UNIVERSITY OF BREMEN

Augmented Virtuality for Trombone Learning: A Prototype for Real-Time Slide Position Guidance and Precision Evaluation

---

Bachelor's Thesis

Author:
Bastian Springer

Supervisor:
Dr. René Weller

First Examiner:
Prof. Dr. Gabriel Zachmann

Second Examiner:
Dr. René Weller

September 23, 2026

Universität Bremen

Eigenständigkeits- und Einverständniserklärung zur Überprüfung mit Plagiatssoftware sowie die Erklärung zur Veröffentlichung bei Bachelor- und Masterarbeiten

Declarations of Authorship and Consent for Checking with Plagiarism Software and the Declaration of Publication for Bachelor's and Master's Thesis

Studierenden-Angaben / Student Information:

Matrikelnr./ Student ID	**6191247**
Nachname / Surname	**Springer**
Vorname / First Name	**Bastian**

Titel der Arbeit / Title of Thesis

Augmented Virtuality for Trombone Learning:
A Prototype for Real-Time Slide Position Guidance and Precision Evaluation

A) Eigenständigkeitserklärung / Declaration of Authorship

Ich versichere, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Teile meiner Arbeit, die wortwörtlich oder dem Sinn nach anderen Werken entnommen sind, wurden unter Angabe der Quelle kenntlich gemacht. Gleiches gilt auch für Zeichnungen, Skizzen, bildliche Darstellungen sowie für Quellen aus dem Internet, dazu zählen auch KI-basierte Anwendungen oder Werkzeuge.
Die Arbeit wurde in gleicher oder hnlicher Form noch nicht als Prüfungsleistung eingereicht.

I hereby affirm that I have written the present work independently and have used no sources or aids other than those indicated. All parts of my work that have been taken from other works, either verbatim or in terms of meaning, have been marked as such, indicating the source. The same applies to drawings, sketches, pictorial representations and sources from the Internet, including AI-based applications or tools. The work has not yet been submitted in the same or a similar form as a final examination paper.

☑ Ich habe KI-basierte Anwendungen und/oder Werkzeuge genutzt und diese im Anhang "Nutzung KI basierte Anwendungen" dokumentiert.

I have used AI-based applications and/or tools and documented them in the appendix "Use of AI-based applications".

B) **Erklärung zur Veröffentlichung von Bachelor- oder Masterarbeiten**
Declaration regarding the publication of bachelor's or master's thesis

Die Abschlussarbeit wird zwei Jahre nach Studienabschluss dem Archiv der Universität Bremen zur dauerhaften Archivierung angeboten. Archiviert werden:

Two years after graduation, the thesis is offered to the archive of the University of Bremen for permanent archiving. The following are archived:

1) Masterarbeiten mit lokalem oder regionalem Bezug sowie pro Studienfach und Studienjahr 10 % aller Masterarbeiten
Master's theses with a local or regional focus, as well as per subject and academic year 10 % of all Master's thesis

2) Bachelorarbeiten des jeweils ersten und letzten Bachelorabschlusses pro Studienfach und Jahr.
Bachelor's thesis for the first and last bachelor's degrees per subject and year.

☑ Ich bin damit einverstanden, dass meine Abschlussarbeit im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.
I agree that my thesis may be viewed by third parties in the university archive for academic purposes.

☑ Ich bin damit einverstanden, dass meine Abschlussarbeit nach 30 Jahren (gem. §7 Abs. 2 BremArchivG) im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.
I agree that my thesis may be viewed by third parties for academic purposes in the university archive after 30 years in accordance with para. 2 BremArchivG).

☐ Ich bin **nicht** damit einverstanden, dass meine Abschlussarbeit im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.
I do not consent to my thesis being made available in the university archive for third parties to view for academic purposes.

C) Einverständniserklärung zur elektronischen Überprüfung der Arbeit auf Plagiate
Declaration of consent for electronic checking of the work for plagiarism

Eingereichte Arbeiten können nach § 18 des Allgemeinen Teil der Bachelor- bzw. der Masterprüfungsordnungen der Universität Bremen mit qualifizierter Software auf Plagiatsvorwürfe untersucht werden.
Zum Zweck der Überprüfung auf Plagiate erfolgt das Hochladen auf den Server der von der Universität Bremen aktuell genutzten Plagiatssoftware.

Submitted papers can be checked for plagiarism using qualified software in accordance with § 18 of the General Section of the Bachelor's or Master's Degree Examination Regulations of the University of Bremen. For the purpose of checking for plagiarism, the upload to the server is done using the plagiarism software currently used by the University of Bremen.

☑ Ich bin damit einverstanden, dass die von mir vorgelegte und verfasste Arbeit zum oben genannten Zweck dauerhaft auf dem externen Server der aktuell von der Universität Bremen genutzten Plagiatssoftware, in einer institutionseigenen Bibliothek (Zugriff nur durch die Universität Bremen), gespeichert wird.

I agree that the work I have submitted and written will be stored permanently on the external server of the plagiarism software currently used by the University of Bremen, in a library belonging to the institution accessed only by the University of Bremen), for the above-mentioned purpose.

☐ Ich bin **nicht** damit einverstanden, dass die von mir vorgelegte und verfasste Arbeit zum o.g. Zweck dauerhaft auf dem externen Server der aktuell von der Universität Bremen genutzten Plagiatssoftware, in einer institutionseigenen Bibliothek (Zugriff nur durch die Universität Bremen), gespeichert wird.

I do not consent to the work I submitted and wrote being permanently stored on the external server of the plagiarism software currently used by the University of Bremen, in a library belonging to the institution accessed only by the University of Bremen), for the above-mentioned purpose.

Das Einverständnis der dauerhaften Speicherung des Textes ist freiwillig. Die Einwilligung kann jederzeit durch Erklärung gegenüber der Universität Bremen, mit Wirkung für die Zukunft, widerrufen werden. Weitere Informationen zur Überprüfung von schriftlichen Arbeiten durch die Plagiatsoftware sind im Nutzungs- und Datenschutzkonzept enthalten. Diese finden Sie auf der Internetseite der Universität Bremen.

Consent to the permanent storage of the text is voluntary. Consent can be withdrawn at any time by making a declaration to this effect to the University of Bremen, with effect for the future. Further information on the checking of written work using plagiarism software can be found in the data protection and usage concept. This can be found on the University of Bremen website.

Mit meiner Unterschrift versichere ich, dass ich die obenstehenden Erklärungen gelesen und verstanden habe und bestätige die Richtigkeit der gemachten Angaben.

With my signature, I confirm that I have read and understood the above explanations and confirm the accuracy of the information provided.

23.9.2026 Barbara Sprungmann

Datum / Date Unterschrift als Scan/ Signature

Acknowledgements

I would like to thank everyone who supported me throughout the process of writing my bachelor thesis.

First and foremost, I would like to thank my parents for their continual encouragement and physical support during the setup of the user-study. Whenever I began to doubt my work, they were always there to support me.

I am especially grateful for a dear friend who gave me a place to calm down and relax, which I struggle with, and provided comfort whenever I felt down. His presence during this particularly difficult period greatly helped me in finishing this thesis.

Furthermore, I also appreciate the support, and occasional play testing from my friends and brother over this time.

My thanks also go to everyone who proofread my thesis and helped me eliminate some of the remaining mistakes.

I would also like to express my gratitude to everyone who contributed to the user study. Therefore, I would like to thank to all participants that took part in this study, as well as Mr. Rohde and Maike Diedrichs, who helped to recruit additional participants.

Finally, I am thankful for Dr. René Weller and Prof. Dr. Gabriel Zachmann for giving me the opportunity to explore this topic for my bachelor thesis.

Abstract

This thesis develops and evaluates a prototype for a slide positioning support system utilizing extended reality. While conventional trombone lessons with an instructor are the best way to advance technically, such instant feedback is not available when practicing independently. The developed prototype addresses this limitation, by providing real-time feedback on slide positioning during practice.

Utilizing six degrees of freedom tracking technology the real trombone gets translated into an XR environment, allowing the trombone itself to become a UI input device to not rely on a controller and increase immersion. A MIDI-based visualization guides the player with both the traditional staff system next to a slide positioning support system. By evaluating the tracking information against a theoretical model of the trombone, the prototype provides instant feedback on the positioning accuracy.

A small user-study has been conducted on this prototype to gain initial indications on the usability of the prototype, possible practice motivation increase, and effectiveness in improving positioning. Participants practiced a piece of music and two scales for ten minutes before performing without assistance from the prototype. The results show that the prototype is highly usable, may have benefits for motivation and increases the precision while practicing. However while performing, the experimental group only showed small improvements over the control group, with no statistically significant difference. Therefore, this study lacks sufficient evidence, that the program creates lasting improvements.

Overall, the results indicate that the developed XR prototype provides a promising environment to support trombone practice outside of conventional instrumental lessons. Though the results did not provide sufficient evidence that this change lasts after 10 minutes of training, they indicate that further investigation of the topic may be worthwhile, possibly after further improvements to the prototype.

Contents

1 Introduction **1**

1.1 Motivation 1

1.2 Goals and Hypotheses 2

2 Background **5**

2.1 XR 5

2.2 The Trombone 6

2.2.1 Overview and Functionality 6

2.2.2 Variants 7

3 Related Work **8**

3.1 Similar Work 8

3.2 Trombone Pedagogy 8

3.2.1 Slide Positioning 9

3.2.2 Additional Challenges 10

3.3 Intuitive Visualization of Trombone Play Instructions 10

3.4 Tracking 11

3.4.1 Distance Tracking 12

3.4.2 6DoF Tracking 13

3.4.3 Embouchure and Air Pressure 16

4 Methodology **17**

4.1 Technology and Software 17

4.2 Functionality 18

4.3 Development 20

4.3.1 Code Structure 20

4.3.2 Trombone Data 22

4.3.3 UI Input System . . . 25
4.3.4 Acoustic Model . . . 26
4.3.5 Play Instructions and Evaluation . . . 35
4.4 Experiment . . . 38
4.4.1 Music Choice . . . 38
4.4.2 Questionnaire . . . 39
4.4.3 Procedure . . . 40
4.4.4 Data Collection . . . 42

5 Evaluation **43**
5.1 Hypothesis 1 . . . 45
5.1.1 Hypothesis test . . . 45
5.1.2 Additional Analyses . . . 45
5.2 Hypothesis 2 . . . 47
5.2.1 Hypothesis test . . . 47
5.2.2 Additional Analyses . . . 48
5.3 Hypothesis 3 . . . 49
5.3.1 Data Preparation . . . 49
5.3.2 Hypothesis test . . . 51
5.3.3 Additional Analyses . . . 52
5.4 Hypothesis 4 . . . 55
5.4.1 Data Preparation . . . 55
5.4.2 Hypothesis test . . . 56
5.4.3 Additional Analyses . . . 57
5.5 General Questionnaire Evaluation . . . 59

6 Discussion **61**
6.1 User Opinion . . . 61
6.2 Trombone Data . . . 63
6.3 Study Reflection . . . 65

7 Conclusion **67**
7.1 Limitations . . . 69
7.2 Future work . . . 70

A Lists **77**
List of Figures . . . 78
List of Tables . . . 79

B "A Theoretical Trombone" Pseudocode **80**

C User-Study Questionnaire **82**
C.1 Part 1 . . . 82
C.1.1 Demographic Information . . . 82
C.1.2 Musical Experience . . . 82
C.2 Part 2 . . . 83
C.2.1 Usability . . . 83
C.2.2 Playing Experience . . . 84
C.2.3 Program Evaluation . . . 84
C.3 Results . . . 86

D Scripts **88**
D.1 H3 Accuracy Calculation . . . 88

E AI Use **90**

Chapter 1

Introduction

1.1 Motivation

The unique component that makes the trombone stand out from a lot of other brass instruments and instruments in general, is the slide, which is the main way of altering the played pitch alongside embouchure. Most other brass instruments instead rely mainly on valves, again in combination with embouchure, to change the pitch. These valves provide a fixed set of tube-length combinations that lower the pitch by discrete intervals.[1] In contrast, the trombone slide allows continuous variation of the tube length enabling the player to freely move between the equivalent valve positions. This also means the pitch of the instrument is directly related to the extent of the physical movement of the arm of the trombonist. While standard slide positions exist, their exact locations vary slightly across different partials. As a result, learning to accurately locate and adjust these positions poses a unique challenge when learning the trombone.

A big part in doing this is to train muscle memory and one's ear to detect when the note is slightly off. Since most players do not have the ability of absolute pitch (/perfect pitch), this relies on using reference points, such as the intonation of preceding notes or playing in an ensemble, where one must adjust to others. This can make the learning process harder for beginners when they have to practice alone. While a tuner can help in establishing accurate pitch, it is primarily designed for achieving and checking single, isolated pitches rather than guiding intonation during musical phrasing. During play, it is difficult to use as guidance while also having to reading sheet music and maintaining the

[1] Many Brass instruments also allow for small adjustments during play with the so-called valve slide(s) comparable to a *mini slide*.

performance. Similarly, ensemble playing and instructor feedback are effective but not always available during individual practice sessions.

This creates a gap between what learners need and what conventional practice environments provide. Extended Reality (XR) offers a potential approach to this problem by directly integrating real-time visual and evaluative feedback into the player's practice sessions. This could provide real-time guidance on slide positioning without diverting the players' attention from playing and additionally introduces the possibility for gamified elements.

While existing approaches explore the use of XR for such visual feedback, the actual instrument is not used. Instead, controllers function as stand-ins for the trombone and therefore lack the tactile and mechanical feedback of a real trombone, making it not suitable for actual practice sessions.

1.2 Goals and Hypotheses

For this Thesis I aim to create a prototype for real-time visual feedback in XR. This prototype shall provide

- a connection between the real instrument and a virtual representation,
- generating intuitive instructions and evaluation for the slide position based on a universal musical file structure, and
- an intuitive user interface to bridge the connection between instrument training and application navigation.

Furthermore, I will conduct a user study testing the usability and for getting first insights into this prototype's potential for learning improvements to slide positioning, outside instrumental lessons. In this study I will mainly focus on trombone beginners, as they are the ones that could benefit the most and with whom a potential learning effect would be the most apparent.

For this prototype I am as such trying to answer the Questions

Q1. Does my XR prototype meet the requirements of usability for trombonists?

Q2. Can my XR prototype support the learning motivation for the trombone through an intuitive learning environment?

Q3. Can the XR prototype support trombonists in learning and memorizing trombone slide positions through an intuitive XR learning environment?

As such for this prototype I do not yet take affordability of the individual technology into account. In conclusion, I want to show if the prototype can create an actual benefit for the progression of musical abilities, while not with a teacher.

From this we can formulate clean hypotheses for this bachelor thesis, which will be explained afterwards, against which the prototype can be evaluated:

H1. The XR prototype achieves a mean System Usability Scale (SUS) score

(a) above the benchmark of 68.

(b) above the benchmark of 80.

H2. Participants with prior trombone experience rate the XR prototype positively regarding the potential to increase their motivation to learn and practice the trombone.

H3. Participants achieve a higher slide-position accuracy – reflected by a smaller distance from the target slide position – while within $10cm$ of the target position when using the XR prototype's slide-position assistance during training, than without it

(a) during training (with the assistance).

(b) after training (without the assistance).

H4. Participants achieve a higher percentage of playing the correct slide position (within $5cm$ radius), when using the XR prototype's slide-position assistance during training, than without it

(a) during training (with the assistance).

(b) after training (without the assistance).

For **H2** only trombonists (participants who already played the trombone before) will be accounted for. If someone did not play the trombone before, they have no experience in

practicing with the trombone without the program, so it would not make sense to regard them here.

For **H4** we want to analyze, if the participants head for the correct slide position, while not regarding precision, as we regard that element with **H3**. The value of $10cm$ in **H3** and $5cm$ in **H4** have some arbitrariness to them. $10cm$ for **H3** was chosen, as it is roughly the distance between two slide positions on the higher end. $5cm$ for **H4** is half that and as such encompasses slightly more that the half point between two slides. Both are as such on the larger end of possible values. Other values will be evaluated as well, but these have been chosen as a representative, to achieve a precise hypothesis definition.

For **H3** and **H4** the content of the parentheses at (a) and (b) is regarding the experimental group.

Chapter 2

Background

2.1 XR

Before going into specifics on how a trombone might be translated into XR, it is important to know the environment that is worked with.

The field of Extended Reality (XR) can roughly be separated into the fields of Augmented Reality (AR), Augmented Virtuality (AV) and Virtual Reality (VR), which gradually increase the amount of virtual elements displayed around you. Augmented Reality is when the surroundings are still visible to, but virtual elements get layered on top. On the opposite side of the spectrum lies a fully Virtual Reality / Environment where your entire surroundings are changed to something completely different and the only *real elements* left might be the location of your hands relative to your body. In the middle lies Augmented Virtuality, where certain "real-world objects or people are integrated into virtual environments" [1] and virtual elements can react to them.

The field of AV itself encompasses another spectrum of concepts as well, that could all be considered AV. One idea is the inversion of AR, where the virtual elements become the predominant environment and only the real elements are all that is visible from the real world, by projecting their image into the virtual environment. A different approach would be a fully virtual environment, in which virtual Elements are represented by a physical haptic element. [1], [2] On one side the exact object is the thing translated, and their visuals are fully preserved, while on the other side the real world object is just the representation of a digital one.

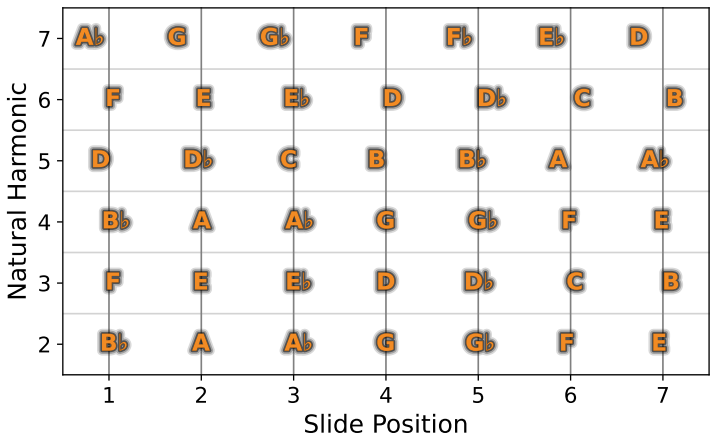
My proposed prototype is right in the middle of these concepts where the real trombone is integrated by position and state (slide) into the virtual world, but gets represented by a virtual stand in.

2.2 The Trombone

2.2.1 Overview and Functionality

The *trombone* is a brass wind instrument that produces sound through the vibration of the player's lips inside a *mouthpiece*. These vibrations generate pressure waves in the instrument's air column. The resulting pitch is determined by the effective length of the air column, set by the slide position, and by the resonant mode being excited.

Unlike most brass instruments that control the effective tube length with valves that redirect the airflow into additional tubing, the trombone expands the tube length with a continuous telescoping *slide* mechanism of roughly $70cm$ [3] for the tenor trombone.[1] This slide consists of two parallel tube sections that can move linearly relative to the main body of the instrument. By varying the position of the slide, the musician changes the total tube length, thereby altering the resonant frequency and produced pitch. Extending the slide increases the air column length and lowers the pitch, while contracting it does the opposite. This physical slide extension does not correspond directly to the instrument's effective acoustic tube length, which is additionally influenced by the bell geometry [4], which later gets discussed in more detail. (See Section 4.3.4)



Values are approximate and visually estimated. The figure is intended for illustrative purposes only. Based on data from *Trombone Lesson: Trombone Slide Chart* [5].

Figure 2.1: Trombone Positions for specific notes across partials.

The slide system is conventionally divided into seven discrete positions called *slide*

[1] There is also a variant that has valves in addition to the slide.

position one to *seven*. These positions variate lightly across different partials to achieve clean notes. [5], [6, p. 6] To get an idea of the extent of this variation, an approximate visualization can be seen in Figure 2.1. To note is, that the first position does not equal the most inward position on a tuned instrument, but with the slide very slightly extended. As the slide position is continuous, it enables the musician to smoothly transition between notes, which is known as a *glissando*.

The way the player's lips are positioned on the mouthpiece – including lip aperture and lower facial muscle tension – is called *embouchure*. Together with controlled air pressure, the embouchure generates lib vibrations that create resonances in the air column of the instrument. By adjusting the embouchure, the player can shift between different resonant modes of the harmonic series which known as the *partial* that the instrument is played at.

2.2.2 Variants

The trombone has multiple variants. The most relevant types that appear in orchestral music are the *alto trombone*, *tenor trombone* and *bass trombone* which all differ in tube length, mouthpiece size, bell diameter, and bore size to a different range of available notes. The tenor trombone and bass trombone are pitched in $B\flat$, while the alto trombone is pitched in $E\flat$, which is a fourth higher than the tenor. While the standard modern tenor trombone has no valves, the modern bass trombone has two valves, with the first lowering the pitch to a F and the second to a D [7] and a similar F-valve attachment variant exists for the tenor trombone. In my study I only refer to the tenor trombone without valves unless stated otherwise.

There exist further sub-variants differentiating in size of the mouthpiece, bell and/or bore of the instrument.

Chapter 3

Related Work

In this chapter I will go over the related and existing work in the different aspects relevant to this thesis.

3.1 Similar Work

A similar paper by Akiyama et al. proposed a "VR-based self-practicing system of a trombone for beginners" [8, p. 1], although their proposed system stayed in the realm of pure VR with only VR hand controllers representing the instrument, similar to *Trombone Champ: Unflattened* [9]. As such this controller was also not locked to be only moved on one axis like the slide, which made it necessary to introduce additional play when determining the position of the slide. [8, p. 2] This may be sufficient for their goal of learning the position numbers for notes and having an idea where that position is on the trombone. But as it, for example, does not reflect the weight of the trombone, and it lacks precision concerning the position evaluation, it is as such not capable of training muscle memory and can not assist as guidance during play.

3.2 Trombone Pedagogy

Learning an instrument is a process during which you are confronted with a variety of challenges. Struggling with these challenges can lead to disappointment. In the art of teaching an instrument it is important that the teacher and student do not obsess over perfection. For healthy encouraging learning it is advised to "think in percentages" [10, p. 5] and the improvements relative to the students previous attempts instead of right and wrong. It is always necessary to know where to improve instead of just knowing that the attempt was

incorrect. If this approach is disregarded it can easily cause loss of interest and turn away young learners.

It is very important to stay relaxed and reduce tension when performing and especially when practicing the instrument.

> "Remember: before a piece can sound good musically, it has to feel good physically."

as Caruso remarks in his paragraph about tension. [10, p. 52]

When it comes to trombone pedagogy, the distinction between *what should be played* and *how should it be played* may come into focus. While the *what* just addresses the correct pitches at the correct intervals with the correct volume, the *how* actually introduces the musicality which introduces texture, but is also often a case of interpretation. This thesis exclusively focuses on the *what*, while the *why* is beyond its scope.

3.2.1 Slide Positioning

It is not indicated on the slide where the correct positions for the individual notes are. Therefore, the musician must memorize them and adjust their hand position based muscle memory [11] and finetune on their acoustic perception. [12, p. 81] it is required that the movement of the slide "stop[s] at the required positions without hesitation" [6, p. 6]. Furthermore, while acoustic perception is a very important skill to learn, in some older orchestra seating arrangements the trombonists may not be able to hear themselves because of the positions of other instruments like the timpani. [13, p. 35] In this case precise muscle memory becomes even more important.

A typical phenomenon learners encounter is what Vanee describes as the "Creeping Slide Syndrome", where due to the heaviness of the slide and applied grip, the targeted slide positions drift from their correct positions resulting in off tune pitches. This of course gets amplified if the musician does not have a reference tone they can correct themselves with or is not able to hear what they play, due to for example the before mentioned suboptimal instrument placement.

As the slide positions are the foundation of playing the trombone, there are a variety of methods and tricks that can help in internalizing them. There are rough visual tricks to find the correct position like the 3rd position being with the wrist just before and the 4th just after the bell. [14] One possible method is the "Mirror Method" where learners

see themselves playing the trombone in a mirror and as such can compare their position to example visuals from a position chart. [11] Generally it can also help to slow down playing, and to use a tuner to see if the player is truly in the correct spot.

But knowing the correct positions is often not enough. Especially in passages with shorter notes, every note has a definite duration and "the slide must stay in the correct position for the duration of the note, however short it may be" [13, p. 35]. When you play every note at its full length – as required – the movement between different positions needs to be quicker, which can cause uncontrolled movement and disturbance of embouchure. To counteract this, "a quick acceleration from rest, and a controlled 'cushioned' braking to stop" [13, p. 35] is advised but must be learned. This braking to stop is also relevant when sliding back to first position. Not doing so may crash the slide into the instrument body which also causes embouchure problems and might cause pain on lips and teeth.

3.2.2 Additional Challenges

Besides finding the correct slide position the trombonist faces the following constant challenges. These are less relevant to the thesis but should be kept in mind.

Regarding learning embouchure and airflow, there are different elements that must be learned:

- Learning the basic technique required to produce a tone,
- Learning how to intentionally play and switch between different partials and
- Shaping that tone to sound softer, sharper, harder, warmer, etc.

Another important part of learning the trombone is correct posture. [15, p. 2] Although the player might be able to play the correct pitch with wrong posture, and general guidelines might not work the exact same way for everyone, correct posture while playing can improve airflow and reduce back strain.

3.3 Intuitive Visualization of Trombone Play Instructions

Another important aspect is the easy, intuitive representation of the music to be played. While notes work as a universal representation of the western tonal system, it does reflect

how instrument has to be played and in what position it has to be in. For example for the piano there is a common, established system as used in *Synthesia* [16] and *JazzHandsXR* [17], where vertical blocks flow from above on the to be played key which might also get highlighted.

The trombone is more complex in it having two different ways of manipulating the pitch: Embouchure and slide position. As embouchure is created through the interaction of lip aperture, facial muscle engagement, and controlled air pressure, it is difficult to create a new intuitive visual representation of it. For Slide Position one established system is the one used in *Trombone Champ* [18] and its VR translation *Trombone Champ: Unflattened* [9]. They use a 2-axis system where the vertical axis represents the slide position – with high being the slide at position 1 and low being position 7 – and the horizontal axis dynamically represents time from present time to n seconds into the future. The vertical axis also has a dynamic marker at the current slide position. Similar to the piano system, it then uses flowing blocks coming from the left at the height of the slide position, to intuitively show the position. Even though the original game (*Trombone Champ* not *unflattened*) is not meant to be played with a trombone, its VR version clearly shows its potential to be understood when using a trombone slide like mechanic.
Another system is the one by Akiyama et al. that still uses conventional sheet music notes, but annotates the slide position as a number – similar to how it can be done on a paper during practice – and then adds an image of a trombone from the side perspective, with the slide positions highlighted. The slide of the image trombone also displays the same slide position of the virtual trombone controlled by the position of the hand controllers. When a note is supposed to be played both the sheet note and the specific slide position location is highlighted on the image in front of the person and the virtual trombone in their hands.

3.4 Tracking

To discuss the process of bringing a real Object into VR and digitalizing their inputs we need to be aware of different approaches to position tracking technology. What kind of tracking system is applicable for a certain use case depends on which elements of the object need to be translated and how they are related to each other. In the case of the tenor trombone the telescoping slide mechanism is such an element. The slide is restricted to a

single degree of freedom (1DoF) where it can move between a minimum and maximum position relative to the instrument's main body. This extends the technology that may be used. In this section I will go over some of the potentially viable technology and ideas for the use case of translating a trombone into VR.

3.4.1 Distance Tracking

For tracking an object across a single degree of freedom with a minimum position, a possible approach is measuring the distance between that origin and the position of the slide.

Ultra Sonic Sensors. The *uSLIDE* by Farwell (See [19]) tried such an approach using a "time-of-flight ultrasound method using a $40KHz$ transmit-receive ceramic transducer pair". [19, p. 359] As the trombone can produce energy around $40kHz$ it may interfere with the devices. Because of that the receiver was mounted on the inner slide in a more stable environment, with the transmitter on the moving part of the slide. Both devices were connected with a cable to a computer, which is "very workable in practice" [19, p. 359] according to the author. If in a lower slide position the cable was hanging between the components attached to the slide and the base. This needs to be accommodated during play and may interfere with playing in some situations. The system uses *continuous-wave transmission* and achieves an accuracy of approximately $3.5mm$ with a refresh rate of $200Hz$.

Laser Sensors. Another tracking technique was chosen by Lemouton et al. (See [20]) to track the seven discrete positions of the trombone slide where a red laser was mounted onto the slide while two photo-electric diodes were attached to the main body at origin. The distance was then measured by "the surface and intensity of the light spot received by the diodes" [20, p. 304]. Due to this it requires careful alignment of the components and further calibration on fixed distances. They further discovered multiple problems: The tracking gets less precise with increasing distance between the laser and sensors. Furthermore, the sensors are subject to "fluctuation" [20, p. 306], although this is not detailed further. There are also no statements on the precision of the tracking in any unit.

Campbell et al. used a similar setup in their experiment using the industrial laser distance sensor *OADM 20I4471 laser sensor* by *Baumer* [21] which promises $< 0.5mm$ precision [22]. Modern versions of the *OADM 20* family bringing precision down to $\pm 3\mu m$

linearity deviation and $0.12\mu m$ repeat accuracy. [23] Important to mention is, that they only used it on a static trombone, where the body was locked into place.

As such the validity of this approach depends on the attachment to the trombone and whether it can withstand the stress of the trombone's movement.

Other Tracking Systems. There are even more interesting approaches which I will not go over in detail in this thesis, but want to mention.

There is an approach by Meister et al. that uses Electromagnetic radio waves with a FMCW Radar mounted to the bell of the trombone to estimate the position of the slide. [24] A mechanical tracking approach is already patented with "Device for indicating trombone slide positions and related systems and methods" that use a mechanical rod next to the slide for precise position tracking. [25]

3.4.2 6DoF Tracking

While 1DoF Tracking is only capable of tracking one specific aspect of an object six degrees of freedom tracking (6DoF) gives you access to the full position of a rigid body in three-dimensional space. This includes the position on three axes (forward/backward, up/down, left/right) and the rotation along around three axes (yaw, roll, pitch). A technology that is able to give us access to these values is as such necessary to represent any physical freely moving object in VR that is bound to its physical counterpart.

It may also be used to calculate the extension of the trombones slide using two tracked locations – one on the trombone body and one on the slide – by translating the slide trackers position into the local space of the body tracker and then calculating the offset to a calibrated slide position start.

Lighthouse Tracking. One of these technologies is lighthouse tracking developed by *Valve*. It is a type of tracking that relies on external so-called *lighthouse base stations* which serve as reference points for the trackers introduced into the system. [26] This enables room-scale tracking for a room set up with this system.

The base stations contain infrared (IR) lasers mounted on one or two[1] rotating drums creating sweeping planes of IR light. These IR light are then detected by the IR receivers of a tracked device. [26] Using precise timing between sweeps and "simple trigonometry" [28]

[1]Depending on the model. The 2.0 Base Station Model by Valve/HTC relies on a single drum creating two planes on different angles. [27, p. 2]

it is possible to calculate the orientation and position of the tracker with sub-millimeter precision with a refresh rate of $1000Hz$. [26], [28]

SteamVR provides an established software for the XR Environment regarding this technology [28], and generic Trackers for this technology have been developed by *HTC* and *Tundra Labs*. As such it is relatively easy to use this technology for prototyping where a 6DoF technology is needed, due to the SDK support and existence of generic trackers.

As the technology uses IR lasers it relies on the trackers having line-of-sight with the base stations which means the tracking stops working if there is an occluding object in-between the tracker and the station. For a limited amount of possible interfering objects in the room this can be mostly circumvented by placing the base station in locations that better oversee the room and possibly introducing more stations mounted at different locations.

Another potential issue with this technology are (especially flat) reflective surfaces, as one can not distinguish between direct and reflected hits. Although this sounds like a big problem concerning a reflective instrument, there is a certain amount of noise the implementation of *SteamVR* in combination with the multiple sensors on one tracker is able to filter out. Still, this would need to be tested if this technology is chosen.

Optical Tracking. Optical Tracking is a technology that uses cameras to reconstruct the position and rotation (pose) of an object. Concerning this method it has to be differentiated between inside-out and outside-in optical tracking and marker-based / markerless approaches.

Generally, **marker-based** approaches use easily identifiable markers, from which a pose can be reconstructed, while **markerless** approaches try to identify the objects' pose from the object itself using computer vision techniques such as SLAM or Visual Inertial Odometry.

In **inside-out optical tracking**, the cameras are placed on the object itself, using the image to reconstruct their own pose from visual features, markers, or other recognizable elements of the environment.

In **outside-in optical tracking** one or multiple cameras are placed in known, often fixed locations of the environment, from which the to be tracked objects or markers pose is reconstructed relative to the camera.

Concerning the use case of trombone tracking in VR the most relevant relative position is the head. If the pose relative to the vision of the player is wrong, it will be the most obvious to the player. As such using the cameras on the head mounted display (HMD) if available, with the HMD becoming the known pose, could be an option that should be considered. This would be similar to how it is already done in many commercial headsets using outside-in optical hand tracking, like the *Meta Quest Pro*. [29]

As this tracking is based on the image-feed of cameras this technology relies on sufficient visible features in the environment and can have problems under poor lighting conditions, extensive occlusion, or feature-poor environments.

Magnetic Tracking. Another technology that gets used more frequently in the realm of XR is magnetic tracking. [30]

Magnetic tracking systems typically consist of "multiple sensors or multiple magnetic sources because it is not observable to estimate the 6-DoF pose by using a single magnetic sensor to measure the magnetic field from a single magnetic dipole source" [30, p. 5063], but there are many variations on the setup. Depending on the number and configuration of sources and magnetic sensors, some systems natively reach 6DoF, while others need additional measurements like from an *inertial measurement unit* (IMU) to reach 6DoF. [30], [31]

One advantage of this technology is, that it does not rely on line-of-sight and therefore unaffected by visual occlusion. This is particularly relevant for the use case of tracking the trombone, where both the player and even the instrument itself, may obstruct the trackers depending on the tracking setup.

A potential problem is the influence of the trombone material on the magnetic field. Most standard trombones are made of yellow or gold brass and/or nickel silver with mouthpieces that may be made of silver- or gold-plated brass. [3], [32], [33] These materials are not ferromagnetic and as such do not produce the strong magnetic response commonly associated with ferromagnetic materials. However, materials are electrically conductive. Time-varying magnetic fields may induce eddy currents that can locally distort the measured magnetic field. [34] Although this must not be the case, some trombone stands are constructed from a ferromagnetic material. [35] The impact of these effects on tracking accuracy would need to be evaluated.

Another limitation of magnetic tracking may be tracking volume. The system proposed by Dong et al. was tested in a workspace of approximately $150 \times 150 \times 80mm$ and was designed for operating distances below $0.5m$. [30] In contrast, the commercial consumer-grade magnetic-inertial system *FluxPose* specifies a substantially larger tracking diameter of $3.4m$ around its wearable beacon with $\pm 1.75mm / \pm 0.5°$ static accuracy verified at $1m$ radius. [36] To be sufficient for this use-case it would at minimum need to be able to track the trombone body during play (near the head) if the slide distance gets tracked by another 1DoF technology (See Section 3.4.1). As such whether this technology is sufficient largely depends on the exact setup used.

3.4.3 Embouchure and Air Pressure

While having the trombone slide position is one thing, it is not the only element determining the resulting frequency. The other important element is embouchure and air pressure. (See definition in 2.2.1.)

An interesting approach is the one by Amann et al. who developed a "digital mouthpiece" to measures air pressure within the mouthpiece of the trombone "via an air channel, directly integrated [...] into the cup" [37, p. 400]. The mouthpiece was already suggested as a learning aid in real-time within the paper. [37, p. 401]

Still missing would be data concerning the lip aperture and other aspects of embouchure. I did not find any study discussion a method to track these during play.

Chapter 4

Methodology

In this chapter I will explain my proposed prototype, its development and the design decisions I made.

A central aspect of this prototype is the intuitive visualization of trombone play instructions for the slide while practicing with a real trombone in VR. But unlike similar programs for piano I do not want to abandon the traditional staff system. The prototype is intended to help the musician improve on positioning. The association with the sheet music must remain, as it is the only thing the player can rely on when outside the system.

Furthermore, it would be nice to not have to rely on controllers. The trombone is held with two hands and it would be cumbersome to continually have to switch between the systems.

4.1 Technology and Software

For the prototype *Unity* (6000.0.34f1) was chosen by me as my game-engine as it is the engine I am most experienced in, and I already had worked with multiple plugin VR-Plugins available for it. Unitys Component System was also a reason for this decision as it allows for easy (re-)connection of different systems. Furthermore, C♯ is a language I am rather comfortable with.

I chose *OpenXR* (1.15.1) as my base VR framework over the other options like the *Meta XR SDKs* and the *OpenVR SDK* (SteamVR). As I wanted to be able to work with my own VR Headset which is a PCVR only headset the Meta XR could not be used. OpenVR was a close contender, but I ultimately chose OpenXR for its easy cross-device development, which would later make it easier to modify the prototype to work with standalone headsets.

As a bonus it already has a few prebuild assets for easy UI interactions and such.

As OpenXR is developed for cross-platform use it also has a SteamVR integration, which made testing the game as easy as pressing the play-button in the editor, while the HMD is connected to the computer.

For tracking, I used three *Vive 3.0* Trackers: Two for the trombone. One for its stand.

Further libraries and plugins that have been used are *NWaves* as a DLL and the free version of the unity plugin *Maestro - Midi Player Tool Kit* (2.16.1). NWaves was used in audio frequency detection, while NWaves was used for Midi integration.

The non primitive meshes used in the prototype other than the controller stand-ins provided by OpenXR were modeled by me in *Blender*. These meshes consist of the trombone itself as it parts, so it can be animated to match the trombone, its stand and the Vive 3.0 Tracker stand-ins.

In addition to that I used the *bravura font*. The font follows the *Standard Music Font Layout* (SMuFL) initiated by *Steinberg* and continued by the *World Wide Web Consortium* (W3C). Fonts that use this layout standard provide thousands of musical symbols as characters for use in the Staff System.

4.2 Functionality

From the user perspective there are three parts of the prototype: Calibration, menu interaction and playing.

The player finds themselves on top of a seemingly infinite plane with a water-like texture, trying to invoke a surreal but calm environment. The environment can be switched to a simple room in the settings.

To calibrate the instrument the physical and digital trombone need to get aligned which still requires a hand controller to adjust the values. First the trombone should be placed on its stand. Then the *stand-tracker* is assigned using the menu and can be laid beneath the stand, which moves the virtual stand and trombone to that location. After that the trombone only needs to be aligned in its height and rotation around its bell. The height can easily be done by placing a hand controller on the physical trombone and adjusting the values to match that. For the rotation it is easiest to rotate the real trombone to how the virtual one is calibrated. Once that is done “Quick Calibrate” can be pressed to finish the calibration.

If the slide has a different lengthwise angle for the slide relative to the base, it can be

automatically adjusted by extending the slide to a higher slide position and pressing the button “Slide Direction”. The adjustable angle on which the slide is attached to the body and with which one can turn the trombone in a left-handed one can be adjusted with a slider. When for any reason[1] the digital slide is not fully in while the real one is, I added a button “Slide Distance Reset” to reset that as well. After the calibration is done, the trombone may also be used as a controller to navigate the UI.

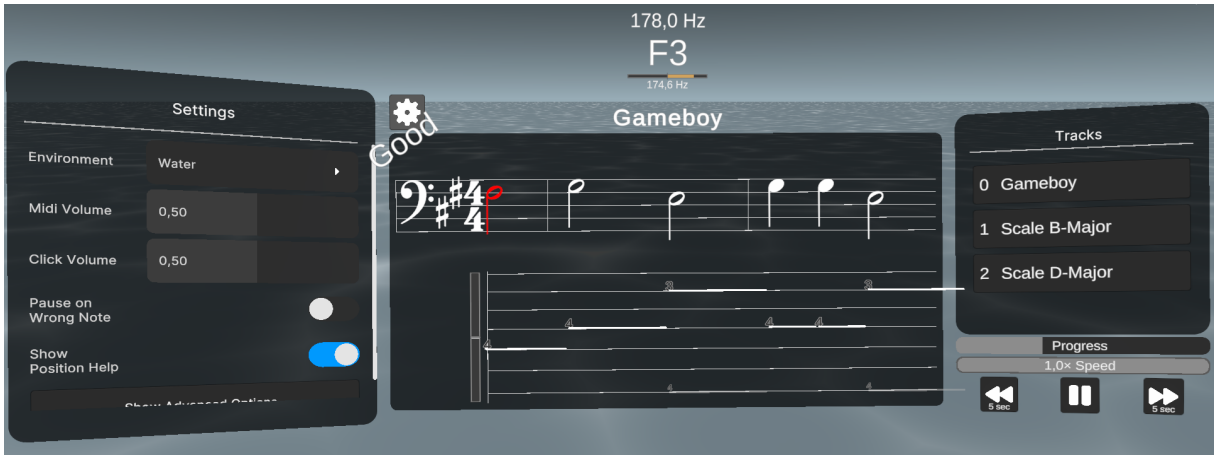


Figure 4.1: UI and Music Track Representations

In the settings menu seen in Figure 4.1 there are options like metronome click volume but also bigger behavior changes. One would be the “Pause on Wrong Note” toggle, where the music piece waits for the player when he plays a wrong note for longer then 0.1 seconds until they correct themselves. Furthermore, the extra slide position help visuals can be toggled here.

On top of the middle panel a tuner is placed to simplify the tuning process of the instrument.

A music piece can be chosen in the list on the right-hand side which the player wants to practice.

The piece will start to play with one into measure. Clicks are added on every measure and beat emulating a metronome. On the central panel, there are two rows, presenting the to be played notes of the music.

The upper is the classical musical notation of a staff system. Notes on this system are not static though and move from the right to left, where once they arrive at the left-hand side, they are supposed to be played.

[1]Why this can happen gets explained in Section 4.3.2

The lower is a visual representation of the slide positions. It consists out of seven bars representing the seven slide-positions of the trombone, with a marker, where the players slide currently is. From the left side, markers flow to the left on the height within the seven slides, where the current note should be played. The markers are often not perfectly on one slide bar, due to the natural small variations of positions across partials. They follow at the same x position as the notes of the staff system to ensure they are viewed as representations of the same thing. On the left-hand side of the lower representation there is one further slider representing how far off the player is from the current note on a ± 1 slide scale. Depending on the percental distance to the goal position within a one slide radius, a text appears in the upper left corner with the feedback text "Wrong Position" ($> 40\%$), "Inaccurate" ($\leq 40\%$), "Okay" ($\leq 30\%$), "Good" ($\leq 12\%$), or "Perfect" ($\leq 2\%$). Words like "bad" have been avoided to not discourage the learner.

There is one further representation directly on the trombone indicating to where the slide should be moved. (See Figure 4.2)

Under the music piece selection list there are also simple controls for the current track enabling to pause or jump back and forth 5 seconds of the track to practice certain bits. There is also a time bar representing how far into the piece you currently are, which also can be interacted with.

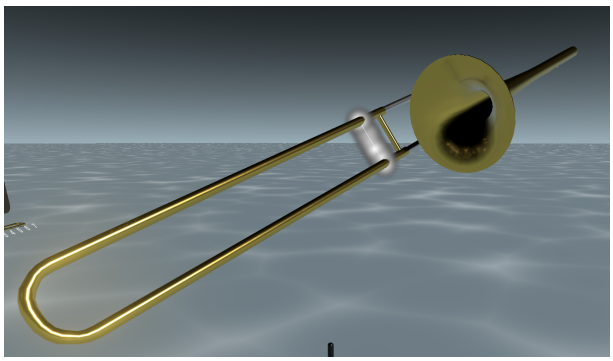


Figure 4.2: Visual Representation with white Particles of the Goal Position on the Trombone Slide

4.3 Development

4.3.1 Code Structure

To start off the section about the development and design decisions made, I want to provide an overview of the code structure, before going into detail on specific elements.

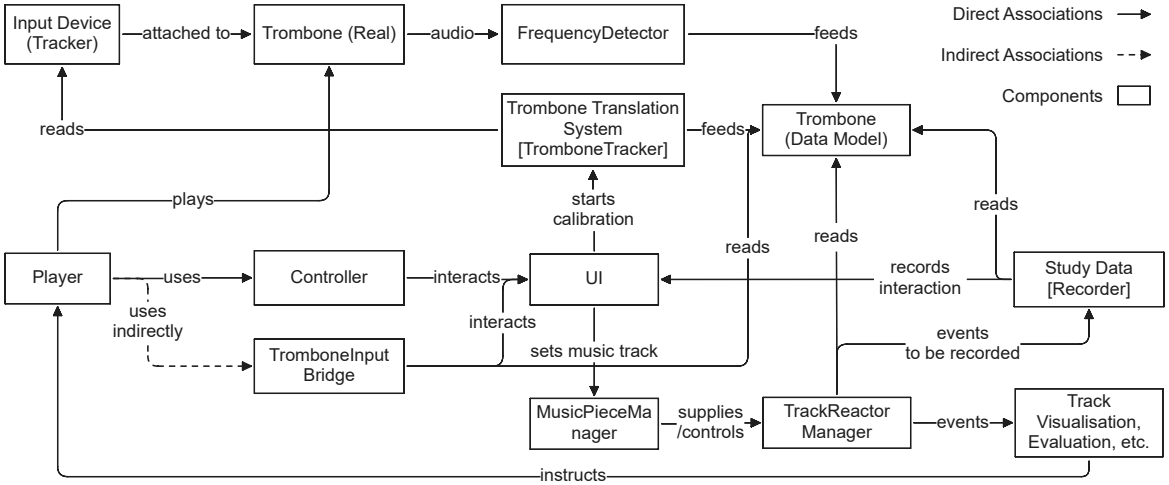


Figure 4.3: Simplified Code Structure Diagram of the System

The implementation is divided into several subsystems that each have a clearly defined responsibility:

- Tracking
- Audio processing
- Trombone state model
- Music track management
- Music Visualization
- Performance evaluation
- User input system

These subsystems communicate primarily through the central `TromboneModel`, which represents the current physical state of the instrument. `TromboneModel` serves as a central source of truth which the other elements should rely on to reduce the need to juggle different classes and methods. Tracking and audio components update this model, while visualization, evaluation, and input systems read its state. This results in a predominantly unidirectional data flow, reducing coupling between the individual subsystems.

Figure 4.3 illustrates the simplified overall architecture of the prototype. Tracking devices and microphone input provide measurements that are processed independently

before updating the shared `TromboneModel`. The music playback subsystem parses MIDI data and `TrackReactorManager` generates and schedules notes from that. Different Visualizers and the Evaluator components then react to those scheduled notes while sometimes also reading the current instrument state of `TromboneModel`.

Static utility classes encapsulate methods for physical/acoustic calculations and musical conventions and are used by multiple subsystems without maintaining their own state. These were chosen to be separate static classes, as they represent the laws the prototype is based upon. `TromboneModel` gives access to methods that serve as bridges to these functions for its data.

4.3.2 Trombone Data

Position Tracking

First I tested the technology I already had, which are the *Vive 3.0* lighthouse trackers, to find out if I would need to build something myself.

To integrate those in OpenXR, a custom Device Profile was needed so that the trackers are registered and become available for use. Such a Device Profile was provided by thep3000 [38] which was further modified by me.
The reflective surface of the trombone could introduce major problems. But as the trombone does not reflect in from a consistent plane and reflected lasers often get scattered – often away from the trackers – rather than focussed, it worked relatively well. After further eliminating reflections with masking tape from my testing it worked as accurately as I am used to.

Most other options (See Section3.4) require custom hardware and all introduce their own error sources, which would need to be tested. In contrast, I am quite familiar with the error sources of this technology. So I decided, that the Vive Trackers are sufficient for the prototype.

When abstracting a simple tenor trombone we have two main elements: The main body and the telescoping slide mechanism. There are further movable parts, but those are usually only adjusted during preparation, such as the angle between the slide and rest of the body. While we need the full position of the body, the slide is locked on one dimension relative to the body.

As such we require two trackers. One tracker was placed on top of the trombone on the tube behind the bell at of the vertical braces. The other was placed on the slide before the

bell when the slide is in its base position. Its connection is bridged across the tube, with the trackers position slightly above one tube to preserve stability. The exact configuration used can be seen in Figure 4.4. This way the small but existent weight of the trackers is distributed roughly equally across the trombone when held, making them less noticeable.

Another tracker was used to mark the position of the trombone stand to make it visible to the player and make calibration easier. Its placed right in the middle under the stand.

In the Section 4.2 I talked about the functionality during calibration of resetting the slide position when it is off. This can happen when there was a bit of obfuscation during the calibration (by for example the bell) or when the trombone is moved/rotated a large amount while there are reflective surfaces in the area.

The trombone calibration system (Described in Section 4.2) is neither user-friendly nor intended to be, as even with instructions, quickly and accurately calibrating the trombone takes a slight bit of learning. While a tutorial for this can absolutely be implemented, I decided for the sake of the user-study to do the calibration myself beforehand to ensure good and as roughly equal calibration. This was deemed more important, as otherwise the evaluation of the main aspect might get muddied by bad calibrations. Furthermore, which tracking system is used and how the calibration process works is not fundamental to the program and could be swapped. At the end of the day I felt like leaving the process was better than creating a half-backed possibly unused tutorial just to check the box.

After the calibration process is finished, the trackers closet to the slide and body get assigned and saved with their offset position and rotation becoming the body-tracker and the slide-tracker. The body-tracker alone is responsible for the position of the instrument, while the rotation is the average of both trackers, as the rotation does not change on the slide, relative to the body. The slide position is calculated as the distance to the calibration position of that tracker along the direction the slide is pointing. The tracking gets slightly smoothed using a circle buffer to lessen the effect of slight tracking errors.



(a) Physical Trombone (b) Digital Trombone

Figure 4.4: Tracker locations on the physical and digital trombone
One tracker on the back of the main body, one on the slide, and one under the trombone stand.

Trombone frequency

A selectable microphone records the audio. Using the YIN algorithm implementation from *NWaves* the fundamental frequency gets estimated and fed into the `TromboneModel` as its pitch.

Embouchure

As I did not find any way to track embouchure, while playing taking it was not feasible for this study. Although we as such can not get every variable of embouchure it is still possible to estimate the partial addressed with embouchure from slide position and audio pitch alone as will be detailed in Section 4.3.4.

4.3.3 UI Input System

At this point I want the player to not need to rely on a controller. Having to switch between the two-handed trombone and a controller would hinder the usability of the prototype.
As requirements for the input system I set myself the goal for it to be intuitive, not require big movements and to not introduce new devices.

The solution I came up with used the trombone as a controller that can interact with the user interface by extending and contracting the slide with a pointer cast from the middle of the slide as can be seen in Figure 4.5. This follows the interaction one would do with the hand controller or a computer mouse, with extending the slide slightly behind the bell as a click/pushing the trigger and moving it back in front of the bell as letting go of the key/trigger. To reduce the filter out accidental clicks there is a small dead area in between the thresholds.

It has to be noted that using the element of the trombone that is often moved quickly during play in this manner obviously introduces a big risk of accidental clicks during play in the area the trombone is pointed. But this can be solved by designing the UI layout with this behavior in mind. I, as such, placed the buttons just on the right and left side of the area where the play instructions appear. the mid was left with only visual elements that can not be interacted with.

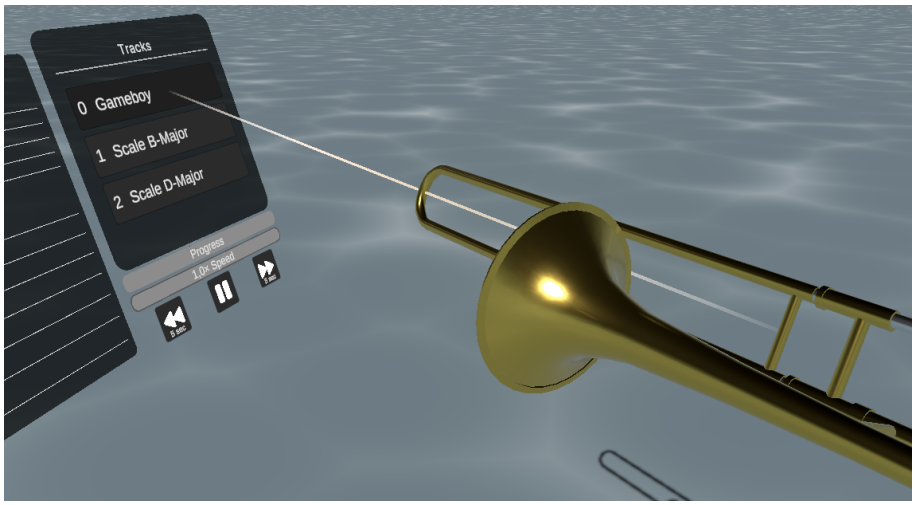


Figure 4.5: Trombone as UI Input Device

As it is similar to existing systems it should increase the intuitiveness. It does not require large motions, just the rotation of the instrument (and probably head) and a small back and forth motion with the arm. As the solution relies on the existing tracking of the trombone it does not introduce new devices. So my personal requirements are fulfilled.

4.3.4 Acoustic Model

When it comes to finding the correct locations for the individual slides on the trombone we have two different options: Hardcode the positions and their variations across different partials or apply a theoretical acoustic model with which the positions can be calculated. As we work with an actual trombone hard-coding the positions is not practical, as slight changes in tuning and embouchure would lead to the positions being wrong. As such I chose the theoretical model option, as – when successful – it would be more reliable to find the actual mathematically correct positions and could be (dynamically) tweaked deepening on the variant of trombone (See Section 2.2.2) and how it is tuned.

For the model we need a function M_f to get from slide position in meters L_s[2] and

[2] $L_s = 0$ being the most inner position, and then how many meters the slide has been moved away from that.

partial n to frequency f

$$\begin{aligned} M_f: \ \mathbb{R} \times \{1, 2, \ldots, 10\} &\rightarrow \mathbb{R} \\ (L_s, n) &\mapsto f \end{aligned}$$

. The other way around – getting possible play positions from the frequency – is slightly more complex, as there may exist multiple partial and slide position combinations, that can result in the same frequency. As the amount of partials is small its easiest to calculate the L result for every partial n in combination with the desired frequency f. As a slide has only a limited distance L_{maxS} until either the slide falls off or the players arm can not reach it anymore, the final step is to exclude every result that is outside the possible slide position range $[0, L_{maxS}]$ resulting in a set that includes every possible combination:

$$\begin{aligned} M_{L_s}: \ \mathbb{R} \times \{1, 2, \ldots, 10\} &\rightarrow \mathbb{R} \\ (f, n) &\mapsto L_s \end{aligned}$$
$$\left\{ (n\,, M_{L_s}(f, n)) \;\middle|\; f \in \mathbb{R},\, n \in [1, 10]\,,\, M_{L_s}(f, n) \in [0, L_{maxS}] \right\}$$

For later evaluation of the correctness of the trombone it is best to also identify the partial the instrument currently plays. Our current model already solves this problem. Using the current recorded pitch of the instrument together with the current slide position, we can evaluate the set of possible partials–position combinations and compare which position the current slide position is the closest to. The partial of the position, is the one the player currently addresses using their embouchure.

Used Model

My first approach to the model was the simplified model described by Seifert [39] who mentioned the model could be approximated using the open-open pipe function

$$f = \frac{n \cdot c}{2 \cdot L}$$

with frequency f, partial n, tube length L and speed of sound c. In our context the full tube length L would be the trombone base length added to two times the slide extension $L_B + 2L_s$ This leads us to our first leading to our first model function M_f

$$M_f(L_s) = \frac{n \cdot c}{2 \cdot (L_B + 2L_s)}$$

This formula is also trivially reversible leading to

$$M_{L_s}(f) = \left(\frac{n \cdot c}{2 \cdot f} - L_B\right)/2$$

. The slight problem regarding this method is, that slide position does not scale linearly with frequency nor semitones. I tested this model against the *DigitalTrombone* Slide chart [5] and found out that with an effective length of $2.92m$ the model looked quite accurate across partials 2-11 as seen in the chart. As such I used it for my early versions to build the base functionality to maybe later switch it out with a more complex, more accurate version.

Such a more complex version was found with "A theoretical trombone" in which LoPresto directly tries to model the augmentation by the bell, which functions like a Bessel-horn, and the mouthpiece correction. [4] The proposed algorithm uses iteration to get a more precise result. The mouthpiece correction depends on frequency from the Bessel-horn, after which the Bessel horn needs to be recalculated with its length adjusted by the mouthpiece correction, which results in a slightly more accurate calculation. These steps need to be repeated to get an accurate result. This leads to the end equation

$$f = \frac{c}{4(L_B + L_{corr} + x_0)}\left((2n-1) + 0.637\sqrt{\gamma(\gamma+1)}\right)$$

with frequency f, trombone pipe length L_B, mouthpiece correction L_{corr}, end correction x_0, harmonic n, speed of sound c and flare-parameter γ. As the algorithm is only described in the paper, not explicitly written down as pseudocode, and I needed a few tries to understand it correctly. I appended my current understanding of the algorithm as pseudocode in the Appendix B as well as the reversed version (from frequency to length) requiring no iteration. I tested this model with the rough true trombone base length of $2.7m$, a mouthpiece volume of $(4 \cdot 4 \cdot 3)cm^3$ mouthpiece cup volume, $5cm$ mouth-pipe length and $1cm$ mouth-pipe radius (as measured from the trombone) and slight variations of these. Over the time of the project I had problems getting the algorithm to work and had problems I did not have with the simple function. I was also not able to recreate their mouthpiece correction results. The study lacks the exact values used for the mouthpiece. Due to this reason I stayed with the simple model.

Another element holding this algorithm back is the reliance on multiple values that would need to be inputted manually by the user. The algorithm is also inaccurate for certain harmonics. [4]

Comparison with the real Trombone

The prototype has been every so often tested by different experienced trombone players who did not notice anything wrong with the evaluated shown positions. As slide position 7 and higher slide positions in general are rarely used abnormalities were not recognized.

As I needed to focus on other elements of the prototype it was forgotten for some time and I began to think the (open-open pipe) model was tested enough. Only late in development I remembered to get actual values shown in Table 4.1. These are average values from repeated testing, but individual tests, from which we still can get a rough idea of the correctness of the model. As stated earlier the values will fluctuate slightly due to tuning and embouchure. "Trombone 1" is a symphonic tenor trombone with F-valve attachment which was the main trombone during development, while "Trombone 2" is the one used in the study, as it is a student model more appropriate for beginners. Trombone 2 Try 1 and Try 2 have been recorded on different days, while Try 2 and Trombone 1 have been recorded on the same day.

As can be seen different types of trombone behave differently and have slightly different slide positions. Due to embouchure the positions may also variate by distances above $1.5cm$ on the same trombone, as can be seen between Try 1 and Try 2 at $C3$. Concerning the model, a pattern can be seen, where it gets less accurate the higher the slide position. While an offset of 1-2cm would still be within an acceptable range, offsets from 7-8cm encompass roughly half a slide position would as such make a quarter-tone-step of a difference is inappropriate for a training program. Even though this is quite big of a difference a trained trombonist can to an extent compensate such a difference with embouchure alone. This may also be a reason why this was not noticed during the earlier testing.

This increases the amount off values that would need to be individually obtained by quite a lot, including bell and bore size, additionally to the for us inaccessible embouchure. To accommodate this an option would be to add a calibration process before or passively during the trombone playing.

Automatic Tuning Adaptation

In this section I will explain how the model can be dynamically adjusted or calibrated. The base idea of the automatic adaptation is to collect the data that is actually being

produced by the instrument. This recorded data being the played partial, slide position and audio frequency as described in Section 4.3.2. Then we approximate variables of the model depending on the partial n to the collected points (frequency f, slide position L) to get the most accurate predicted results, so that we end up with one set of variables for every partial. When evaluating pipe length or frequency we then use those variables depending on the partial.

Note	Partial	Slide	Offset from model in *cm*		
			Trombone 1	Trombone 2 (Try 1)	Trombone 2 (Try 2)
$B4$	6	7	-3,94	–	–
$C4$	6	6	-2,04	–	-1,62
$D\flat4$	6	5	-2,02	–	-0,75
$D4$	6	4	-1,72	–	-1,85
$E\flat4$	6	3	-0,29	-1,56	-0,56
$E4$	6	2	-0,18	0,11	0,29
$F4$	6	1	0,92	-0,11	0,09
$A\flat3$	5	7	-3,70	–	-3,54
$A3$	5	6	-2,79	–	-4,33
$B\flat3$	5	5	-0,68	–	-2,84
$B3$	5	4	-0,35	-1,31	-1,61
$C4$	5	3	0,65	0,82	0,10
$D\flat4$	5	2	1,33	1,32	0,94
$D4$	5	1	0,40	0,66	0,88
$E3$	4	7	-5,07	–	-8,48
$F3$	4	6	-4,24	–	-3,47
$G\flat3$	4	5	-3,07	-2,05	-2,70
$G3$	4	4	-2,18	-3,28	-2,14
$A\flat3$	4	3	0,54	-2,42	-1,28
$A3$	4	2	0,96	-0,79	0,32
$B\flat3$	4	1	-0,56	-0,92	-1,24
$B2$	3	7	-4,70	-8,80	-8,46
$C3$	3	6	-5,61	-6,76	-4,59
$D\flat3$	3	5	-1,82	-5,76	-4,07
$D3$	3	4	-1,01	-3,70	-2,58
$E\flat3$	3	3	-0,16	-1,92	-0,62
$E3$	3	2	0,35	-0,62	-0,72
$F3$	3	1	-0,96	-0,71	-0,93
$E2$	2	7	-2,57	-6,89	-7,49
$F2$	2	6	-0,12	-3,21	-4,39
$G\flat2$	2	5	-3,73	-3,62	-3,19
$G2$	2	4	-1,54	-2,31	-1,74
$A\flat2$	2	3	-0,51	-0,42	0,12
$A2$	2	2	-0,01	0,76	0,21
$B\flat2$	2	1	-0,83	-0,30	-0,23

(with extreme cases: orange $\geq 3cm$, red $\geq 5cm$ absolute offset)

Table 4.1: Trombone $\leftrightarrow$ Model Comparison

Function Approximation

Although I was not able to replicate the results of the of "A theoretical trombone," as we are now approximating variables anyway we may as well base it on the Bessel-horn to approximate its missing[3] values. Some values are unchanging, like the flare-parameter γ, allowing us to simplify the equation to

$$\begin{aligned} f_n &= \frac{c\left((2n-1)+0.637\sqrt{\gamma(\gamma+1)}\right)}{4(L+x_0)} \\ &= \frac{c\left(2(n-\frac{1}{2})+0.637\sqrt{\gamma(\gamma+1)}\right)}{4(L+x_0)} \\ &= \frac{2c\left(n-\frac{1}{2}+\frac{0.637\sqrt{\gamma(\gamma+1)}}{2}\right)}{4(L+x_0)} \\ &= \frac{c\left(n-\frac{1+0.637\sqrt{\gamma(\gamma+1)}}{2}\right)}{2(L+x_0)} \\ &= \frac{c\,(n+a)}{2(L+x_0)} \\ &= \frac{c(n+a)}{2(L+b)} \end{aligned} \quad \left| \begin{array}{l} a := -\frac{1+0.637\sqrt{\gamma(\gamma+1)}}{2} \\ \\ b := x_0 \end{array} \right.$$

, with the initial values $a = 0$ and $b = 0$. The reversed equation $f_n \to L$ becomes

$$\begin{aligned} \frac{c(n+a)}{2(L+b)} &= f_n \\ \frac{2(L+b)}{c(n+a)} &= \frac{1}{f_n} \\ 2(L+b) &= \frac{c(n+a)}{f_n} \\ L+b &= \frac{c(n+a)}{2f_n} \\ L &= \frac{c(n+a)}{2f_n} - b \end{aligned}$$

To be able to easily approximate the variables we want to linearize it, so that we can preform a linear regression. We can do this by creating the reciprocal of the points

[3] values which would need to be measured and inputted manually

becoming $(f, \frac{1}{L})$ and the function:

$$\begin{aligned}\frac{1}{f} &= \frac{2(L+b)}{c(n+a)} \\ &= \frac{2L+2b}{c(n+a)} \\ &= \frac{2}{c(n+a)}L + \frac{2b}{c(n+a)} \\ &= \underline{\underline{mL+q}}\end{aligned} \quad \Bigg| \quad m := \frac{2}{c(n+a)},\ q := \frac{2b}{c(n+a)}$$

We can extract the values a, b by evaluating $\frac{2}{mc} - n$ and $b = \frac{q}{m}$:

$$\begin{aligned}\tfrac{2}{c(n+a)} &= m \\ \tfrac{c(n+a)}{2} &= \tfrac{1}{m} \\ \tfrac{n+a}{2} &= \tfrac{1}{mc} \\ n+a &= \tfrac{2}{mc} \\ a &= \underline{\underline{\frac{2}{mc} - n}}\end{aligned}$$

$$\begin{aligned}\tfrac{2b}{c(n+a)} &= q \\ mb &= q \\ b &= \underline{\underline{\frac{q}{m}}}\end{aligned}$$

Point collection

To collect the data points I suggest the following approach: A point should be added to the set once the frequency or slide position have passed a significant threshold, such as a quarter-note change in pitch and $1cm$ of movement. The points should be saved in circle buffers, with newer values overwriting old ones. This has the benefit that we limit the amount of memory used, as well as automatically get rid of old values that might not be accurate anymore due to changes in embouchure. The circle buffers should not only be separated by partial but also by discretized position, as we otherwise might lose data for the higher slide positions as they are used less often. The distance between these discretized positions should be less than a slide change, so that points for different slide positions can not be in the same buffer. Such a distance could be about $5cm$. Within this Buffer the points should still be saved with their frequency and precise position $(f, \frac{1}{L})$

Data Pollution

A major problem during development is, that the microphone is not exclusively picking up sound from the trombone, but also the environment. As the YIN algorithm estimates the fundamental and not the average frequency and the trombone is usually the loudest recorded sound there are most of the time no problems with detecting the correct frequency. Other loud noises can of course still disrupt it. Moreover, when the trombonist is not playing the trombone, but still a sound is recorded (by for example the trombonist speaking) detecting that this frequency is not the trombone can become a problem.

This especially becomes a problem, when not only the evaluation at that moment becomes wrong, but more importantly wrong values is saved to the dataset and as such the acoustic model gets misaligned.

There are a few things that can be done to reduce this data pollution:

- **Note Stability**: Another possible option is to only record frequencies that have been stable for x amount of time, under the assumption that non trombone sounds have more fluctuation to them. A major disadvantage might be that the model is loosing data for notes, only played a short amount of time depending on the implementation. Furthermore, the notes of a young learner might tremble a lot more, possibly causing the acoustic model not being able to calibrate. Finally background noises that have a consistent fundamental frequency are not obstructed at all. This option as such partially hides the problem, but is not able to solve it.

- **Mouth – Mouthpiece Distance**: When the digital mouthpiece location is further away than $10cm$ from Xcm below the viewport in its local space (guessed/selected distance from eyes to the mouth, $\sim$ 6-7cm), discard the pitch, as a trombone can not be the one recorded when the instrument is not at the mouth.

- **Slide Position Distance**: While our model will always just be approximated, some values are so far off our current approximation, that they must be wrong. As such some frequencies can be discarded, if their slide position is off by more than a slide position difference from any of the possible positions.

- **Recording**: Placing a dynamic microphone which has a shorter pickup range or a shotgun microphone which is directional near the bell. Using a microphone in this way would change the sound balance between instrument and background in a way where background sounds might be able to be mostly avoided with a minimum

volume the trombone needs to surpass. To note is that such a microphone must be able to withstand the high sound-pressure level produced near the bell. This also introduces more necessary hardware components.

- **Note – Slide Position Logic**: As a fail-safe this should always be part of this calibration. We know which notes can and can not be played on a given partial n. (See Figure 2.1) As such we can validate the newly evaluated a, b against this. These validation tests should include whether the highest note on that partial is still possible, whether the note a halftone above one is not in the possible range and whether the note on slide position 7 is within $0.7m$ extension of the slide. Otherwise, the new values a, b need to be rejected.

  If expanding the model to work with multiple trombone types, types as the $E\flat$ alto trombone needs a different set of possible notes for each partial, due to the different fundamental.

4.3.5 Play Instructions and Evaluation

I wanted to implement multiple ways of visualizing the play instructions in intuitive and traditional ways. As the traditional staff system is the main way music is notated for trombone, it should be represented here to not lose the association.

File Type

Before going over the exact representations I had to decide on a format to use. File formats that come to mind are *MIDI* and *MusicXML*. Musical Instrument Digital Interface (MIDI) is an event based format for electronic instruments with exact information on when a specific note starts and ends. MusicXML, on the other hand, stores all information to create a visual in staff notation with all its annotations. As such, it includes a lot more information for the staff system MIDI lacks. For example, it includes articulations, tuplets, slurs, dynamics, and grace notes. Some elements can be emulated with MIDI such as glissandi as pitch-shifts, while others are practically impossible[4] with its normal events. Although a correct implementation using MusicXML could lead to a much feature richer product, there is a lot more to interpret leading to a whole array of possible edge cases and errors. Using MIDI on the other hand is much more accessible for users to add their own tracks to the program. As MusicXML is also not intended for playback it does not

[4]There is of course the option to encode additional information as markers, but that makes their creation very cumbersome and would almost require a custom builder for the file creation to be accessible.

seem like there are any existing libraries, so I would have to implement everything myself. With MIDI on the other side there do exist plugins like *Maestro - Midi Player Tool Kit* which can serve as a foundation for the implementation. Due to this I ultimately decided on utilizing MIDI for this prototype.

Visualization and Evaluation

My first idea of visualizing the slide position was sending *Synthesia*-like bars directly to the slide positions on the trombone itself. This concept has been scraped as the bars were not sufficiently visible through the headset before they had to be played. As such the player would barely be able to react to them which defeats the purpose of the idea. This system still has remanents in the project in the form of a smaller indicator that just shows the current to be played position on the slide itself.

For the full slide bars I ended up utilizing a system closer to the system used in *Trombone Champ*, but instead of the bars flying towards you, they are flat in front of you similar to how the sheet musics staff system appears on a stand in front of you. The system has seven bars representing the seven slide-positions of the trombone to give some orientation to the player. On the left side of that area there was a position marker added with the current extension of the slide. To it is left there is an even more zoomed-in representation in the form of a slider that is in the middle when the position is perfectly matched and fully at the top or bottom once the current position is one slide-position away from the correct position.

The current form of the conventional staff systems implementation is a lot more bare bones, but functions as a workable foundation for this prototype. Features that have been implemented include:

- The pitch being correctly mapped to a position on or between the lines.
- Displaying the note lengths correctly as a Whole, Half, Quarter, Eighth, or Sixteenth.
- Showing the correct current time signature and accidentals according to the key signature at the clef.
- Showing the correct accidentals at the notes dependent on the current key signature of the track.
- The correct amount of ledger lines being added if the musical note is outside the 5 main lines of the staff system.

- Bar lines being added between the measure of the track dependent on the current time signature. (Displayed slightly in front of the actual first beat of the measure, as that space is quite often occupied by the first note)

It obviously lacks the features that MIDI does not supply as well as certain note lengths that could be created with ligatures or augmentation dots and rests where there is no note to be played.

In general every note is saved with its midi-value and gets evaluated on the current acoustic model of the trombone when visible to get the correct slide positions of that midi note. This evaluation happens if the note first gets generated, first becomes visible and anytime the trombone model changes (in the case that it can adapt, See Section 4.3.4).

A feature I found necessary to implement for the trombone was some sort of glissando functionality. This was made possible by translating the pitch shift events within a note into a `AnimationCurve`, which can be evaluated at any time to smoothly calculate the required slide positions and evaluate the player at any required time. A requirement of that calculation is of course that every pitch must be possible on the same partial. This fits our existing system, where alternative positions can easily be found so that all needed positions for the glissando are on the same partial. This has been also visually implemented for the bar representation utilizing line renderers, but in the sheet music it is not visible yet. As such also the slide positions on different partials are saved as `AnimationCurve`s once evaluated.

The position of the trombone is instantly compared to the position it is supposed to be. If a note is also played, the partial is reconstructed and the trombone is evaluated on that partial if an alternative position exists there. The error is evaluated in percent with an 100% error being roughly the next slide position. This error gets displayed as a small Feedback text in the top-left corner of the visualization window. On a maximum error defined as 50% (The halfway point between two consecutive slide-positions) “Wrong Position” gets displayed. Then in different intervals the texts “Inaccurate” (40%), “Okay” (30%), “Good”(12%), and “Perfect” (2%) get displayed. I tried not to be too negative in the choice of the feedback so as not to discourage the learner and as such avoided judgmental labels like “Bad”.

4.4 Experiment

I conducted a user study to test the possible benefit for learning and usability and quality of this prototype and its different features. As I only had limited time a limited timeframe was available for a single participant. The study was designed with more participants in mind, then I ultimately got. In this section I describe the design of the study, its order of events and exactly what the participants were asked to do.

For this study the participants were split into experimental group (EG) and control group (CG). For both groups the study had a practical part, where they were trying out the prototype in a controlled environment and a questionnaire part. The locations of the study were a school in Lower Saxony, a Bremen based music school and my own home. The target participants included trombone players of different skill levels (but mostly slightly advanced beginners, but also some more experienced players) as well as some brass instrument players that try out the trombone for the first time. The main target for this user-study were the advanced beginners.

I decided that brass players should be the only non-trombone player group to participate, as they know how to handle and not damage this kind of instrument, with two exceptions. They also may be able to actually play notes, as the mouthpiece is similar.

Another question I had to face was the choice for the headset. Ultimately I decided to do it with a valve index for its wide field of view of about 108° horizontally and up to 120° vertically. This may enable the player to better see more of their instrument, which would increase the similarity to looking at the real one.

4.4.1 Music Choice

For such a user-study I needed to decide which piece of music should be used. It had to be something simple that can be learned quickly but also uses different slide positions, including several medium-sized jumps between positions. Moreover, it would be good if the music piece avoids the insufficiencies of the theoretical acoustic model. (See Section 4.3.4) I chose "Gameboy" (Track 9) from the school for trombone *Hören, Lesen & Spielen 2* [40] as this piece included all that. The piece with twelve measures is divided into three phrases of four measures each. The phrases all have a simple equal rhythmic pattern to not add an additional difficulty. Furthermore, the piece has the key signature D-major, which is quite uncommon for beginners, leading to notes they do not play as often – being $C\sharp$ and $F\sharp$ – where assistance like the program can show its effect.

In addition to the piece the participants were to play the common scale of B-major and the uncommon scale of D-major. These have easy consistent steps to have a reference frame when evaluating. D-major was also added, so that the participants can practice the key-signature of the main piece.

4.4.2 Questionnaire

The questionnaire was designed to detect the usability and possible problems of the program. It also includes questions that might explain abnormalities in the data. The full questionnaire can be found in the Appendix C.

The questionnaire is split into two parts: One to be filled before and one after testing the program.

The first part (See Appendix C.1) asked them about their age, trombone skills and when the last training session was. It also includes their play range (highest/lowest notes they can play). With this data, I can judge the rough skill level and later compare, growth depending on that level.

The second part (See Appendix C.2) asked the participant about their experience and problems they encountered. To separate struggles with the instrument from struggles with the program there are a few questions that directly ask for typical problems when playing a brass instrument. This includes problems with embouchure which might happen when they already played their instrument for a while before participating. The usability of the program was queried using the System Usability Scale (SUS). To separate the position help from other problems, I added questions like "Was it clear which notes had to be played and when?". I also asked them if the real trombone and its digital counterpart felt in sync with each other. I distinctly asked about the positioning help in a slide position precision level – learning where the 7 slides positions are on the trombone – and on a sub slide position precision level – learning the variations across partials – to differentiate between the two.

Near the end of the questionnaire, I asked the participant whether they themselves think that the application succeeds in its use case and whether they would use it themselves. At the bottom of the questionnaire the participants could also leave additional feedback in text form to inform me about additional quirks that might not have come up in the questions.

4.4.3 Procedure

Before the participants could partake in the Study I had to prepare the environment. This included the calibration of the trombone, to achieve a similar calibration between all runs. When in a remote location, I also needed to sett up the VR-Environment including its base stations. In such cases the base stations were mounted on cymbal stands with 1/4-20 UNC attachment screws allowing for the stable attachment of base stations. I made sure of that there were no reflective surfaces in the immediate tracking area to reduce tracking errors.

The sequence of events during the studies individual participant's run in outlined below:

First the participants were asked to fill out the first part of the questionnaire (See Appendix C.1).

After they finished this, I explained the rough idea of the program and gave them an overview of following procedure. I also gave a short instruction on the VR headsets including things like how to adjust the eye distance, so the participant could properly see.

Inside the VR Headset the first task was to navigate the system with the trombone as a controller. After I made sure, they successfully understood this, the 10-minute **Training Phase** began, where they could test out the different functionalities and were supposed to practice the three music tracks. During this phase the EG had access to the advanced position help, where the slide positions were displayed as bars in front of them and at their trombone itself. The CG did not get access to these features. Instead, a reference sheet with the positions of the notes and a small trombone model with its slides 1-7[5] indicated (See Figure 4.6), like it may appear in a learning book, was supplied.

Once the 10 minutes had passed or the participant had enough and had played every piece, it was time for the **Perform Phase**. In this phase the participant had to play through every piece once with "Pause on Wrong Note" (See 4.2) and the *advanced position help* both deactivated without pausing or trying again. I found it suitable that they could keep their set speed, as young learners may not play music at full pace when it is too fast for them at that point. As the goal was position rather than speed, this did not interfere with the study objective. The Tuner was also left visible, as such a device can be used in traditional play.

[5] as calculated average from different partials

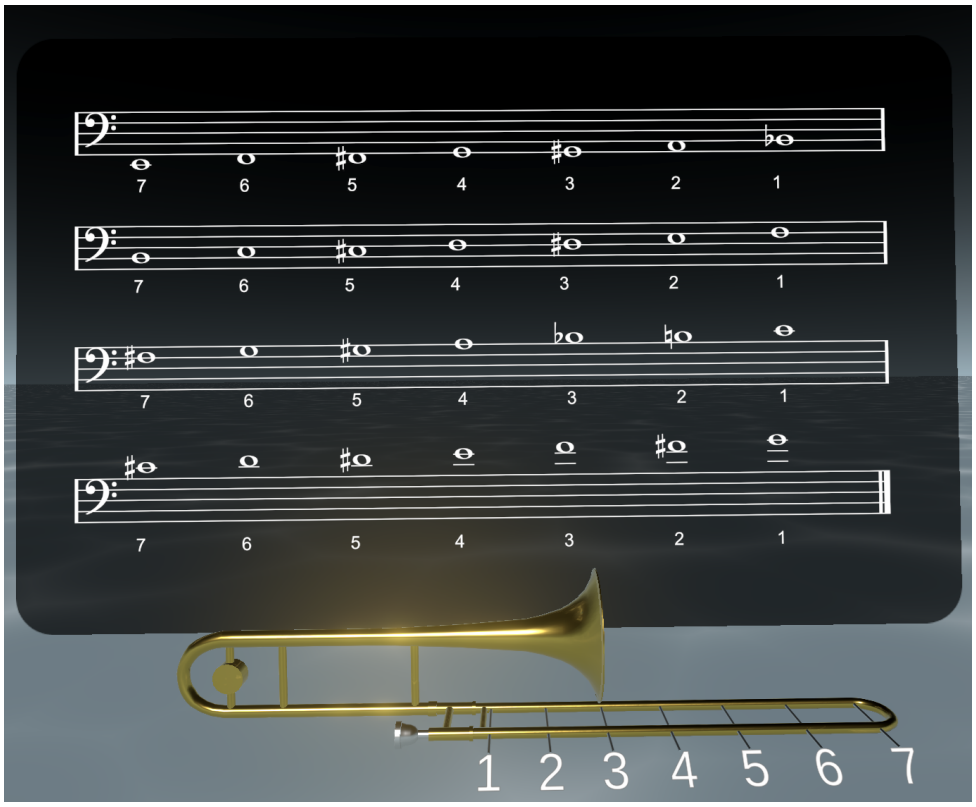


Figure 4.6: Slide position reference sheet for the CG

Now the last task for the participant was to fill out the second part of the questionnaire (See Appendix C.2), after which I thanked them for their participation.

4.4.4 Data Collection

To be able to later evaluate what was played some data were recorded during playing for both the EG and CG.

These include single values recorded once, being the set length of the virtual instrument, the maximum slide distance, start/end time (to match the Questionnaire) and the time the **Perform Phase** started as well as values that are recorded on event being menu interaction and finally data that is recorded every frame. The values that are recorded every frame are: The time since last frame, if the advanced position help is active (in case something goes wrong), the current tracked trombone frequency and slide distance, (when present) the current midi note number, and slide position they are supposed to play, the current position within the active track and finally the slide position tracker in local space of the trombone.

These data allow the accuracy of the two groups to be compared across the two phases.

To prevent recordings from accidentally being deleted, the time between participants, during which the program was still running, was also recorded and saved with the suffix "discarded". For this purpose the audio was also recorded and saved. Although not often this sometimes included the voice of the participant. This data includes as such includes personal data, and was only saved on a local protected device and discarded once the correct recordings have been identified.

The participants were informed that such data would be recorded if they participate in this study.

Chapter 5

Evaluation

The user study had 17 participants across three locations split into the experimental group (EG) with 9 members and control group (CG) with 8 members. Most participants were between 11 and 25 years old, with two older participants over the age of 50. Of the participants 10 had prior experience with the trombone and under the remaining 7, 2 did not have any experience with any brass instrument. (See Figure 5.1)

Regarding prior VR Experience, most participants had tried it at some point, but 4 had no experience. Only two participants use it “once in a while” or “occasionally”. (See Figure 5.2)

Analyses directly addressing the predefined Hypotheses are reported as tests to confirm the hypotheses. Additional analyses, including analyses with alternative thresholds and subgroup comparisons, are considered exploratory and to be interpreted as such. For the Evaluation statistical tests were chosen based on the measurement scale, study design and characteristics of the data. A p-value $< .05$ was considered statistically significant and p-value $< .001$ highly statistically significant.

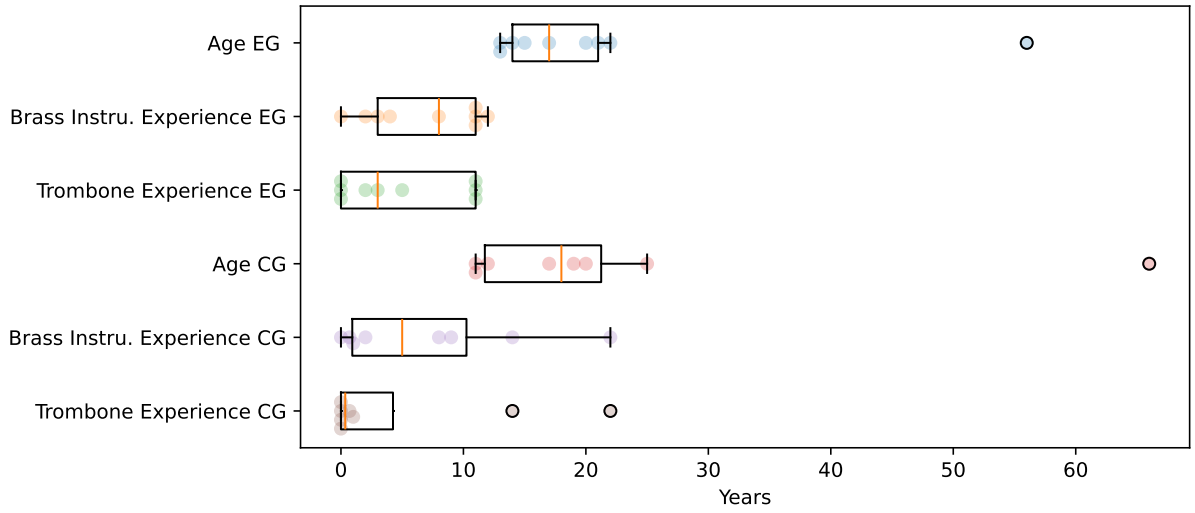


Figure 5.1: Age and Years of Experience of Participants

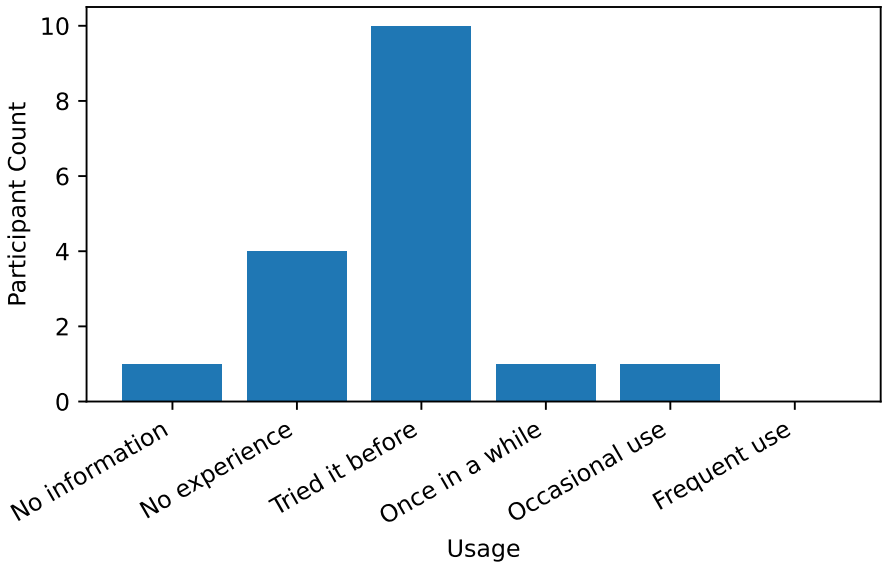


Figure 5.2: VR Experience of Participants

5.1 Hypothesis 1

H1: The XR prototype achieves a mean System Usability Scale (SUS) score

(a) above the benchmark of 68

(b) above the benchmark of 80

5.1.1 Hypothesis test

The XR prototype achieved a mean SUS score of 82.06 with a standard deviation of 10.40 and a 95% confidence interval of $[76.71, 87.41]$. The median of the score is 82.5. (See Figure 5.3) The full distribution is shown in Figure 5.4. Two one-sided one-sample t-tests were conducted to find out whether the mean SUS scores exceeds the benchmarks of 68 and 80.

The mean SUS score was significantly higher for the benchmark of 68. ($t(16) = 5.40, p < .001$) **H1a** is as such supported within this sample.

Regarding the second benchmark of 80 there was no sufficient evidence ($t(16) = 0.79, p = .22$) to conclude that the mean SUS score was significantly higher than 80. which results in **H1b** not being supported.

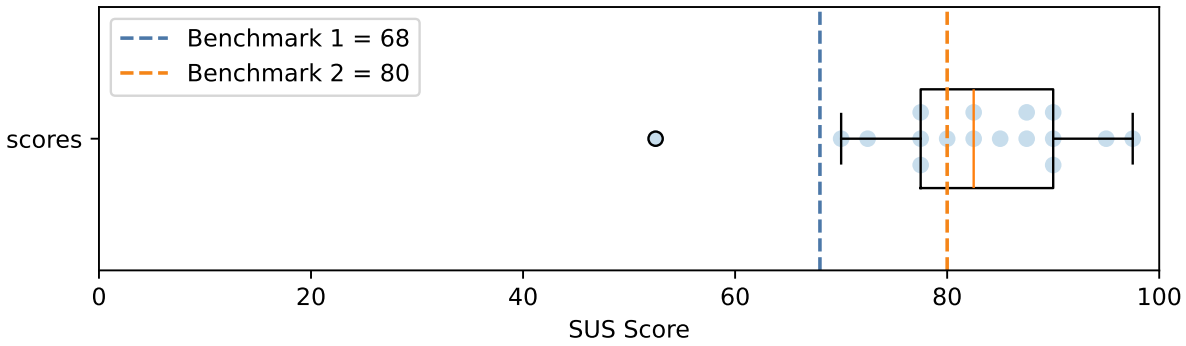


Figure 5.3: System Usability Scale Score Results

5.1.2 Additional Analyses

The EG achieved a slightly higher mean SUS score ($M = 84.44, SD = 7.34$) than the CG ($M = 79.38, SD = 12.48$). As such the EG scored 5.06 SUS points higher on average. A Welch's independent-samples t-test found no statistically significant difference between the

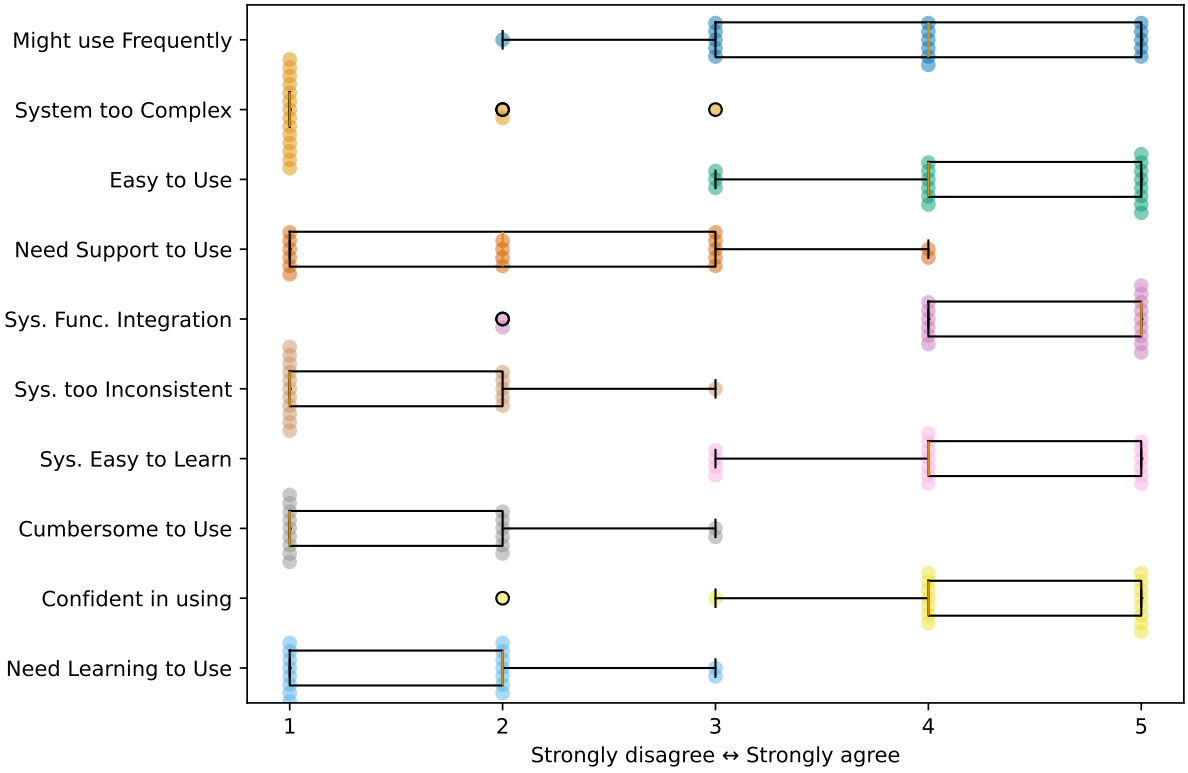


Figure 5.4: System Usability Scale Answer Distribution

two groups. ($t(10.99) = 0.94, p = .36$) This test was chosen, as it does not assume equal population variances and the EG and CG are independent groups.

One notable case is the outlier with a SUS Score of 52.5 (As shown in Figure 5.3), despite reporting a high level of enjoyment (5/5). They were a member of the control group, and had as such no access to the advanced positioning help. They are an experienced trombonist and had no problems with the real trombone. It was stated, that they would need support when using the system (4/5, SUS). The participant had no prior VR experience. They rated the learnability of the prototype relatively positively (4/5, SUS) and that the prototype was not too cumbersome to use (1/5, SUS). On the other side they were of the opinion, that the features were not well integrated (2/5, SUS) Otherwise, they were just very undecided on the answers of the SUS (3/5). They also reported no problems in understanding which note has to be played when (1/5, Question at C.2.3) while reporting problems in finding the correct slide position during training (4/5, Question at C.2.3) and performing (4/5, See Question at C.2.3).

5.2 Hypothesis 2

> **H2:** Participants with prior trombone experience rate the XR prototype positively regarding the potential to increase their motivation to learn and practice the trombone.

$H2_0$: The motivation ratings do not exceed the neutral midpoint of 3.

$H2_1$: The motivation ratings exceed the neutral midpoint of 3.

5.2.1 Hypothesis test

The participants with prior trombone experience ($n = 10$) rated the potential of the XR prototype to increase their motivation positively with a mean of 3.80 ($SD = 0.75, 95\% CI = [3.26, 4.34]$) and a median of 4.00 on the five-point Likert scale. (Data in Figure 5.5) As the motivation rating was measured using a five-point Likert scale, a non-parametric test was used. A one-sample Wilcoxon signed-rank test against the neutral midpoint of 3 indicated that the ratings were significantly higher ($W = 74.0, p = .002$) than the neutral midpoint. As such, **H2** is supported.

This data has to be taken with caution, as due to the very small sample size of 10.

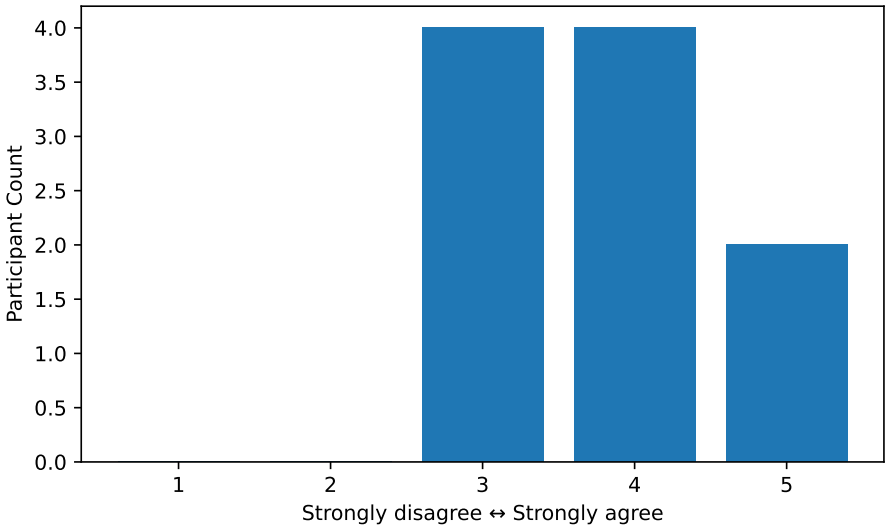


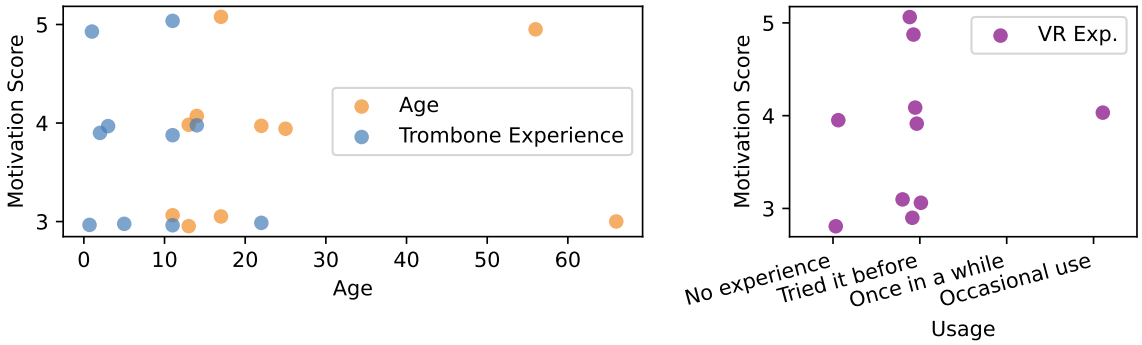
Figure 5.5: Trombonists Perceived Potential to Increase their Motivation

5.2.2 Additional Analyses

To investigate whether trombone experience, age, or VR experience could be associated with the participants' perception of the prototype's motivational potential (See Figure 5.6), an exploratory Spearman rank correlation was conducted. No significant correlation ($\rho_s = -0.13, p = .71$) was found between trombone experience and motivation rating, indicating a very weak negative association in the sample. Additionally, due to the small sample size, an independent permutation test[1] was conducted ($p = .004$) leading to the same conclusion. The same can be said for the Spearman rank on age ($\rho_s = 0.29, p = .41$), but with a weak positive association, and the additional permutation test ($p = .41$). Concerning VR experience we have a similar result with the single test ($\rho_s = 0.2, p = .56$) with a weak positive association and the permutation test ($p = .71$). But due to most participants having tried VR it at some point (6), only two having no experience and just one person using the technology occasionally together with the experience amount being very vague, especially this result is to be taken extremely cautiously.

Therefore, in this sample, trombone experience nor age or VR experience was not associated with participants' ratings of the prototype's motivational potential.

[1]The permutation procedure is described in detail in Section 5.3.3.



(slightly randomized position along the y-axis) (slightly randomized position)

Figure 5.6: Motivation Score in Relation to Age, Trombone Experience and VR Experience

5.3 Hypothesis 3

H3: Participants achieve a higher slide-position accuracy – reflected by a smaller distance from the target slide position – while within $10cm$ of the target position when using the XR prototype's slide-position assistance during training, than without it

(a) during training (with the assistance).

(b) after training (without the assistance).

5.3.1 Data Preparation

For this hypothesis the evaluation becomes more complicated. Obviously we compare EG against CG for a significant difference, but there are further subgroups to be considered: experienced trombonists and newcomers.

The values for each participant were calculated as follows:

1. The samples (recorded each frame of their performance) where no target note is defined (`isnan(record_goal_midi)`) or where the intended slide position is further away than $50cm$ were removed.

   Those positions are harder to reach for younger players and that is also where the model is most inaccurate (See Section 4.3.4). The slide distance of $45cm$ is the distance between the calculated slide position 5 and 6. As such notes on position 6 and 7 are excluded.

2. The remaining samples were divided into groups according to changes in the target note.

   This group's consecutive notes of equal MIDI value, as intended, since the slide position does not need to change between them. Grouping by MIDI value instead of target slide position avoids inconsistencies due to brief alternative-position suggestions in the record data. The MIDI values serve the same purpose as different notes – although maybe on the same slide position – always have slight variation between them.

3. For each sample the absolute distance to the target slide position was calculated.

4. Values over the distance of $10cm$ are filtered, as the player is not yet at the correct slide position.

   This Hypothesis is about the accuracy, not if they found the correct rough slide position. This cutoff point is to some extent arbitrary. For the following presentation of the data the cutoff point has been set at $10cm$ in text and graphics. This point was chosen, as it is the rough average distance between two slide positions of the model. Within this area the participant might still have intended to hit the correct slide position. But as this value is still arbitrary see Table 5.2 for the results with different cutoff points and their distributions in Figure 5.9.

5. The samples within a group were combined using a weighted average under a bell curve $f(x) = e^{-(x/0.5)^2}$.

   This is done to reduce the effect of the start and end of the notes, where the player is still finding the correct position or is already preparing for the next note. This is intended to make the evaluation reflect the position at which the participant actually plays the note.

6. Finally, the mean of the groups is the value used for each participant.

This process results in one value for each participant, calculated as the mean of the weighted mean distances belonging to each (timeframe separated by differences in) target note.

The python function for this calculation can be found in the Appendix Listing D.1. The resulting participant-level accuracy scores are summarized in Table 5.1 and visually displayed in Figure 5.7.

Group	n	Mean in cm of Training Distance	Perform Distance	Change
EG	9	2.12	3.59	4.98
CG	8	4.08	4.42	1.73
EG (No Trombone Exp.)	3	2.00	5.21	12.58
CG (No Trombone Exp.)	4	3.94	5.22	3.16
EG (Trombone Exp.)	6	2.18	2.79	1.17
CG (Trombone Exp.)	4	4.21	3.62	0.29

(Change is evaluated for every person from which the mean is calculated)

Table 5.1: Mean Distance from the Target Slide Position for the Different Participant Groups

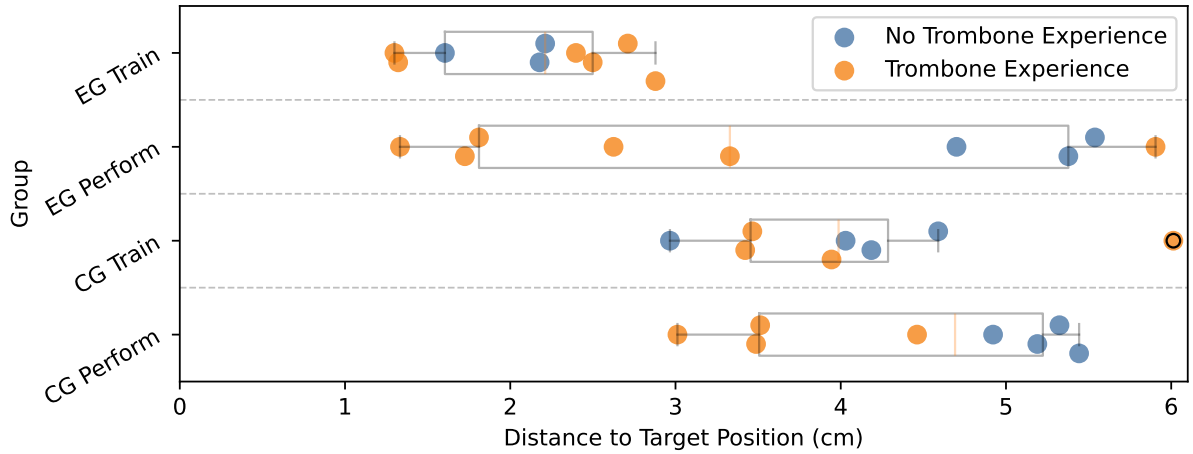


Figure 5.7: Accuracy of the Individual Participants

5.3.2 Hypothesis test

As the groups are independent, and the accuracy scores can not be expected to be normally distributed due to the small sample size of EG ($n = 9$) and CG ($n = 8$), a one-sided Mann–Whitney U test was conducted to compare these groups.

For the first sub-hypothesis **a** regarding the **Training Phase** the results of that test were statistically highly significant. ($U = 0, p < .001$) The U statistic of 0 indicates complete separation of the two groups, so that every EG participant had a smaller distance from the target position than every CG participant. (This can also be seen in Figure 5.7.) This result supports **H3a** for these samples.

Regarding the sub-hypothesis **b** about the **Perform Phase** the difference was not statistically significant. ($U = 28, p = .24$) This result does not support **H3b** in this experiment.

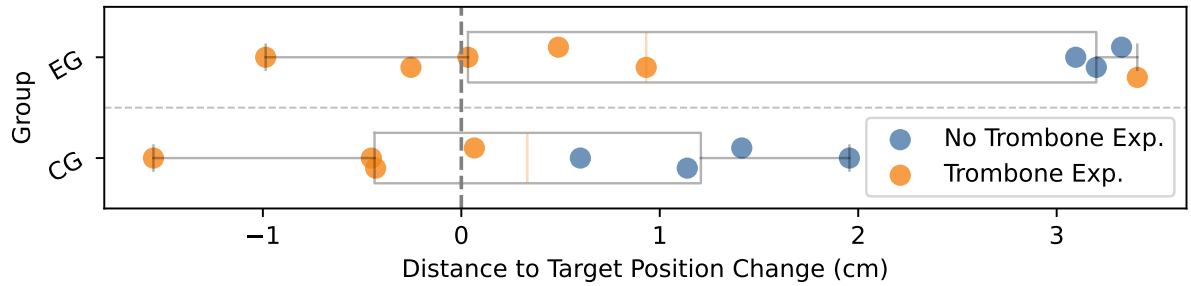


Figure 5.8: Change in Accuracy from Training to Perform Phase of the Individual Participants. Negative values indicate improvements from the Training Phase.

Due to the arbitrariness of the cutoff point, additional exploratory tests have been conducted with different values (See Table 5.2) to assess the robustness of these findings. There the samples were analyzed with values within a reasonable range of 5-10cm with 1cm steps for the cutoff. Setting the cutoff above 10cm we are fully in the area of a different slide position, while setting it below 5cm we would start only filtering for already roughly correct data, as it is already closer to the correct slide position than any other slide position.

As all results for the **Training Phase** within that area are significant, **H3a** is still supported by this experiment. For the **Perform Phase** we do have 2 significant results out of the six at a cutoff after 7cm and 8cm. Due to this not being the clear majority[2] **H3b** continues to be not supported for the samples.

5.3.3 Additional Analyses

I also repeated these tests (with cutoff 10cm) for the groups separated according to prior trombone experience. As the sample size for this is even smaller the results have to be treated with caution. For the **Training Phase** the experienced members ($n_{EG} = 6, n_{CG} = 4, U = 0, p = .004$) as well as the inexperienced ones ($n_{EG} = 3, n_{CG} = 4, U = 0, p = .02$) had a significant difference. On the other hand for the **Perform Phase** both the experienced ($n_{EG} = 6, n_{CG} = 4, U = 5, p = .08$) and the inexperienced ($n_{EG} = 3, n_{CG} = 4, U = 7, p = .68$) had no significant difference.

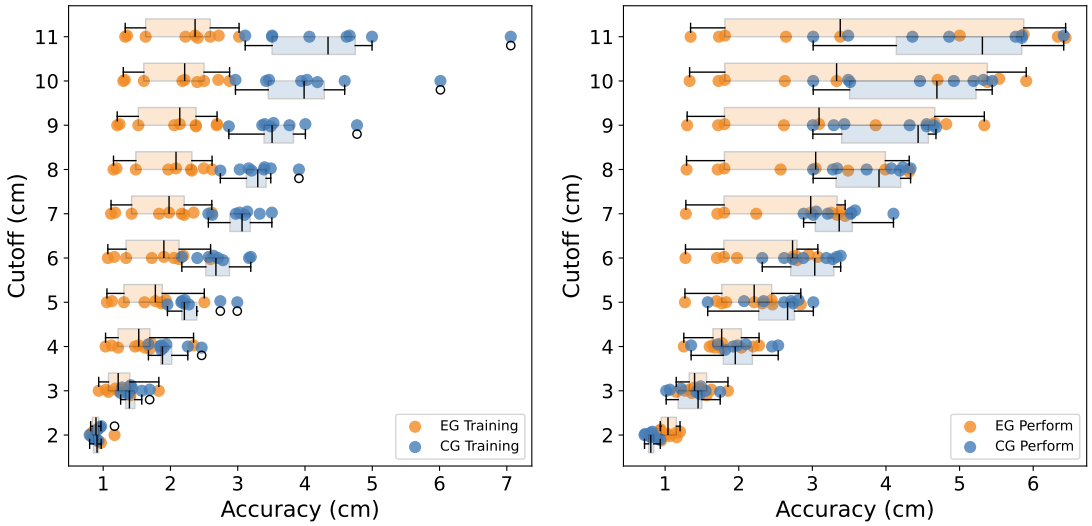
This prompts to investigate the change between the training and perform phase. Furthering motivation for this analysis came from the four of the nine EG participants, who formed a visually distinct group (See Figure 5.7), with their values separated from

[2]The majority would not have confirmed anything, just would have prompted more tests.

Cutoff	Train U	Train p	Perform U	Perform p	
never	4	$< .001$	23	.11	
$11cm$	0	$< .001$	27	.21	
$10cm$	0	$< .001$	28	.24	Reasonable Cutoff Ranges
$9cm$	0	$< .001$	27	.21	
$8cm$	0	$< .001$	22	.11	
$7cm$	1	$< .001$	17	.03	
$6cm$	4	$< .001$	16	.02	
$5cm$	6	.001	20	.06	
$4cm$	9	.003	24	.13	
$3cm$	18	.04	40	.66	
$2cm$	35	0.48	71	0.99	

(Significant results are highlighted in green. The reasonable cutoff range is discussed in Section 5.3.2.)

Table 5.2: Hypothesis Test with different Cutoff Points as described in Section 5.3.1 Step 4.



(a) During the Training Phase

(b) During the Perform Phase

Figure 5.9: Distributions with different Cutoff Points as described in Section 5.3.1 Step 4. No cutoff is left out in this Depiction, as it would make the Figure unreadable.

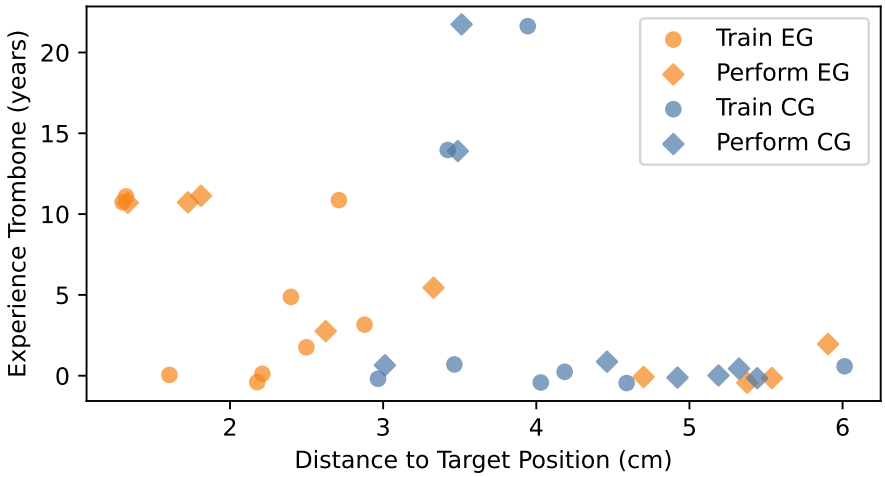


Figure 5.10: Scatter plot for the participant values in relation. Experience years randomly altered by up to ∓ 0.5 years for better readability.

the remaining participants by approximately 1-3cm. To analyze if the deviation between phases amongst participants (Seen in Figure 5.8) increases significantly, a one-sided Wilcoxon signed-rank test was conducted on the paired values. The result for the EG ($W = 38, p = .037$) is statistically significant while the same test is not significant for the CG ($W = 24, p = .23$).

Comparing this deviation between the EG and CG have no significant difference according to a two-sided Mann–Whitney U test. ($U = 46, p = .27$) There is also no significant difference when we run the same test on the groups with only the participants with experience ($U = 19, p = .17$) or without ($U = 12, p = .057$). But analyzing if the participants with experience have a smaller deviation compared to the ones without experience regardless of group leads to a significant result according to a one-sided Mann–Whitney U test. ($U = 8, p = .003$)

Another interesting analysis is the possible association between trombone experience and distance to the target position. (See Figure 5.10) Due to the small sample sizes the statistical significance of Spearman's rank correlations was assessed using a permutation test for each phase and group, as well as for both groups combined. For each test, the observations of one variable have been randomly permuted relative to the other, and Spearman's ρ_s was recalculated for each permutation. The final p-value is the fraction of permutations that produced a correlation at least as extreme as the observed correlation in absolute value. A total of 100,000 permutations were performed for each test. The

Phase	Group	n	Spearman ρ_s	Permutation p
Train	EG	9	-0.069	.87
Perform	EG	9	-0.84	.008
Train	CG	8	-0.26	.51
Perform	CG	8	-0.83	.016
Train	EG + CG	17	-0.18	.46
Perform	EG + CG	17	-0.7	.002

Table 5.3: Results of the Spearman's rank correlation with a permutation test for the association between trombone experience and distance to the target position

resulting values are listed in Table 5.3.

The result for this dataset indicate a strong negative association between distance and experience for all groups (EG, CG, combined) during the performance phase. In contrast, no such association could be detected in the training phase. ($p > .05$)

Repeating the same analysis using brass instrument experience instead of trombone experience yielded no significant associations, but for the combined groups during the perform phase, where a negative association is indicated. ($\rho_s = -0.54, p = .025$)

5.4 Hypothesis 4

> **H4:** Participants achieve a higher percentage of playing the correct slide position (within $5cm$ radius), when using the XR prototype's slide-position assistance during training, than without it
>
> (a) during training (with the assistance).
>
> (b) after training (without the assistance).

5.4.1 Data Preparation

For this we want to analyze if the person headed for roughly the correct slide position not regarding their precision. We do this by testing if the player was proximity, which we can do similar to Section 5.3.1, with the following key modifications:

3. Each sample gets tested if the absolute distance to the target slide position is smaller than the distance $5cm$. If this is true the value is 1, if false 0.

The Distance of $5cm$ is the distance at which we do not consider it the correct slide position anymore. Similar to the cutoff point discussed in Section 5.3.1 Step 4 this distance has some arbitrariness to it. Additional reasonable values will be evaluated.

4. The Samples are not filtered.

This process results in one value for each participant, calculated as the mean percentage of the weighted mean correctness belonging to each (timeframe separated by differences in) target note.

5.4.2 Hypothesis test

The primary analysis for **H4** used the distance of $5cm$, as defined in the hypothesis. To test **H4** a one-sided Mann–Whitney U test is used just like for the previous hypothesis to determine whether the EG ($n = 9$) had a significantly greater proportion of correct slide positions (within $5cm$ of the target) than the CG ($n = 8$).

For the first sub-hypothesis **a** regarding the **Training Phase**, the results of that test were statistically highly significant. ($U = 68, p < 0.001$) This result supports **H4a** for these samples.

For the second sub-hypothesis **b**, inspecting the **Perform Phase**, the difference was not statistically significant ($U = 52, p = .069$) and therefore does not support **H4b** within this experiment.

Additional exploratory tests for different values, after which a position was considered wrong, were conducted to assess the robustness of these findings. Their results are in Table 5.4 and the participant values visualized in Figure 5.11. The range of 3-6cm was considered reasonable with $1cm$ steps for each test. (white rows in the table) With a distance of $> 6cm$ the slide position is far closer to a different slide position then the target. With a distance of $5cm$ this is already slightly the case for some positions, but $6cm$ was still considered reasonable with a small extra buffer. $3cm$ was deemed as the smallest position as otherwise the errors of the model (See Section 4.3.4) can greatly affect participants (Especially CG) who, for example, worked with the tuner.

Every result considering the **Training Phase** within the reasonable range are significant, so **H4a** continues to be supported by the samples. For the **Perform Phase** the data is more nuanced where 2 out of 4 versions have significant results. Significant results have been observed at $3cm$ ($U = 56, p = .029$) and $4cm$ ($U = 56, p = .029$). To explore this

Wrong after	Train U	Train p	Perform U	Perform p	
$1cm$	69	$< .001$	43	.27	
$2cm$	71	$< .001$	53	.056	
$3cm$	71	$< .001$	56	.029	Reasonable Wrong Distance Values
$3.25cm$	71	$< .001$	55	.037	
$3.5cm$	71	$< .001$	55	.037	
$3.75cm$	70	$< .001$	56	.02	
$4cm$	70	$< .001$	56	.029	
$4.25cm$	70	$< .001$	53	.056	
$4.5cm$	69	$< .001$	52	.069	
$4.75cm$	69	$< .001$	52	.069	
$5cm$	68	$< .001$	52	.069	
$5.25cm$	68	$< .001$	49	.11	
$5.5cm$	68	$< .001$	49	.11	
$5.75cm$	68	$< .001$	49	.11	
$6cm$	67	$< .001$	49	.11	
$7cm$	66	.001	47	.16	
$8cm$	64	.002	49	.11	
$9cm$	66	.001	48	.13	
$10cm$	66	.001	53	.056	

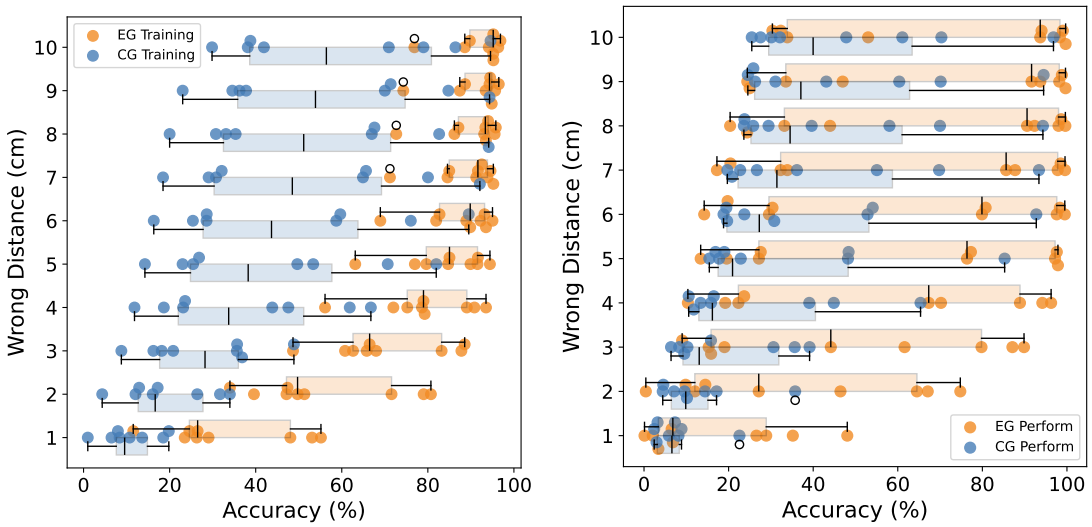
(Significant results are highlighted in green.
The reasonable wrong distance values are discussed in Section 5.4.2.)

Table 5.4: Hypothesis 4 Test with different Distances at which a slide is wrong

further test were conducted with a step size of $0.25cm$. (See Table 5.4, gray rows) With these additional results there are 5 significant ones from 3 to $4cm$, while the other 8 are not significant.

5.4.3 Additional Analyses

Especially due to the inaccuracies of the model, further exploratory tests regarding the pitch make sense to rule out that these results are only due to those inaccuracies. For this experiment participants with no experience will be disregarded, as learning to also produce the correct pitch using embouchure is a skill not addressed within the prototype. Due to that we work with a sample size of 6 in the EG and 4 in the CG.



(a) During the Training Phase (b) During the Perform Phase

Figure 5.11: Distributions with different Distances at which a Slide Position is wrong.

Similarly to how **H4** was evaluated, further one-sided Mann–Whitney U tests have been conducted on the data with the following changes in the data preparation from Section 5.4.1:

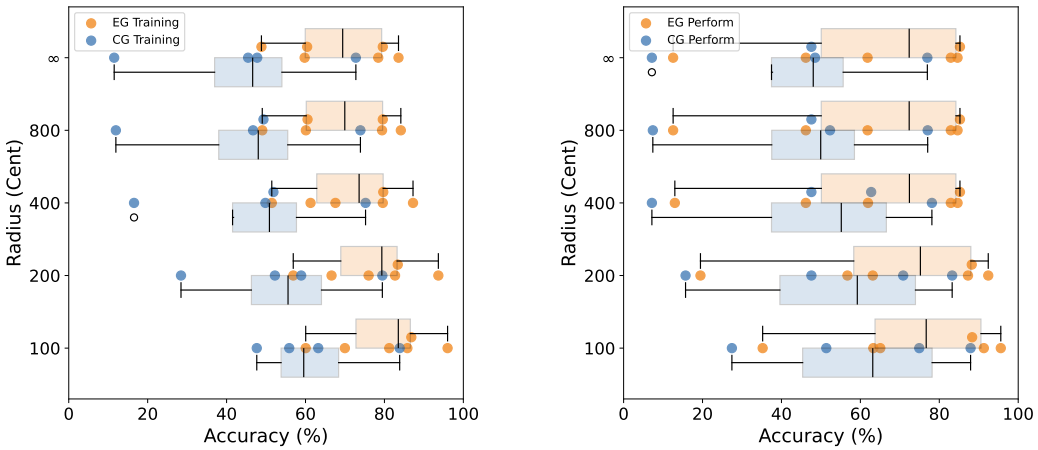
- After grouping them as notes, we filter away all values, where no frequency was detected. (30.37% of the Data)
- A frequency is correct, when the difference is between the goal and detected frequency as cent[3] is smaller than 50 cent. 50 cent being the midpoint between two Semitones.
- Additionally a filter for noise where the difference is too far off is logical. The results will be presented for a filter radius from the goal cents of 100, 200 and 400 cent as well for no filter.

The resulting weighed average for the different participants can be seen in Figure 5.12.

Within a 100 cent radius (24.83% of samples rejected) and 200 cent radius (17.46% of samples rejected) filter the results are insignificant for both the training ($U = 20, p = .057$) and perform ($U = 17, p = .17$) phase.[4] For the filter of 400 cent this stays to be the case. For a 400 cent radius (7.95% of samples rejected) this stays to be the case for both training ($U = 20, p = .057$) and perform ($U = 16, p = .23$) phase. When rejecting no

[3]Unit of measurement. 1 semitone = 100 cent

[4]Values are the same for at least up to the seventeenth decimal place for both phases.



(a) During the Training Phase

(b) During the Perform Phase

Figure 5.12: Distributions with different Cent Radius Filters.

further samples the result for the trainings phase becomes significant ($U = 21, p = .033$), while during the perform phase it stays insignificant ($U = 17, p = .17$).

These result have to be treated with caution. Especially the one with no filter.

5.5 General Questionnaire Evaluation

In this section I will descriptively analyze over parts of the remaining questionnaire. The question answers, that follow a Likert scale like pattern can be found in Appendix C.3.

Regarding the feedback option only 5 were submitted from which only 2 are relevant, as two were simple praise and one did not fully understand the reason for the perform phase. One participant described that slide-position 4 was more difficult than the others. They also mentioned that they were not used to the bell position of the (digital) trombone, as it differs from their own, while also giving the synchronization between real and digital trombone a good (4/5) rating.

The second one mentioned that the Program might be more suited for AR than VR/AV, so that one can fully see your own trombone. Doing so would also remove the difficulty of accurate position tracking. They mentioned tracking errors that happened during their use of the program, specifically regarding inaccuracies in the position detection.

Starting from the largest group, eight gave the synchronization between the real and digital trombone a good rating while four participants diverged to both very good and average rating. Only one person gave a bad rating. The person who gave that feedback

still evaluated the synchronization with a good score. A different person gave the negative score.

I also received some verbal feedback, sadly not written down in the questionnaire, by one of the older participants expressing interest in the program and that – if further developed – it could help in specifically training precision.

A fundamental aspect of such a prototype is that it is understood what has to be played when. The question of if the notes were clear was answered all around positively to very positively, with two neutral responses in the CG but no negative ones.

Regarding the feeling of preparedness before the Perform Phase, the EG's mean leaned towards the positive, while the CG's mean was slightly negative. The participants with experience gave answers from very prepared (1) to not prepared (4), while the inexperienced just answered with prepared (2) or neutral (3) but for one participant of the CG who felt very unprepared (5).

Another interesting topic is how difficult the slide positions were perceived by the participants. For both the training and the perform phase, the answers are widely distributed, with all five answers represented. Especially the subgroups are difficult to compare due to the sample size, and these results needs to be treated with caution. During the training phase EG the scores were generally slightly positive, with new players actually scoring it easier than the experienced players. Half of the experienced players of EG actually rated it at mid-difficulty. The CG generally scored the piece harder than the EG group and subgroups. The experienced players had a mid-difficulty mean score, while the inexperienced mostly rated it as difficult.

Interestingly the mean scoring for the Perform Phase tends to be rated easier for all groups than the Training Phase, except for the mean score of the inexperienced players of the EG, who evaluated it as difficult or very difficult. The experienced players of the EG only rated the Perform Phase slightly easier than for the Training Phase, the mean position is roughly mid-difficult.

Chapter 6

Discussion

With the previous user-study I evaluated my XR prototype designed to support trombone learning through slide position assistance and investigated its usability, perceived motivational potential and tried to obtain first insights into the effectiveness of the slide-position assistance. The study was performed with a small sample size that had to be split into experimental (EG) and control group (CG) and worked with a very short training window. As such, the results should be interpreted with caution, as they may differ when evaluated with a larger sample size and under more realistic training conditions. In general, the results indicate that the prototype was well received by the participants and that the slide-position assistance was effective in improving accuracy during the training phase of the experiment. However once the assistance was removed in the following perform phase, the difference to the control group was not statistically significant. This suggests that the prototype can help with effective immediate guidance, while calling into question of whether it can improve slide-positioning independent of the program.

6.1 User Opinion

Regarding the SUS values the program overall seams to be highly usable with a mean of 82.06 and significantly above the 68 Benchmark, resulting in a general above average user experience. While the mean was over the Benchmark of 80, the distribution had no significant distance to it. As such there is still uncertainty if most users would call the prototype great. But this result has to be treated with caution, regarding that most participants have only tried VR before or never at all. Due to this the new medium might have affected their judgment. New technology can be something interesting and fascinating

but also can be quite overwhelming. As all participants reported enjoying the experience (see Appendix C.3, fun) and generally did not perceive the program as cumbersome to use (see Figure 5.4), it is possible, that the environment of XR resulted in more positive evaluations. However, the magnitude of such an effect can not be determined from this study.

While the experimental group mean SUS score was slightly higher, there was no significant difference found between their scores and the control group scores. This means a positive effect of the slide assistance for usability could not be detected.

This generally positive result matches the users opinion the notes to be played were generally clear, speaking for an understandable UI layout.

An interesting case for the SUS is the outlier who was the only one whose SUS score was below the 68 benchmark with a score of 52.2. One possible explanation is the participant's lack of prior VR experience and the previously mentioned potential overwhelmingness with an entirely new system. Their responses further indicate difficulties with finding the correct slide position and concerns regarding the integration of features. Despite that they found the system potentially easy to learn, not particularly cumbersome and even had a high level of enjoyment. This suggests that enjoyment of the system did not necessarily correspond to perceived usability. It could of course also be the case that they just did not know how to respond to those questions.

It should also be taken into account that firstly, the experiment only worked with a small sample size of 17 on a very subjective topic and secondly, the participants were mostly within the age range of 11-25 and as such can only reflect that age group.

Concerning the potential of increased motivation during practice, the trombonists of the participants all gave the prototype a relatively high rating. The result was also significantly higher than the neutral midpoint, leading to a positive assumption about motivation under the participants. No association was found between these opinions and trombone experience, age or VR Experience. As such the lack of year-long experience, or the familiarity to the technology probably did not influence their decisions.

If this potential would hold true in practice for those participants can not be inferred from this, but it indicates that the idea of this prototype is not unwanted within this small sample.

Whether or not the digital and real trombone were in sync is more nuanced. Seemingly all but four participants have encountered some form of tracking error or misalignment,

as they did not give a perfect score. Outside one person the participants did not report major tracking problems. Although I would not make an assumption of this being the case, it could also be that the difference between their own trombone and the trombone they used during the study was confusing for some, as one participant reported, which might have affected that scoring.

6.2 Trombone Data

But it is not only important that the prototype is usable and that the trombonists are of the opinion, that it could benefit their practice motivation, but also that the playing correctness improves.

Regarding the tests for accuracy (where wrong slide-positions are discarded), during the training phase where EG had additional access to the advanced position help, there was a significant difference between the two groups. Even the worst participant in EG was better than the best participant in CG, resulting in no overlap are between the areas. Similarly, the EG were also significantly better at targeting the correct slide position (from 1-5) than the CG, although the difference is not that big. So while using the prototype, the participants' performance increased overall.

But being a lot more precise with the help is not relevant, if you can not maintain this improvement when performing without it. The result regarding both tests (precision & correct slide) did not show that the difference between the distributions was not because of chance, even though the EG was on average slightly better on both. By tweaking the variables one can find significant differences, but the large majority shows no proof for such a difference. As such there is not enough evidence, to support that the prototype helped them to learn within the 10 minutes of preparation.

A very important aspect is, that unlike the CG some members of the EG got a lot worse (more than $3cm$ less precise) between the training to the perform phase in their precision. Those participants were usually the ones with no or only a few years of experience. (See Figure 5.10 and Figure 5.8) This could be explained by them becoming overreliant on the support system to achieve good results, which when suddenly removed they were not able to replicate. This phenomenon that the trombonist with experience were a lot better at keeping their accuracy compared to the ones without experience was observed for both groups. This also reflects in participants with more trombone experience having better results during the end performance overall, while none such association was visible during

the training phase. But there was also no significant observable difference between the groups, when disregarding the ones without experience. Due to the sample size becoming quite small when doing so, this observation could differ a lot, when more participants are observed.

Interestingly the members of the EG did not feel unprepared and even felt very slightly better prepared than the CG, while getting significantly worse. Therefore, the advanced position help might have instilled a false sense of security within the participants. This false sense of security is also reflected in the difficulty score of the EG newcomers who rated the Performance as difficult to very difficult and more difficult than the training. The experienced members did not feel as if the performance was more difficult, which makes sense, as they also were a lot better at keeping their position.

The general result was not unexpected, as 10 minutes within a new environment might make it difficult to properly learn and internalize the practiced material. When then the help mechanism suddenly stops being available while one might have relied too much on it before, it can easily lead to such results. A study regarding the changes of using this prototype over a longer period might be interesting. But, although never significant, there are differences between the groups that could become significant if more participants follow this pattern.

Although I tried to eliminate the worst offenders, these observations might change to an extent with a more accurate model without the errors. (See Section 4.3.4)

Although no significant differences regarding the actually played pitch have been found, the difference that was there still followed the pattern of the EG being slightly better than the CG and that the difference lessened in the perform phase. While having no meaningfulness on their own, they do not contradict the pattern. This indicates, that the results of this study may not be fully discredited by the errors of the model.

As a brass instrumentalist, I find it odd, that so many players were under the 60% accuracy mark based on their pitch. This could just be due to there being still a lot of beginners under the participants, but there also is the possibility of an error in the pitch detection or too much background noise resulting in incorrect samples.

Regarding the worsening between the train and perform phase for EG, a feature might be useful, where the advanced position help gets slowly faded out the longer one practices.

6.3 Study Reflection

Now let us regard the study design in retrospective, to see how it holds up. While the data tracking and the general procedure worked there are multiple elements too that could be improved upon.

Performing the study in different locations I also had to set up the VR environment each time. This and the fact that participants had major height differences can result in a very different environment, leading to new reasons for tracking errors unknown to me. But overall this will have affected both the CG and EG in equal measures. The only problem could be, that differences between the groups were not that apparent as they could have been. Although a good technology in the correct environment, for user-studies like this another tracking method, might improve upon this problem.

Additionally, I left the tuner active during the perform phase as it is a physical device one can also have access to in practical settings. In afterthought, I realize that this further may have equalized or even benefitted the CG as it is placed on the opposite side of the staff system than the position help, leading to members of EG not paying attention to it and as such not being able to utilize it during the perform phase. Potential learning effects might get more apparent when both groups lose all their help outside the notes itself.

As mentioned earlier the study would majorly additionally benefit from a much longer use period before the performance.

Regarding the choice of music for the study, I think it was well suited, as reflected in the difficulty ranking by the participants. It was not regarded as super easy or difficult, so that the participants can improve on playing, while not being an impossible task.

Due to the feedback of older participants I am also now aware of the interest these older groups might have. Previously to this study I was of the opinion that a possible effect would be most visible in early learners Now that I know of this interest, and how the experienced players are better at keeping their accuracy between phases, a study a broader spectrum might have been useful. This was already a thought prior to the study, but sadly the contact to an orchestra failed.

In this context it is also relevant to have enough participants of a certain level of experience, if the experience levels are spread out as far as in this study (0-22 years of experience). This allows for better analysis of these subgroups and outliers will have less of an effect.

Due to it being my first user-study I still lacked a lot of experience which definitely impeded on the study design and flow. I also delivered the beginning explanation myself, which did at times lead me to forget minor details which I needed to add subsequently. While acceptable, this could have been made more efficient and less headache inducing (for me) by automating this explanation inside the prototype, instead of relying on my thoughts to not drift off. This may have introduced minor irritation for the participants but should not have relevant effects on results, as I made sure the participants knew how to navigate the program, before starting the training phase.

Chapter 7

Conclusion

For this bachelor thesis, I set out to build a prototype for real-time visual feedback in XR that provides assistance for learning the trombone slide positions outside of conventional instrumental lessons.

The prototype connects the real trombone with a virtual representation and environment to provide visual guidance and feedback to support the position finding process. For displaying its MIDI-based music tracks it combines conventional staff notation with a visual slide position representation and directly indicates whether the current slide position is too far in or too far out. By tracking the physical slide position and pitch, the program can estimate the partial and thereby adapts its feedback when the same pitch is played using a different partial and slide position combination.

The currently implemented theoretical model has certain accuracy issues, as it is relying on a simplified model while also being static which means it lacks the ability to adapt to embouchure/environment changes. An improved currently unimplemented model has been suggested, that allows for the approximation of the fully correct positions and that can adapt to the previously mentioned changes.

A small user-study with 17 participants – mainly regarding beginners and advanced beginners – was conducted to evaluate the prototypes' usability, motivational potential, and ability to support the learning process of the trombone slide positions.

Firstly, the prototype needs to be practical and usable for its target audience of trombone learners (**H1**). The results indicate that, from their interaction with prototype, it is highly usable, with a mean System Usability Scale score significantly above the benchmark of 68. However, the results did not significantly surpass the more difficult benchmark of 80. This indicates, while the prototype provides a usable foundation, there

is still room for improvement before it can be considered a highly polished application. This rating is also limited to the parts of the program they interacted with, which does not include the calibration process, excluded to create a more equal environment for the participants.

Secondly, the target audience needs to perceive the prototype as potentially beneficial for their motivation to practice the trombone (**H2**). Participants with prior trombone experience rated the XR prototype positively in regard to its potential to increase their practice motivation. Due to the limited scale of the study, no conclusions can be made concerning the actual long-term motivation increase.

Thirdly, the Program needs to provide effective assistance for learning and practicing trombone slide positions (**H3** + **H4**). The results show that participants achieved significantly higher location of and accuracy in slide-positioning while using the prototypes advanced position help. This demonstrates that the provided feedback can support participants in locating the indicated slide position and even increases fine precision. However, the study did not provide sufficient evidence that this improvement persisted after this assistance was removed. Across multiple analyses, the results showed tendencies in the direction of improvement, but these differences did not reach statistical significance. A larger sample size or longer training period may therefore be necessary to determine if those tendencies represent an actual learning effect or were just an anomaly. Additionally, the ability to retain the trained positions appears to be influenced by the participants' previous trombone experience. The interpretation of the slide-position results is also limited by inaccuracies in the implemented theoretical trombone model, which affect the accuracy of the provided feedback and evaluation of the results.

Overall the results show that the prototype can provide usable and effective real-time assistance for trombone slide positioning and precision. The positive usability and motivational evaluations, together with the significant improvement during training, indicate that the XR prototype can provide a promising environment outside of conventional instrumental lessons. However, this study could not prove that this assistance can lead to lasting improvements in slide positioning or precision memorization. The inaccuracies of the theoretical trombone model may affect the accuracy evaluation results and the conclusions that can be drawn from it.

7.1 Limitations

The results of this study should be interpreted in regard to several limitations. These limitations generally concern the size and composition of the participant sample, the limited duration of the user studies training phase, decisions made for the study, as well as technical aspects of the implemented prototype, which I will address in the following text.

The user-study only had 17 participants. This small group had to be further divided into experimental (EG) and control group (CG), making them even smaller. Because of that it limits the statistical power of the study and universal applicability of its results. So may some small effects have not been detected, subgroups have not been properly reflected, or new samples that actually show an effect is smaller than this results makes it seem.

Furthermore, the training time of 10 minutes was not properly suited for an evaluation concerning the effect of it. A proper evaluation would need to evaluate the progress during a week or more using the prototype. As such this study can only serve as a first indicator and certain results may change when regarded on a large scale.

There are also decisions made for the user-study that in afterthought limiting its results. It was decided that the EG and CG would get a conventional help device in form of a tuner. This device was not removed during the end evaluation (perform phase) and as the EG paid less intention to the tuner, the CG may have gained an advantage. This might have lessened the effect that the advanced position help relative to not having it and as such in the evaluation.

Another big limitation is the inaccuracy of the theoretical trombone model as described in Section 4.3.4, which has impacted the correctness in the precise position of a slide position. This therefore affected the feedback of the position help, as well as the evaluation. An improved automatically adapting model addressing these inaccuracies has been suggested in Section 4.3.4 but was not yet implemented. The user study therefor only evaluates on the original model, and it is unclear to which degree this new model might influence the results.

A further limitation concerns potential tracking errors in the Lighthouse Tracking system. Those errors will directly lead wrong samples within the record data of a participant. Since the study did not include a validated method for detecting and correcting or removing such samples, their influence could not be quantified. However, determining their effect would require additional criteria and variables to identify tracking errors, which is beyond the scope of this thesis.

One element to mention are the chosen arbitrary cutoff thresholds regarding from when a position is too far off to consider it for accuracy ($10cm$ in H3) and from when a position is considered as the correct slide position ($5cm$ in H4). Additional exploratory tests have been made with different logical values to secure that it is not only those values coming to this conclusion.

The accuracy of the slide position is further limited by the relationship between slide and pitch is not equivalent for every player, instruments and environmental conditions. The pitch produced by the instrument at a specific time can be influenced by factors like the players' embouchure, the specific construction and tuning of the instrument, and conditions like temperature. Because of these factors the slide position for a target pitch/note can vary between participants by multiple centimeters. (See Table 4.1)

This together with the errors of the model become very important to the $10cm$ and $5cm$ thresholds chosen in the evaluation. As the distance between the theoretical model and the true position for that participant increases so does the percentage of wrongly filtered/evaluated values. For example positions which would have been correct can be labeled as wrong and conversely a wrong distance can still be counted as correct.

These variating factors introduce uncertainty for the evaluated slide position/accuracy and will have influenced the results of **H3** and **H4**. The extent of that influence could not be quantified within the extent of the study.

7.2 Future work

The results of this thesis point towards multiple directions for future work. An immediate priority should be to improve the technical accuracy of the prototype followed by a new more extensive evaluation of its potential to support learning improvements.

One immediate focus should be the implementation of a more accurate theoretical model that accounts for the specific embouchure and environmental conditions. This could be achieved by approximating the model function to the frequencies recorded from the instrument, as described in Section 4.3.4. However, for such an approach to be effective the data pollution problem in the recorded audio would need to be addressed. By, for example improving the location of the microphone in a better location to reduce interference from external sounds, this problem could be improved.

Further improvements can be made by creating a kind of system that fades out the

position help with longer use, so that the players need to memorize the positions.

Concerning an evaluation that can securely evaluate the capabilities of the program, it would need to be used for at least a week to be able to observe and analyze a normal learning progress. The length of a week would be fitting, as this is a typical distance between instrument lessons. Further improvements could be made to my user-study design discussed in Section 6.3.

A topic not considered in my thesis is affordability of the system. This becomes a real concern, at least when a significant persistent learning advantage is observed in a longer study. While Lighthouse Tracking is a good option for prototyping it is relatively expensive solution. Different tracking systems like camera- or magnet tracking combined with a laser distance sensor for the slide (See Section 3.4), could therefore be investigated. However, these systems most likely will require custom hardware. Furthermore, these systems may prove more reliable in concerning accurate tracking.

Additionally, Lighthouse Tracking may not be the best option due to potential tracking errors. The data collected in this study could further be analyzed to investigate the occurrence of such errors. One approach would be to analyze local slide tracker data for deviations from the average axis and change of it over time. Another approach would be finding moments where the slide position significantly changed, while there was no significant pitch change. Both have limitations: the first one may fail if an error causes the position to deviate along the correct axis, while the second one may fail in situations where the player did not produce any pitch. While these limitations are not unsolvable, addressing them would require defining too many additional variables, making a proper analysis beyond the scope for this study.

I already mentioned some systems to improve the user experience and feature catalogue concerning study relevant improvements. As the usability score evaluated during the study has room for improvement, there are further improvements to be made for a better usability.

Generally the prototype lacks a tutorial on how it works. This becomes very important for the trombone clicking System (See Section 4.3.3) and the calibration (See Section 4.2), as these systems, might not be that intuitive. The calibration process in general needs to be made user-friendly, as it currently requires specific menu input one needs to know beforehand and should be replaced by a guiding system.

One idea that could be experimented with is layering the traditional staff system and the slide position system on the front UI on top of each other. I mentioned in the beginning, that it is important to not remove the staff system, as this is the system a trombonist usually has to work with outside this prototype. Layering the systems would be unifying them, and it would be interesting, how such changes can impact readability and the association between sheet note and the shown position.

An additional quality of life upgrade might be to introduce physical floor buttons for pausing and skipping forward and backwards. Due to the positive SUS score I am convinced that the trombone UI input system is usable non-interruptive system, navigating the UI while playing a track gets cumbersome and having to push a simple button with your foot is way easier.

Bibliography

[1] Unity. "Augmented virtuality (av)," Accessed: Sep. 17, 2026. [Online]. Available: https://unity.com/glossary/augmented-virtuality

[2] P. Milgram, D. Drascic, J. Grodski, A. Restogi, S. Zhai, and C. Zhou, "Merging real and virtual worlds," *Proceedings of IMAGINA'95*, Jan. 1995.

[3] Yamaha Corporation. "How a trombone is made: The slide must be straight-no exceptions!" Accessed: Sep. 17, 2026. [Online]. Available: https://www.yamaha.com/en/musical_instrument_guide/trombone/manufacturing/

[4] M. C. LoPresto, "A theoretical trombone," *Physics Education*, vol. 49, pp. 580–585, Sep. 2014. DOI: 10.1088/0031-9120/49/5/580

[5] A. Larson. "Trombone lesson: Trombone slide chart," DigitalTrombone, Accessed: May 31, 2026. [Online]. Available: https://www.digitaltrombone.com/trombone-slide-chart.html

[6] S. Walling, *Technical Studies for the Slide Trombone*. Melbourne, Australia: Allan and Co. Pty. Ltd.

[7] Vienna Symphonic Library. "Bass trombone," Accessed: May 21, 2026. [Online]. Available: https://www.vsl.co.at/academy/brass/bass-trombone

[8] M. Akiyama and M. Makino, "A vr system of slide position self-learning with controllers position detection and tactile feedback for trombone beginners," in *2024 21st International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2024, pp. 1–4. DOI: 10.1109/ECTI-CON60892.2024.10594849

[9] Flat2VR Studios and Holy Wow Studios LLC, *Trombone champ: Unflattened*, Nov. 2024. Accessed: May 16, 2026. [Online]. Available: https://store.steampowered.com/app/3151670/Trombone_Champ_Unflattened

[10] C. Caruso, *Carmine Caruso - Musical Calisthenics for Brass*. Milwaukee, WI: Hal Leonard Corporation, 2002.

[11] Vanee. "Master trombone slide positions – start learning today! [blog post]," Accessed: May 21, 2026. [Online]. Available: https://tromboneLessons.net/master-trombone-slide-positions-start-learning-today

[12] W. Heinitz, *Instrumentenkunde*, W. Heinitz, R. Lachmann, C. Sachs, and P. P., Eds. Wildpark-Potsdam: Akademische Verlagsgesellschaft Athenaion, 1928, Originally published in parts.

[13] D. Wick, *TROMBONE TECHNIQUE*, 2011 Revised Edition. 2011.

[14] Old Dominion University. "Slide tips," Accessed: May 21, 2026. [Online]. Available: https://fs.wp.odu.edu/jhall/performance-clinic/slide-tips/

[15] T. Lautzenheiser, J. Higgins, C. Menghini, P. Lavender, T. C. Rhodes, and D. Bierschenk, *Essential Elements for Band - Trombone Book 1 with Eei (Book/Online Audio)*. Hal Leonard Publishing Corporation, 1999.

[16] N. Piegdon, *Synthesia*, Piano training and MIDI visualization software, Synthesia LLC, 2006. Accessed: May 16, 2026. [Online]. Available: https://synthesiagame.com/

[17] D. Tschernobai, "Jazzhandsxr: A mixed reality application for visualising hand structure in piano playing," Bachelor Thesis, Universität Bremen, Feb. 2025. Accessed: May 16, 2026. [Online]. Available: https://cgvr.cs.uni-bremen.de/theses/finishedtheses/jazz_hands_xr/thesis.pdf

[18] Holy Wow Studios LLC, *Trombone champ*, Sep. 2022. Accessed: Jun. 11, 2026. [Online]. Available: https://store.steampowered.com/app/1059990/Trombone_Champ

[19] N. Farwell, "Adapting the trombone: A suite of electro-acoustic interventions for the piece," in *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME 2006)*, Zenodo, 2006, pp. 358–363. DOI: 10.5281/zenodo.1176895

[20] S. Lemouton, M. Stroppa, and B. Sluchin, "Using the augmented trombone in 'I Will Not Kiss Your F.ing Flag'," in *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME 2006)*, Zenodo, 2006, pp. 304–307. DOI: 10.5281/zenodo.1176963

[21] D. Campbell, A. López-c, and J. Kemp, "The effect of fast trombone slide glissandi on the mechanics of artificial lips," in *Proceedings of the 32rd International Conference on Acoustics*, Sep. 2019. Accessed: Jun. 11, 2026. [Online]. Available: https://pub.dega-akustik.de/ICA2019/data/articles/001333.pdf

[22] Baumer Electric AG, *Oadm 20i4471/s14c laser distance sensor – datasheet*. Accessed: Jun. 11, 2026. [Online]. Available: https://media.baumer.com/OADM_20I4471_S14C_web_DE.pdf

[23] Baumer Electric AG, *Laser distance sensors – product overview*. Accessed: Jun. 11, 2026. [Online]. Available: https://www.baumer.com/de/en/product-overview/distance-measurement/laser-distance-sensors/c/289

[24] M. Meister, A. B. Carman, and C. Li, "Using fmcw radar and range-gated clutter suppression for trombone slide position detection," in *2025 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNeT)*, 2025, pp. 4–7. DOI: 10.1109/WiSNeT63956.2025.10905011

[25] R. W. Stroetz et al., "Device for indicating trombone slide positions and related systems and methods," US10347230B2, Jun. 2019. Accessed: Jun. 13, 2026. [Online]. Available: https://patents.google.com/patent/US10347230B2/en

[26] Valve Corporation. "Base stations," Valve Corporation, Accessed: Sep. 12, 2026. [Online]. Available: https://www.valvesoftware.com/en/index/base-stations

[27] A. Taffanel, B. Rousselot, J. Danielsson, K. McGuire, K. Richardsson, M. Eliasson, T. Antonsson, and W. Hönig, "Lighthouse positioning system: Dataset, accuracy, and precision for UAV research," *CoRR*, vol. abs/2104.11523, 2021. arXiv: 2104.11523. [Online]. Available: https://arxiv.org/abs/2104.11523

[28] Valve Corporation. "Steamvr tracking licensing program," Accessed: Jun. 20, 2026. [Online]. Available: https://partner.steamgames.com/vrlicensing

[29] Meta. "Learn about hand and body tracking on meta quest," Accessed: May 23, 2026. [Online]. Available: https://www.meta.com/help/quest/290147772643252/

[30] S. Dong and H. Wang, "A compact and robust 6-dof pose tracking system using magnetic-inertial sensors and a single uniaxial electromagnetic coil," *IEEE Sensors Journal*, vol. 24, no. 4, pp. 5062–5072, 2024. DOI: 10.1109/JSEN.2023.3348243

[31] Polhemus. "Our technology," Accessed: Sep. 13, 2026. [Online]. Available: https://polhemus.com/technology

[32] Thomann. "5. materials," Thomann Music, Accessed: Jun. 21, 2026. [Online]. Available: https://www.thomannmusic.com/onlineexpert_page_trombones_materials.html

[33] Yamaha Corporation. "Choosing a trombone: Choosing by material," Accessed: Jun. 21, 2026. [Online]. Available: https://www.yamaha.com/en/musical_instrument_guide/trombone/selection/

[34] M. Li, T. Bien, and G. Rose, "Construction of a conductive distortion reduced electromagnetic tracking system for computer assisted image-guided interventions," *Medical Engineering & Physics*, vol. 36, no. 11, pp. 1496–1501, Nov. 2014. DOI: 10.1016/j.medengphy.2014.07.018

[35] König & Meyer. "14985 posaunenständer," Accessed: Sep. 13, 2026. [Online]. Available: https://www.k-m.de/produkte/staender-und-zubehoer-fuer-instrumente/blasinstrumente/blechblasinstrumente/posaune/14985-posaunenstaender-schwarz

[36] FluxPose. "Tech specs," Accessed: Sep. 13, 2026. [Online]. Available: https://www.fluxpose.com/specs

[37] P. Amann, V. Chatziioannou, P. Nicolay, and R. Willmann, "On the capabilities of a 'digital mouthpiece' for real-time recording of embouchure parameters for brass musicians," in *10th Convention of the European Acoustics Association*, Sep. 2023, pp. 399–403. DOI: 10.61782/fa.2023.0840

[38] thep3000. "Reply to 'OpenXR and OpenVR together?'" Unity Discussions forum post, Accessed: Jul. 2, 2026. [Online]. Available: https://discussions.unity.com/t/openxr-and-openvr-together/841675/21

[39] M. Seifert, *Answer to "Does a trumpet operate using an open air column or a closed air column"*, Physics Stack Exchange, Nov. 2016. Accessed: Jul. 28, 2026. [Online]. Available: https://physics.stackexchange.com/q/292724

[40] M. Oldenkamp and J. Kastelein, *Hören, Lesen & Spielen 2, Schule für Posaune in C (BC)*. De Haske Publications, 2000, Track 9.

Appendix A

Lists

List of Figures

2.1 Trombone Positions for specific notes across partials. 6

4.1 UI and Music Track Representations 19
4.2 Visual Representation with white Particles of the Goal Position on the Trombone Slide 20
4.3 Simplified Code Structure Diagram of the System 21
4.4 Tracker locations on the physical and digital trombone 24
4.5 Trombone as UI Input Device 26
4.6 Slide position reference sheet for the CG 41

5.1 Age and Years of Experience of Participants 44
5.2 VR Experience of Participants 44
5.3 System Usability Scale Score Results 45
5.4 System Usability Scale Answer Distribution 46
5.5 Trombonists Perceived Potential to Increase their Motivation 48
5.6 Motivation Score in Relation to Age, Trombone Experience and VR Experience 49
5.7 Accuracy of the Individual Participants 51
5.8 Change in Accuracy from Training to Perform Phase of the Individual Participants. Negative values indicate improvements from the Training Phase. 52
5.9 Distributions with different Cutoff Points as described in Section 5.3.1 Step 4. No cutoff is left out in this Depiction, as it would make the Figure unreadable. 53
5.10 Scatter plot for the participant values in relation. Experience years randomly altered by up to ∓ 0.5 years for better readability. 54
5.11 Distributions with different Distances at which a Slide Position is wrong. . 58
5.12 Distributions with different Cent Radius Filters. 59

List of Tables

4.1 Trombone ↔ Model Comparison 31

5.1 Mean Distance from the Target Slide Position for the Different Participant Groups 51
5.2 Hypothesis Test with different Cutoff Points as described in Section 5.3.1 Step 4. 53
5.3 Results of the Spearman's rank correlation with a permutation test for the association between trombone experience and distance to the target position 55
5.4 Hypothesis 4 Test with different Distances at which a slide is wrong 57

C.1 Results for the SUS Questions 86
C.2 Learning Speed 86
C.3 Rating Questions with Likert Scale 87
C.4 Rating Questions with Negative Likert Scale 87

E.1 AI Use in this Bachelor Thesis 91

Appendix B

"A Theoretical Trombone" Pseudocode

Based on Algorithm in "A theoretical trombone"[4].

Algorithm 1 Calculate the Playing Frequency

Require: Tube length L, end correction x_0, flare parameter γ, mouthpiece volume V, mouthpipe length L_m, mouthpipe diameter S, tube radius r, harmonic number n, tolerance ε speed of sound c

$\beta \leftarrow 0.637$

$f \leftarrow \dfrac{c}{4(L+x_0)}\left((2n-1)+\beta\sqrt{\gamma(\gamma+1)}\right)$

$L_e \leftarrow L_m + 0.8\sqrt{S}$

$f_p \leftarrow \dfrac{1}{2\pi}\sqrt{\dfrac{c^2 S}{L_e V}}$

$L_0 \leftarrow \dfrac{V}{\pi r^2}$

$L_1 \leftarrow \dfrac{c}{4f_p}$

repeat

  $L_{\text{corr}} \leftarrow \dfrac{c}{2\pi f}\tan^{-1}\left(\dfrac{(2\pi f/c)L_0}{1-(4fL_1/c)^2}\right)$

  $L_{\text{total}} \leftarrow L + L_{\text{corr}}$

  $f_{\text{new}} \leftarrow \dfrac{c}{4(L_{\text{total}}+x_0)}\left((2n-1)+\beta\sqrt{\gamma(\gamma+1)}\right)$

  $\Delta \leftarrow |f_{\text{new}} - f|$

  $f \leftarrow f_{\text{new}}$

until $\Delta < \varepsilon$

return f

Algorithm 2 Calculate the Instrument Length (Algorithm 1 reversed)

Require: frequency f, end correction x_0, flare parameter γ, mouthpiece volume V, mouthpipe length L_m, mouthpipe diameter S, tube radius r, harmonic number n, tolerance ε speed of sound c

$\beta \leftarrow 0.637$

$L_{\text{effective}} \leftarrow \dfrac{c}{4f}\left((2n-1)+\beta\sqrt{\gamma(\gamma+1)}\right)$

$L_{\text{total}} \leftarrow L_{\text{effective}} - x_0$

$L_e \leftarrow L_m + 0.8\sqrt{S}$

$f_p \leftarrow \dfrac{1}{2\pi}\sqrt{\dfrac{c^2 S}{L_e V}}$

$L_0 \leftarrow \dfrac{V}{\pi r^2}$

$L_1 \leftarrow \dfrac{c}{4f_p}$

$L_{\text{corr}} \leftarrow \dfrac{c}{2\pi f}\tan^{-1}\left(\dfrac{(2\pi f/c)L_0}{1-(4fL_1/c)^2}\right)$

$L \leftarrow L_{\text{total}} - L_{\text{corr}}$

return L

Appendix C

User-Study Questionnaire

C.1 Part 1

C.1.1 Demographic Information

Age (years):

C.1.2 Musical Experience

Experience with brass instruments (years):

Experience with trombone (years):

Playable pitch range:

The participant may select the lowest and highest playable notes.

When did you practice last?

- No information
- Today
- 1–2 days ago
- 3–7 days ago
- 1–2 weeks ago
- More than one month ago

Experience with virtual reality:

- No information
- No experience
- Tried it before
- Occasional use
- Regular use
- Frequent use

C.2 Part 2

Unless stated otherwise, the following questions use a five-point Likert scale.

1 = Strongly disagree 5 = Strongly agree

C.2.1 Usability

Please rate the following statements:

1. I can imagine using the system regularly.
2. I found the system unnecessarily complex.
3. I found the system easy to use.
4. I think I would need technical support to use the system.
5. I found the different functions of the system were well integrated.
6. I found there was too much inconsistency in the system.
7. I imagine that most people would learn to use the system quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn many things before I could use the system.

C.2.2 Playing Experience

How many problems did you have with ...

1. The mouthpiece.
2. The embouchure / lip tension.

Scale:

1 = No problems 5 = Many problems

How prepared did you feel before the performance phase?
Scale:

1 = Very confident 5 = Very uncertain

C.2.3 Program Evaluation

How difficult was it to...
Scale:

1 = Very easy 5 = Very difficult

1. Play the piece.
2. Hit the correct slide positions during the training phase.
3. Hit the correct slide positions during the performance phase.
4. Play the correct notes during the training phase.
5. Play the correct notes during the performance phase.

How clear was it which notes had to be played and when?
Scale:

1 = Very clear 5 = Very unclear

What was unclear about the notes?

How synchronized did the digital and physical trombone feel?
(Overall position, slide position, etc.)
Scale:

1 = Perfect synchronization 5 = No synchronization

How large was the effect of the program on your playing?
Positive effect:
Scale:

1 = No change 5 = Very large effect

Negative effect:
Scale:

1 = No change 5 = Very large effect

What was the effect of the program on your learning speed?
Scale:

1 = Strongly accelerated 7 = Strongly slowed down

Do you think a program of this type would...
Scale:

1 = No / not at all 5 = Yes / very much

1. Make learning slide positions easier?
2. Make learning variations of slide positions on different natural tones easier?
3. Increase your motivation to practice?

Did you have fun?
Scale:

1 = No 5 = Yes

Additional comments / feedback:

C.3 Results

	Likert Scale				
	1	2	3	4	5
I can imagine using the system regularly.	0/0	0/1	3/2	3/3	3/2
found the system unnecessarily complex.	7/7	2/0	0/1	0/0	0/0
I found the system easy to use.	0/0	0/0	0/3	4/2	5/3
I think I would need technical support to use the system.	5/1	1/3	3/2	0/2	0/0
I found the different functions of the system were well integrated.	0/0	1/1	0/0	5/1	3/6
I found there was too much inconsistency in the system	5/6	4/1	0/1	0/0	0/0
I imagine that most people would learn to use the system quickly.	0/0	0/0	1/3	5/2	3/3
I found the system very cumbersome to use.	7/2	1/5	1/1	0/0	0/0
I felt very confident using the system.	0/0	1/0	0/1	3/4	5/3
I needed to learn many things before I could use the system.	4/4	4/3	1/1	0/0	0/0

Answers from EG / CG

Table C.1: Results for the SUS Questions

	accelerated ↔ decelerated						
What was the effect of the program on your learning speed?	1/0	4/3	0/3	3/2	1/0	0/0	0/0

Answers from EG / CG

Table C.2: Learning Speed

	Likert Scale				
	1	2	3	4	5
How large was the positive effect of the program on your playing?	0/0	2/3	4/1	3/3	0/1
How large was the negative effect of the program on your playing?	6/8	3/0	0/0	0/0	0/0
Make learning slide positions easier?	0/0	0/1	0/0	6/3	3/4
Make learning variations of slide positions on different natural tones easier?	0/0	0/1	1/2	4/4	4/1
Increase your motivation to practice?	0/0	1/0	2/3	4/2	2/3
Did you have fun?	0/0	0/0	0/0	2/1	7/7

Answers from EG / CG

Table C.3: Rating Questions with Likert Scale

	Negative Likert Scale				
	5	4	3	2	1
How many problems did you have with the mouthpiece?	6/6	2/0	1/2	0/0	0/0
How many problems did you have with the embouchure?	5/4	0/0	2/3	2/0	0/1
How prepared did you feel before the performance phase?	2/1	4/1	2/3	1/2	0/1
How difficult was it to play the piece?	2/2	5/1	2/4	0/0	0/1
How difficult was it to hit the correct slide positions during the training phase?	2/0	4/1	3/3	0/3	0/1
How difficult was it to hit the correct slide positions during the perform phase?	2/1	2/2	1/1	3/3	1/1
How difficult was it to play the correct notes during the training phase?	1/1	7/3	0/2	0/1	1/1
How difficult was it to play the correct notes during the perform phase?	2/1	2/2	2/4	2/0	1/1
How clear was it which notes had to be played and when?	6/5	3/1	0/2	0/0	0/0
How synchronized did the digital and physical trombone feel?	1/3	5/3	3/1	0/1	0/0

Answers from EG / CG

Table C.4: Rating Questions with Negative Likert Scale

Appendix D

Scripts

D.1 H3 Accuracy Calculation

Listing D.1: Accuracy Calculation for H3 in Python

```
import numpy as np

def calculate_accuracy_distance(
    samples:NPArray[np.object_],
    slide_distance_max:float,
    slide_difference_max:float
):
    notes_means = []
    midi = np.asarray(samples["record_goal_midi"],
                      dtype=float
                     )
    valid = ~np.isnan(midi)

    samples = samples[valid]
    midi = midi[valid]
    change_points = np.flatnonzero(midi[1:] != midi[:-1]) + 1

    data_notes = np.split(samples, change_points)

    for note in data_notes:
        midi_value = note[0]["record_goal_midi"]
        slide_distance = note[0]["record_goal_slideDistance"]
```

```
23
24          if np.isnan(midi_value) or slide_distance >
               slide_distance_max:
25              continue
26
27          values = [
28              diff
29              for d in note
30              if (diff := abs(
31                  d["record_goal_slideDistance"]
32                  - d["record_trombone_slideDistance"]
33              )) < slide_difference_max
34          ]
35
36          if values and len(values):
37              n = len(values)
38              x = np.linspace(-1, 1, n)
39
40              sigma = 0.5
41              weights = np.exp(-(x / sigma) ** 2)
42
43              note_mean = np.average(values, weights=weights)
44              notes_means.append(note_mean)
45      return np.mean(notes_means)
```

Appendix E

AI Use

Tool	Topic	Use Description	Reason
Consensus	Literature	Overview of the relevant literature	Find literature I may have missed on the topic
ChatGPT	Reflection	Presenting it an idea as is. If my reaction to the ChatGPT's reframing the idea sounds bad to me, I need to overthink the idea/approach.	ChatGPT frequently reformulates such ideas positively (when not directly asked otherwise). Even further, I dislike the style it provides this reframing in. This puts me already in a state where I am more critical. If this can not make my idea sounds bad to me, the idea seems to be a good approach.
ChatGPT	Tests	Discussing which evaluation methods could be appropriate for the data.	Inexperienced Topic. ChatGPT helped to identify possible approaches, that could be evaluated by me.

ChatGPT	Feedback	Presenting it a section of my bachelor thesis asking for errors (especially gaps in reasoning) and awkward/hard to read wording. The suggestions were used to identify sections I need to rework. While not copying any sentences it provided, sometimes small sentence components (especially sentence starters) were incorporated into the text. This was only done, if I at that point was so uptight, that I could not think of an alternative to my own wording. I also did not use any small sentence components, I would not have used myself.	To identify issues I overlook while proofreading especially on the same day. At the same time I don't want something in my thesis, that is not my work. But the text needs to be readable, and I sometimes get headaches from being too uptight, so while this is a solution I don't like, it's the better choice among the options.
ChatGPT	Code Snippets	Find a feature of a language or a small specific code snippet using natural language.	I am a lot less familiar with Python and LaTeX than I am with C#. When I fail to formulate an effective search query to find what I am looking for on *stack overflow* or similar, LLMs usually can point me directly at what I tried to find.

Table E.1: AI Use in this Bachelor Thesis