmented reality, Computer-assisted instruction, Graphical user interfaces

1 Introduction

Early learning of musical instruments is traditionally conducted under the guidance of a more experienced person or teacher [8] and begins with understanding the basics of playing an instrument and with notation literacy. Today's beginners often choose a more informal method of self-teaching from community-created material. A vast amount of such context exists online ranging from video tutorials to sheet music or guitar tablature. However, these approaches come with different drawbacks. Following video tutorials often takes a long time and the quality of the lesson material can vary. Reading sheet music is a skill in itself, and abstract charts disregard complex details such as hand posture and motion in 3D space.

Devices like the *Optek Fretlight*[1] highlight the locations to press down on the guitar neck with embedded LED lights, circumventing the need to learn how to interpret traditional instructions. This has been shown to increase the interaction with the device, learning speed, as well as retention of learned material [22]. At the same time, these devices typically cannot evaluate what mistakes users make. A major drawback in self-teaching, specifically compared to formal learning methods, is the lack of feedback, which can make learning barriers harder to overcome for beginners. Correctly playing a note or chord on the guitar requires familiarity with its components, such as the frets. These separate the guitar neck into segments where the strings can be pressed down to produce a particular note. By pressing a string behind a fret with one hand and strumming one or multiple strings with the other hand, players can produce individual notes or play chords. This complex action introduces a large number of possible error sources, which makes learning by doing particularly difficult for beginners.

Augmented Reality (AR) systems have been proposed as interactive alternatives that offer solutions to the aforementioned problems. In such systems, instructions can be presented in a magic mirror [6] style approach, where cameras capture the view of the guitar and overlay instructions onto the video stream before showing the result on a monitor [19, 21]. Another option is an in situ approach, where virtual content is integrated within a user's natural view using a head-mounted display (HMD) [28, 29]. Furthermore, lesson content can be presented in a number of ways, for example, by highlighting the frets that should be pressed [25], or by animating a virtual hand [21] for users to mimic. Existing solutions implement guidance visualizations, but lack direct visual feedback, require special hardware, or cannot track the instrument. See a summary of limitations in Table 1.

- *Lucchas Ribeiro Skreinig, Ana Stanescu, Peter Mohr, Shohei Mori, Dieter Schmalstieg and Alexander Plopski are with Graz University of Technology. E-mail: {lucchas.ribeiroskreinig | ana.stanescu | mohr | shohei.mori | schmalstieg | alexander.plopski}@icg.tugraz.at*
- *Denis Kalkofen is with Flinders University and Graz University of Technology. E-mail: denis.kalkofen@flinders.edu.au*
- *Frank Heyen and Michael Sedlmair are with the University of Stuttgart. E-mail: {frank.heyen | michael.sedlmair}@visus.uni-stuttgart.de*

[1] https://www.fretlight.com/

Table 1: Comparison of related work with regard to display method, instrument tracking method, guidance technique, feedback capabilities, and specialized instrument use.

Study/product	Display Method	Instrument Tracking	Guidance Visualization	Feedback	Special instrument
Fretlight [12, 22]	LED board	N/A	Embedded markers	No	Yes
LOOG Guitar	Monitor	No	Virtual markers	Yes	Yes
Marky et al. [18]	LED board	N/A	Embedded markers	Yes	Yes
Motokawa/Saito [21]	Monitor	Marker tracking	Virtual 3D hand	No	No
Rio-Guerra et al. [25]	Monitor	Marker tracking	Virtual markers	No	No
Liarokapis et al. [13]	Monitor	No	Chord diagram	No	No
Gutierrez et al. [19]	Monitor	No	Virtual markers/3D hand	No	No
Wang et al. [31]	Monitor	No	Waveform	Yes	No
Kumaravel et al. [28]	HMD	No	Point cloud	No	No
Torres/Figueroa [29]	HMD	Marker tracking	Real images/avatar/virtual markers	No	No
***guitARhero* (ours)**	**Monitor HMD**	**N/A Lighthouse**	**Virtual markers/3D hand**	**Yes**	**No**

We present *guitARhero*[2], which can be used with virtually any acoustic or electric guitar, requiring only a pair of webcams mounted onto the guitar. Similarly to previous approaches [24], guitARhero allows users to easily create interactive lessons, and it provides feedback to users by analyzing the played notes. *guitARhero* can visualize interactive instructions using *fret highlighting* and a *virtual hand* on a desktop monitor using an AR magic mirror and a video see-through (VST) HMD. While the design of our system is based on *AR Hero* [24] this paper introduces an emphasized hand visualization and contributes an evaluation on a VST HMD (Figure 1) that provides a high-fidelity and wide field of view of the surroundings. We conducted a user study with 20 participants to better understand how well users can follow instructions presented on an AR magic mirror and in situ using the two aforementioned guidance methods, compared to using video instructions.

Our findings have implications beyond guitar playing. We show that for fine-grained tasks, hardware capabilities play a major factor in the appeal of naturally integrated instructions, e.g., due to the placement of instructions, low fidelity or a narrow field of view of the headset. We also show that a simple overlay may be preferable over a more detailed visualization that can result in clutter or affect the visibility of the real environment. Overall, we make the following contributions:

- We present *guitARhero*, a system for novice guitar players that allows learners to create their own lessons from online material, receive feedback on their performance, and view instructions using different guidance techniques, as well as various display methods.
- We conducted a user study comparing the *fret highlighting* and *virtual hand* guidance techniques, visualized on a *magic mirror* display or *HMD*. We show that compared to a baseline video, *fret highlighting* on a magic mirror resulted in fewer errors and a better overall user preference.

2 Related Work

guitARhero draws inspiration from a broad spectrum of prior explorations of interactive guitar lessons as well as embedded guitar lessons. A commercial example is the *LOOG Guitar*[3], which can be connected to an application that overlays guidance on the user's instrument. However, the application lacks any tracking ability and requires the guitar to be held in place instead. When defining the characteristics of AR, Azuma [1] emphasized that, without proper registration of real and virtual objects, the illusion that they coexist cannot be sustained. Registration remains one of the major challenges in AR, particularly in developing a system that adequately conveys the learning material to the user while being integrated with the real environment. In the following section, we discuss the interactive and embedded aspects of guitar lessons and highlight how we expand on existing findings.

2.1 Interactive guitar visualization

Interactive guitar education is not a new concept, with commercial applications and games such as *Yousician*[4] and *Rocksmith*[5] providing such services. However, these applications lack integration of the lesson content with the real instrument. Liarokapis and Anderson [14] describe education scenarios in which AR can be used to assist students by visualizing the content of the lesson in more compelling ways. The authors comment that technology encourages discussion between participants and that user-controlled repetition of information is beneficial to the learning process. However, they emphasize that the success of an AR learning scenario strongly depends on its implementation. In prior research, Liarokapis [13] proposed a theoretical system in which users can see where to place their fingers by following a chord diagram displayed on top of their guitar. Rio-Guerra et al. [25] implemented a similar application that highlights spots on the guitar that are color-coded to match stickers on the user's fingers. The authors compared the time taken to learn basic chords using traditional or AR methods and conclude that there is no significant difference in the effectiveness of the two approaches.

Motokawa and Saito [21] proposed a visual guidance method that uses a virtual hand instead of markers on the guitar neck. Users can see themselves and a 3D hand overlaid onto their own guitar on a desktop monitor and imitate the pose by overlapping their hand with the model. However, they did not evaluate whether users were able to follow instructions correctly and focused on the quality of tracking the guitar. Gutierrez et al. [19] implemented a similar application that displays instructions to a user in the form of an animated hand on a monitor. The participants could then adjust the virtual representation of the instructions to obtain a different viewpoint. This process required users to manually align the guitar with the visualization shown on the monitor. They compared how well novices and those with experience could follow the instructions and found that those with experience could follow the instructions faster than beginners.

Wang et al. [31] developed *Soloist*, a system that extracts musical information from guitar tutorial videos to generate interactive lessons. User performances were then captured by a recording interface and compared to the extracted notes. Comparing the performances to the video content allowed them to visualize the users' learning progression and provide instant feedback to the user. They found that the quality of the lessons extracted correlates with the quality of the video content, which can vary, specifically when using free content.

2.2 Embedded guitar lessons

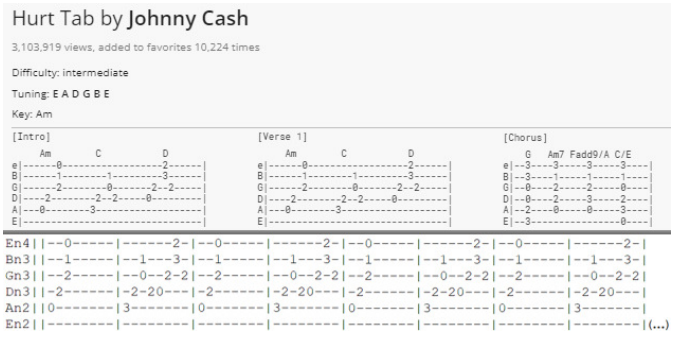
Patzer et al. [22] experimented with a commercial product, the *Fretlight* guitar, an instrument with LED beacons integrated into the fretboard. The beacons are illuminated according to instructions sent by a connected device. The authors suggest that integrating the learning material

[2]The name is meant as an homage to the seminal 2005 rhythm game *Guitar Hero*, developed by *Harmonix* and published by *Activision*.

[3]https://www.loogguitars.com/

[4]https://www.yousician.com/

[5]https://www.ubisoft.com/en-gb/game/rocksmith/



Hurt Tab by Johnny Cash

3,103,919 views, added to favorites 10,224 times

Difficulty: intermediate

Tuning: E A D G B E

Key: Am

```
[Intro]                    [Verse 1]                  [Chorus]
    Am      C       D          Am      C       D          G   Am7 Fadd9/A C/E
e|------0-----------------2------|e|------0-----------------2------|e|--3----3-----3-----3----|
B|------1--------1--------3------|B|------1--------1--------3------|B|--3----1-----1-----1----|
G|------2--------0------2--2-----|G|------2--------0------2--2-----|G|--0----2-----2-----0----|
D|----2--------2--2-----0--------|D|----2--------2--2-----0--------|D|--0----2-----3-----2----|
A|---0--------3------------------|A|---0--------3------------------|A|--2----0-----0-----3----|
E|-------------------------------|E|-------------------------------|E|--3----------------0----|

En4||--0-----|------2-|--0-----|------2-|--0-----|------2-|--0-----|------2-|--0-----|------2-|
Bn3||--1-----|--1---3-|--1-----|--1---3-|--1-----|--1---3-|--1-----|--1---3-|--1-----|--1---3-|
Gn3||--2-----|--0--2-2|--2-----|--0--2-2|--2-----|--0--2-2|--2-----|--0--2-2|--2-----|--0--2-2|
Dn3||-2------|-2-20---|-2------|-2-20---|-2------|-2-20---|-2------|-2-20---|-2------|-2-20---|
An2||0-------|3-------|0-------|3-------|0-------|3-------|0-------|3-------|0-------|3-------|
En2||--------|--------|--------|--------|--------|--------|--------|--------|--------|(...)
```

Fig. 2: (Top) Online lessons in the form of "Ultimate Guitar" tablature can quickly be rewritten as (bottom) text files, which our system parses into interactive lessons.

into a real object reduces the need for users to interpret charts or diagrams, thereby increasing interaction with the device itself, which in turn can boost knowledge retention. Keebler et al. [12] support their findings, stating that *Fretlight* lowers the entry barrier for beginners. Löchtefeld et al. [17] developed a system which uses a mobile projector attached to the guitar head to overlay guidance onto guitars; however, they reported that users' hands blocked the projections in some instances. Marky et al. [18] implemented a similar method that visually guides users with beacons integrated into the fretboard of a custom-built guitar. They use touch sensors to capture the position of the user's finger and stream the captured data to a mobile application that displays the lesson. Using a combination of both visualization methods yielded favorable results in terms of user preference and accuracy. While these methods of embedding learning content into a real guitar are immersive and effective choices of visual guidance, they require access to a custom instrument, reducing their appeal to users who prefer to play their own instrument.

Torres and Figueroa [29] propose a system that guides a user by animating a 3D avatar according to the recorded movements. Their application runs on the *Microsoft HoloLens*[6] HMD which tracks the user's position. A marker attached to the guitar allows tracking the instrument and overlaying virtual charts. The major disadvantage of their method was that the chosen HMD has a narrow field of view, requiring users to adjust their body into uncomfortable positions to view the augmentations, making it difficult to use the system for prolonged periods of time. Despite user-reported discomfort using the HMD, they reported high levels of satisfaction with the application.

Another method of visualizing instructions in 3D is by visualizing captured point clouds in real space. For example, Kumaravel et al. [28] present a system that streams a user's real surroundings as well as virtual annotations to another user. The authors specifically highlight one-on-one guitar lessons as a potential application of their system.

Ribeiro Skreinig et al. [24] introduce a method for AR guitar tutorials using a magic mirror display and a VST HMD. They highlight the shortcomings of existing work with a focus on the visualization of guitar lessons, the embedding of content into real instruments, and the authoring of additional material. While the approach proposes multiple display and guidance alternatives, it does not offer any kind of user-based evaluation. We aim to use the principles they propose and explore further in terms of usability. We present our findings on improving AR tutorials for guitar learning, and integrate user feedback into our own AR visualizations.

3 Requirements

The existing methods address many of the issues associated with AR education strategies, but each has its specific shortcomings (Table 1). From these observations, we distill five design goals, specifically selected to address these issues.

[6] https://docs.microsoft.com/en-us/hololens/

User input and live feedback Most of the described systems aim to visualize lesson content without considering the user's performance. However, since AR is interactive, it can support automatic guidance if the user's performance can be detected at run-time [27]. Thus, we expand upon this one-directional teaching method by capturing the user's played notes. The performance measurements enable us to adapt the lesson speed to the user's skill level and give feedback to the user, potentially reducing the time needed to search for and correct one's mistakes.

Compatibility with standard instruments Existing approaches that rely on customized instruments [12, 18, 22] succeed in integrating lessons with real objects, but at the cost of requiring purpose-built instruments. An AR system that can easily be integrated with a standard instrument has a much broader appeal.

Integration with real objects Previously mentioned solutions often suffer from hardware limitations, such as a narrow field of view, unnatural viewing angles, or a lack of spatial registration [28, 29]. We prefer a solution that is effectively integrated with the user's guitar and prioritizes user comfort.

Guidance visualization Learners may have different preferences with respect to visualization styles, as the design of guidance elements can vary [13, 21, 25]. Hence, we would like to offer flexibility in learning by giving the user a variety of guidance and visualization options to choose from.

Authoring from online sources Previous work often does not discuss how additional content can be added. The ability to utilize existing content greatly increases the appeal of an AR system, if generating new content can be reduced to a quick conversion of existing materials. Guitar beginners can find learning resources on websites such as the popular "Ultimate Guitar"[7] that features free guitar playing instructions (tablature, chords) for a large collection of songs, as well as diagrams and charts designed to help beginners learn guitar basics. An easy way to take advantage of this rich source of learning materials would be desirable for many learners. However, some interpretation or prior music knowledge is required to use these digital tabs, as they lack standardization and may not include timing information.

4 Method

In the following, we describe our methodological choices in terms of visual design and display setup. During the design, we had primarily a self-teaching scenario in mind.

4.1 Visual guidance

Following our design goals, we implement methods to generate lessons and intuitively visualize these abstract instructions. We detail the approaches chosen to address these problems and discuss how our system reacts to a user's performance, as well as the options provided to control the playback of a lesson.

Authoring In order to address our design goal of making the creation of AR learning content as easy as possible, we opted to support tablature files (short tabs). These tabs are text files that use ASCII characters to encode guitar music: six lines indicate the strings, and numbers on the strings encode the fret position (the open string is denoted by zero). Tabs are by far the most common form of sharing guitar music online. By deriving our guitar lessons from existing online materials, they become not only human-readable but also allow users to convert lessons they may already be familiar with into AR tutorials. Since these lessons are stored in text files, they can also be easily shared with other users.

Figure 2 shows an example of guitar tablature, as can be found online on the website "Ultimate Guitar"[8]. The provided tabs at the top of the figure are displayed in sequence. However, as is commonly the case, the instructions for the intro section, the verse, and the chorus are displayed separately and without information as to how often they need

[7] https://www.ultimate-guitar.com/

[8] https://tabs.ultimate-guitar.com/tab/johnny-cash/hurt-tabs-47818/

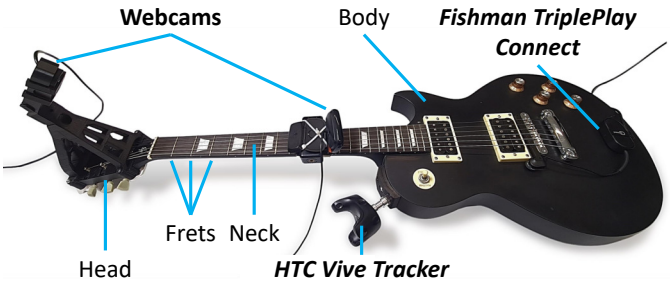


Fig. 3: Key components of a guitar, including our 3D printed attachments. Cameras are rigidly mounted on the guitar's head and neck. An *HTC Vive Tracker* captures the position and rotation of the guitar body, while a *Fishman TriplePlay Connect* device captures string vibrations.

to be repeated. We combined the sections into a single lesson file, as seen at the bottom of the figure, by using the online tabs as a template and copying sections to complete the repeating patterns. Our tool takes such text-based tabs as input and parses them into lessons, but requires instructions to be extended with timing information, which is done by separating notes, represented by numbers, with dashes that determine their timing (i.e., horizontal space encodes time).

In some cases, songs may require the tuning of guitar strings to be adjusted, which can be done by prefacing each line with the tuning of its corresponding guitar string in scientific pitch notation [34]. As an alternative to encoding a song one note at a time, it can be expressed with chord names. Following standard notation, these are written using uppercase and lowercase letters for major and minor chords (e.g., A, C, e, etc., each chord refers to a standard, well-known finger placement).

Interactive visualization As we are targeting a beginner self-teaching scenario, we wanted to enable users to learn the guitar without having to learn any musical notation first. Well-established methods of visualizing musical instructions, such as staff notation, tablature diagrams, and fretboard charts are abstract visualizations and require the user to translate them into actions on the instrument. In contrast, presenting this information through AR intuitively indicates where users should place their fingers on the fretboard.

We chose to investigate two previously proposed highlighting techniques. We implemented a method to guide learners which highlights the fret to be pressed [25, 29], such as the *Fretlight* guitars, for example. Small markers on the virtual fretboard are superimposed between the frets in advance of the note, which gives the user time to place their fingers before strumming each string. As this visualization does not indicate which fingers to use or how to place the hands on the guitar, we also implemented the overlay of a virtual hand model [19, 21] on the neck of the guitar as an alternative approach. The idea of this approach is that by replicating the presented hand pose with their own hand, users can more intuitively understand how to hold the guitar for a particular chord. These guidance methods can be seen in in Figure 4.

We generated and exported a collection of hand models from the 3D human model library, MakeHuman[9]. These models were generated with combinations of varying sizes and finger shapes, giving users the option to customize the appearance of the virtual hand according to their preference or to match the appearance of their own hand, as suggested in previous work [15]. The hand models were rigged for articulation with a hand-shaped armature using the 3D modeling software *Blender*[10]. We apply an inverse kinematics (IK) model to the hand to procedurally animate the finger placement based on the location of the fret; however, the animation may result in unrealistic behavior of the fingers when IK fails to produce a lifelike deformation of the hand model. For this reason, we include a list of poses for the most well-known finger placements of the most commonly used chords by manually adjusting the IK anchors to make the hand pose deform more realistically. Since the lesson material only contains information on which note to play at which time, our system assigns which finger should press which fret depending on their relative location. For each fret in ascending order, the note on the lowest string (E2, in standard tuning) is assigned to the index finger, then the note on the next string is assigned to the middle finger, etc. Creating high-quality finger-to-string assignments from sheet music is a musical expert task in and of itself, which goes beyond the scope of this work.

To provide feedback on the user's performance, it is important to determine which note was played. There exist a variety of methods to detect notes acoustically [7]; however, their reliability remains imperfect for real-time applications. To ensure high-quality multi-pitch detection we opted to use a *Fishman TriplePlay Connect*[11] MIDI controller attached to the guitar body (Figure 3), which is equipped with a hexaphonic pickup that captures string vibrations when the user plucks a string. The MIDI messages inform the AR system about the note played on each string. By comparing the input from the MIDI device with the notes described in the lesson file, we can give feedback on the user's performance. When the user plays the wrong note, a text in the error region informs the user about the mistake, indicating the direction and distance from the correct fret. Moreover, the intended note is marked with a colored highlight when using the highlighted fret guidance. This visualization allows errors to be easily identified.

User-controlled playback The basic interface places a tablature diagram showing the selected practice session and controls either on the monitor or in close proximity to the guitar, depending on the selected mode of guidance. While practicing, especially when starting out, it is important to allow users to practice at their own pace. The playback speed can be adjusted using the interface before starting a lesson or during playback. The rate of new notes is displayed in beats per minute in the user interface, which can help users evaluate their own performance, or restart lessons at known speeds. The user also has the option to restart playback from the beginning, from the current measure, or from a manually chosen interval. We implemented a responsive playback mode in which each beat in the lesson is halted until the user plays the correct combination of notes on the guitar. This procedure adjusts the content to the speed of the user, allowing a newcomer to learn the lesson at their own pace and ultimately helping them to familiarize themselves with the content [23].

4.2 Display

Throughout the development, we applied iterative design to identify and improve important aspects of the overall experience. We implemented two methods of displaying AR content to the user: (1) a magic mirror setup that presents a mirrored video view with AR overlays on a desktop monitor, and (2) a display method that utilizes a VST HMD to integrate AR content into the user's point of view.

Magic mirror display The first display method that we implemented is a desktop setup featuring a magic mirror [6]. A magic mirror is a simple and effective way to combine real and virtual content with high fidelity. It is also widely utilized in prior AR-supported guitar training [19, 21]. One design concern with this method of visualization is the need to track the instrument. Prior implementations require users to align the guitar with a static visualization on the monitor [19] or attach optical markers to track the guitar [21]. In addition, placing a camera at a distance from the user means that the guitar neck will occupy only a small portion of its view, making it difficult to discern which frets should be pressed. Instead, we rigidly attach a 3D printed support to the head of the guitar and mount a lightweight webcam on it so that it captures the neck of the guitar (Figure 3). This approach ensures that the guitar does not move relative to the camera, eliminating the need for any tracking of the instrument. To capture a view of the guitar body, we attach a second webcam to the guitar with a view of the bridge and display the two images side by side to emulate a "panoramic view" of the guitar. This enables the user to see instructions on which strings to strum with their right hand. Furthermore, showing the user

[9] http://www.makehumancommunity.org/
[10] https://www.blender.org/
[11] https://www.fishman.com/portfolio/tripleplay-connect-midi-guitar-controller/

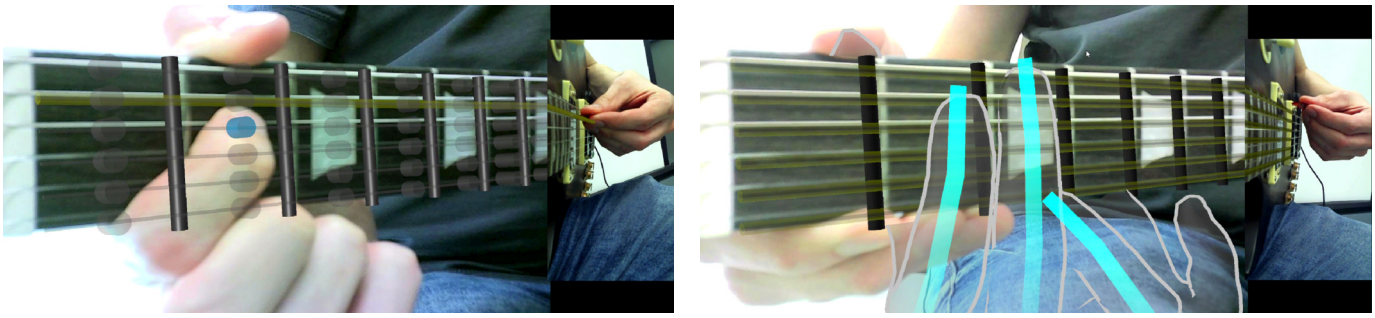


Fig. 4: Guitar instructions are visualized on the guitar neck using (left) virtual fret markers or (right) a virtual hand model to indicate finger positions. We differentiate between (left) plucking or (right) strumming instructions by highlighting one or many strings.

more of their own body in this way can further integrate their feeling of presence and immersion [30].

Head-mounted display Our second display method embeds virtual content within the user's perspective using a VST HMD that lets the user experience the AR instructions from a first-person view. This choice was made due to the observation that the limited field of view of optical see-through (OST) HMDs resulted in user discomfort [29]. A smaller field of view can also result in missed instructions, potentially low visibility due to overlap with the background, and can result in excessive fatigue as learners try to find a good viewpoint [10, 19].

The presentation of visual guidance on a real-world scale has previously been reported to be particularly intuitive [35]. The magic mirror display aligns the lesson content with the real instrument but still represents the instructions separately from the user's own perspective. Integrating the augmentations into the user's own view of the guitar may improve the user experience [2]. Using an HMD, the relative position and rotation between the user's head and the instrument must be tracked. We attached an *HTC Vive Tracker* to the body of the guitar by screwing a 3D printed attachment into the hole typically used to attach a guitar strap (Figure 3), thereby requiring no major modification of the instrument. The controls and tablature are presented in panels floating in front of the user, and their location can be adjusted to better suit the user's preferences. Users can interact with the interface with an *HTC Vive Controller*.

5 Pilot Studies

To evaluate our system's overall ease of use as well as the learnability of its user interface, we performed qualitative evaluations during the development of our prototype. Evaluations were carried out similarly to a cognitive walkthrough [26, 32], whereby participants were asked to play music pieces of different lengths while commenting on their experience. The users' actions during the completion of the given tasks were noted, with a specific focus on whether they knew what to do at each step and whether they could tell if they were making progress toward their goal.

5.1 First evaluation

In the first evaluation, users employed the desktop display for some tasks and an *HTC Vive Pro*[12] headset for others (Figure 5). We recruited ten participants from local universities. Only one of the participants had considerable prior experience playing the guitar, the rest either had no musical experience at all or had some experience with another instrument. Two participants had previously tried to teach themselves the guitar but had quickly abandoned the effort. Therefore, we consider them as having no significant experience. The participants were asked to play chords and a short song using the desktop display and a scale over two octaves using the HMD. Upon completion of each task, we conducted a short discussion in which participants could address any potential difficulties they encountered completing the task and comment on their enjoyment of the system.

Results Participants reported difficulties following the instructions in the *virtual hand* guidance method. These problems were attributed in part to the fingers of the virtual hand partially occluding each other, making it difficult to determine which frets to press down upon. In addition, the hand was rendered using a semitransparent gray color, which made it hard to see against the image of the guitar fretboard. A participant also commented that the placement of the virtual hand did not match their own hand perfectly. Some participants suggested highlighting the fingertips to more clearly indicate the location on which to press and that adjusting the hand to match the size of their own hand more closely was desirable. Despite these shortcomings, the participants who had previously tried to teach themselves with online material commented that the augmentation made learning easier in comparison.

In the *fret highlighting* guidance method, we only highlighted the spots between frets and not the strings, requiring users to simultaneously use the augmented visualization and the tablature. This form of visualization initially confused the participants, as without looking at the tabs, they were unsure which note should be played next; however, they quickly became more comfortable with the guidance method. All ten participants remarked that the *fret highlighting* method of visual guidance was much clearer than the virtual hand and that it helped them find the locations to press on the fretboard more quickly.

When using the HMD with the *fret highlighting* guidance method, all ten participants criticized the low resolution of the *HTC Vive Pro*'s outward-facing cameras. They stated that they could not distinguish the string on which the highlight was shown and relied on the tablature and their feeling to understand what fret to press. Interestingly, two participants still preferred to use the HMD. These results highlighted the need to utilize an HMD with higher visual quality to effectively integrate the tutorials within the real instrument.

In summary, the results of the qualitative evaluation were in line with our expectations. Despite some initial uncertainty, all participants understood the visualizations without requiring much explanation and could tell when they were making progress toward the task goal shortly after starting the tutorial. Whenever the augmentations failed to provide clear instructions due to insufficient visual clarity, all participants resorted to the provided tablature, even those who were previously unfamiliar with that form of musical notation. The lack of visual clarity when using the HMD resulted in participants focusing on the virtual tablature instead of the AR augmentations, highlighting the headset's lack of fidelity for visual integration.

5.2 Second evaluation

During our first evaluation, users encountered difficulties while using the HMD due to the low resolution of its video capture. Therefore, we decided to replace the *HTC Vive Pro* with a *Varjo XR-3*[13] headset that captures the world with cameras with higher resolution than the previous HMD, significantly improving fidelity. We also adjusted the way the instructions were shown to the users. To reduce the reported confusion regarding what strings to strum, we decided to highlight the strings

[12] https://www.vive.com/eu/product/vive-pro/

[13] https://varjo.com/products/xr-3/

Fig. 5: Participants in the first evaluation used (left) the desktop magic mirror display or (center left) the *HTC Vive Pro* head-mounted display. Instead, the second evaluation used (center right) the *Varjo XR-3* head-mounted display. Alternatives to augmented guidance included (right) a tablature view and a virtual mirror.

in addition to the frets (Figure 4). Participants generally preferred to use the *highlighted frets* over the *virtual hands* guidance method due to difficulties in understanding the described hand placement. They also had difficulty identifying the shapes of the individual fingers, so we added an outline to the virtual fingers, as shown in Figure 4 (right). We added a separate virtual window that shows a mirror view of the fret highlights and the virtual hand, which can be seen in Figure 5 (right). This visualization approximates what users previously could see in the magic mirror display method, only without the live webcam video.

The purpose of the second evaluation was to evaluate the performance of the new visualizations on the high-resolution *Varjo XR-3* headset. Ten participants were again asked to perform tasks using the HMD and to comment on their experience. They were asked to play chords using the *virtual hand* guidance method, play a short song intro using the highlighted frets, and play a scale over two octaves using both methods of guidance simultaneously. Upon concluding the tasks, the participants were asked a series of questions aimed at evaluating their impression of individual guidance methods and their experience using the HMD.

Results Tracking the virtual hand on the real guitar helped users to quickly produce the chords by mimicking the model with their own hand. Two of the users specifically commented that seeing how their fingers were meant to be curved helped them avoid accidentally pressing down on other strings while preparing the chord. Two participants, who had previously taken part in the first evaluation round, positively commented on the changes to the visuals of the hand, stating that the outlines made it much easier to tell individual fingers apart.

Highlighting the frets separately from the strings helped the ten participants interpret the markers more quickly on the guitar and led to less confusion compared to the first evaluation. Only one participant stated that using this guidance method was more difficult compared to the virtual hand, as it took them longer to see where to place their fingers using the smaller visualizations. Most of the participants commented that using the augmentations helped them to identify more quickly where their mistakes were.

Using both guidance methods at the same time in the third task meant that the guitar neck was significantly covered by augmentations. Despite this occlusion, six out of the ten participants stated that they found where to press the fretboard more quickly than in the prior tasks. The other four participants commented that the augmentations hid too much of the real image, making it more difficult to see the guitar neck or their own hand.

5.3 Discussion

All six participants who had previously taken part in the first evaluation commented on the improvement in clarity that the new headset offered. Higher visual quality increased their confidence when following the guidance, resulting in faster overall task completion compared to the first evaluation. Overall, when asked which method of guidance they preferred, five participants stated preference for the highlighted frets, while the other five users stated that the combination of both methods helped them more than either of the two individually. Many users specifically commented that they appreciated being able to observe coarse information such as hand placement from the virtual hand, while finer details such as exact finger positions were easier to see using the virtual fretboard. When asked about their impression of the augmented information, five participants commented that the augmentations made errors easier to find compared to traditional self-teaching, and one participant specifically stated that they preferred this method of learning over an instructional video.

6 Evaluation

While our results were encouraging, we could not conclude which of the display and guidance techniques was better overall and whether an HMD presents a benefit over a magic mirror-like system. To this end, we designed a more comprehensive third study that evaluated user performance and perceptual impact of the different techniques.

6.1 Study Design

Our study was inspired by Marky et al. [18], who compared their purpose-built guitar to fretboard charts. We determined two independent variables: guidance technique (*virtual hand*, *highlighted frets*) and display method (*magic mirror*, *HMD*). We also added a *baseline* condition that simulated a typical self-teaching setting, in which participants were presented with a video instructing them how to play a chord without augmentations, resulting in a total of five conditions:

`BL`: *video instructions (baseline)*

`MM-F`: *magic mirror + highlighted frets*

`MM-V`: *magic mirror + virtual hand*

`HMD-F`: *HMD + highlighted frets*

`HMD-V`: *HMD + virtual hand*

We selected ten chords that were deemed equally complex by an experienced guitar player, and each required pressing on three frets simultaneously: *A, Am, Bm7/E, C, D7, D, Dm, E, Fmaj7/E, G*. In each condition, the participants played two of these chords. We utilized a 5×10 balanced Latin square design to avoid order effects. We also balanced the chords and the conditions under which they were played. Whenever participants finished performing the task, they were asked to vocalize the completion.

We rated performance using an error scoring system adapted from Marky et al. [18], where chords were judged on three conditions: incorrect finger position, incorrectly strummed strings, and chord quality

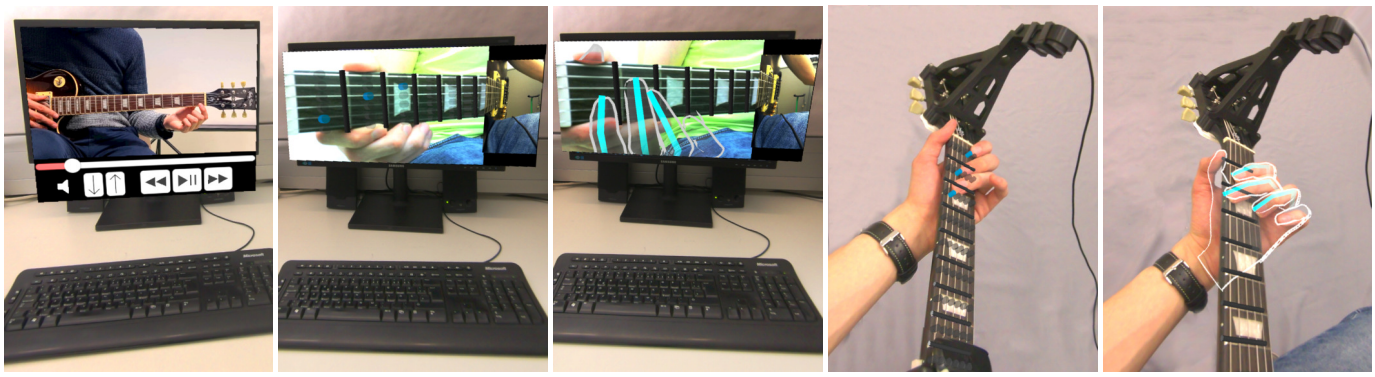


Fig. 6: The five conditions we determined for our user study were: (Left) *Video Instructions*, (center left) *magic mirror + highlighted frets*, (center) *magic mirror + virtual hand*, (center right) *HMD + highlighted frets*, and (right) *HMD + virtual hand*. Screenshots were captured from the video feed of the *Varjo XR-3* for all conditions.

(e.g., fingers accidentally touching multiple strings, strings pressed too softly against the fretboard). A chord was rated between zero and three errors. The completion time was given as the time between presenting the task and vocalizing the completion. After each condition, we collected subjective feedback on the user interface and the workload through SUS [4] and NASA-TLX [9] questionnaires, as well as an informal interview at the end of the experiment. The study did not require approval from the institutional review board.

6.2 Hypotheses

Following our pilot studies and prior literature on guitar learning assistance, we reached the following hypotheses:

- H1: Participants will be faster using guidance augmentations than using video instructions.
- H2: Participants will make fewer mistakes using guidance augmentations than using video instructions.
- H3: Participants will prefer the highlighted frets over the virtual hand visualization.
- H4: Participants will prefer the HMD over the magic mirror.

6.3 Technical Setup

Despite conditions `BL`, `MM-F`, and `MM-V` not requiring an HMD, we used the *Varjo-XR-3* for all experiments (Figure 6), which ensured that users always saw the same scene. Presenting all conditions under the same circumstances prevented biasing the results of the evaluation, e.g. the weight of the headset from influencing users' perception of individual conditions. We attached an *HTC Vive Tracker* and rigidly mounted cameras to the guitar used in all conditions.

As we opted to perform the entire experiment using the HMD, we implemented a virtual monitor on which the AR instructions and captured camera images were presented for both magic mirror conditions and the `BL` condition. We placed the virtual monitor in the location of the physical monitor in the room to create the illusion that the instructions were shown on the monitor. For condition `BL`, we recorded short videos (29 seconds on average) of a guitarist who explained vocally and visually where to place the individual fingers and which strings to strum to produce a chord, followed by a clear audible strum of the chord. Participants could pause and replay the video using a simple virtual button interface (Figure 6, left), and an *HTC Vive Controller*, which simulated self-teaching by video instruction.

Given the positive feedback we received on the visualizations in our second pilot study, we utilized the same representations for the magic mirror condition as well, meaning the virtual hand was displayed with a highlighted outline, and the strings lit up to indicate which strings should be played.

6.4 Participants

We recruited 20 participants, all of whom had not previously participated in the pilot studies, through several methods including social media, forum posts, and email advertisements, among others. The participants were between 19 and 31 years old (mean=24, standard deviation (SD)=3.324), eight of them identified as female, and twelve identified as male. 19 participants indicated that they had little or no prior experience playing the guitar, while one stated that they had learned some chords years before.

6.5 Procedure

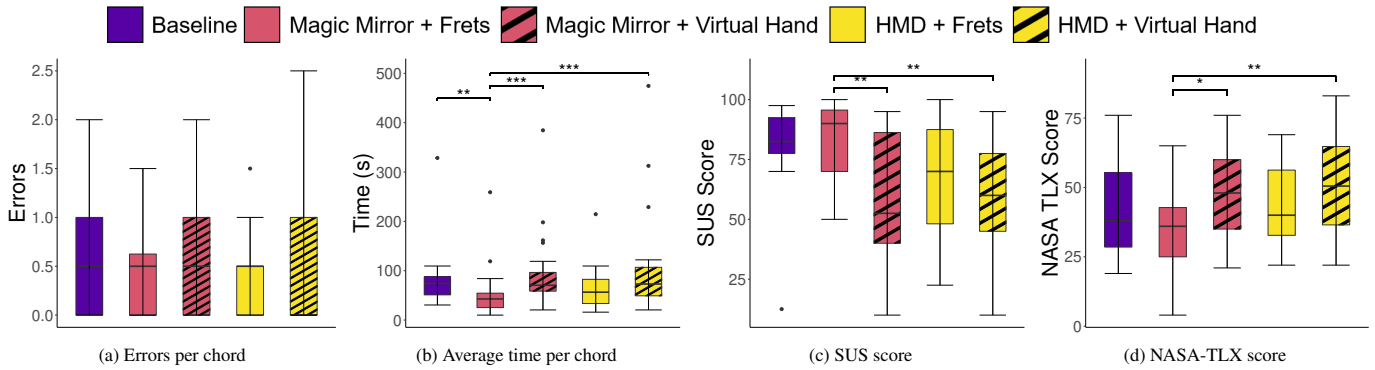
The participants first received a brief explanation of the objective of the experiment and the task procedure. They were asked to sign a consent form and were informed that they could stop the experiment at any moment. After signing the consent form, the participants were asked to sit in a chair in front of a monitor. The *Varjo XR-3* headset was fitted to the participant and an eye-gaze calibration was performed, after which they were handed the guitar.

Each condition began with a familiarization process, in which participants were presented with an example chord (*Em*) which was not included in the actual experiment. Participants could spend time familiarizing themselves with the guidance method and the interface and asking questions. Once satisfied, they could begin the proper experiment by saying "start".

The users were informed that they had as much time as they needed to practice and could end the practice phase by saying "stop" as soon as they felt confident enough to perform the chord. The experiment began as soon as the participants received the first chord. After verbally confirming the completion of the practice phase, they would perform the chord once. This performance was rated using the aforementioned error metric without informing the user of their score. The participants were then informed that the second round would begin and the second chord appeared. Once participants confirmed completion of the second practice phase and performed the second chord, they were shown the questions from the SUS (5-point Likert scale) and NASA-TLX (21-point Likert scale) questionnaires individually. We opted to show the questions on a virtual monitor and allow participants to answer each question verbally using the provided scales. This procedure simplified the questionnaire process and allowed users to hold the guitar for the duration of the experiment. Once all questions were answered, the experiment process was repeated using the next condition. After completing all five conditions, the guitar and the HMD were collected and participants were asked free-form questions about their experience with the individual conditions. This dialogue was also the opportunity for participants to provide unstructured feedback about their experience.

7 Results

As our experiment did not follow a fully factorial design, we performed a two-step analysis in R. As our data was not normally distributed, we



(a) Errors per chord (b) Average time per chord (c) SUS score (d) NASA-TLX score

Fig. 7: Results of our study: (a) The number of errors participants made, (b) how long it took them to learn each chord on average, (c) the combined SUS scores, and (d) the raw NASA-TLX.

first applied aligned rank transform followed by ANOVA [33] using the ARTool package [11]. If significant differences were found, we conducted contrast tests as a post-hoc analysis [5]. We then excluded the data for the baseline conditions and analyzed the remaining data as a 2-way repeated-measures study. As the data also violated normality, we again applied aligned rank transform followed by ANOVA analysis. For questionnaire data analysis we used a Friedman's test with a Nemenyi post-hoc test if any significant differences were detected. We define the level of significance at α=0.05 for all evaluations.

7.1 Effectiveness

To evaluate effectiveness, we compared the number of mistakes that participants made after they finished the practice session (Figure 7a). The number of possible errors ranged from zero to three (chord quality, finger placement, strummed strings) for each tested chord. On average, participants made 0.55 (SD=0.56) errors in condition `BL`, 0.4 (SD=0.447) in `HMD-F`, 0.55 (SD=0.793) in `HMD-V`, 0.475 (SD=0.444) in `MM-F`, and 0.575 (SD=0.568) in `MM-V`. We did not find significant differences in the total number of errors. This result was also the case when we considered the first and second chords played by the participants separately. We also did not find any significant differences after excluding the baseline condition.

7.2 Efficiency

To determine how efficiently the participants learned to perform a chord, we measured the time between presentation of the task and verbal confirmation of practice completion (Figure 7b). On average, participants took 83.1 s (SD=62.3) to learn a chord in condition `BL`, 66.2 s (SD=45.9) in `HMD-F`, 109 s (SD=110) in `HMD-V`, 54.1 s (SD=54.6) in `MM-F`, and 96.2 s (SD=82.3) in `MM-V`. We found that the average learning time was significantly different between the conditions ($F(4,76)=7.2278$, $p<0.001$), with participants completing the learning sessions significantly faster in condition `MM-F` than in `BL`, `HMD-V`, and `MM-V`.

After excluding the baseline, we found that the guidance technique significantly affected users' efficiency ($F(1,19)=29.06141$, $p<0.001$), with conditions that included *highlighted frets* resulting in significantly faster learning times than those that included the *virtual hand*. We found no interaction between guidance technique and display method.

7.3 Satisfaction and Workload

To determine how satisfied the participants were with their experience, we compared the combined SUS score (Figure 7c) and the raw NASA-TLX scores (Figure 7d).

For SUS, the `MM-V` condition was rated with a score of 59.8 (equivalent to a D according to Bangor et al. [3], `HMD-V` a score of 60.5 (equivalent to a D), `HMD-F` a score of 65.5 (equivalent to a D), `BL` a score of 80.4 (equivalent to a B), and `MM-F` a score of 83.5 (equivalent to a B). A Friedman test showed that there were significant differences in the overall scores (χ^2=20.66, $p<0.001$), with `MM-F` being rated significantly higher than `HMD-V` ($p=0.001$) and `MM-V` ($p=0.003$).

Similarly, for NASA-TLX `HMD-V` was rated with 50.4 points, `MM-V` 47.3 points, `HMD-F` 43.7 points, `BL` 41.6 points, and `MM-F` 35.5 points. A Friedman test showed statistically significant differences between the scores (χ^2=15.529, $p=0.004$), with `MM-F` being rated significantly easier than `HMD-V` ($p=0.0035$) and `MM-V` ($p=0.03$).

7.4 Hypotheses

We compared the average speed of the tasks for each condition and found that the participants significantly outperformed the other conditions when using condition `MM-F`, closely followed by condition `HMD-F`. This finding partially supports hypothesis H1 ("Participants will be faster using guidance augmentations than using video instructions."), since users completed their tasks faster using the *highlighted frets*, although they were comparatively slower than the baseline condition using both display methods guided by the *virtual hand*.

Hypothesis H2 ("Participants will make fewer mistakes using guidance augmentations than using video instructions.") could not be supported by our findings, as our evaluation did not find significant differences in the average number of errors between the five conditions. Users tended to respond equally well to AR instructions as to the baseline condition, which presented both video and audio information to the user.

The results of the SUS and raw NASA-TLX questionnaires show that, when comparing the AR-enabled conditions (discounting the baseline), users prefer the *highlighted fret* guidance method over the *virtual hand* guidance, which supports hypothesis H3 ("Participants will prefer the highlighted frets over the virtual hand visualization."). Although some users commented that they enjoyed the virtual hand model and the fact that it helped them select the correct fingers for each fret, it is not sufficient to guide users. Our second pilot study indicated that users enjoyed a combination of the virtual hand and the highlighted frets, which could be worth exploring in the future.

The same questionnaire results refuted our hypothesis H4 ("Participants will prefer the HMD over the magic mirror.") as conditions that included the magic mirror display had higher overall satisfaction results compared to HMD conditions. The fine-grained task of placing one's fingers on a guitar fretboard was shown to be more engaging when displayed from a closer point of view on a static, high-resolution display, rather than integrated within the user's perspective. This finding is logically consistent, as the guitar is at a significant distance from the user's head during use, and the augmentations appear significantly smaller on the HMD than on the magic mirror. A possible implication beyond guitar playing is that, despite high-fidelity hardware, the placement of instructions remains an important consideration in AR design.

7.5 Discussion

Despite our `HMD-F` condition being similar to existing solutions that highlight the locations to be pressed on the guitar, e.g., *Fretlight*, this condition was not preferred to the magic mirror view. We believe that this is mainly because the augmentations appear quite small when seen

at the distance the guitar is held. In addition, users had to rotate the guitar and adjust their viewpoint to see the augmentations. On the other hand, the magic mirror view allowed users to remain in a comfortable, forward-facing position, while still seeing their own hand and the locations on the guitar neck to press their fingers. Subjective comments by the participants were collected during and after the experiment and provided anecdotal explanations as to why users tended to prefer certain conditions. Users commented, for example, that virtual fingers were hard to see from their perspective. Another explanation could be that, due to tracking noise and capture latency of the *Varjo XR-3*, participants experienced minor drift of the augmentations, making it more difficult to interpret the augmented instructions. This result shows that, while one may expect that direct augmentations are preferable, a remote view may be the better choice for precise tasks.

The questionnaire results showed that users were less satisfied with the guidance in the form of a virtual hand, which was supported by some participants' comments that the virtual hand was too big and distracted from the lesson. Users also commented that the highlighted frets were more precise and less confusing overall. One possible explanation may be that in the *HMD* conditions, the augmentation and the participant's own hand needed to be in the same place, introducing visual clutter and making it difficult for participants to determine whether they were pressing on the correct frets. Another explanation could be that the virtual hand did not correctly match the participant's own hands or was placed in a way that participants found uncomfortable, as indicated by some of the participants in our first pilot study. In the second evaluation, some participants chose to adjust the size of the hand, suggesting that customization of this guidance approach could make it clearer.

Compared to the findings of Marky et al. [18], our baseline condition showed a video of the instructions with detailed audio descriptions of the intended hand poses, not only a fretboard chart. This meant that participants not only received step-by-step instructions on their finger placement, but also could hear the sound of the chord. The ability to match the chord acoustically was rated very positively by the participants and was named as the primary reason for their preference for this condition. Adding the possibility to hear the correct sound, e.g., triggered through voice commands, would likely improve the usability and satisfaction of our system.

We found that visualizations based on highlighting frets are generally easy to understand but still require a concentrated effort. The drawback of the alternative method of animating a virtual hand is its inherent ambiguity, since the application must approximate which finger should be assigned to which note. Other issues can occur when IK calculations result in unrealistic hand poses. This is especially true when it comes to the animation of the hand's movement, as the IK model currently simply interpolates between poses. This can produce unrealistic or impossible motions of the hand and fingers, which could cause user frustration as the instructions become difficult to follow accurately. This confusion may further decrease user satisfaction with the guidance method.

8 Conclusions and Future Work

This paper reports on a method to generate teaching content that is not only fast and easy but requires little effort from the user. Deriving the format of our instruction files from online content makes our system appealing to users who want to expand their library of lessons, as well as to users who are already familiar with the commonly used musical notation. Giving the user the choice between an abstract visualization of the instructions, a 3D animated virtual hand, or a combination of both allows them to customize their learning experience and can give them information beyond what common notations can offer, such as hand poses.

Using our system, we empirically examined users' performance and acceptance when using combinations of different AR display and guidance methods. Based on our qualitative analysis, we conclude that these features contribute to the user's enjoyment of the system, as well as the system's ease of use. Interestingly, integrating virtual content within a user's natural view did not result in the highest user satisfaction. Although some users commented that they enjoyed seeing the position of instructions on the instrument itself, the highest overall satisfaction score went to the condition that visualized augmentations on a magic mirror, even when compared to audiovisual tutorials. We conclude that integrating instructions within the manipulated object makes logical sense as long as it is not done at the expense of visual clarity. This is further supported by the results of our pilot studies, given the positive response to the improvements implemented after the first evaluation, most notably the visual upgrade of the HMD.

When comparing the five conditions, the lack of a significant difference in the overall number of errors indicates that our proposed system has the potential to instruct newcomers who have absolutely no prior experience playing the guitar. To better gauge the efficacy of our system, future work is planned to include a longitudinal user study to better determine the effects of the implemented features on its teaching capabilities. We plan on measuring students' knowledge retention by observing whether their performance scores increase over multiple iterations of playing the same chord in the same conditions. Long-term observations are needed to give a clearer picture of the benefits of one form of visual guidance over another. We also wish to observe whether user customization of the instructions improves satisfaction scores.

The results of our evaluations show that AR training for fine-grained motor tasks, such as playing the guitar, is feasible, but requires sufficiently capable technology. Apart from the improved user interface on a higher-fidelity VST HMD, additional improvements need to be made to increase user enjoyment and address some of the reported issues found during our evaluation. One option to introduce a magic mirror-like view of the instructions into the user's natural view could be to place an additional view close to the augmentations on the guitar. This would allow users to comfortably see the instructions while also confirming them on the guitar if needed. Participants in our second pilot study enjoyed this alternative visualization.

We expect that the user interface can be optimized and the overall application be made more intuitive. Such an overhaul may reduce confusion and increase the user's focus on learning the guitar. The users of the pilot studies positively remarked on the ability to choose how the content of the lessons was presented. By implementing alternative guidance methods, such as audio cues, we can cater to more diverse learning styles. Currently, the guidance visualizations highlight the next note to be played, requiring users to search for the next augmentation as they play. We suggest that visualizing instructions further in advance of the next note may decrease the time needed to search for the next augmentation. This may increase the efficiency of our guidance techniques, as precueing augmented instructions has been shown to benefit user performance [16]. In addition, the system only tracks which notes the user is playing so far. By implementing additional feedback, we could allow for a more comprehensive evaluation of the user's performance. Examples may include detecting a user's hand pose or the tempo of their playback.

Our user study demonstrated the system's ability to successfully create AR lessons from text files. However, we believe that the potential of our system goes beyond just parsing tablature. The drawback of the current source of instructions is its lack of supplementary information, particularly when it comes to finger techniques, as it cannot assign the fingers to specific notes or include more complicated instructions than when to play which note. The implemented guidance techniques and display methods could be used to visualize the instructions captured by a remote instructor [20]. This could enable users to receive more detailed instructions or learn advanced techniques, thereby expanding the scope of the system. Furthermore, our system could be used for remote streaming of another user's captured performance. This would allow users to collaborate with other musicians or receive feedback from others, further enhancing the overall learning experience.

Acknowledgments

This work was partly supported by the Austrian Science Fund *FWF* (grant no. P30694), by Deutsche Forschungsgemeinschaft *DFG* under the Excellence Strategy (EXC 2075 – 390740016), the Cyber Valley Research Fund, and by Snap, Inc.

References

[1] R. Azuma. A Survey of Augmented Reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385, 1997. doi: 10.1162/pres.1997.6.4.355 2

[2] Z. Bai and A. F. Blackwell. See-through window vs. magic mirror: A comparison in supporting visual-motor tasks. In *Proc. International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 239–240, 2013. doi: 10.1109/ISMAR.2013.6671784 5

[3] A. Bangor, P. Kortum, and J. Miller. An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction*, 24(6):574–594, 2008. doi: 10.1080/10447310802205776 8

[4] J. Brooke. SUS: A quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194):4–7, 1996. 7

[5] L. A. Elkin, M. Kay, J. J. Higgins, and J. O. Wobbrock. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *Proc. ACM Symposium on User Interface Software and Technology (UIST)*, p. 754–768, 2021. doi: 10.1145/3472749.3474784 8

[6] M. Fiala. Magic Mirror System with Handheld and Wearable Augmentations. In *Proc. IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 251–254, 2007. doi: 10.1109/VR.2007.352493 1, 4

[7] T. Goodman and I. Batten. Real-Time Polyphonic Pitch Detection on Acoustic Musical Signals. In *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 1–6, 12 2018. doi: 10.1109/ISSPIT.2018.8642626 4

[8] E. Harrison. Challenges Facing Guitar Education. In *Music Educators Journal*, vol. 97, pp. 50–55, 2010. doi: 10.1177/0027432109334421 1

[9] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, vol. 52 of *Advances in Psychology*, pp. 139–183. North-Holland, 1988. doi: 10.1016/S0166-4115(08)62386-9 7

[10] F. Heyen and M. Sedlmair. Augmented Reality Visualization for Musical Instrument Learning. In *International Society for Music Information Retrieval Conference (ISMIR)*, 2022. 5

[11] M. Kay, L. A. Elkin, J. J. Higgins, and J. O. Wobbrock. ARTool: R Package for the Aligned Rank Transform for Nonparametric Factorial ANOVAs. R Package Version 0.11.1, 2021. doi: 10.5281/zenodo.594511 8

[12] J. R. Keebler, T. J. Wiltshire, D. C. Smith, S. M. Fiore, and J. S. Bedwell. Shifting the Paradigm of Music Instruction: Implications of Embodiment Stemming From an Augmented Reality Guitar Learning System. In *Frontiers in Psychology*, vol. 5, 2014. doi: 10.3389/fpsyg.2014.00471 2, 3

[13] F. Liarokapis. Augmented Reality Scenarios for Guitar Learning. In *Proc. of Theory and Practice of Computer Graphics*, pp. 163–170, 2005. 2, 3

[14] F. Liarokapis and E. F. Anderson. Using Augmented Reality as a Medium to Assist Teaching in Higher Education. In *Eurographics 2010 - Education Papers*. The Eurographics Association, 2010. doi: 10.2312/eged.20101010 2

[15] L. Lin, A. Normoyle, A. Adkins, Y. Sun, A. Robb, Y. Ye, M. Di Luca, and S. Jörg. The Effect of Hand Size and Interaction Modality on the Virtual Hand Illusion. In *Proc. IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 510–518, 3 2019. doi: 10.1109/VR.2019.8797787 4

[16] J. Liu, B. Tversky, and S. Feiner. Precueing Object Placement and Orientation for Manual Tasks in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 28:3799–3809, 2022. doi: 10.1109/TVCG.2022.3203111 9

[17] M. Löchtefeld, S. Gehring, R. Jung, and A. Krüger. Using Mobile Projection to Support Guitar Learning. In *Smart Graphics*, vol. 6815, pp. 103–114, 2011. doi: 10.1007/978-3-642-22571-0_9 3

[18] K. Marky, A. Weiß, A. Matviienko, F. Brandherm, S. Wolf, M. Schmitz, F. Krell, F. Müller, M. Mühlhäuser, and T. Kosch. Let's Frets! Assisting Guitar Students During Practice via Capacitive Sensing. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 1–12, 05 2021. doi: 10.1145/3411764.3445595 2, 3, 6, 9

[19] J. Martin-Gutierrez, M. S. D. R. Guerra, V. Lopez-Chao, R. H. S. Gastelum, and J. F. V. Bojórquez. Augmented reality to facilitate learning of the acoustic guitar. In *Applied Sciences*, vol. 10(7), 4 2020. doi: 10.3390/app10072425 1, 2, 4, 5

[20] P. Mohr, S. Mori, T. Langlotz, B. H. Thomas, D. Schmalstieg, and D. Kalkofen. Mixed Reality Light Fields for Interactive Remote Assistance. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 1—12, 2020. doi: 10.1145/3313831.3376289 9

[21] Y. Motokawa and H. Saito. Support System for Guitar Playing Using Augmented Reality Display. In *Proc. International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 243–244, 2006. doi: 10.1109/ISMAR.2006.297825 1, 2, 3, 4

[22] B. Patzer, D. Smith, and J. Keebler. Novelty and Retention for Two Augmented Reality Learning Systems. In *Human Factors and Ergonomics Society Annual Meeting*, vol. 58(1), pp. 1164–1168, 2014. doi: 10.1177/1541931214581243 1, 2, 3

[23] S. Pongnumkul, M. Dontcheva, W. Li, J. Wang, L. Bourdev, S. Avidan, and M. F. Cohen. Pause-and-Play: Automatically Linking Screencast Video Tutorials with Applications. In *Proc. ACM Symposium on User Interface Software and Technology (UIST)*, pp. 135–144, 2011. doi: 10.1145/2047196.2047213 4

[24] L. Ribeiro Skreinig, A. Stanescu, S. Mori, F. Heyen, P. Mohr-Ziak, M. Sedlmair, D. Schmalstieg, and D. Kalkofen. AR Hero: Generating Interactive Augmented Reality Guitar Tutorials. In *Proc. IEEE Virtual Reality and 3D User Interfaces Abstracts and Workshops*, pp. 395–401, 2022. doi: 10.1109/VRW55335.2022.00086 2, 3

[25] D. Rio-Guerra, M. Sylvia, J. Martin-Gutierrez, V. A. Lopez-Chao, R. Flores Parra, and M. A. Ramirez Sosa. AR Graphic Representation of Musical Notes for Self-learning on Guitar. *Applied Sciences*, 9(21):4527, 2019. doi: 10.3390/app9214527 1, 2, 3, 4

[26] R. Spencer. The Streamlined Cognitive Walkthrough Method, Working around Social Constraints Encountered in a Software Development Company. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI)*, p. 353–359, 2000. doi: 10.1145/332040.332456 5

[27] A. Stanescu, P. Mohr, D. Schmalstieg, and D. Kalkofen. Model-free authoring by demonstration of assembly instructions in augmented reality. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 28(11):3821–3831, 2022. doi: 10.1109/TVCG.2022.3203104 3

[28] B. Thoravi Kumaravel, F. Anderson, G. Fitzmaurice, B. Hartmann, and T. Grossman. Loki: Facilitating Remote Instruction of Physical Tasks Using Bi-directional Mixed-reality Telepresence. In *Proc. ACM Symposium on User Interface Software and Technology (UIST)*, pp. 161–174, 10 2019. doi: 10.1145/3332165.3347872 1, 2, 3

[29] C. Torres and P. Figueroa. Learning How to Play a Guitar with the HoloLens: A Case Study. In *Proc. XLIV Latin American Computer Conference (CLEI)*, pp. 606–611, 2018. doi: 10.1109/CLEI.2018.00078 1, 2, 3, 4, 5

[30] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The Impact of Avatar Personalization and Immersion on Virtual Body Ownership, Presence, and Emotional Response. In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 24, pp. 1643–1652, 2018. doi: 10.1109/TVCG.2018.2794629 5

[31] B. Wang, M. Y. Yang, and T. Grossman. Soloist: Generating Mixed-Initiative Tutorials from Existing Guitar Instructional Videos Through Audio Processing. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 1–14, 5 2021. doi: 10.1145/3411764.3445162 2

[32] C. Wharton, J. Rieman, C. Lewis, and P. Polson. *The Cognitive Walkthrough Method: A Practitioner's Guide*, p. 105–140. John Wiley & Sons, Inc., 1994. 5

[33] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI)*, p. 143–146, 2011. doi: 10.1145/1978942.1978963 8

[34] R. W. Young. Terminology for Logarithmic Frequency Units. *The Journal of the Acoustical Society of America*, 11(1):134–139, 1939. doi: 10.1121/1.1902141 4

[35] X. Yu, K. Angerbauer, P. Mohr, D. Kalkofen, and M. Sedlmair. Perspective Matters: Design Implications for Motion Guidance in Mixed Reality. In *Proc. International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 577–587, 2020. doi: 10.1109/ISMAR50242.2020.00085 5