CHARACTERIZING MIDAIR HANDWRITING IN VIRTUAL REALITY

by

MATTHEW CHAN

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Master of Science
in
Computer Science
Faculty of Science
University of Ontario Institute of Technology (Ontario Tech University)
Oshawa, Ontario, Canada
April, 2024

1

THESIS EXAMINATION INFORMATION

An oral defense of this thesis took place on March 11, 2024 in front of the following examining committee:

CHAIR OF EXAMINING COMMITTEE:
Dr. Li Yang

SUPERVISOR:
Dr. Christopher Collins

CO-SUPERVISOR
Dr. Faisal Qureshi

EXAMINING COMMITTEE MEMBER
Dr. Alvaro Quevedo

THESIS EXAMINER
Dr. Ali Neshati

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

2

ABSTRACT

Midair handwriting poses challenges due to the lack of a physical plane to press against while writing, making it difficult to determine when ink should be placed. In this thesis, we gathered midair handwriting data from 24 participants in an environment that allowed them to write freely. We compared writing with a pen-like object and writing using a finger across two writing methods (writing freely versus on a virtual whiteboard). Using our data, we trained a neural network to detect when ink should be placed during midair handwriting, achieving an overall 85% accuracy. We developed a data-viewing application to recreate sentences for visual analysis. Participant feedback favoured the pen-like object as a writing utensil, with equal preference for both writing methods.

Our contributions include a midair handwriting Virtual Reality (VR) application for data collection, a dataset containing 480 sentences of frame-by-frame midair handwriting data, and 20 unique prompts used in participant trials.

Keywords: midair handwriting, virtual reality, machine learning, visualization, user preferences

3

DECLARATION

I, Matthew K. Chan, herby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

The research work in this thesis that was performed in compliance with the regulations of Research Ethics Board under REB Certificate file #17222.

Oshawa, Ontario, Canada, April, 2024

Matthew Chan

4

STATEMENT OF CONTRIBUTIONS

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication. I have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others. Furthermore, I hereby certify that I am the sole source of the creative works and/or inventive knowledge described in this thesis.

5 ACKNOWLEDGEMENTS

My biggest acknowledgement goes to Dr. Christopher Collins for providing me with the opportunity to pursue this master's degree. This project has opened doors to the biggest challenges I have encountered in my life and resulted in a learning experience that I'll carry on forever. His guidance has been crucial in putting me on the right track in order to reach my objectives. Secondly is my co-supervisor, Dr. Faisal Qureshi, who also assisted me by providing guidance in solving problems outside of my knowledge. Without both professors, I would not be where I am today.

Additionally, I acknowledge Julia Micallef and Artur Arutiunian from maxSIMhealth lab for helping me in designing and 3D printing the trackers that were essential for my data gathering application to work.

I would also like to acknowledge Professor Mariana Shimabukuro for helping me throughout my thesis, as I would always turn to her for advice, and she always provided the extra guidance I needed.

I must also thank my friends as well as the new friends I made over these years, as the experiences and memories we shared were what kept me moving forward. They were there for me when times got rough and I can't thank them enough for the advice they have given me, the laughs we had and everything they have done.

My thanks also goes out to my parents who provided me hospitality and care. I am grateful to have been able to come home on countless late nights to a warm dinner plate.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

HCI Human Computer Interaction

VR Virtual Reality

AR Augmented Reality

HUD Heads Up Display

SVD Singular Value Decomposition

FPS Frames per second

LSTM Long Short-Term Memory

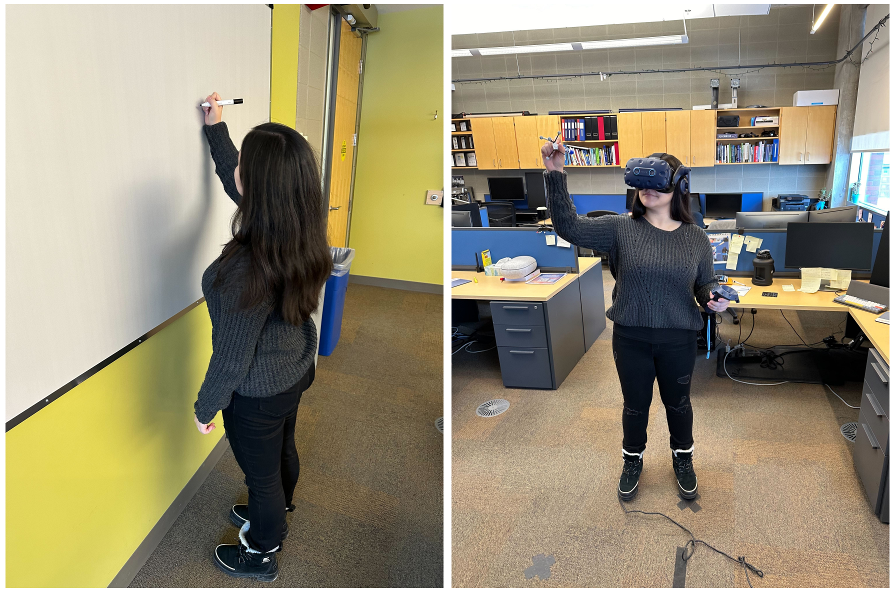
EOS End-Of-Sequence

ROC-AUC Receiver Operating Characteristic Area Under the Curve

6 INTRODUCTION

In recent years, digital writing has been a ubiquitous practice, significantly contributing to enhancing workplace productivity. It has revolutionized the way we can perform writing, conveying our ideas in a readable form for self-interest or sharing ideas with an audience. While the reliance on digital platforms has grown exponentially in common everyday workplaces, the importance of efficient text input has grown alongside. For example, a teacher may want to quickly write down notes for their students while giving a lecture, or a group of people could write their thoughts down while brainstorming. Writing things down swiftly and conveniently is key in various workplaces, ranging from individual to collaborative settings. The practice of digital writing in the modern day resembles traditional writing in the same way that any writer would have to approach a digital surface to write. While this method is straightforward, there lies much room for improvement, such as removing the need of a physical writing plane which can allow people to write anywhere they wish. This can be beneficial for people who require fast and efficient writing, such as writing short notes at a desk or in a collaborative setting, or writing things down for an audience. Therefore, a promising solution that addresses this gap in writing technology is the ability to write in midair, as it is a versatile approach to writing down ideas and thoughts in a digital environment.

Writing in the midair using a digital medium can be advantageous in numerous settings, such as working in augmented reality (AR) work environment where multiple users can see the same handwriting as well as sketches, public displays and sterile conditions [36]. Alongside is the benefit of ease of transfer from traditional writing skills. A writer capable of writing proficiently on various mediums, including writing on paper, chalkboard, or whiteboard, also implies that they have the necessary skills to write on a digital interface. Another benefit of midair writing includes the limitless space due to the virtual environment. Regarding spatial capacity, the space a user can write in is bound by the 3D space around the user, which can be con-



(a) (b)

Figure 1: Left: Person writing on a physical whiteboard. Right: Person writing in midair without constraints on where to write and how to write.

stantly expanded. Moreover, a digital text input interface can equip a user with creative freedom with various features, such as the ability to manipulate ink colours, font sizes and other formatting options. Additionally, writing in an alternate reality such as AR or virtual reality (VR) allows the possibility of providing visual feedback of the ink being placed in the air after writing.

To make midair handwriting optimal for users, it must be a fluid and seamless experience. This means the ink must be placed whenever the user intends to do so, and should not be placed whenever the user does not intend to. Currently, midair handwriting is a relatively new technology. A commercially available method of midair handwriting is the Meta Horizon Workrooms ' [37] virtual blackboard. This method of midair handwriting requires the user to be close to the virtual blackboard, which adds extra steps for the user to perform a simple task. For example, the user has to physically move to the blackboard or manually select buttons to move their avatar to the blackboard. When reaching the blackboard, the writer must move their writing utensil close to the virtual surface for the ink to be placed. This ultimately reflects a real-life scenario of a writer walking up to a chalkboard and bringing the piece of chalk close to the board to begin writing. Compared to the type of midair handwriting we propose, we intend to allow midair handwriting without the need of bringing the writing utensil close to any surface - virtual or physical.

The design goal we have in mind is to remove the steps needed to begin writing - simplifying midair handwriting by allowing the writing to begin as soon as the writing utensil is picked up. Additionally, our research aims to avoid using explicit methods of switching between writing modes, such as button presses. This requires detecting the intention of midair handwriting, as it is also present in past research, using various methods to track the writer's hand movements to classify the letters being formed [4, 5, 11, 12]. However, many constraints on how writing must be performed are present in these works; some include only being able to write capital letters, having a constrained writing area and a required stroke order for different letters. While these constraints can provide advantages to midair writing, such as the limited writing area allowing users to write in a smaller private space, these constraints can take away from an individual's traditional method of writing. With minimal tracking technology, we cannot write in midair without constraints. Therefore, we propose to investigate the feasibility of autonomous midair handwriting and the required tracking to do so.

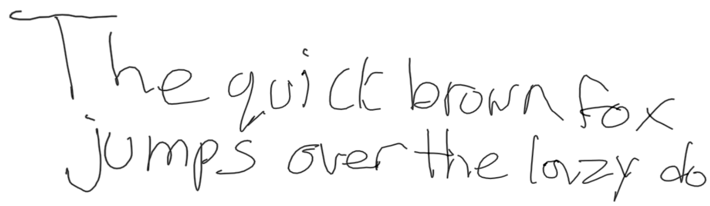
The long-term goal of this research is to allow the user to write anywhere freely by automatically placing the ink whenever the user intends to write without explicit commands such as button presses, given that they are in a setting with the appropriate tracking technology. Hypothetically, one can raise one's hand and create strokes while holding a writing utensil or using one's finger. Then, they can proceed to perform the writing by making the strokes in midair. To do this, we must be able to detect whether or not the user intends to write in the open 3D space. This is a challenge many research papers have addressed and have solved by using a variety of AI models, but alongside, it has been applied with many constraints on how writing must be performed. These constraints are elaborated further in Chapter 7. However, many share a similar methodology of using the acceleration and angular velocities of the writer's hand and/or finger as features for their models as time series data to determine the intent of writing.

In this thesis, we have adopted a similar strategy of using the movement and rotations of trackable objects. These objects include the writer's writing utensil, wrist and head. Additionally, we are interested in using eye tracking, which includes the gaze position and pupil dilation, as these have been linked to the intent of human activities in past research [1, 21, 30, 31]. Past research implemented gaze-based tasks [14, 32] for explicit mode switching while using a pen-based system.

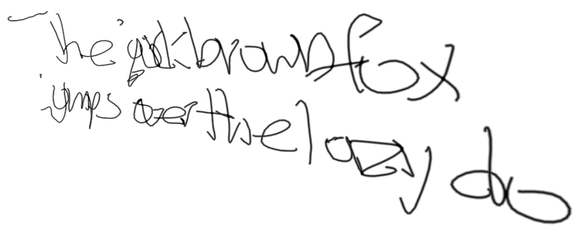
We have also taken an HCI approach as we investigate what users prefer regarding midair handwriting. We compare two different methods of writing — one where the writing utensil is a pen-like object in which we use a whiteboard marker, and one where the writer's finger can be used to write as it will act as a lightweight writing utensil. The purpose is to investigate whether people prefer to write in midair while holding onto a writing utensil or not holding onto anything. The finger is also defined as a "tool-less" technology, which can be advantageous in a variety of settings [28] such as one for midair writing for reasons such as providing ease of use and cleanliness. In addition to writing methods, we also compare writing freely in midair and writing with the assistance of a virtual whiteboard, as we intend to investigate whether a virtual writing plane can assist people in midair writing. The virtual whiteboard technique has been implemented in various ways, especially for collaborative work. For example, the researchers who developed CollaboVR [25] explored different layouts for presenting virtual writing, such as mirroring writing in front of an audience and projecting writing for remote users to view. Petrowski et al. [41] implemented a VR prototype where users can add writing and sticky notes to a virtual blackboard. Lastly, we have chosen to implement this application in VR, allowing the writer to see the digital ink as writing is performed, improving the naturalness of the handwriting experience.

6.1 MOTIVATION

When writing in midair with no constraints, there are two main challenges: firstly, the absence of a physical surface to apply pressure while writing, and secondly, the inherent untidiness of the writing. In the first challenge, the lack of a physical surface to write on makes it impossible to perform a 'pen lift' to stop the act of writing. Normally, that is the only way to distinguish between intentional and non-intentional writing. However, in midair writing, the lack of distinction between the intent of writing introduces the challenge of de-



(a) Midair writing seen from the front



(b) The same sentence viewed from the right side. A degree of curvature can be seen from this side of the same sentence

Figure 2: A sentence written in the air by a participant, viewed from the front and right side. The front view shows the handwriting normally, but there is a degree of curvature from the top-right side.

termining the precise moments when writing is intended to begin and end. In the second challenge, writing in midair will also introduce an extra degree of freedom, which invites errors within the extra dimension, the z-axis. This can result in handwriting appearing messy, as seen in figure Figure 2, where the writing appears clear when viewed from the front but appears to be performed over a curve when viewed from another angle. This phenomenon describes the challenge of performing traditional writing in a midair setting without constraints.

Detecting intentional and non-intentional writing strokes in midair is not a novel concept. Past research has used Hidden Markov Models and Support Vector Machines [4, 5, 11–13] to determine the intention of writing based on various features such as angular velocities and accelerations of the writing hand/finger. Amma et al. [5] were able to distinguish the act of writing from performing everyday activities using the sudden increase in angular velocities and accelerations. They could classify handwritten gestures as letters by determining when midair writing is being performed. Although these works achieved character and word recognition at high accuracies, there were several constraints, some recurring across many papers, which negatively impacted the naturalness and fluidity of the writing process. Some of the constraints include writing only in capital letters, writing with letters that overlap instead of left to right, specific stroke orders, fixed wrist movement, and requiring users to write with specific heights. These constraints can serve as drawbacks, as writing should ideally be devoid of any limitations.

To offer a way to perform midair writing without constraints, we intended to find the most suitable techniques for this practice. Since there is currently no data concerning this unconstrained method of midair handwriting and the ideal handwriting techniques for individuals, we aimed to fill this research gap. Therefore, we have two research questions:

RQ1. Can we predict the intention of midair handwriting without constraints on how people write?

RQ2. How do people write in midair?

- **RQ2.1** Does the presence of a virtual whiteboard affect how people write in the air?
- **RQ2.2** What writing utensil do people prefer to write in midair with, and do they prefer to write with or without a virtual whiteboard?

To answer our research questions, we gather data using a VR application in a participant study described in Chapter 8 and Chapter 9. **RQ1** is answered by performing tests on our data using a neural network in Chapter 11. The second research question is two-fold as we explore midair handwriting from an HCI perspective. **RQ2.1** investigates midair handwriting patterns and how they are affected by a virtual whiteboard in Chapter 12. **RQ2.2** investigates how people prefer to write in midair through subjective feedback in Chapter 13. In this scope of this thesis, we focused on English printed writing, excluding cursive and other languages.

6.2 CONTRIBUTIONS

We collected and provided a new dataset containing the positional and rotational values of all tracked objects, such as the writing utensil and the participant's head and wrist, from 24 participants who wrote sentences in a user study. This number was chosen to encompass all possible combinations of the order of writing utensils and methods in our experiments. Additionally, each frame includes eye-tracking data such as pupil diameter and gaze position. Using our data, we predicted the intention of midair handwriting through a series of tests by using a recurrent neural network composed of long short-term memory layers (LSTM). The long-term goal of the data and analysis is to provide a foundation that describes the requirements to perform unconstrained midair handwriting in any 3D space and not just in VR, and for people of all backgrounds.

The contributions of this work are as follows:

- A dataset containing frame-by-frame positional and rotational data of tracked devices used in our experiments, along with eye-tracking data. Additionally, we provide the stimuli used for the data-gathering experiment.
- A Unity VR application that allows midair handwriting and is used as a data-gathering tool.
- A Unity application that uses recorded data from experiments to reconstruct sentences written by participants. Tracked objects including the writing utensil, wrist, head, and off-hand controller can be viewed in the editor frame-by-frame.
- Machine learning tests using our recorded data to predict the intention of midair handwriting for all conditions using a recurrent neural network composed of long short-term memory layers
- Analysis of midair handwriting patterns.
- Qualitative analysis of preferred writing methods and conditions, introducing suggestions for design considerations for future research.

7

BACKGROUND

There have been numerous approaches to performing midair handwriting, ranging from past research to consumer products. This chapter discusses past works and the techniques that have been implemented to support midair handwriting. Additionally, Section 7.4 explores papers that have combined eye tracking with writing.

7.1 SEGMENTING AND CLASSIFYING MIDAIR HANDWRITING

A common problem many researchers have encountered is that midair handwriting is a constant stream of data, no matter how the writing utensil is tracked. If the writing utensil is in midair, it will always be constantly in one mode and, therefore, unable to switch between writing and not writing. Unlike having a surface to write on, the mode is switched easily just by performing a pen-up or pen-down motion. While there are explicit methods of switching modes, such as using buttons to turn the ink on or off, this research aims to avoid explicit methods and only implicitly switch between writing modes.

Past research commonly used computer vision methods by tracking the writer's hand and/or finger while performing midair writing and used movement as a feature to detect when writing was intended. Chen et al. [11, 12] used features including positions, velocities, orientations, accelerations and angular speeds to classify the writing motions into letters and words. Then they tested different combinations of features to evaluate the performance using Hidden Markov Models (HMM), resulting in average error rates as low as 1.9% for word detection and 0.66% for characters in their 1k-Word Vocabulary derived from Web 1T 5-gram [7]. Their HMMs were used in their second paper [12], utilizing an Leap Motion Controller [51] to perform hand tracking. This introduced an incredibly lightweight design of midair handwriting. Similarly, Amma et al. [4, 5] used average angular velocity and mean shifted acceleration to segment the midair handwriting motions apart from writing and not writing. A comprehensive overview performed by Singh et al. [48] mentioned several other methods of detecting handwriting. Such features include the coordinates of the hand, writing direction, curvature, and even shape-based features such as deviation from an ellipse [13] for gesture recognition. Huang et al. [27] utilized Microsoft Kinect's SDK [34] and its RGB camera and depth sensor to extract depth, using path coordinates to classify numbers from tracked hand motions. All of these papers share a common goal of recognizing midair handwriting by using various positional and rotational features of the hand or any object used to perform tracking, such as a motion-sensing glove [5].

(a) Leap Motion Controller attached to the HTC Vive Headset

(b) Optitrack cameras (8 total) used in our experiments

Figure 3: Two methods of tracking. (Left) is the Leap Motion Controller for performing hand tracking. (Right) is one of the Optitrack cameras.

While these past works scored very high accuracies in segmenting and classifying movements into letters and words, many constraints naturally restricted the writer from writing. Many of these share a constraint of requiring only capital letters to be written [4, 5, 11, 12, 27, 47], some others requiring the writer to perform specific stroke orders for letters, as well as having to write in large letters. This ultimately took away from the instinctive nature of writing in an individual's way. A writer should be able to write however they want, especially in their stroke order for different letters. Additional constraints include having to write the letters in the same location as opposed to traditional left-to-right when writing English and lacking visual feedback, as the writing can only be viewed on a 2D screen away from the original writing. This limitation can be double-edged, as the constrained space of writing and lack of visual feedback can be beneficial in settings where confidentiality is required. However, in our case, where we propose a tool for workplace settings, writing freely would be advantageous as usability is our higher priority. Also, users may have intentions to share their handwriting with others.

Past research that has utilized left-to-right HMM [46] apply additional constraints such as writing slowly on an average of 3 seconds per character, though also due to tech limitations. However, this project aims to approach midair handwriting in an HCI fashion, as we aim to provide the means to perform midair handwriting in its natural left-to-right form and with the ability to see the ink in 3D space. Lastly, we want to provide freedom of movement while writing, as past research that used stationary sensors required writing in a designated area, which could restrict the range of movement of a user. As a result, we opted not to use these types of sensors.

7.1.1 *Tracking Methods*

There were a variety of methods for tracking motions for writing, including using a prototype glove with inertial sensors [5], WiFi signals [18, 22] and even millimetre-wave radio [45]. However, they also share the same limitations described in Section 7.1. As our goal is to allow writers to perform midair writing in an unconstrained area, the Leap Motion Controller seemed promising. Only the Leap Motion Controller 1 was available during this research. While some research has used the Leap Motion to perform hand-tracking, Niechwiej et al. [39] has shown that the tracking will begin to lose accuracy when performing hand tracking past 225mm in front of the sensor. At around 300 mm away from the controller, tracking can deviate at around 6 mm. This has been experimented with during the development of this project, and the deviation was observed, as our setup is shown in Figure 3a. According to Gordon et al. [20]'s anthropometric survey, the mean forearm-to-hand length for males is 48.4 cm and 44.3 cm for females. Therefore, the Leap Motion Controller is unreliable as a tracking device for our type of midair handwriting, as the writer is most likely to extend their arm past the point where deviation becomes noticeable. As a result, we resorted to tracking using Optitrack with one of the cameras shown in Figure 3b. The implementation of performing the tracking is described in Section 8.1.1.

7.2 WRITING METHODS

Kern et al. [33] conducted a study comparing multiple writing methods for writing and sketching on virtual surfaces. They investigated different ways to hold an Oculus Quest controller to use as a writing utensil as it can be easily flipped around and held similarly to a pen. We initially adopted this method for the HTC Vive controller. However, the controller is significantly longer than the Oculus Quest remote, and holding it like a writing utensil shifts all the weight toward the back, causing arm fatigue from participants during pilot studies. As a result, we adopted another similar method in their study, which is using a stylus. We used a whiteboard marker so it would closely resemble a traditional writing utensil. To perform tracking, we used an Optitrack tracker taped on the tip for improved weight distribution.

7.3 VIRTUAL WHITEBOARDS

An area of interest in our research is exploring whether or not midair handwriting can be assisted by providing a non-tangible virtual whiteboard in the 3D space. A commercially available solution includes the Horizon Workrooms' virtual blackboard [37]. As mentioned earlier in Chapter 6, while this method is intuitive, it requires the user to move themselves close to the virtual blackboard to write physically. In contrast, our goal lies in implementing a method of midair handwriting that doesn't require the writer to be close to a writing surface.

CollaboVR [25] explored multiple implementations of virtual whiteboards for collaborative settings. These implementations include different layouts, such as writing on a virtual whiteboard in a face-to-face configuration, where the content on the whiteboard is reflected to users on each side. Another configuration is eyes-free, where the writer can rest their hand on a surface while drawing and have their drawing projected for an audience. This method is particularly useful when writing or drawing for long periods and having an audience. Earlier iterations of virtual whiteboards include Chan et al. [9]'s Magic Pad, where a wireless pen can be used on any flat surface while images are projected from a projector. This allowed interaction with 3D visualizations without requiring wearing any devices. Similar work includes the virtual whiteboard of Lech et al. [35] "virtual" whiteboard, which also used projections, and Petrykowski et al. [41]'s collaborative VR whiteboard.

The use of virtual whiteboards in VR space seems intuitive as it reflects the natural method of writing on large surfaces, and we have seen interesting implementations of these to compensate for the lack of a physical writing surface. While the usability and collaborative aspect of a virtual whiteboard isn't the primary focus of this research, we have taken design considerations when creating the virtual whiteboard in our data gathering application, further described in Section 8.1.1

7.4 EYE TRACKING AND WRITING

Hacker et al. [24] introduced TRAKTEXT, an investigation using eye-tracking technology while writing for problem-solving. They were interested in performing eye tracking, recording oculomotor features, which include eye fixations and pupil diameter, and attempting to find a correlation between these behaviours and problem-solving. The researchers performed data analysis, classifying gaze and writing behaviours such as "formative area of review," meaning when their participants reviewed an area of text beyond 12 characters or spaces from the last point of writing a word. They found that these writing behaviours may be unique to the writing tasks. Depending on the purpose of writing, it can strongly impact writing. The researchers also attempted to manipulate writing behaviours by introducing problem-solving tasks in their experiments. They could find differences, albeit the differences were obscure at the time. The researchers recommended further future research to investigate these behaviours, including planning or even daydreaming, suggesting that processing time and cognitive effort should be considered.

The premise of this research doesn't investigate the intent of writing, but the fact that they found out that writing behaviours could change depending on what needs to be written introduced the curiosity of whether the same can be applied to midair handwriting. This sparked a question of whether writing behaviour could be different when a writer is writing something they are being instructed to write versus something they write off the top of their head. To record data on different writing behaviours, we have added the **imagination** sentence type further elaborated in Chapter 9.

Karaman et al. [14, 32] investigated the use of eye gaze combined with writing as a means of creating a predictive model for performing manipulation commands such as *drag, maximize, minimize and scroll*. They used two features related to gaze-based task prediction: the distance between the sketch and gaze positions and the "Within-Cluster Variance of Gaze Positions" which is a feature that measures how eye gaze position is clustered and spread out while performing a task. In addition, their system could detect when a user intended to perform a sketch or one of the manipulation commands. On top of using eye gaze positions for mode switching, past research also includes using pupil diameter, showing that cognitive demand is associated with pupil dilation [1, 21, 30, 31]. A study in 2007 conducted by Alamargot et al. [3] stated that it is difficult to say for certain that when a user's eye gaze is fixated on a piece of visual information such as text, information is being processed. Therefore, Hacker et al. [24] also incorporated the use of pupil diameter in their study.

Based on findings from past research regarding how eye gaze and pupil diameter can be correlated to writing behaviours, we decided to include both eye gaze position and pupil diameter as features to predict the intent of when ink should be placed during midair handwriting.

7.5 MIDAIR TEXT INPUT USING VIRTUAL KEYBOARDS

Another area of research related to midair handwriting is using virtual keyboards as the text input method. Shoemaker et al. [47] compared three different methods of midair text input in a large display. They compared a circular keyboard, a traditional QWERTY keyboard and a cubic keyboard, all of which are inputted by pointing a Wii remote and using the trigger button. Their results show that the QWERTY keyboard was overall better quantitatively and qualitatively. Vulture [36] is a midair word gesture keyboard with the goal of fast text entry. Users can draw shapes of a word on the input surface and control when drawing begins by using a pinch gesture and stop drawing by releasing the pinching gesture. Text-entry was reported to be 20.6 words per minute (WPM) in their first study and 28.1 WPM in their second study. Similar to this method of input is Rotoswype [23], which uses a ring to swipe on a virtual keyboard in VR. Other methods include PizzaText [53], which uses a circular keyboard layout controlled by dual thumbsticks on a hand-held game controller.

Virtual keyboards provide a surefire way to select text in 3D space, removing the need to detect and classify handwritten text. While easy to use, our HCI approach aims to provide support for natural handwriting, as the aesthetics of handwriting, as well as flexibility, can be preserved, processed and edited in digital text [2].

8

DATA COLLECTION APPLICATION

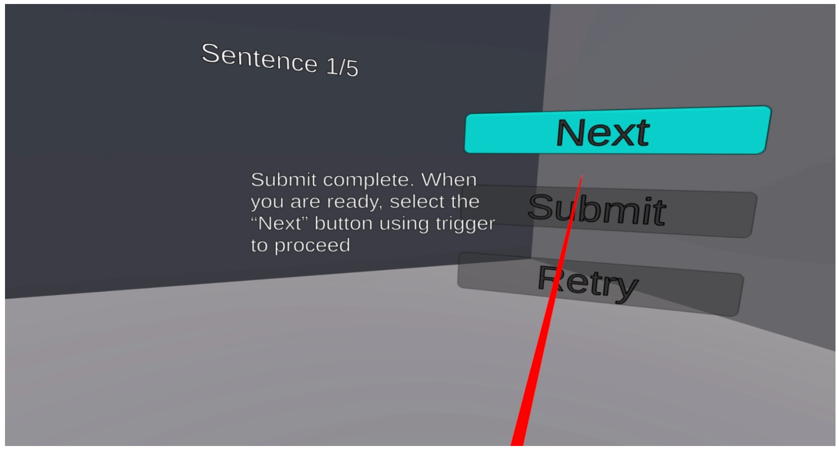
8.1 APPARATUS

To meet the requirements of unconstrained handwriting and performing the tracking we desire while being able to provide visual feedback, we chose to perform midair handwriting in a VR environment with an HTC Vive Pro Eye [26] The application was developed in Unity [52] version 2021.3.8f1. To implement tracking for the writing utensil and wrist, we used 8 Optitrack [40] Prime_x 41 cameras placed around the perimeter of the room to track rigid bodies composed of infrared markers, shown in Figure 6. The process of aligning the Optitrack and Unity coordinate systems is described in Section 8.2.1.1. The participant was instructed to wear the tracking devices for the wrist on their dominant hand while holding onto the Vive controller in their non-dominant hand as it was used to activate/deactivate the ink.

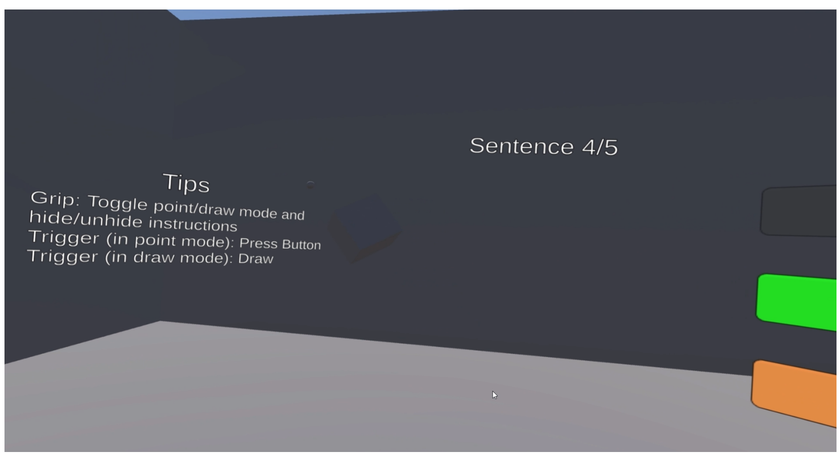
8.1.1 *VR Application*

The application logs data in each frame, including the positions and rotations of all tracked objects described in Table 7, such as the writing utensil and headset. Eye tracking data such as pupil diameter and eye gaze position were also included. We used a push-to-activate method of placing the ink, so a boolean labelled **`ink_activated`** is recorded as **True** whenever the ink was activated and **False** otherwise. The participant is guided through the study, receiving a series of prompts instructing the participant on what sentences to write, with an option of a break in between.

The application has a ***break phase*** shown in Figure 4a and a ***write phase*** shown in Figure 5a and Figure 5b. Please note that the HUD text in the screenshots does not accurately reflect what the participant saw through the VR headset lenses. The screenshots represent a flattened version of the VR scene, which lacks the stereoscopic depth and parallax effects from the participant's head movement. From the

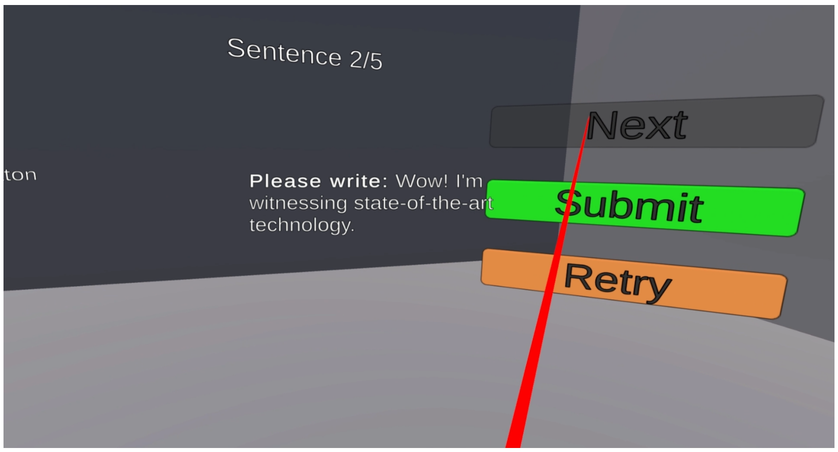


(a) Main buttons used for navigating the VR scene. The pointing ray, which is used to select the buttons, can be visibly seen. Currently, the participant is in the ***break phase*** and can select "Next" to proceed to the next sentence by pointing and using the trigger with the HTC Vive remote. The HUD text can be seen in the middle of the screen.

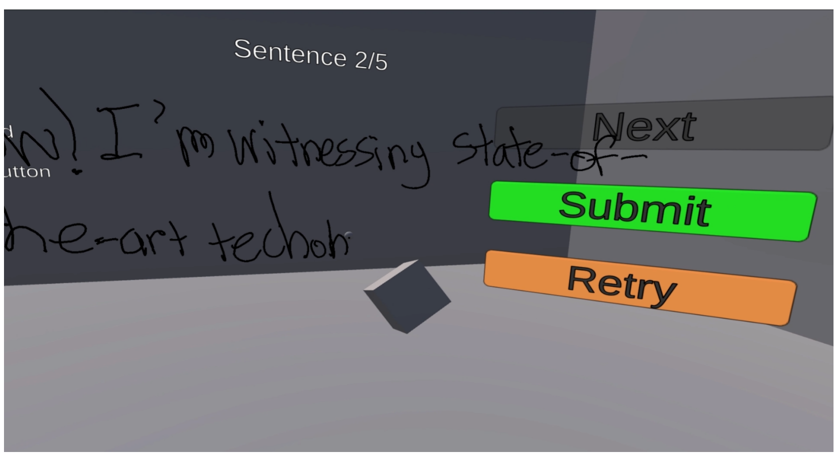


(b) A set of tips are listed on the side that serve as reminders of the buttons and functions for the participant.

Figure 4: First half of screenshots of a participant navigating through a condition and writing a sentence.



(a) ***Write phase***, where the prompt can be seen on the HUD. The Submit and Retry buttons are also available for the participant to select.



(b) ***Write phase***, The sentence "Wow! I'm witnessing state-of-the-art technology!" is being written. The pointing ray can not be seen as it has been disabled. The trigger is now used to activate the ink instead.

Figure 5: Second half of screenshots of a participant navigating through a condition and writing a sentence.

participant's point of view, the text will appear slightly stretched, and the additional line breaks were added to prevent text from nearing the edges of the headset lenses, as it would result in blurriness. Additionally, during the study, the participants were instructed to ignore the new lines in the instructions. Participants were encouraged to write the sentences to their preferences.

If the participant has been guided to the ***write phase***, the heads-up-display (HUD) will display the sentence to write as shown in Figure 5a. Additionally, the actual act of writing can only be performed during the ***write phase***. This is to avoid recording unnecessary data while the participant approaches the different phases throughout the study. The participant can use three buttons in the VR scene shown in Figure 5. If the participant is in the ***break phase***, the "Next" button can be selected to proceed to the ***write phase*** to receive the prompt for the next sentence to write. If the participant is already in ***write phase***, the "Next" button can only be selected to proceed to the ***break phase*** after the "Submit" button has been selected.

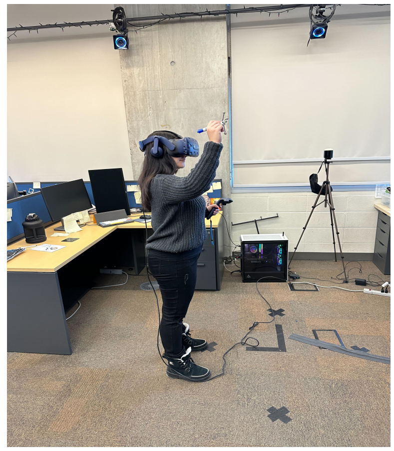
The "Submit" button can only be selected during the ***write phase*** and after at least one stroke has been made. Selecting "Submit" will export all of the recorded data for the participant's sentence. Lastly, the "Reset" button can only be used under the same conditions as the "Submit" button. Its function is to remove all visible strokes in the scene and save the data for the erased sentence in a separate file, but a new file will be created to record data for the new sentence. The VR buttons and writing area were carefully designed to minimize any background interference with the writing while also providing a 3D scenery.

While the participants used the application, they could toggle between two modes by using the grip buttons on the controller. By default, the participant will be in ***point mode***, where the pointing rays can be visibly seen in Figure 4.

In ***point mode***, the participant can only select the in-scene VR buttons by pointing with the Vive controller held in their non-dominant hand and pressing the trigger. The other mode is ***write mode***, where the pointing rays will disappear and pressing the trigger will instead place the ink. In this mode, the participant cannot select the VR buttons and can only use the trigger to perform writing.

Another consideration was minimizing any potential effects on eye-tracking data while displaying the sentence for the participant to

(a) Tracker on tip of the marker, which is referred to as the "Pen"



(b) Using the Pen

(c) Finger tracker

(d) Using the finger tracker

(e) Wrist tracker

(f) Using the wrist tracker

Figure 6: The tracker used for the pen and finger writing utensils, and the wrist. The left column is the trackers, and the right column is the trackers in use.

write. The goal was to have the writer's eye gaze unaffected during writing. One solution was to display the sentence on the HUD during ***point mode*** and remove it during ***write mode*** mode. That way, we could separate the two behaviours of eye gaze of writing and not writing. Another solution was to read out the sentence for the participant, which may also separate the two behaviours but with the possibility of overlapping if the participant decided to look at the text in the VR scene while the sentence was being read out loud. We opted for the former solution, where the system will only place the ink and record **`ink_activated`** to be true in ***write mode***, whereas pressing the trigger in ***point mode*** will not record true nor place the ink. We believe this can separate the two writing behaviours' effect on eye gaze during writing and not writing more effectively.

The hand participants used to hold the writing utensil was switched for participants accordingly to accommodate their dominant hand. In the Unity application, a setting was flipped so a variable representing right-handedness would be recorded as true for right-handed participants and false for left-handed participants. This was not used in our features but is included in all of our recorded data shown in Table 40 in our appendix.

An overview of the processes the data gathering application performs is shown in Figure 7.

8.2 WRITING UTENSILS AND METHODS

We want to compare two writing utensils and two writing methods in our tests. The first utensil, called **"Pen"**, uses a traditional pen-like object as the writing utensil, which we opted for as a whiteboard marker. The second writing utensil, called **"Finger"**, used the participant's finger as the utensil, as it closely reflects using touchless technology [28, 29]. The two writing methods were to investigate whether or not the presence of a virtual whiteboard would affect the writing done by participants. The first method, called **"No Whiteboard"**, will not have any form of assistance, meaning the ink will be placed at the position of wherever the tip of the writing utensil is at any moment of writing. The second method, called **"Virtual Whiteboard"**, is a virtual whiteboard which acts as a method of assistance for midair handwriting. This virtual whiteboard adopts a simple "snap" effect, where the z-value of the writer's ink will be set to the z-value of the

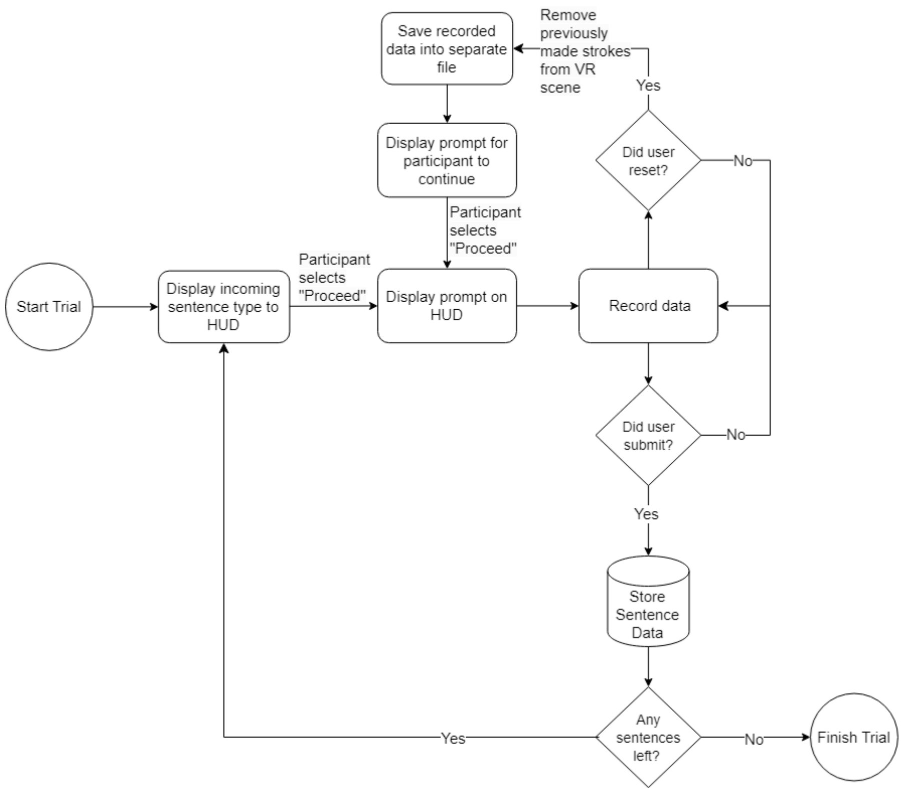


Figure 7: A flowchart that models the process used for data collection in the data gathering application. This is the entire process for one condition, where the participant will be guided through writing one block of sentences in the current condition.

whiteboard. The position of the virtual whiteboard is set manually by the participant during the training phase described in Section 9.1. The combination of writing utensils and methods are combined into **writing conditions**, shown in Table 1.

Condition Number	Writing Condition
1	Pen/No Whiteboard
2	Pen/Virtual Whiteboard
3	Finger/No Whiteboard
4	Finger/Virtual Whiteboard

Table 1: Conditions used in participant study

8.2.1 *Optitrack Tracking*

We used the Optitrack Prime$_x$ 41 cameras to track objects for our writing methods and Motive [38] version 1.10.3 to send the tracking data to our Unity application. The objects were recognized by using rigid bodies, which are infrared markers set in a specific configuration shown in Figure 6. The finger tracker was weighed at 7 grams, and the entire marker was weighed at 21 grams. The cameras were calibrated to 0.633 mm mean 3D error for the overall projection, 0.563 mm mean 3D error for the worst camera, and overall wand error with a mean of 0.386 mm.

8.2.1.1 *Coordinate Misalignment*

The positions of the tracked objects were manually aligned due to a misalignment when the positional data recorded by Optitrack was recreated in Unity. As a result, we aligned the coordinate systems manually. The full details of the alignment process are described in our appendix in Section A.1. The coordinate system in Unity relative to the real world is shown in Figure 8.

8.2.1.2 *Signal Interference*

Another problem with running Optitrack alongside the HTC Vive was the interference with the infrared rays used for tracking in both systems. By default, the Optitrack cameras ran at 180 Hz, while the HTC Vive lighthouses track at a frequency of 90 Hz. Because of these

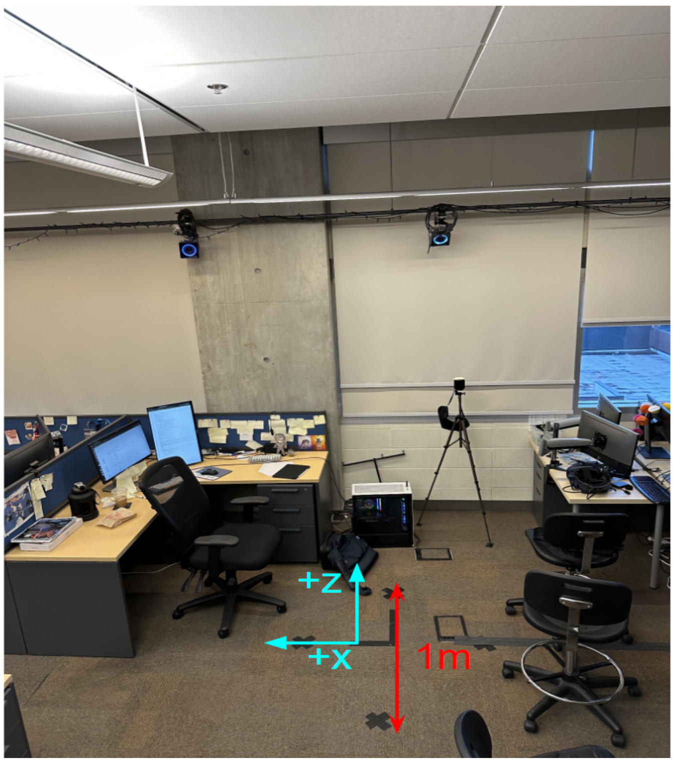


Figure 8: Blue arrows show the positive x and z coordinates in the Unity world relative to the real world. The double-ended red arrow shows 1m in the Unity world relative to the real world. Please note that all objects potentially obstructing movement were removed, such as the chairs.

high sampling rates, interference between the two systems caused the Vive to be unable to perform tracking and rendered the VR application unable to run. The solution was to lower the Optitrack cameras' sampling rate to 60 Hz as it was the highest frequency which allowed both systems to run simultaneously. This came at the cost of data since the Unity application ran at 90 fps, and we could not lower the application's framerate without resulting in a low framerate VR experience. As the data gathering application ran, Optitrack would send samples to the application. Since the application logs data every frame, this resulted in logging a duplicate sample from Optitrack every 3rd sample. The details of this problem and our solution are described in Section 10.2.1.

The last problem with tracking was interference from sunlight, as the lab room was directed towards the sun during afternoon hours, which also caused the application to fail. We resolved this by blocking the windows using construction paper.

During pilot studies, we tested using the marker and finger tracker. We asked participants if they felt that the positions for both the writing utensils were accurate, to which they responded, "Yes." When asked if they felt that there was any lag while using the writing utensils, they responded, "No."

8.2.2 *3D Eye Gaze Position*

One of the features we recorded was the 3D position of eye gaze. We used the SRanipal toolkit [17] to retrieve the left & right eye gaze origin and the respective gaze directions. These were transformed into the headset's basis by using Unity's functions: *transform.TransformPoint()* to transform the eye gaze origins and *transform.TransformPoint()* to transform the eye gaze directions. Our goal was to find the gaze intersection position, but rays in 3D generally do not have a point of intersection [16], as they are either parallel or coincidental. Duchowski et al. [16] provided a modified derivation of gaze depth in C#, which was originally made for finding the intersection at a 2D plane for eye-trackers placed at monitors [39]. Since there are no obvious planes in VR, they wrote a function that used the gaze origin and direction to calculate the length t for both gaze rays, which allowed us to find the point at which both gaze rays end. By finding the midpoint between the two endpoints, we calculated the gaze intersection in 3D.

The eye tracking was calibrated for each participant before the study, and after any time, the headset was removed to ensure that the eye tracking was accurate.

8.3 Usage in Participant Study

As a data gathering tool, this application was used in a participant study, which recorded the data previously mentioned frame by frame while a participant wrote each sentence. Appendix Table 40 contains all of the recorded data, including data not used for our features. To collect midair handwriting data in a practical context, we presented participants with stimuli, prompting them to write sentences in Chapter 9

9 STUDY DESIGN

One of our main contributions is midair handwriting data. This chapter describes how we conducted our participant study to gather the data.

We conducted a study following a within-subjects design and gathered 24 participants from Ontario Tech University and Durham College to use our application through recruitment posters and email. Each participant went through a 1-hour study and was compensated $20 after the study. This study was approved by the institutional research ethics board (REB FILE #17222).

What is your experience with VR?

Level	**# of Participants**
Experienced	2
Intermediate	3
Little	11
None	8

Table 2: Data regarding participant's experience using VR gathered from pre-study questionnaire.

How often do you write on physical whiteboards?

Level	**# of Participants**
Very often	8
Sometimes	11
Very rarely	5

Table 3: Data regarding participant's experience writing on whiteboards gathered from pre-study questionnaire.

Prior to beginning a study for each participant, we asked for their past experience using VR and writing on physical whiteboards. The questions and answers are shown in Table 2 and Table 3.

Each participant was assigned 4 combinations of writing utensils and writing methods. Each combination is grouped under the term **writing condition**:

1. Pen/No Whiteboard
2. Pen/Whiteboard
3. Finger/No Whiteboard
4. Finger/whiteboard

Please note that "Pen" refers to a pen-like writing utensil, which is an Optitrack tracker with a whiteboard marker attached to it, as shown in Figure 6a.

Each participant was assigned a unique order of conditions to avoid order effects. Table 4 describes each participant's specific order of conditions.

9.1 TRAINING AND WRITING PHASE

After signing the consent form, each participant was asked to complete a pre-study questionnaire to gather their prior experience in using any extended realities (virtual reality, augmented reality, mixed reality). Afterwards, they were instructed on how to use the HTC Vive headset and controller before being positioned to stand in the location marked to be the center of the Unity scene. After putting on the HTC Vive headset, they were verbally instructed to use the built-in eye-tracking calibration.

Participants were told to inform us if they were experiencing any discomfort, such as nausea or dizziness, as well as fatigue at any moment so we could pause the study and allow them to take a break when needed. Also, they were told to write as they would naturally without worrying about writing more slowly or quickly than they normally would.

For each combination of conditions, the participant was brought into the application's training mode, where they were instructed on how to use the application to write and navigate through the sentences. In training mode, there are 3 simple sentences designed to familiarize the participant with the current combination of writing tools and conditions. The instructions are:

1. Please write your name
2. Please write the numbers 1, 2, 3, 4, 5, 6, 7, 8, 9
3. Please write today's date

Participant	Condition Order
1	1234
2	1243
3	1324
4	1342
5	1423
6	1432
7	2134
8	2143
9	2314
10	2341
11	2413
12	2431
13	3124
14	3142
15	3214
16	3241
17	3412
18	3421
19	4213
20	4231
21	4312
22	4123
23	4321
24	4132

Table 4: Condition combination order for individual participants. 1 = Pen/No Whiteboard 2 = Pen/Virtual Whiteboard, 3 = Finger/No Whiteboard, 4 = Finger/Virtual Whiteboard

	Sentence Type				
Combination	**Pangram**	**PHA**	**QNC**	**CB**	**Imagination**
1	The quick brown fox jumps over the lazy dog	He's a jack-of-all-trades!	There were 6 signs that read, "do not enter".	Two ingredients: flour and butter (preferably unsalted).	Please write a sentence describing your favourite food.
2	The jay, pig, fox, zebra and my wolves quack!	Did you hear his can't-believe-it's-true story?	"Remember, leave at 12:35 AM" she said.	Concert is at 9:00 PM (assuming it's not cancelled).	Please list 2 or 3 things you would like to do for vacation
3	The five boxing wizards jump quickly.	Wow! I'm seeing state-of-the-art technology.	It was 4 against 9, but coach said "let's do this".	(According to this study), exercise is good.	Please write a sentence describing your favourite hobby.
4	Six big juicy steaks sizzled in a pan.	Is the well-known actor at tonight's event?	He said, "We sold about 7 or 8 copies".	A message: (Stay positive and keep pushing forward).	Please list your favourite season or weather, and list 2 or 3 reasons why.

Table 5: Stimuli used to instruct participants on what sentences to write. The combination number refers to the writing condition combination the participant.

The training sentences were designed to familiarize the participants with basic writing tasks, writing letters and numbers. For training sentence 2, we asked that the participants include writing the commas; for sentence 3, the participant was free to write in any date format. If it was the participant's first time writing using a condition that included the virtual whiteboard, they were also given the option to adjust the distance of their whiteboard to their liking. After it was set, the positional values for the virtual whiteboard were recorded and re-used whenever the participant came across the virtual whiteboard condition the second time.

ID	Sentence Type
1	Pangram
2	Punctuations / Hyphens / Apostrophes (PHA)
3	Quotes / Numbers/ Commas (QNC)
4	Colons / Brackets (CB)
5	Imagination

Table 6: Sentence types written by participants. These were also used to identify the sentence types used in training and testing.

In the test phase, the participant was instructed to write a total of 5 sentences. These sentences and their designs are described in Section 9.2. A HUD message described the next sentence's details before the participant proceeded to the ***write phase*** to receive the prompt. For example, if the participant submitted their **Pangram** sentence, the message will inform them that the next sentence will include punctuations, hyphens and apostrophes. There are 5 different types of sentences shown in Table 6, and for each combination of conditions, the participant would receive prompts to write sentences in the order of these categories:

The sentences used as prompts throughout the experiment are shown in Table 5. After completing writing all the sentences for either the training or writing phase, we asked the participants if they would like to proceed to the next combination or take a break. As mentioned in Section 8.2.2, eye tracking calibration was repeated if the headset was removed. After calibration was finished, we conducted a quick test by bringing the participant into the VR scene where a transparent sphere visualized their eye gaze position. We asked the participant to focus on different objects in the scene while monitoring the VR view to ensure the accuracy was still performing properly.

Lastly, a post-technique questionnaire was conducted after the participant completed each condition. They were allowed to answer verbally or by themselves on a laptop. Upon completing all conditions, participants were asked to complete a post-study questionnaire.

9.2 WRITING SENTENCES

The Beery-Buktenica developmental test for visual-motor integration [49] has typically been used to investigate handwriting in past research [19]. While this standardized test effectively measured handwriting skills, such as visuomotor and graphomotor skills, our goal was to measure natural handwriting patterns instead.

We narrowed our scope to investigate writing patterns solely on writing different **letters**, **punctuations** and **numbers**. As a result, we created a total of 20 sentences, with 5 sentences for each writing condition.

The **Pangram** category was designed to encompass every letter of the alphabet, while the **PHA**, **QNC**, and **CB** were to encompass the most common punctuations written in English [42]. We created these sentences using a combination of sentences retrieved from Purdue Online Writing Lab [43], our own original ideas and with the help of ChatGPT [10]. Appendix Table 38 shows the prompts used to generate the sentences before applying modifications to ensure the length of the sentences was suitable for our study.

The last sentence category is the **Imagination** sentences, which asks the participant to write something using their imagination rather than writing something they are instructed to write. These sentences were made to capture eye-tracking behaviours described in Section 7.4. Table 5 shows the prompts used for each sentence type for all of the writing conditions.

10 DATA PREPARATION

10.1 DATA GATHERING

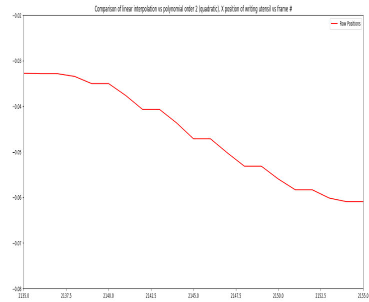
We recorded data in Unity at a rate of 90 Hz, where every frame includes timestamps, positions and rotations of tracked objects, eye tracking data and a boolean of whether the ink was turned on or off. Table 7 describes the features and the data type they were recorded as. A full table is in Appendix Table 40, which includes all recorded data, including data that was not used as part of data processing as they may be useful in future research.

Feature	Data Type
Timestamp	String
Head position	3D Vector (x, y, z)
Writing utensil position	3D Vector (x, y, z)
Wrist position	3D Vector (x, y, z)
Gaze position	3D Vector (x, y, z)
Head rotation	Quaternion (w, x, y, z)
Writing utensil rotation	Quaternion (w, x, y, z)
Wrist rotation	Quaternion (w, x, y, z)
Left pupil diameter	Float
Right pupil diameter	Float
Ink Activated	Boolean

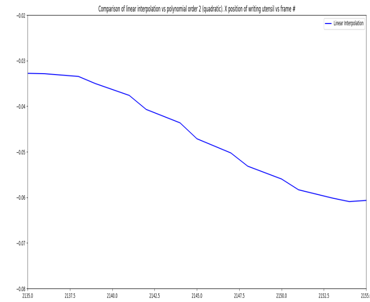
Table 7: Recorded features gathered per sentence written by participants used for data processing.

10.2 DATA CLEANUP

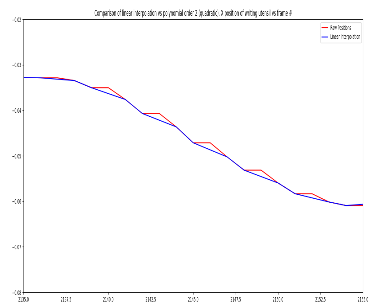
This section outlines the procedures taken to prepare the data before being put into features for model training.



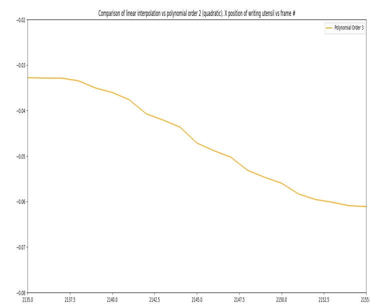
(a) Raw positions



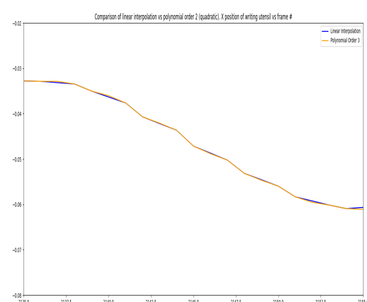
(b) Positions after performing linear interpolation



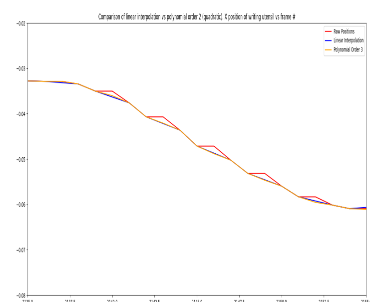
(c) Linear interpolated positions together with raw positions



(d) Polynomial order 3 interpolation



(e) Polynomial order 3 interpolation together with linear interpolation. It can be seen that the linear interpolation still has some plateaus.



(f) All 3 together

Figure 9: Comparison of different methods of smoothing out data. Red: raw positions, blue: linear interpolation, yellow: polynomial order 2. a) shows the raw positions. b) are the raw positions after performing linear interpolation. c) shows linear interpolation and the raw positions on the same plot to show the effect of smoothing out data. d) are the raw positions after performing polynomial order 3 interpolation. e) plots both polynomial order 2 and linear interpolation together. f) is everything together

10.2.1 *Different Sampling Rates*

As mentioned in Section 8.2.1.2, the sampling rate of the Optitrack system was lowered to 60 Hz during recording. Since the HTC Vive was running at 90 Hz, there was a difference in the sampling rates between the two systems. As a result, Unity would record a new sample every frame, while Optitrack would send a new sample once every third frame. Figure 9a shows a sample of the x positions of the writing utensil for a sentence segment, where it can be seen in several instances that the same position appears twice. This would cause a problem when calculating velocities and angular velocities, as every third sample from an object tracked by Optitrack would be calculated as 0. Therefore, we performed polynomial order 3 interpolation to interpolate between values, and the results are shown in Figure 9e. While this method was more computationally expensive than linear interpolation, our initial attempt at using linear interpolation is shown in Figure 9b. It can be seen that there are still some flat areas where some samples were too close to 0 and, thus, ended up being rounded to 0 instead.

10.3 KINEMATICS CALCULATIONS

This section goes over the calculations performed to obtain the velocities, accelerations, and angular velocities and accelerations.

10.3.1 *Velocities and Accelerations*

Let $\mathbf{P_n} = \begin{bmatrix} x_n & y_n & z_n \end{bmatrix}^T$ represent the positions of all tracked objects as well as gaze position, and $\mathbf{t_n} = \begin{bmatrix} t_n \end{bmatrix}$ represent the timestamps for each frame. The x, y and z velocity for each time step was calculated using the change in position over time encompassed by c is {x, y, z}, and N represents the size of the data set: $\mathbf{V}^c(\mathbf{t}) = \begin{bmatrix} \frac{p^c_{i+1} - p^c_i}{\Delta t} \end{bmatrix} \quad \text{for } i = 0, \dots, N-1$.

The accelerations were calculated using the change in velocities over time: $\mathbf{A}^c(\mathbf{t}) = \begin{bmatrix} \frac{V^c_{i+1} - V^c_i}{\Delta t} \end{bmatrix} \quad \text{for } i = 0, \dots, N-1$.

10.3.2 *Angular Velocities and Angular Accelerations*

Let $\mathbf{Q_t} = \begin{bmatrix} w_t & x_t & y_t & z_t \end{bmatrix}^T$ represent the quaternions for all tracked objects. We used the ***numpy-quaternion*** library to compute the angular velocities. Since the quaternions in Unity are unit quaternions, we calculated the difference between each quaternion by multiplying the second quaternion by the conjugate of the first quaternion. This was achieved using ***numpy*** conjugate method: $\mathbf{Q_{diff}} = [q_{i+1} * \overline{q_i}] \quad \text{for } i = 0, \ldots, N-1$.

The differences were converted to rotation vectors by multiplying the logarithm of the normalized quaternions by 2. This was achieved by using the ***quaternion.as_rotation_vector()*** method:

$\mathbf{Q_{rotation}} = [2 \log \|Q_{diff_i}\|] \quad \text{for } i = 0, \ldots, N-1$.

With $\mathbf{Q_{rotation}}$, we calculated the angular velocities **AV**: $\mathbf{AV^c(t)} = \left[\frac{Q_{rotation_i}}{\Delta t}\right] \quad \text{for } i = 0, \ldots, N-1$.

Using **AV**, we were able to calculate the angular accelerations **AA**: $\mathbf{AA^c(t)} = \left[\frac{AV^c_{i+1} - AV^c_i}{\Delta t}\right] \quad \text{for } i = 0, \ldots, N-1$.

10.3.3 *Unit Vector Between Head and Wrist*

To encode each participant's left or right-handedness, we calculated the unit vector from their headset to their wrist. Since the unit vector provides the direction from the head to the hand, it should generally be different for left and right-handed writers. Therefore, we hypothesized that the neural network is able to distinguish between left and right-handed writers using this unit vector. Let $\mathbf{HP} = \begin{bmatrix} x & y & z \end{bmatrix}^T$ represent the positions of the VR headset and $\mathbf{WP} = \begin{bmatrix} x & y & z \end{bmatrix}^T$ represent the positions of the wrist. The x, y, z components of the unit vector for each timestep is then calculated by dividing the difference between each component and the magnitude of the difference vector. Let U represent the unit vector between the head and wrist at time t. $\mathbf{U_t} = \left[\frac{HP_t - WP_t}{|HP_t - WP_t|}\right]$

10.4 REMOVING OUTLIERS AND STANDARDIZING DATA

In some instances, an object's velocity or acceleration would reach an extremely high value $X > 2\sigma$ or low value $X < -2\sigma$, where X represents an outlier. We attributed these occurrences to moments where there was a loss of tracking. This was confirmed when we found these occurrences in our Unity data viewing application (application is described in Chapter 12. When the sentences were recreated in our Unity data viewing application, the frames where these outliers occurred were found to be moments when an object jumped from one position to another. Additionally, the position of the eye gaze would be inside of the headset, which was suspected to be also due to loss of tracking. Since these samples rarely occur, we decided to use the IQR (interquartile range) method to remove samples above the 95^{th} percentile and below the 5^{th} percentile. Afterwards, we performed the same interpolation method on the removed values.

To standardize data, we used ***StandardScaler*** from the ***sklearn*** class to remove the mean and scale all data to unit variance. We performed checks to ensure no null values were present and that all data was mean-centred and had unit variance.

11

PREDICTING THE INTENT OF WRITING DURING MIDAIR HANDWRITING

RQ1 was to predict the intention of writing during midair handwriting. We used the features recorded from Table 7 and the features calculated in Section 10.3 to predict when the ink was on/off. Table 8 shows the features and labels used for model training and testing. Please note that the vectors were divided into separate x, y, and z components into individual features. To do so, we used machine learning models to train and make predictions in our tests to answer this research question.

Feature	**Data Type**
Writing Utensil Velocity	3D Vector (x, y, z)
Wrist Velocity	3D Vector (x, y, z)
Head Velocity	3D Vector (x, y, z)
Gaze Velocity	3D Vector (x, y, z)
Writing Utensil Acceleration	3D Vector (x, y, z)
Wrist Acceleration	3D Vector (x, y, z)
Head Acceleration	3D Vector (x, y, z)
Gaze Acceleration	3D Vector (x, y, z)
Writing Utensil Angular Velocity	3D Vector (x, y, z)
Wrist Angular Velocity	3D Vector (x, y, z)
Head Angular Velocity	3D Vector (x, y, z)
Writing Utensil Angular Acceleration	3D Vector (x, y, z)
Wrist Angular Acceleration	3D Vector (x, y, z)
Head Angular Acceleration	3D Vector (x, y, z)
Left Pupil Diameter	Float
Right Pupil Diameter	Float
Mean Pupil Diameter	Float
Label	**Data Type**
Ink Activated	Boolean

Table 8: Features and labels used for model training and testing. Please note that the 3D vectors were divided into separate x, y and z components into individual features.

11.1 PRELIMINARY TESTS

Before creating our neural network to perform tests to predict the intention of midair handwriting, we performed a series of tests using alternate machine learning methods. These methods include using logistic regression, and a neural network composed of two dense layers. Based on our results, these methods could not pick up on the patterns within the data. Please view our results in Table 41 and Table 42 in our appendix.

11.2 SEQUENCING DATA

Since our data is in the form of a time series, a more effective approach would be to use data from previous time steps to make predictions for the current time step. Therefore, we performed data transformations for our next testing method to use a neural network that included (Long Short-Term Memory) LSTM layers.

Before preparing our data, we split the data into individual sentences per participant, which can later be selected individually for training and testing purposes. This is further elaborated in Section 11.4. We turned our data from each sentence into sequences to prepare our data to be read by an LSTM network. Each sequence corresponds to its original sentence, meaning sequences do not overlap between different participants or sentences. To elaborate on this process, each sentence written by a participant contains a certain number of frames or time steps of data. At each time step, the data is then sequenced with a window size of 19, meaning each sequence contains 19 time steps of data. Thus, in the sequenced data, a time step now contains a sequence from a sentence containing the features and labels for the next 19 time steps. (up until n-19th frame to ensure all sequences contain 19 frames, where n is the number of frames for each sentence).

11.2.1 *Label Selection Methods*

Each sequence was then labelled with a Boolean, representing whether the ink was turned on or off. In these tests, we compare two methods of determining the label. First, we select the Boolean based on most appearances within each sequence. The Boolean that appeared the most will be the label for said sequences, and due to the odd number for the window size, there would always be a majority Boolean greater than the other. The second method of selecting the label is selecting the final Boolean in the sequence. The former method is more effective at capturing the overall trend (if any) in each sequence, as well as being more robust to noise. The latter might be better if the end of the sequence is the most important part of the sequence.

11.3 NEURAL NETWORK ARCHITECTURE

Our neural network architecture is shown in Figure 10 and is implemented in Tensorflow. The final choice for the parameters and layers of the network was determined through several training iterations. The first two layers consisted of LSTM layers with 25 units each, which enabled the network to learn effectively. Second, a dropout layer with a strength of 0.2 was added as a method to perform regularization. A dense layer of 25 units with *relu* activation follows to bring non-linearity. Lastly, a hidden layer with 1 unit and sigmoid activation is the output layer for the binary classification task. Due to the nature of LSTM and its proneness to over-fitting, many measures were taken to minimize over-fitting while attempting to help the network learn as much as possible. To do so, we used our previously mentioned dropout layer and applied L2 regularization to the 2 LSTM layers with a strength of 0.001. These values for L2 regularization were determined during the iterative training process to find the optimal values. Additionally, we applied class weights due to an imbalance of data where there were more false (ink off) instances than true (ink on) instances in all of our sentences. We followed the documentation from Tensorflow [15] to calculate the class weights and apply them as arguments in our *model.fit()* call.

During training, we used a batch size of 32, a learning rate of 0.00001 and a maximum of 150 epochs. However, not all epochs were used for all of our models since training with that many epochs would certainly cause the model to lead to over-fitting. To stop training at an optimal result, we applied early stopping by monitoring validation accuracy and set *patience* to 7 epochs. To do so, we used 2 participants inside of our training folds for validation. This will stop the model from training if the validation accuracy does not improve for 7 epochs.

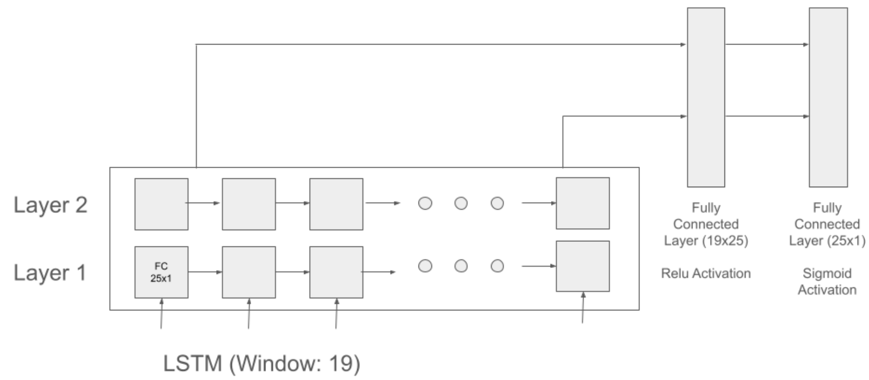


Figure 10: The network used to train models to predict the intent of writing.

Our PC specifications for the neural network were trained on an AMD Ryzen 7 5800X CPU with 16 GB of RAM and an NVIDIA RTX 3070 Ti GPU.

11.4 TRAIN AND TEST PLAN

Training and testing were performed for each condition through k-fold cross-validation, where k=6. The data was divided into folds based on the participants, where each fold contained 4 participants. The exact participants that were used for each fold are described in Table 9, and the sentences each participant wrote for each condition are described in Table 5. During training and validation, only sentences 1 and 2 (Pangrams and PHA) were used to train the models, so the remaining sentences can used for testing. This allowed us to leave a combination of sentences written by participants and sentence types the model has seen and not seen for testing. The specific participants used for training and testing depend on the current fold. To be more precise, at each fold, the remaining folds were used for training, and in these folds, the last 2 participants were used for validation for early stopping. For testing, both the participants in the current fold and the remaining folds were used for different test cases, which we elaborate on in Section 11.4.1.

We also added 2 additional conditions using our existing data, which combined the pen & finger into the no whiteboard and virtual whiteboard methods, as the results from these tests could deter-

Participants Per Fold

Fold	Participants
1	1,2,3,4
2	5,6,7,8
3	9,10,11,12
4	13,14,15,16
5	17,18,19,20
6	21,22,23,24

Table 9: Participants used per fold in k-fold cross validation

mine whether or not the virtual whiteboard would affect prediction accuracy. The conditions in our tests are listed in Table 10. It is important to note that training for these new conditions includes training on data in the combination of both conditions, and the test results are the aggregated results of testing on the same participant for both conditions in the combination. For example, training would include data from both Pen/No Whiteboard + Finger/No Whiteboard conditions. Afterwards, a test would involve testing on a participant's data from Pen/No Whiteboard and then the same participant's data from Finger/No Whiteboard.

Condition Number	Writing Condition
1	Pen/No Whiteboard
2	Pen/Virtual Whiteboard
3	Finger/No Whiteboard
4	Finger/Virtual Whiteboard
5	Pen/No Whiteboard + Finger/No Whiteboard
6	Pen/Virtual Whiteboard + Finger/Virtual Whiteboard

Table 10: Conditions used for training and testing

Hence, the final plan for training will iterate over a total number of 6 different combinations of writing utensils and conditions, with 2 methods of label selection (majority and last), across 6 folds - each producing one model. During training, data from the respective conditions were used for training and testing (e.g. In condition Pen/No Whiteboard, we used data from this condition to perform training and testing). This resulted in a total of 72 models which we saved in H5 format so we can use them for further testing. The final plan for testing included testing each model under 3 different test cases, which is elaborated in Section 11.4.1.

The entire process of training and testing the different models is described in Algorithm 1. Note that due to the sequencing of data mentioned in Section 11.2, having different participants and/or sentences within the same data sequence will not overlap.

Algorithm 1: Model training and testing algorithm

```
for i in range (all writing conditions) do
    for j in range (label selection method) do
        for k in range (all folds) do
            C = sequenced data from current writing
              condition
            L = current label selection method
            F = current fold
            if condition is a combination then
                S_i = sentence type from first condition
                T_i = sentence type from second condition
                model = train(C, L, !F, [S_1, S_2, T_1, T_2])
                First Condition Test Results = {
                  test case 1 result = test(model, F, [S_1, S_2])
                  test case 2 result = test(model, F, [S_3, S_4, S_5])
                  test case 3 result = test(model, !F, [S_3, S_4, S_5])
                }
                Second Condition Test Results = {
                  test case 1 result = test(model, F, [T_1, T_2])
                  test case 2 result = test(model, F, [T_3, T_4, T_5])
                  test case 3 result = test(model, !F, [T_3, T_4, T_5])
                }
            else
                S_i = sentence type
                model = train(C, L, !F, [S_1,S_2])
                Test Results = {
                  test case 1 result = test(model, F, [S_1, S_2])
                  test case 2 result = test(model, F, [S_3, S_4, S_5])
                  test case 3 result = test(model, !F, [S_3, S_4, S_5])
                }
```

11.4.1 *Test Cases*

To test our models, we tested them on every single sentence that was not part of the training and validation process. The test cases encompass the sentences into different categories, which describe how the model performs when predicting different types of data that the model has never seen. The test cases are described as follows:

1. Test on **new** participants and sentences types the model **has seen**.
2. Test on **new** participants and sentences types the model **has not seen**.
3. Test on **the same training** participants and sentences types the model **has not seen**.

To elaborate on the test cases, **S** denotes the sentence type, for example, S1 = Pangram. Table 11 describes an example of how the sentences were divided for training, validation and the different test cases. In the table, the test fold is fold 6, while the training folds are folds 1-5. The green cells are the sentences used for training, while the grey cells are sentences used for validation. The test cases examine the robustness of our models in different ways. **Test case 1** (shown in red cells) tested the model's performance on the same type of sentences the model was trained on (S1, S2), but written by participants the model has never seen (participants in the test fold). **Test case 2** (shown in orange cells) tested the model on new sentence types (S3, S4, S5) as well as new participants or participants in the test fold. Lastly, **Test case 3** (shown in blue cells) also tested on new sentence types, but the participants in the training folds wrote these sentences. All of these test cases include sentences that the model has never seen.

11.5 TRAINING RESULTS

The aggregated results from our models during training are shown in Table 12 for when the sequences were labelled based on the majority Boolean and Table 13 for when the sequences were labelled based on the last Boolean. From our training results, our models trained by the majority label selection method had lower training and validation loss for all conditions. However, regarding training and validation

Participant	Sentence Types				
1	Pangram	PHA	QNC	CB	Imagination
2	Pangram	PHA	QNC	CB	Imagination
3	Pangram	PHA	QNC	CB	Imagination
...	Pangram	PHA	QNC	CB	Imagination
19	Pangram	PHA	QNC	CB	Imagination
20	Pangram	PHA	QNC	CB	Imagination
21	Pangram	PHA	QNC	CB	Imagination
22	Pangram	PHA	QNC	CB	Imagination
23	Pangram	PHA	QNC	CB	Imagination
24	Pangram	PHA	QNC	CB	Imagination

Table 11: Colourized example of test cases for one test fold. Green cells are sentences used for training, grey cells are sentences used for validation, red cells are sentences used for test case 1, orange cells are sentences used for test case 2, and blue cells are sentences used for test case 3.

accuracy, there are instances where they are higher when the majority Boolean was the selection method for the labels and other instances where the last Boolean scored higher accuracies.

Condition	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy
Pen/No Whiteboard	0.25145	0.35175	0.91176	0.86951
Pen/Virtual Whiteboard	0.26796	0.37331	0.90478	0.86250
Finger/No Whiteboard	0.27436	0.36613	0.90158	0.86623
Finger/Virtual Whiteboard	0.27871	0.35545	0.89926	0.86735
Pen/No Whiteboard + Finger/No Whiteboard	0.24961	0.32933	0.90805	0.87781
Pen/Virtual Whiteboard + Finger/Virtual Whiteboard	0.26633	0.32915	0.90181	0.8748

Table 12: Aggregated training and validation results for the different conditions using majority method for label selection.

Condition	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy
Pen/No Whiteboard	0.30505	0.37820	0.89285	0.84656
Pen/Virtual Whiteboard	0.31535	0.41536	0.88050	0.83598
Finger/No Whiteboard	0.30538	0.39610	0.88375	0.84350
Finger/Virtual Whiteboard	0.30691	0.38623	0.88378	0.84625
Pen/No Whiteboard + Finger/No Whiteboard	0.28805	0.36876	0.88733	0.85278
Pen/Virtual Whiteboard + Finger/Virtual Whiteboard	0.28768	0.37101	0.88786	0.85056

Table 13: Aggregated training and validation results for the different conditions using last label selection method.

11.6 TEST RESULTS

The aggregated results for the majority label selection method are shown in Tables 14 to 21 and Tables 22 to 29 with the last boolean label selection method. A full spreadsheet, including results for individual sentences, is available in our supplementary material, along with all of the confusion matrices and areas under the receiving operating characteristic (AUC-ROC) graphs.

Models trained on Pen/No Whiteboard sentences

Tested on Pen/No Whiteboard sentences		**Label Selection Method: Majority**			
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.86610	0.81704	0.86406	0.83711	0.94368
New Participants New Sentence Types	0.85410	0.80800	0.84640	0.82272	0.93366
Same Participant New Sentence Types	0.86576	0.81680	0.86694	0.83807	0.94295

Table 14: Aggregated results from tests for models trained on data from Pen/No Whiteboard sentences, then tested on data from Pen/No Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Pen/Virtual Whiteboard sentences

Tested on Pen/Virtual Whiteboard sentences		**Label Selection Method: Majority**			
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85526	0.81090	0.842270	0.82202	0.93450
New Participants New Sentence Types	0.85615	0.81414	0.84076	0.82385	0.93320
Same Participant New Sentence Types	0.86697	0.82488	0.85679	0.83832	0.94037

Table 15: Aggregated results from tests for models trained on data from Pen/Virtual Whiteboard sentences, then tested on data from Pen/Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Finger/No Whiteboard sentences

Tested on Finger/No Whiteboard sentences			**Label Selection Method: Majority**		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85443	0.79812	0.85586	0.82019	0.93422
New Participants New Sentence Types	0.85421	0.80582	0.85030	0.82286	0.93200
Same Participant New Sentence Types	0.86583	0.81359	0.86989	0.83864	0.94016

Table 16: Aggregated results from tests for models trained on data from Finger/No Whiteboard sentences, then tested on data from Finger/No Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Finger/Virtual Whiteboard sentences

Tested on Finger/Virtual Whiteboard sentences			**Label Selection Method: Majority**		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85671	0.80880	0.86237	0.83079	0.93886
New Participants New Sentence Types	0.84964	0.81285	0.84259	0.82164	0.93214
Same Participant New Sentence Types	0.86467	0.82869	0.85879	0.83976	0.94152

Table 17: Aggregated results from tests for models trained on Finger/Whiteboard sentences, then tested on Finger/Whiteboard sentences data. The majority Boolean of the sequence was selected as the label.

Models trained on Pen/No Whiteboard + Finger/No Whiteboard Tested on Pen/No Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.87548	0.83150	0.86936	0.84858	0.89511
New Participants New Sentence Types	0.86420	0.82226	0.85691	0.83542	0.94173
Same Participant New Sentence Types	0.87733	0.83211	0.87767	0.85218	0.93729

Table 18: Aggregated results from tests for models trained on data from Pen/No Whiteboard and Finger/No Whiteboard sentences, then tested on data from Pen/No Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Pen/No Whiteboard + Finger/No Whiteboard Tested on Finger/No Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.86905	0.81756	0.86880	0.83808	0.9371
New Participants New Sentence Types	0.86698	0.821271	0.86511	0.83949	0.94161
Same Participant New Sentence Types	0.87630	0.82792	0.87846	0.85042	0.94772

Table 19: Aggregated results from tests for models trained on data from Pen/No Whiteboard and Finger/No Whiteboard sentences, then tested on data from Finger/No Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Pen/Virtual Whiteboard + Finger/Virtual Whiteboard					
Tested on Pen/Virtual Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.87019	0.82579	0.86087	0.84134	0.89849
New Participants New Sentence Types	0.86665	0.82731	0.85069	0.83596	0.94045
Same Participant New Sentence Types	0.87666	0.83822	0.86510	0.84922	0.93593

Table 20: Aggregated results from tests for models trained on data from Pen/Whiteboard and Finger/Whiteboard sentences, then tested on data from Pen/Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Pen/Virtual Whiteboard + Finger/Virtual Whiteboard					
Tested on Finger/Virtual Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.87051	0.82673	0.87322	0.84636	0.94716
New Participants New Sentence Types	0.86223	0.82672	0.85594	0.83689	0.94014
Same Participant New Sentence Types	0.87204	0.83550	0.86785	0.84829	0.94626

Table 21: Aggregated results from tests for models trained on data from Pen/Whiteboard and Finger/Whiteboard sentences, then tested on data from Finger/Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Pen/No Whiteboard sentences Tested on Pen/No Whiteboard sentences			Label Selection Method: Last		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.86555	0.79408	0.89542	0.83969	0.94633
New Participants New Sentence Types	0.83960	0.76831	0.86289	0.80882	0.92638
Same Participant New Sentence Types	0.84854	0.78531	0.86008	0.81783	0.93156

Table 22: Aggregated results from tests for models trained on data from Pen/No Whiteboard sentences, then tested on data from Pen/No Whiteboard sentences. The last boolean of the sequence was selected as the label.

Models trained on Finger/No Whiteboard sentences Tested on Finger/No Whiteboard sentences			Label Selection Method: Last		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.83465	0.77238	0.83755	0.79889	0.92102
New Participants New Sentence Types	0.83411	0.77138	0.83703	0.79914	0.91939
Same Participant New Sentence Types	0.84399	0.77913	0.85586	0.81252	0.92694

Table 23: Aggregated results from tests for models trained on data from Finger/No Whiteboard sentences, then tested on data from Finger/No Whiteboard sentences. The last Boolean of the sequence was selected as the label.

Models trained on Pen/Virtual Whiteboard sentences

Tested on Pen/Virtual Whiteboard sentences			Label Selection Method: Last		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.83241	0.76475	0.84097	0.79218	0.92269
New Participants New Sentence Types	0.83595	0.77657	0.83931	0.80136	0.91952
Same Participant New Sentence Types	0.84897	0.78398	0.86181	0.81867	0.92851

Table 24: Aggregated results from tests for models trained on data from Pen/Whiteboard sentences, then tested on data from Pen/Whiteboard sentences. The last Boolean of the sequence was selected as the label.

Models trained on Finger/Virtual Whiteboard sentences

Tested on Finger/Virtual Whiteboard sentences			Label Selection Method: Last		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.83790	0.77959	0.85217	0.80885	0.92831
New Participants New Sentence Types	0.83169	0.78346	0.83409	0.80063	0.92093
Same Participant New Sentence Types	0.84703	0.79472	0.85541	0.82004	0.93099

Table 25: Aggregated results from tests for models trained on data from Finger/Whiteboard sentences, then tested on data from Finger/Whiteboard sentences. The last Boolean of the sequence was selected as the label.

Models trained on Pen/No Whiteboard + Finger/No Whiteboard					
Tested on Pen/No Whiteboard			Label Selection Method: Last		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85421	0.79822	0.85257	0.82156	0.88630
New Participants New Sentence Types	0.84407	0.78979	0.84047	0.80997	0.92704
Same Participants New Sentence Types	0.85636	0.79726	0.86330	0.82639	0.92342

Table 26: Aggregated results from tests for models trained on data from Pen/No Whiteboard and Finger/No Whiteboard sentences, then tested on data from Pen/No Whiteboard sentences. The last Boolean of the sequence was selected as the label.

Models trained on Pen/No Whiteboard + Finger/No Whiteboard					
Tested on Finger/No Whiteboard			Label Selection Method: Last		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.84466	0.77755	0.85628	0.80824	0.93182
New Participants New Sentence Types	0.84712	0.79177	0.84692	0.81481	0.92762
Same Participants New Sentence Types	0.85820	0.79880	0.86472	0.82832	0.93583

Table 27: Aggregated results from tests for models trained on data from Pen/No Whiteboard and Finger/No Whiteboard sentences, then tested on data from Finger/No Whiteboard sentences. The last Boolean of the sequence was selected as the label.

Models trained on Pen/Virtual Whiteboard + Finger/Virtual Whiteboard

Tested on Pen/Virtual Whiteboard			**Label Selection Method: Last**		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
Same Participants New Sentence Types	0.84654	0.78806	0.84691	0.81239	0.89047
New Participants New Sentence Types	0.84322	0.78014	0.84663	0.80828	0.92739
Same Participants New Sentence Types	0.85654	0.79693	0.86442	0.82649	0.92367

Table 28: Aggregated results from tests for models trained on data from Pen/Whiteboard and Finger/Whiteboard sentences, then tested on data from Pen/Whiteboard sentences. The last Boolean of the sequence was selected as the label.

Models trained on Pen/Virtual Whiteboard + Finger/Virtual Whiteboard

Tested on Finger/Virtual Whiteboard			**Label Selection Method: Last**		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
Same Participants New Sentence Types	0.84934	0.79453	0.85695	0.82073	0.93451
New Participants New Sentence Types	0.84305	0.79158	0.84939	0.81485	0.92778
Same Participants New Sentence Types	0.85283	0.80023	0.86194	0.82659	0.93465

Table 29: Aggregated results from tests for models trained on data from Pen/Whiteboard and Finger/Whiteboard sentences, then tested on data from Finger/Whiteboard sentences. The last Boolean of the sequence was selected as the label.

To our surprise, the metrics were all relatively close across all conditions, even though there were many differences in terms of the z-positions, or depth during midair writing described in Chapter 12. The closeness in our results across different writing utensils and methods could indicate that the models could not distinguish patterns between writing with and without a virtual whiteboard. We conducted further investigation and found that this was likely due to the small window size of 19 frames, where there was little change in the z-positions within that time window. This can be seen in our plots of the writing utensil's z position when plotted in a 19-frame window and a 300-frame window in Figure 11. This sample is chosen from a random participant's sentence during writing with the Pen/No Whiteboard condition.

To compare results from the two label selection methods, we averaged them from their respective tests. The results are shown in Table 30. It is abundantly clear that selecting the majority Boolean in each sequence outperforms selecting the last Boolean in each sequence concerning every metric.

Label Selection Method	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Majority	0.86489	0.82052	0.86029	0.83659	0.93570
Last	0.84569	0.78892	0.85384	0.81613	0.9236

Table 30: Aggregated results from all the tests comparing the two label selection methods.

An example of test results from one of the iterations in the tests performed for the Pen/No Whiteboard condition is shown in the following confusion matrices and AUC-ROC graphs. The confusion matrices for the 3 test cases are shown in Figures 12, 14, 16 and 17, and their respective AUC-ROC graphs are shown in Figures 13, 15, 18 and 19

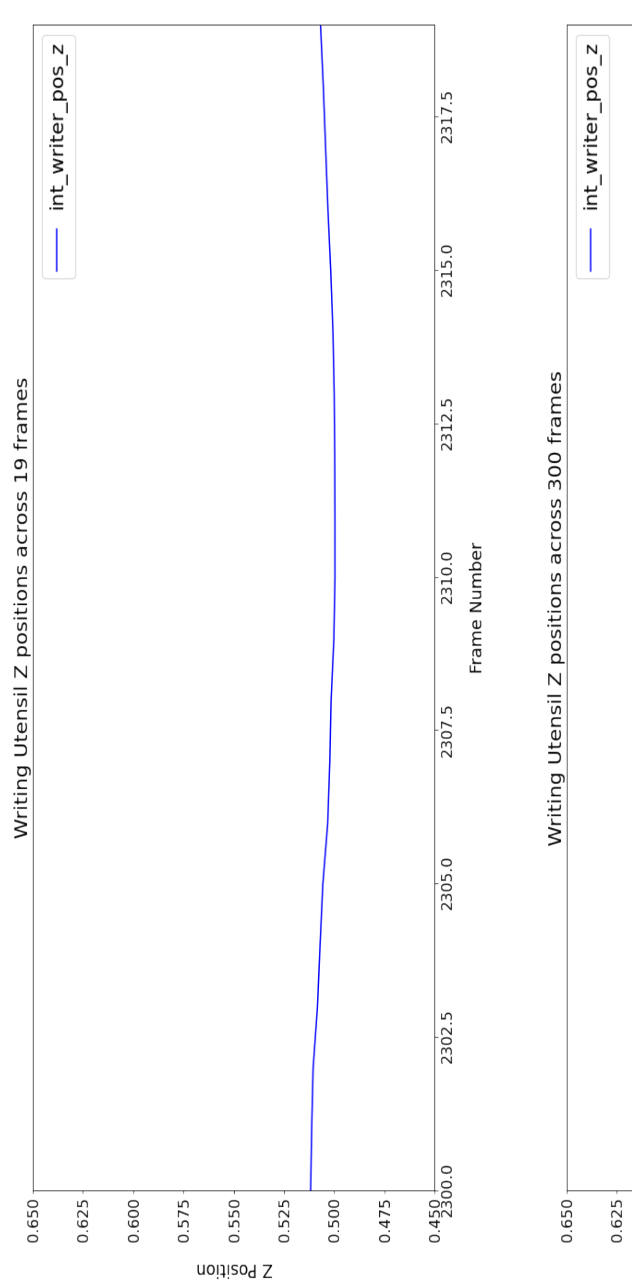


Figure 11: Z-positions of writing utensil for a sentence, 19 frames (top) vs 300 frames (bottom).

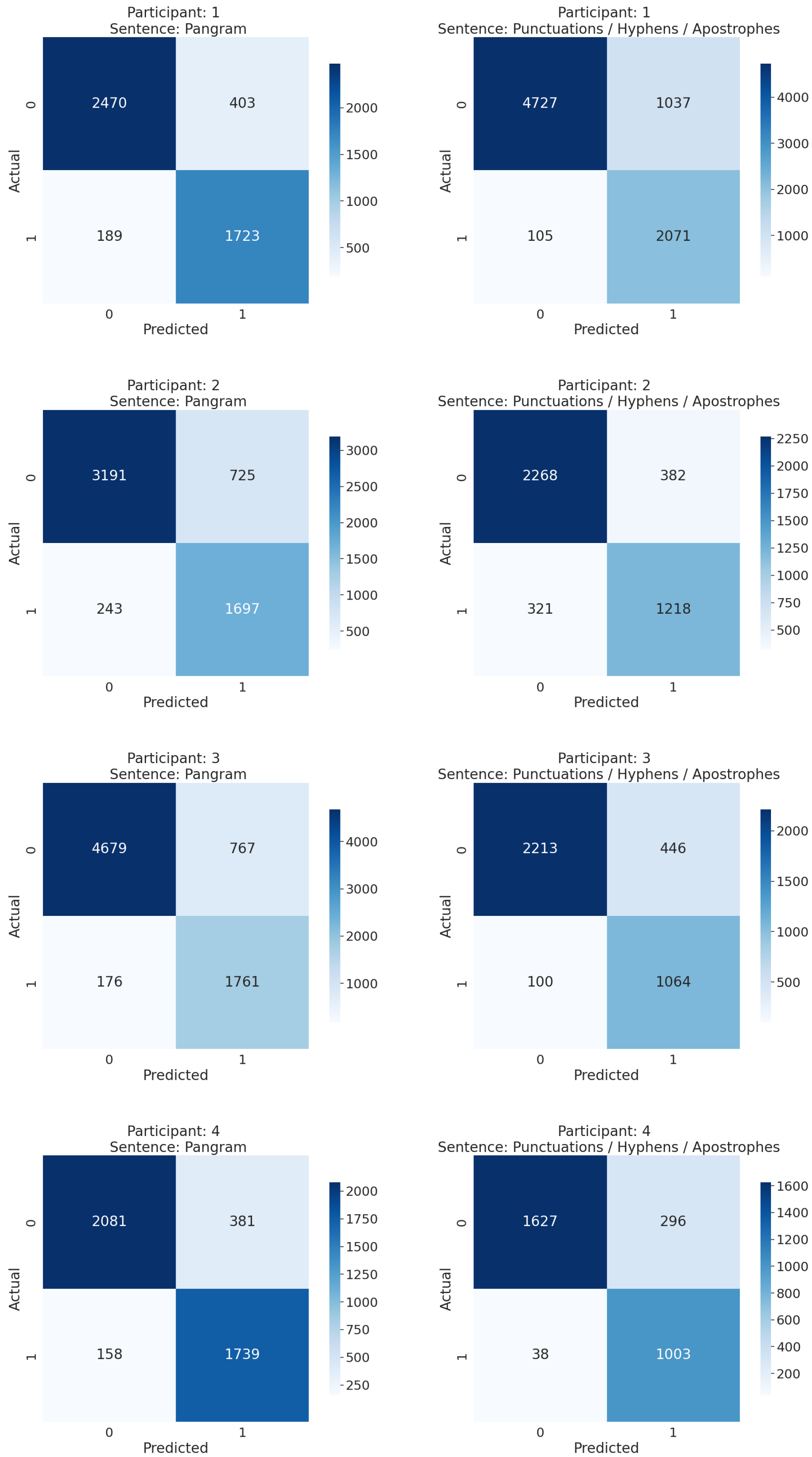


Figure 12: Confusion matrices for test case 1, Pen/No Whiteboard condition. Model was trained on folds 2-6, tested on fold 1, sentences Pangram and PHA.

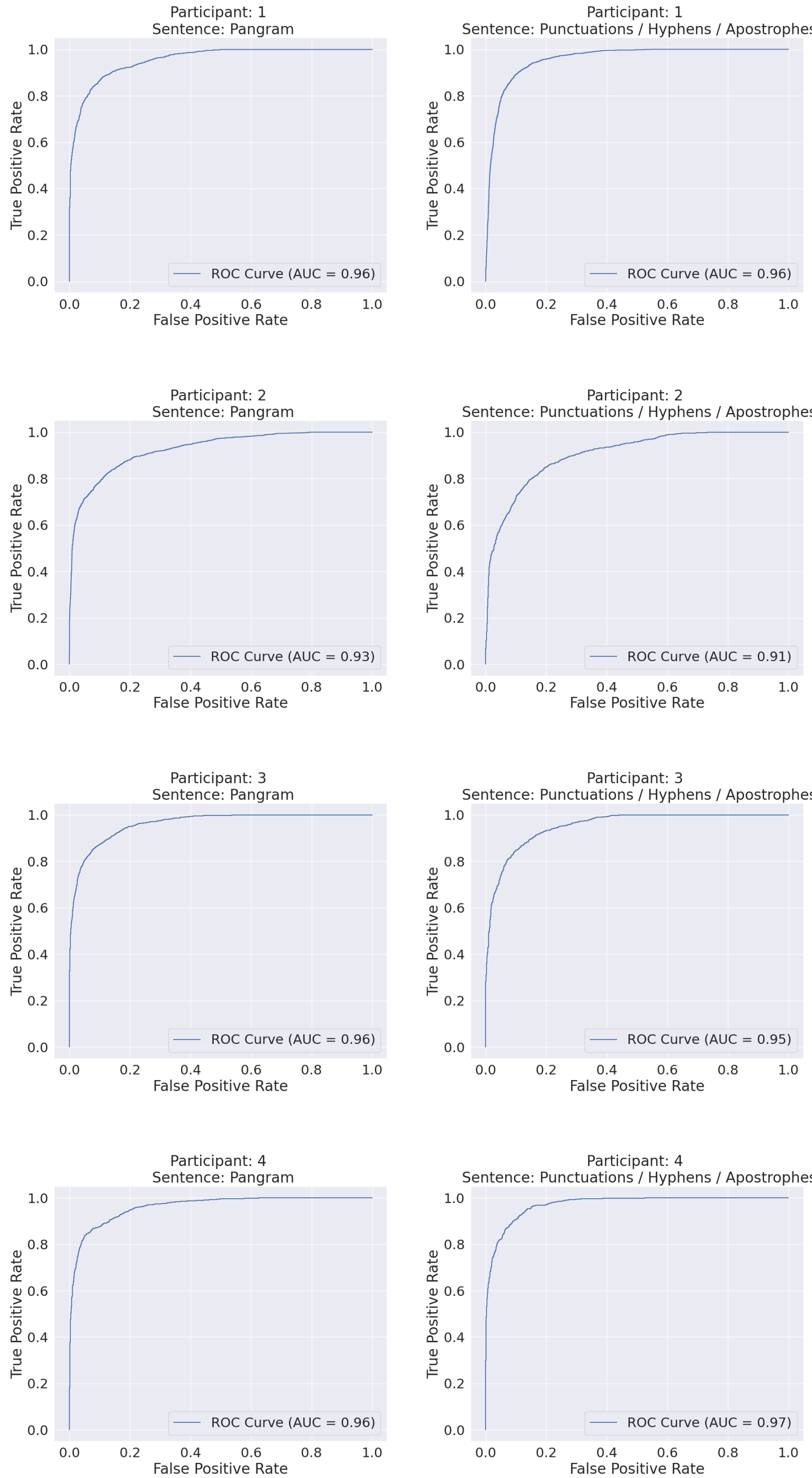


Figure 13: ROC-AUC graphs for test case 1, Pen/No Whiteboard condition. Model was trained on folds 2-6, tested on fold 1, sentences Pangram and PHA.

Figure 14: Confusion matrices for test case 2, Pen/No Whiteboard condition. Model was trained on folds 2-6, tested on fold 1, sentence QNC, CB, Prompt.

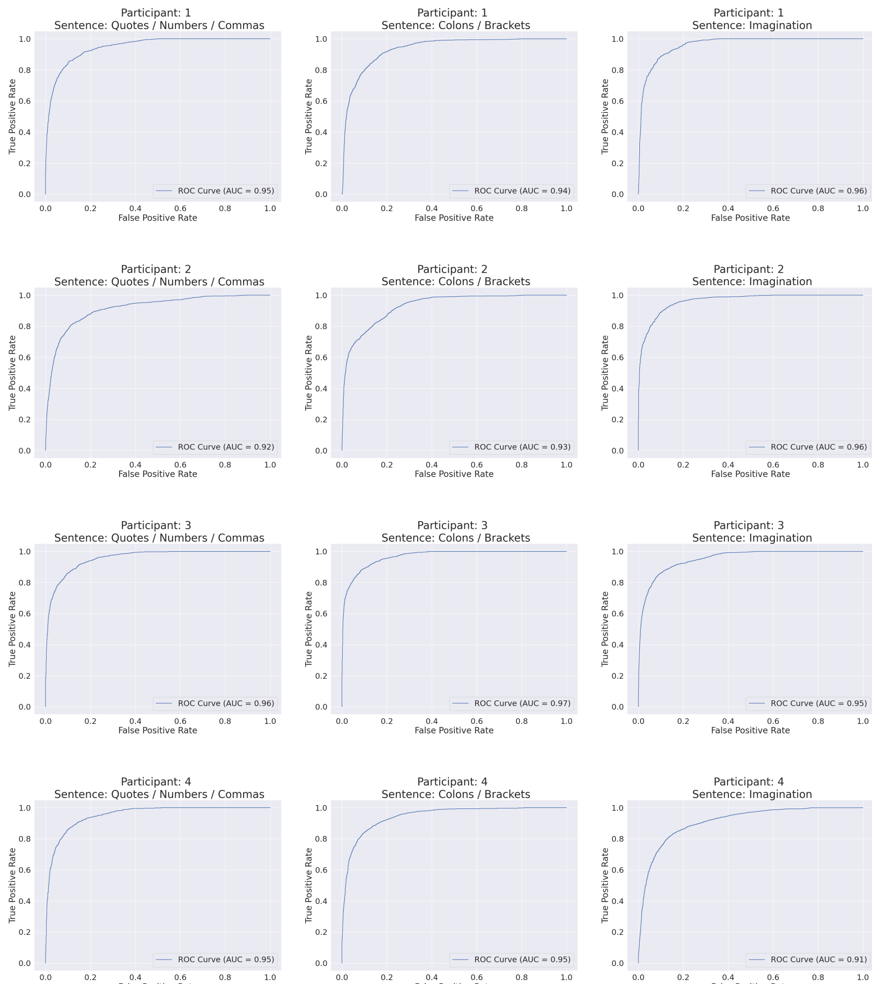


Figure 15: ROC-AUC graphs for test case 2, Pen/No Whiteboard condition. Model was trained on folds 2-6, tested on fold 1, sentences QNC, CB, Prompt.

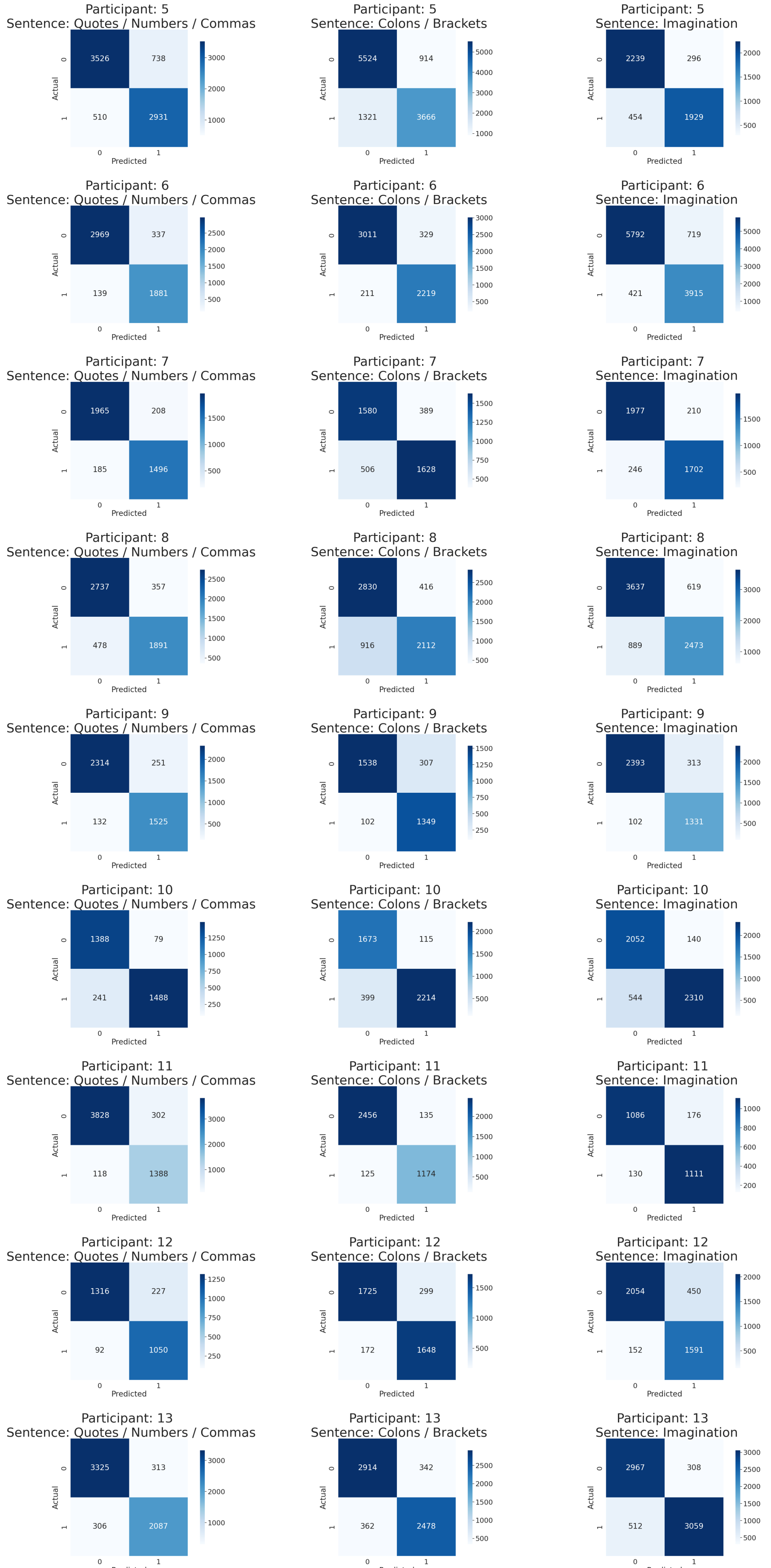


Figure 16: First half of confusion matrices for test case 3, Pen/No Whiteboard condition. Model was trained on folds 2-6, tested on folds 2-6, sentences QNC, CB, Prompt.

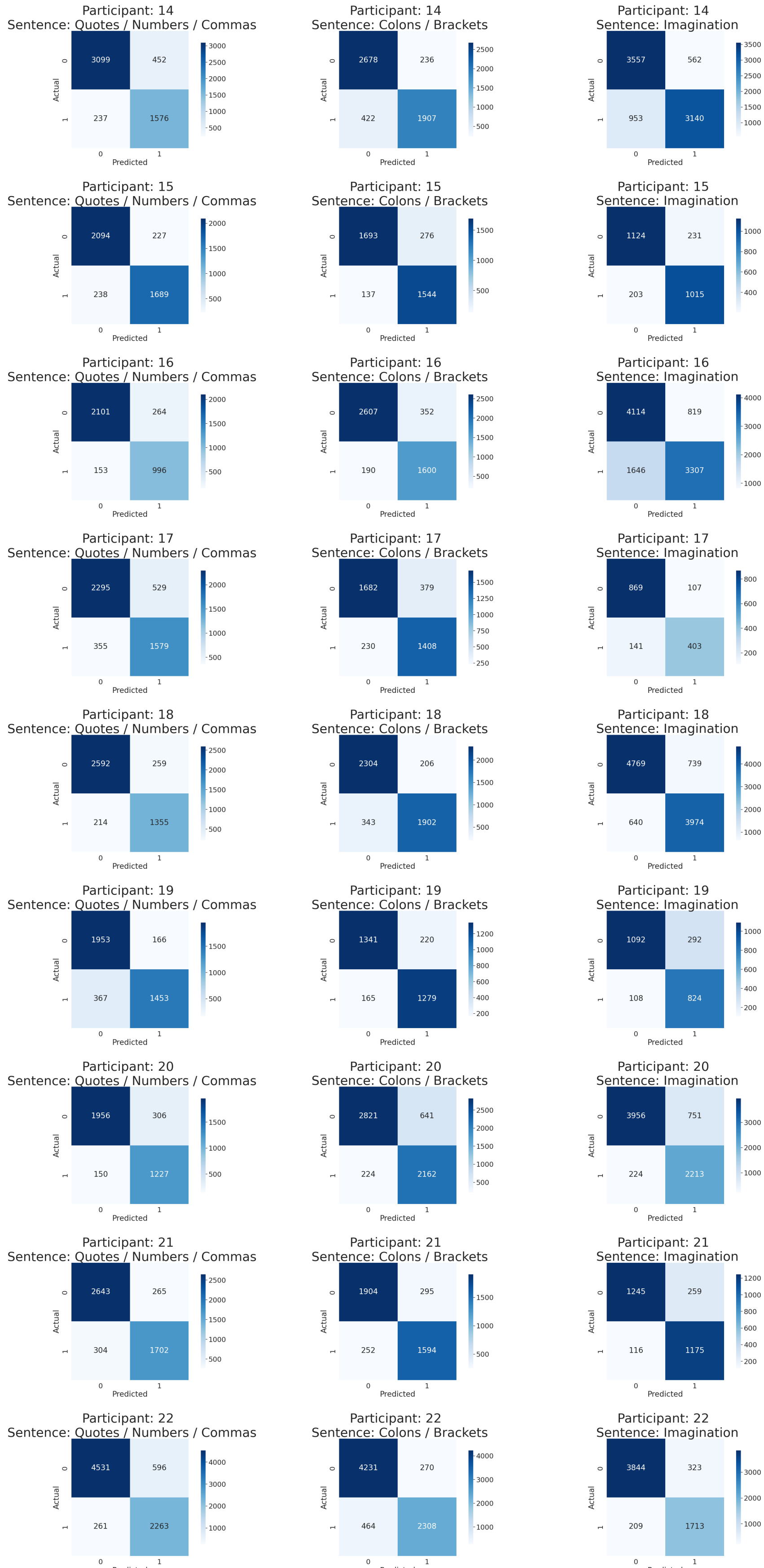


Figure 17: Second half of confusion matrices for test case 3, Pen/No Whiteboard condition. Model was trained on folds 2-6, tested on folds 2-6, sentences QNC, CB, Prompt.

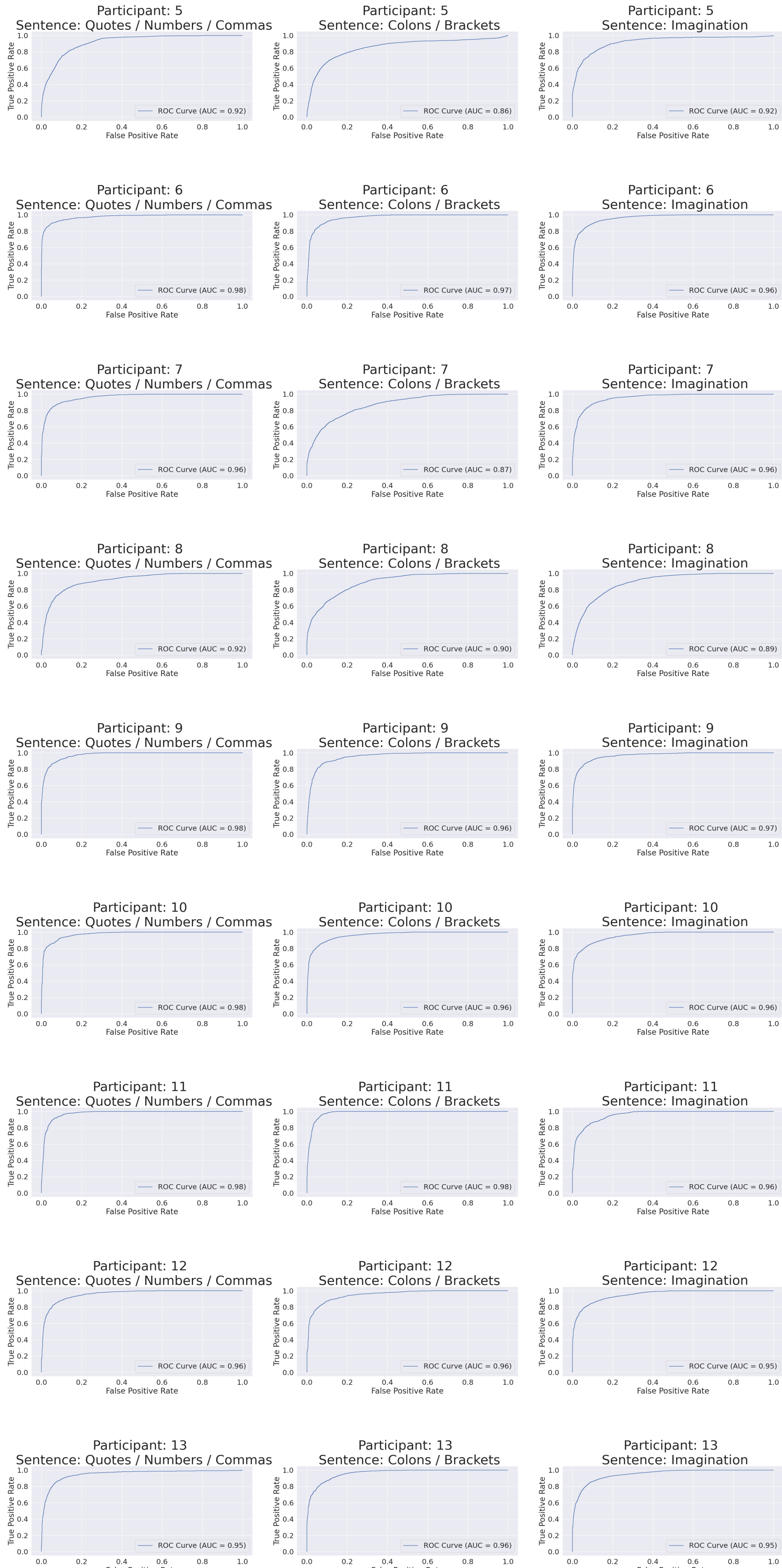


Figure 18: First half of ROC-AUC graphs for test case 3, Pen/ No Whiteboard condition. Model was trained on folds 2-6, tested on folds 2-6, sentences QNC, CB, Prompt.

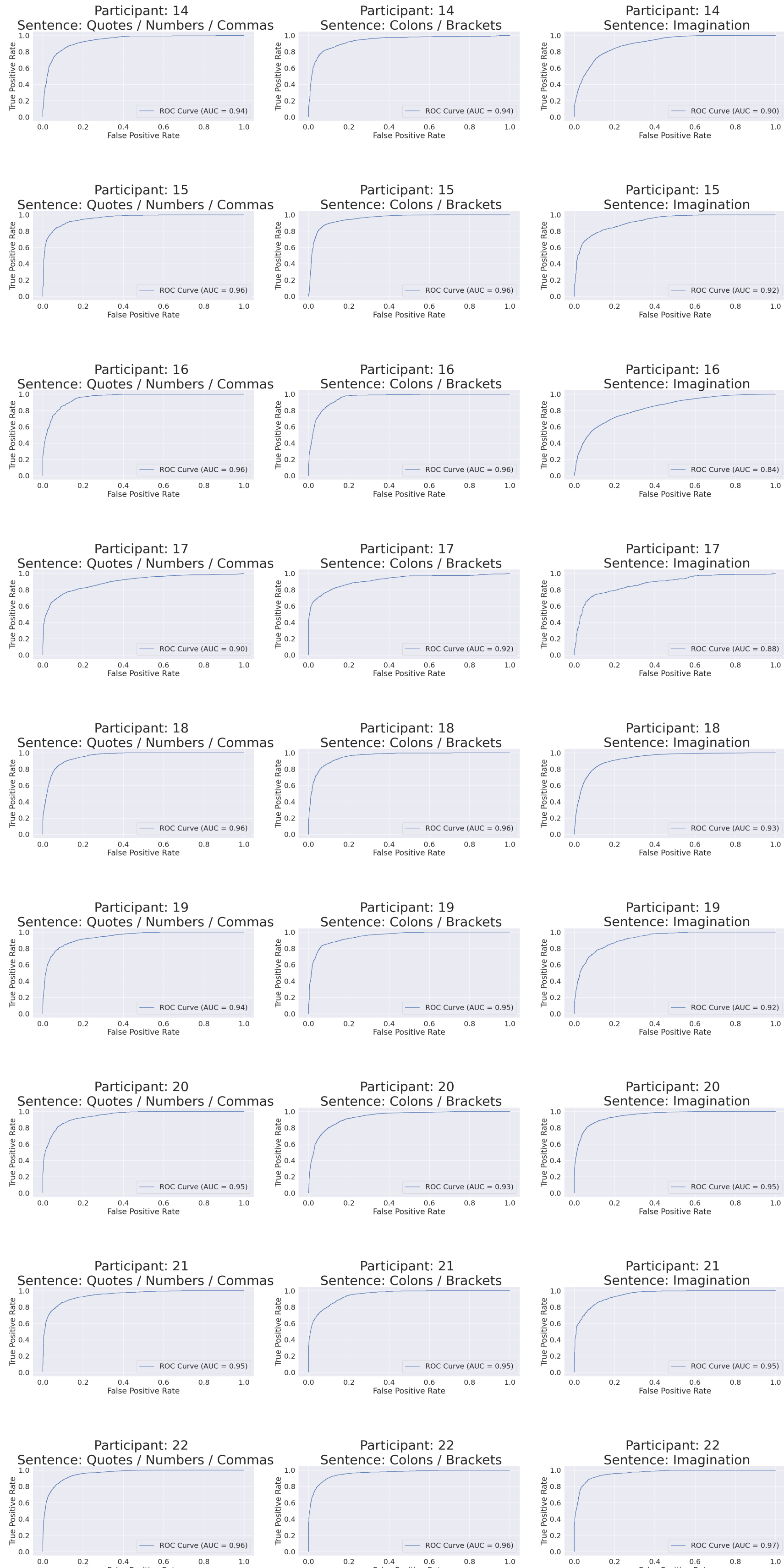


Figure 19: Second half of ROC-AUC graphs for test case 3, Pen/No Whiteboard condition. Model was trained on folds 2-6, tested on folds 2-6, sentences QNC, CB, Prompt.

11.7 ADDITIONAL TESTS

After testing the models on the same conditions they were trained for, our next objective was to test the robustness of the model. Therefore, we tested the models against the writing methods they were not trained for. For example, the models trained on data from the Pen/No Whiteboard condition were tested on data from the Pen/Virtual Whiteboard condition. We only tested on the models that were trained with the majority Boolean label selection method as per our findings in Table 30.

11.7.1 *Testing Model Robustness*

The details of these additional test conditions are described in the following list:

- Condition 7: Train on sentences from Pen/No Whiteboard, Test on sentences from Pen/Virtual Whiteboard
- Condition 8: Train on sentences from Pen/Virtual Whiteboard, Test on sentences from Pen/No Whiteboard
- Condition 9: Train on sentences from Finger/No Whiteboard, Test on sentences from Finger/Virtual Whiteboard
- Condition 10: Train on sentences from Finger/Virtual Whiteboard, Test on sentences from Finger/No Whiteboard

The aggregated results from these tests are shown in Tables 31 to 34.

11.7.2 *Comparison of Test Cases*

Along with evaluating the predictive ability of the models, another goal was to investigate if the models would perform differently when being tested on a mix of data from participants and sentences the model has seen and not seen during training. Specifically, these refer to the test cases, which include new data in the form of sentences written by new participants & same sentences, new participants & new sentences, and same participants & new sentences. Table 35 shows the results for the models trained and tested on the same conditions.

Table 36 shows the results for models where the writing utensil was combined (e.g. Pen/No Whiteboard + Finger/No Whiteboard) and Table 37 shows the results for models trained and tested on opposite conditions (e.g. Train on Pen/Virtual No Whiteboard, test on Pen/Virtual Whiteboard). All the tables we have seen from our test results follow the same trend where test case 2 is consistently lower than test cases 1 and 3.

We perform a detailed discussion of all of our results in Chapter 14.

Models trained on Pen/No Whiteboard Tested on Pen/Virtual Whiteboard		Label Selection Method: Majority			
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85849	0.82460	0.83238	0.82464	0.93479
New Participants New Sentence Types	0.85337	0.82435	0.82114	0.81878	0.93168
Same Participant New Sentence Types	0.85787	0.82719	0.83087	0.82534	0.93405

Table 31: Aggregated results from tests for models trained on data from Pen/No Whiteboard, then tested on data from Pen/Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Pen/Virtual Whiteboard Tested on Pen/No Whiteboard		Label Selection Method: Majority			
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.86216	0.81461	0.85755	0.83233	0.94040
New Participants New Sentence Types	0.85072	0.80402	0.84579	0.82010	0.93084
Same Participant New Sentence Types	0.85697	0.81247	0.84847	0.82662	0.93346

Table 32: Aggregated results from tests for models trained on data from Pen/Whiteboard, then tested on data from Pen/No Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Finger/No Whiteboard

Tested on Finger/Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85412	0.82077	0.83360	0.82121	0.93476
New Participants New Sentence Types	0.84276	0.82020	0.81221	0.80646	0.92551
Same Participant New Sentence Types	0.85533	0.82974	0.83125	0.82526	0.93363

Table 33: Aggregated results from tests for models trained on data from Finger/No Whiteboard, then tested on data from Finger/Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Models trained on Finger/Whiteboard

Tested on Finger/No Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85383	0.79689	0.85234	0.81970	0.93192
New Participants New Sentence Types	0.85465	0.80369	0.85607	0.82500	0.92548
Same Participant New Sentence Types	0.86079	0.80775	0.86294	0.83188	0.93493

Table 34: Aggregated results from tests for models trained on data from Finger/Whiteboard, then tested on data from Finger/No Whiteboard sentences. The majority Boolean of the sequence was selected as the label.

Results from models trained and tested on same writing conditions			**Label Selection Method: Majority**		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85812	0.80871	0.85614	0.82753	0.93794
New Participants New Sentence Types	0.85353	0.81020	0.84501	0.82277	0.93275
Same Participant New Sentence Types	0.86581	0.82099	0.86310	0.83870	0.94125

Table 35: Aggregated test results for models trained and tested on the same writing conditions.

Results from models trained and tested when writing utensil was combined under the same condition			**Label Selection Method: Majority**		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.87131	0.82539	0.86806	0.84359	0.91946
New Participants New Sentence Types	0.86501	0.82439	0.85716	0.83694	0.94098
Same Participant New Sentence Types	0.87559	0.83344	0.87227	0.85003	0.94180

Table 36: Aggregated results from tests for models trained and tested on the same writing utensil but combined with and without whiteboard, such as Pen/No Whiteboard + Pen/Whiteboard and Finger/No Whiteboard + Finger/Whiteboard.

Results from models trained and tested on opposite writing conditions

Test Case	Label Selection Method: Majority				
	Accuracy	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
New Participants Same Sentence Types	0.85715	0.81422	0.84397	0.82447	0.93547
New Participants New Sentence Types	0.85037	0.81306	0.83380	0.81759	0.92838
Same Participant New Sentence Types	0.85774	0.81929	0.84338	0.82727	0.93402

Table 37: Aggregated results from tests for models trained and tested on the opposite writing conditions, such as when the model was trained on Pen/No whiteboard and tested on Pen/Whiteboard.

12

WRITING PATTERNS

RQ2.1 was to investigate whether the presence of a virtual whiteboard affected midair handwriting, which we refer to as the curvature of midair handwriting. We hypothesized that the presence of a virtual whiteboard could help the writer keep their midair handwriting restricted in a plane to some extent. During pilot studies, we noticed that midair handwriting generally appeared in a curve or in a sphere around the writer where the writer is at the center and the writer's arm is the sphere's radius.

12.1 DATA VIEWING APPLICATION

We developed a separate Unity application that can read the recorded data files and recreate the strokes that were made for each sentence. This was used to visualize the entire sentence in a 3D environment while traversing in 6 degrees of freedom using a first-person point of view camera. Users can use the mouse to look around, WASD keys to move the camera, and arrow keys to scroll back and forth between frames. A menu can be used to select files containing data for each sentence and a specific frame to jump to.

The following figures provide 6 different views (front, back, left, right, top, bottom) for each sentence. Our supplementary material contains the 6 different views for every recorded sentence from our experiments. It is important to note that the Unity application took these views manually via screenshots. The **true** positions of the writing utensil for each sentence were visualized, meaning these are the actual positions of the writing utensil in both writing methods. Even if the sentences were written using the virtual whiteboard method, the actual positions of the writing utensil were visualized and **not** the ink after being affected by the virtual whiteboard. The positions are visualized by spheres, where the blue spheres represent when the ink was activated, and the grey spheres represent when the ink was not activated.

12.2 WRITING CURVATURE

P denotes a participant when followed by a number, such as P1 representing participant 1. Visualizing the positions of the writing utensil enabled us to get a clear picture of various levels of curvature for different participants while writing different sentences. We discovered that midair writing generally appeared in a curve or a sphere around the participants when they wrote without the virtual whiteboard, which aligns with our hypothesis. Whereas when participants wrote with the virtual whiteboard, their writing appeared to follow a flat plane, as if a physical whiteboard existed. The curvature of the writing can be seen more prominently in the top and bottom views. There were varying degrees to the amount of curvature across different sentences, which occurred within the same participant under different conditions and sentences. An example of this case is **P1**'s sentences, such as their Imagination & PHA sentences in Figures 20 and 21 where a curve is more visible, and their Pangram sentence in Figure 22 which has a much lesser curve. These sentences were written under the Pen/No Whiteboard condition. The same observations were made even using a different writing utensil, such as P1's Colons/Bracket sentence under the Finger/No Whiteboard condition in Figure 23.

In contrast, when a virtual whiteboard was present, the writing consistently appeared to be on a flat plane. An example is **P1**'s Colons/Bracket sentence written under the Pen/Virtual Whiteboard condition in Figure 24 and their QNC sentence when they wrote under the Finger/Virtual Whiteboard condition in Figure 25. This type of writing behaviour appeared consistently across all participants, where the curvature is noticeable in the No Whiteboard methods and where there is much less curvature in the Virtual Whiteboard methods. An example of an extreme case where curvature was present is **P5**, where their QNC sentence written under the Pen/No Whiteboard condition is shown in Figure 26, and their PHA sentence written under the Pen/Virtual Whiteboard condition is shown in Figure 27. Based on these observations, we can conclude that the presence of a virtual whiteboard can affect how a writer writes in midair. Writing can be made to be flatter as if there were an actual whiteboard, despite the whiteboard being completely virtual.

Figure 20: Visualized writing utensil positions for the sentence written by Participant 1, under the Pen/No Whiteboard condition. The Sentence type is Colons/Brackets.

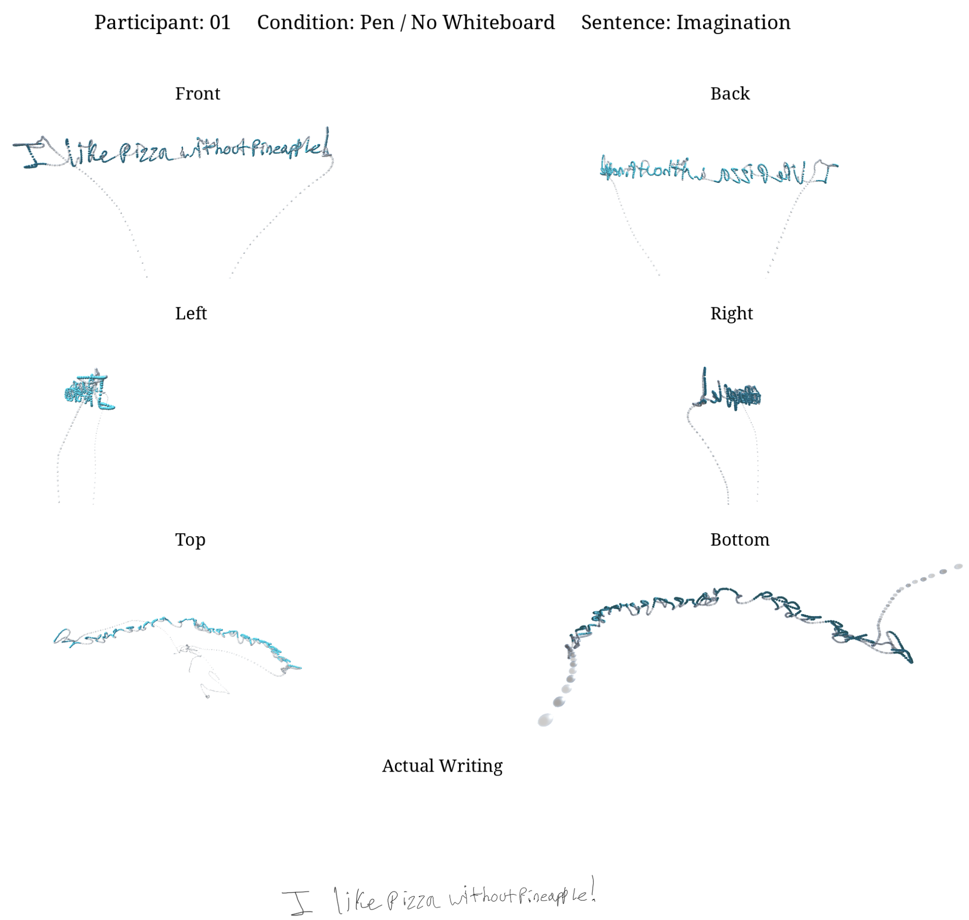


Figure 21: Visualized writing utensil positions for the sentence written by Participant 1, under the Pen/No Whiteboard condition. The Sentence type is Imagination.

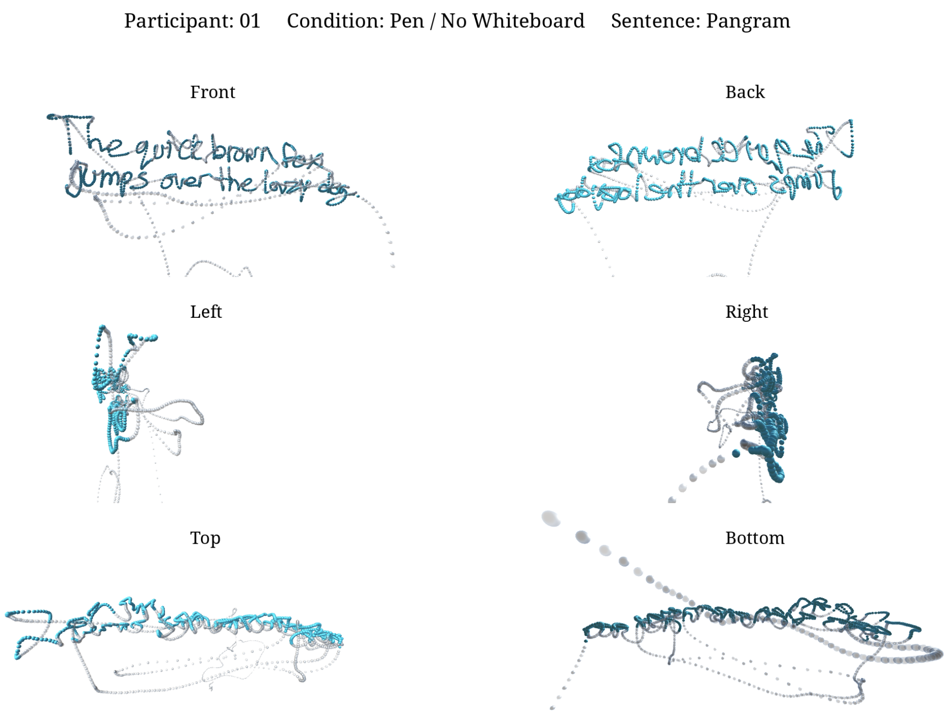


Figure 22: Visualized writing utensil positions for the sentence written by Participant 1 under the Pen/No Whiteboard condition. The Sentence type is Pangrams.

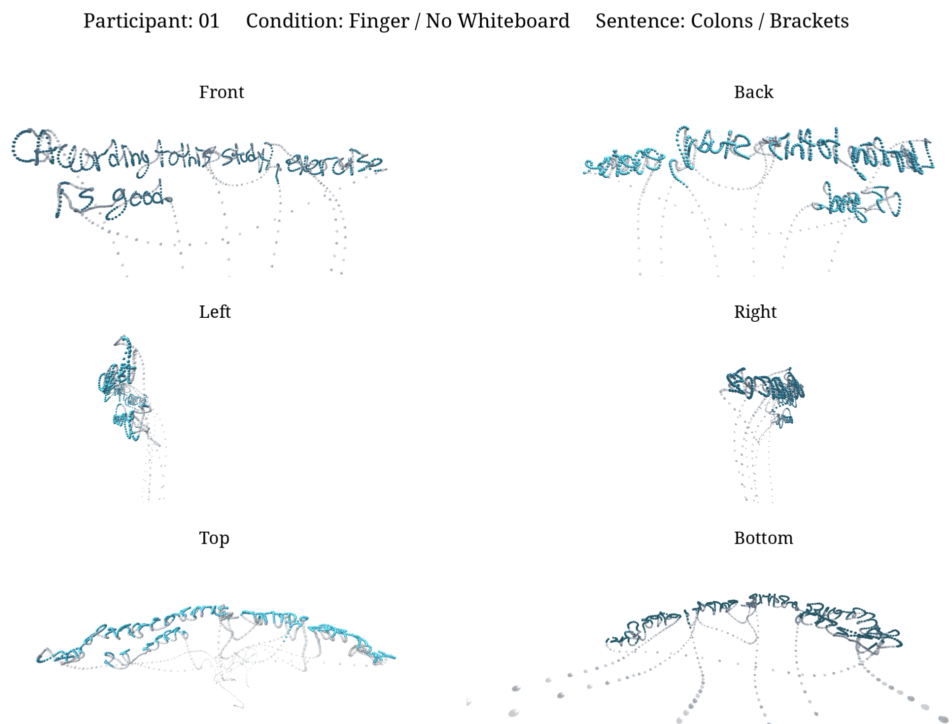


Figure 23: Visualized writing utensil positions for the sentence written by Participant 1, under the Finger/No Whiteboard condition. The Sentence type is Colons/Brackets.

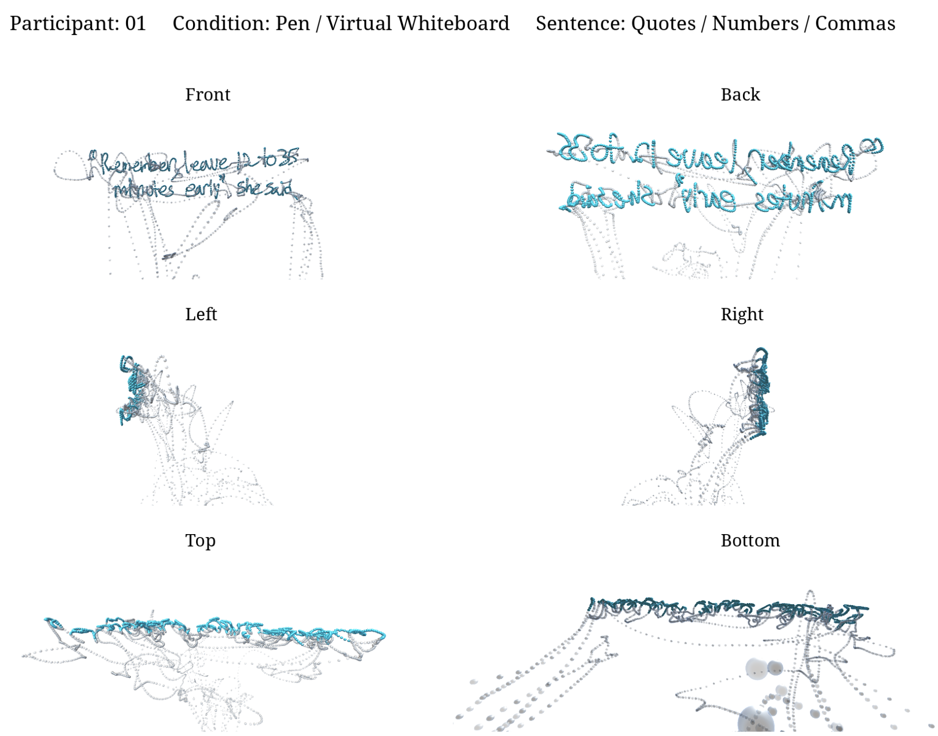


Figure 24: Visualized writing utensil positions for the sentence written by Participant 1, under the Pen/Virtual Whiteboard condition. The Sentence type is Quotes/Numbers/Commas.

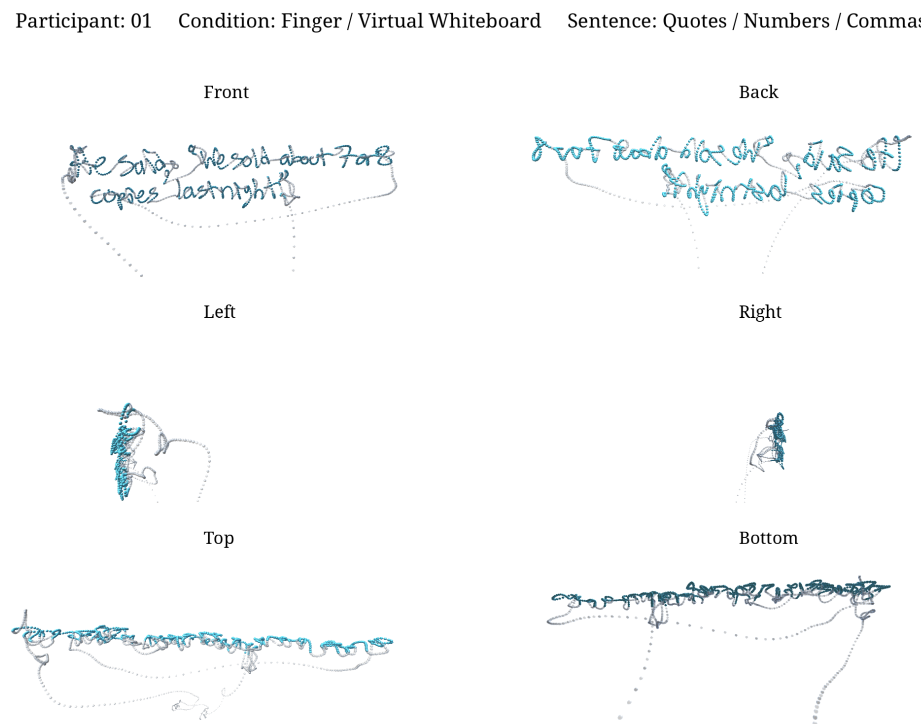


Figure 25: Visualized writing utensil positions for the sentence written by Participant 1, under the Finger/Virtual Whiteboard condition. The Sentence type is Quotes/Numbers/Commas.

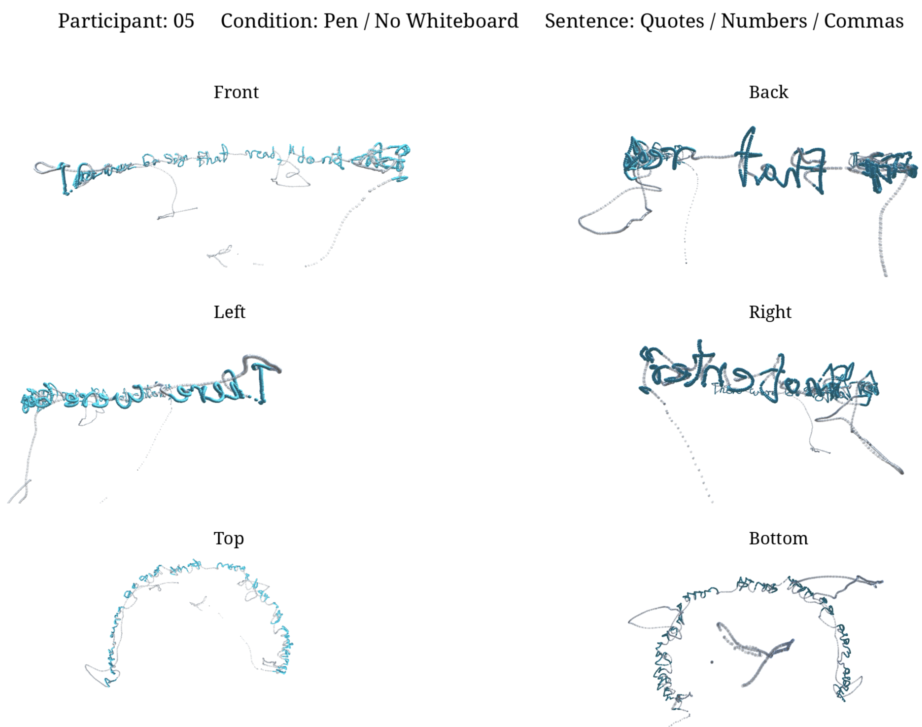


Figure 26: Visualized writing utensil positions for the sentence written by Participant 5, under the Pen/No Whiteboard condition. The Sentence type is Quotes/Numbers/Commas.

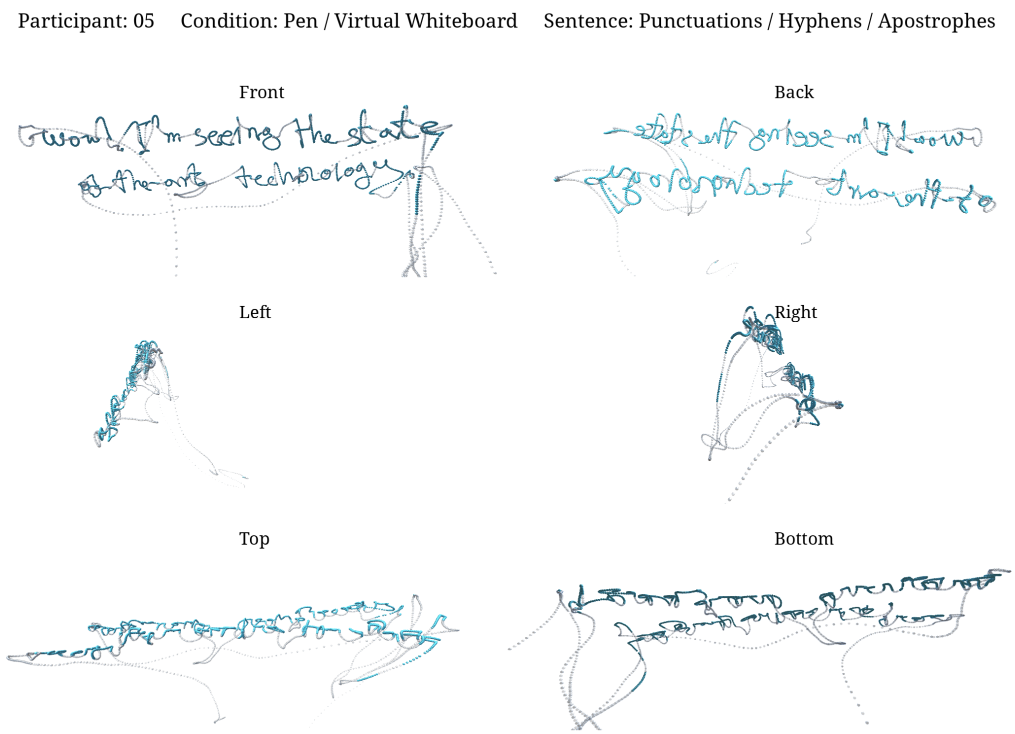


Figure 27: Visualized writing utensil positions for the sentence written by Participant 6, under the Pen/Virtual Whiteboard condition. The Sentence type is Punctuations/Hyphens/Apostrophes.

12.2.1 *Stair-casing Effect*

Another interesting pattern we observed was when participants wrote on a new line while writing under the conditions without the virtual whiteboard, where their new line of handwritten text was written **closer** to the participant as if they appeared in a staircase fashion. This effect can be examined in the **left** and **right** views of the sentences. An example is **P2**'s Imagination sentence written under the Finger/No Whiteboard condition shown in Figure 28, where there are noticeable changes in depth for each new line. On the other hand, the new lines of text written **P2** for their Imagination sentence under the Finger/Whiteboard condition are shown in Figure 29, where the handwriting appears to be inline. This can also be seen in the previous figures, although less prominent. As with the writing curvature effect, there are also different degrees to this stair-casing effect across different participants. Based on these observations, we can conclude that while performing midair handwriting, it is likely a writer's handwriting will appear in a sphere-like shape around the writer, with the writer in the center of the sphere. Additionally, the presence of a virtual whiteboard can also play a role in the stair-casing effect by reducing the amount of stair-casing.

12.3 WRITING SPEEDS

From the recordings, we made observations of varying writing speeds. Some participants wrote faster than others, which may be attributed to factors such as their natural handwriting speed & habits and writing on an unfamiliar medium in midair for the first time. The sentence written by a fast writer is visualized in our application in Figure 30. We observed that the spheres visualizing the faster writer's sentence appeared more dispersed than those in the slowly written sentence. This also introduces the possibility of affecting the data and models due to higher velocities and accelerations, although this was not investigated in this study.

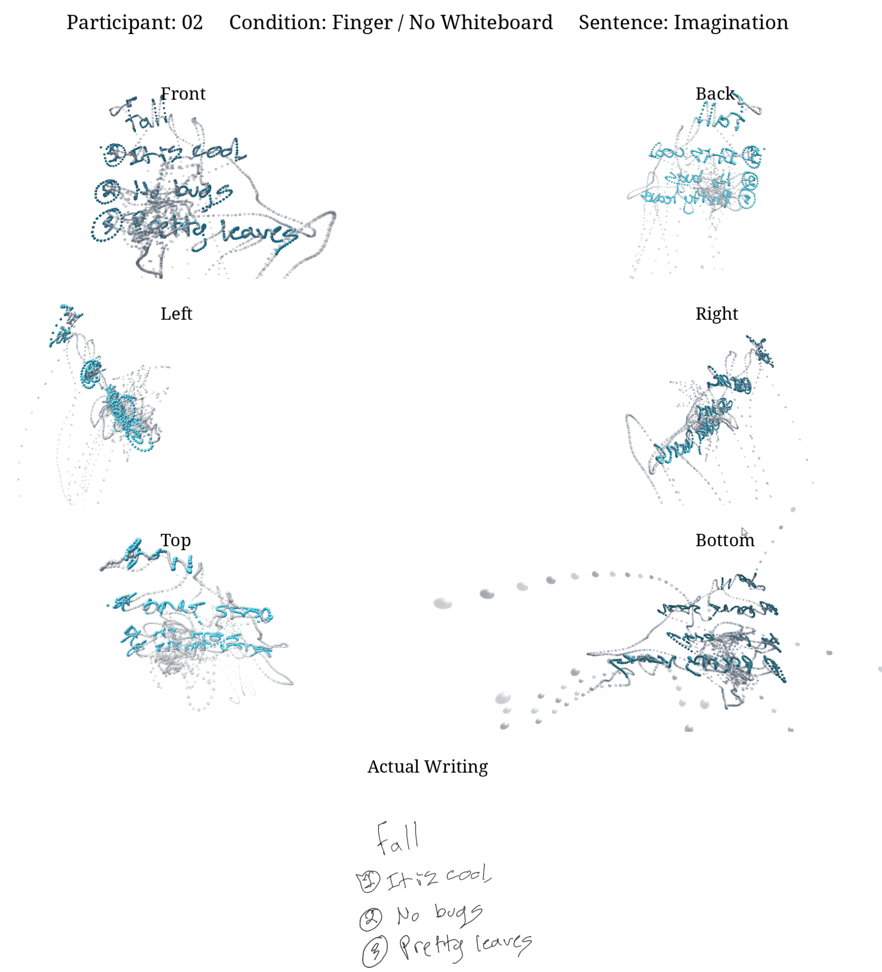


Figure 28: Visualized writing utensil positions for the sentence written by Participant 2, under the Finger/No Whiteboard condition. The Sentence type is Imagination.

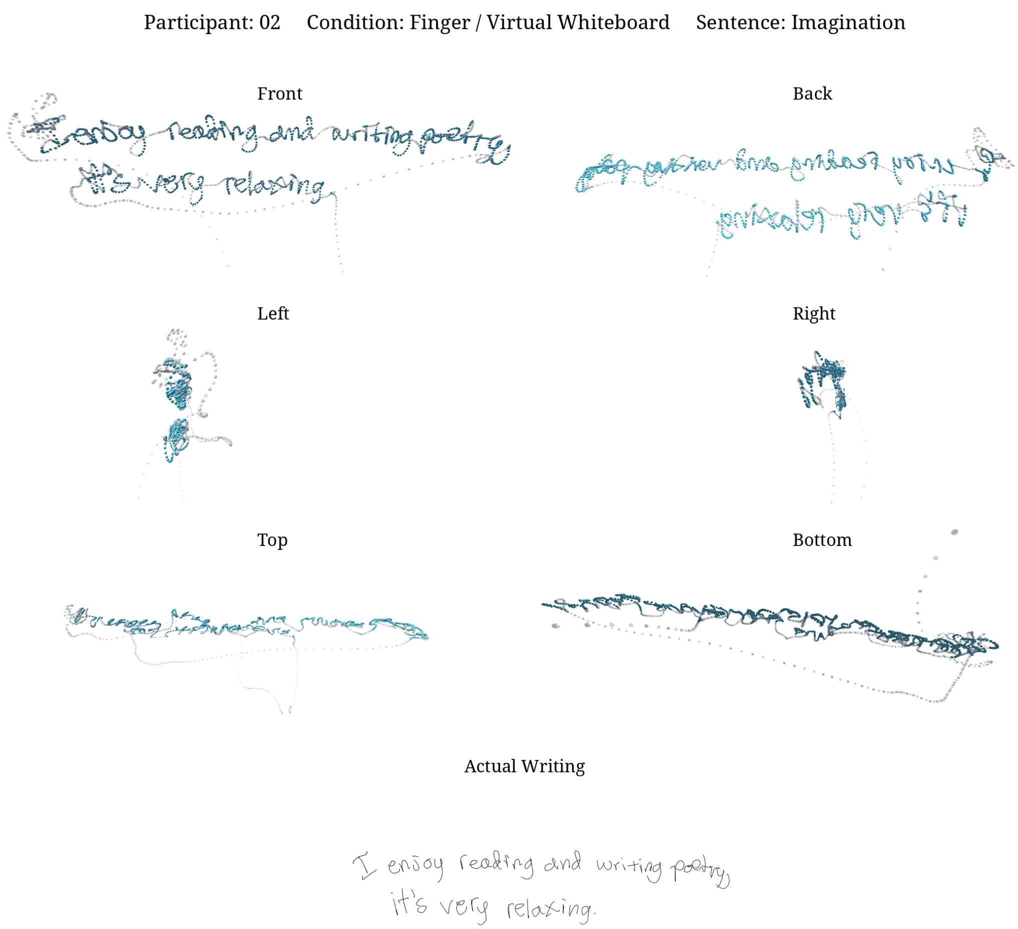


Figure 29: Visualized writing utensil positions for the sentence written by Participant 2, under the Pen/Virtual Whiteboard condition. The Sentence type is Imagination.

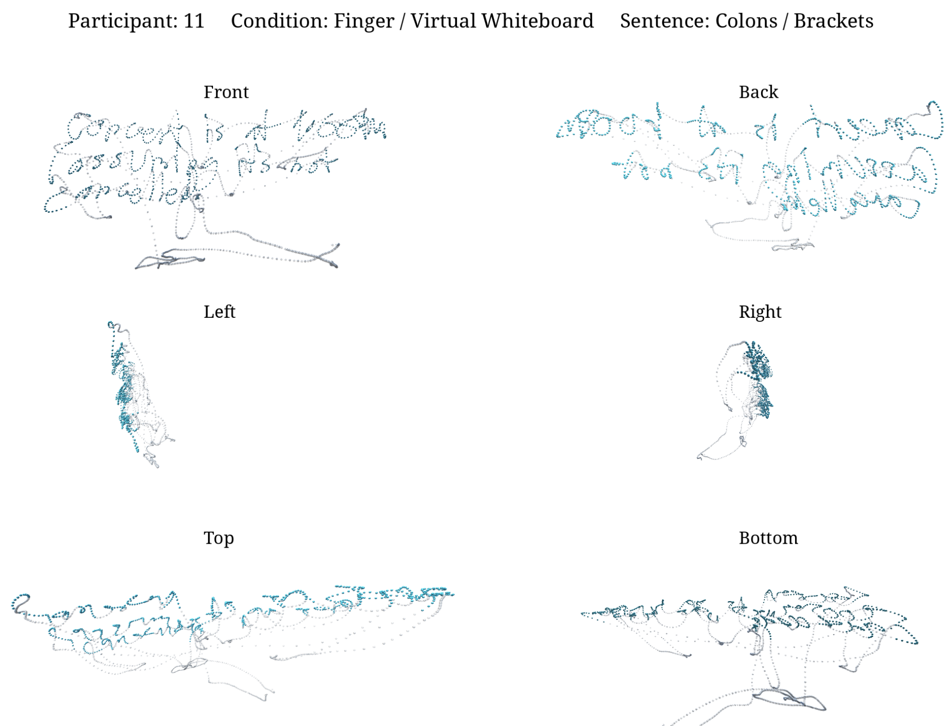


Figure 30: Visualized writing utensil positions for the sentence written by Participant 11, under the Pen/Virtual Whiteboard condition. The Sentence type is Colons/Brackets.

12.4 OTHER OBSERVATIONS

Many other characteristics of midair handwriting were observed when we reviewed the recordings, such as different writing ranges, writing speeds, heights and depths. The following observations are in our supplementary material containing the screen recordings. One observation was that participants wrote with distance between themselves and the virtual whiteboard. This was their decision, as they were informed to set the virtual whiteboard at a distance they preferred. Most participants kept the virtual whiteboard close to them while writing, while a small number (P7, P11, P17) kept a gap (~30 cm) between themselves and the virtual whiteboard. One participant (P20) wrote with the virtual whiteboard far away (~50 cm). When asked why, they responded that they prefer to be able to maintain a wide view of the scene while writing. They also added that it was related to how they write on a whiteboard while teaching.

Another observation was that participants wrote at different levels of elevation, where writing would occur in a range between their eye level and above their head. Most participants wrote at elevations between their eye level and slightly above their heads and were unaffected by the presence of the virtual whiteboard. P20 wrote with their wrist and writing utensils above their head level as if they were projecting the writing from their location. The elevation may affect rotations such as head and wrist rotation but was not investigated in this study. Lastly, different writers also had varying distances between their heads and hands. For example, P6 wrote close to their face while writing in the Pen/No Whiteboard condition but wrote slightly further away for the rest of the conditions. The condition combination order for P6 was (1, 4, 3, 2), so it is possible it was due to an order effect.

13 SUBJECTIVE RESULTS

Our final research question is **RQ2.2**, where our goal was to investigate which writing utensils our participants preferred, as well as whether they preferred writing with or without the virtual whiteboard. To answer this question, we gathered subjective results based on preferred writing utensils and methods through the use of questionnaires. We have two different types of these questionnaires, one for post-technique and one for post-study. The post-study technique surveys were conducted after each participant completed writing sentences for one condition, and the post-study survey was conducted at the end of the trial.

13.1 POST-TECHNIQUE RESULTS

The questions in the post-study questionnaire were in the form of Likert scales, asking different levels of qualitative feedback and an optional field for additional comments. Table 38 describes the questions in the post-technique questionnaire. The subjective responses for all conditions are shown in Figure 31 for the **Pen/No Whiteboard** and **Pen/Virtual Whiteboard** conditions, and Figure 32 for the **Finger/No Whiteboard** and **Finger/Virtual Whiteboard** conditions.

13.1.1 *Survey Results*

Overall, the results show positive feedback across all conditions. In terms of ease of use, most of the responses were in the "Agree" or "Strongly Agree", with no responses in "Strongly disagree". This may be attributed to the transferable skill of midair handwriting, which is something that generally comes naturally for people as long as they're able to write. The same can be said when participants were asked if they found the technique to be comfortable. The pen conditions stood out when no participants answered "Strongly Disagree" when asked if they found any of the conditions with the pen as the writing utensil to be natural. Compared to the conditions where the writing utensil was the finger, a small number of responses said "Strongly Disagree." This may be due to the pen conditions resembling traditional writing while writing with the finger is uncommon in everyday use. Many participants answered "Agree" with some answering "Strongly agree" when asked about how tiring the technique was. This is a downside for midair handwriting in general, as any writer who holds out their arm for extended periods will easily begin to experience fatigue. Nevertheless, there were still many participants that selected "Strongly disagree" and "Disagree", which may be due to different participants having varying limitations in their physical ability. We discuss this in more detail in Chapter 14.

Question	**Response Type**
This technique and condition was easy to use	5-point Likert scale
This technique and condition was tiring	5-point Likert scale
This technique and condition was comfortable	5-point Likert scale
This technique and condition felt natural	5-point Likert scale

Table 38: Questions used for post-technique questionnaire. Note that at the time of the participant study, we referred to "technique" as the writing utensil (Pen/Finger) and "condition" as the method (No Whiteboard/Virtual whiteboard)

13.1.2 *Participant Feedback*

Many participants commented on some difficulty when using the virtual whiteboard writing methods, mainly due to the uncertainty of where the ink would appear when the writing utensil was held a distance away. **P5:** "When starting writing again, it was hard to write at the correct spot.". **P6:** "When writing on whiteboard, you don't see how much your hand moves away.", **P28:** "I kept finding myself trying to register where the whiteboard was, so I tried orienting my hand." There were also comments addressing the issue of the simplistic design of the virtual whiteboard's "ink snap" effect, which caused discomfort for some participants. **P14:** "I expected when I point at the whiteboard for the ink to be directional rather than flat projection" and **P12:** "If I'm drawing in midair, I want the ink to project onto a surface. Something like a dome or even a flat whiteboard with a slight amount of curvature. That could probably make things a bit

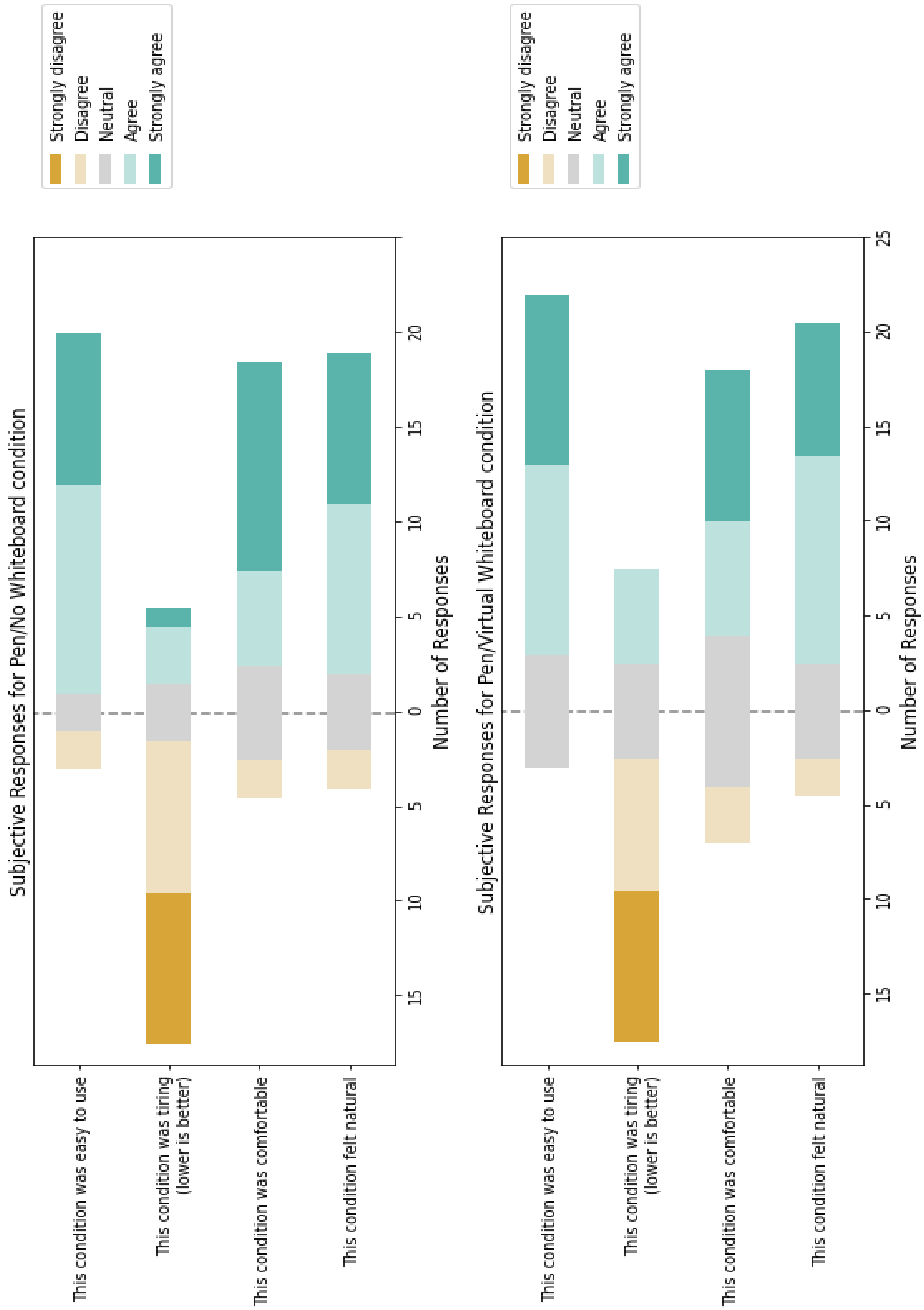


Figure 31: Subjective results for pen writing utensil in both conditions

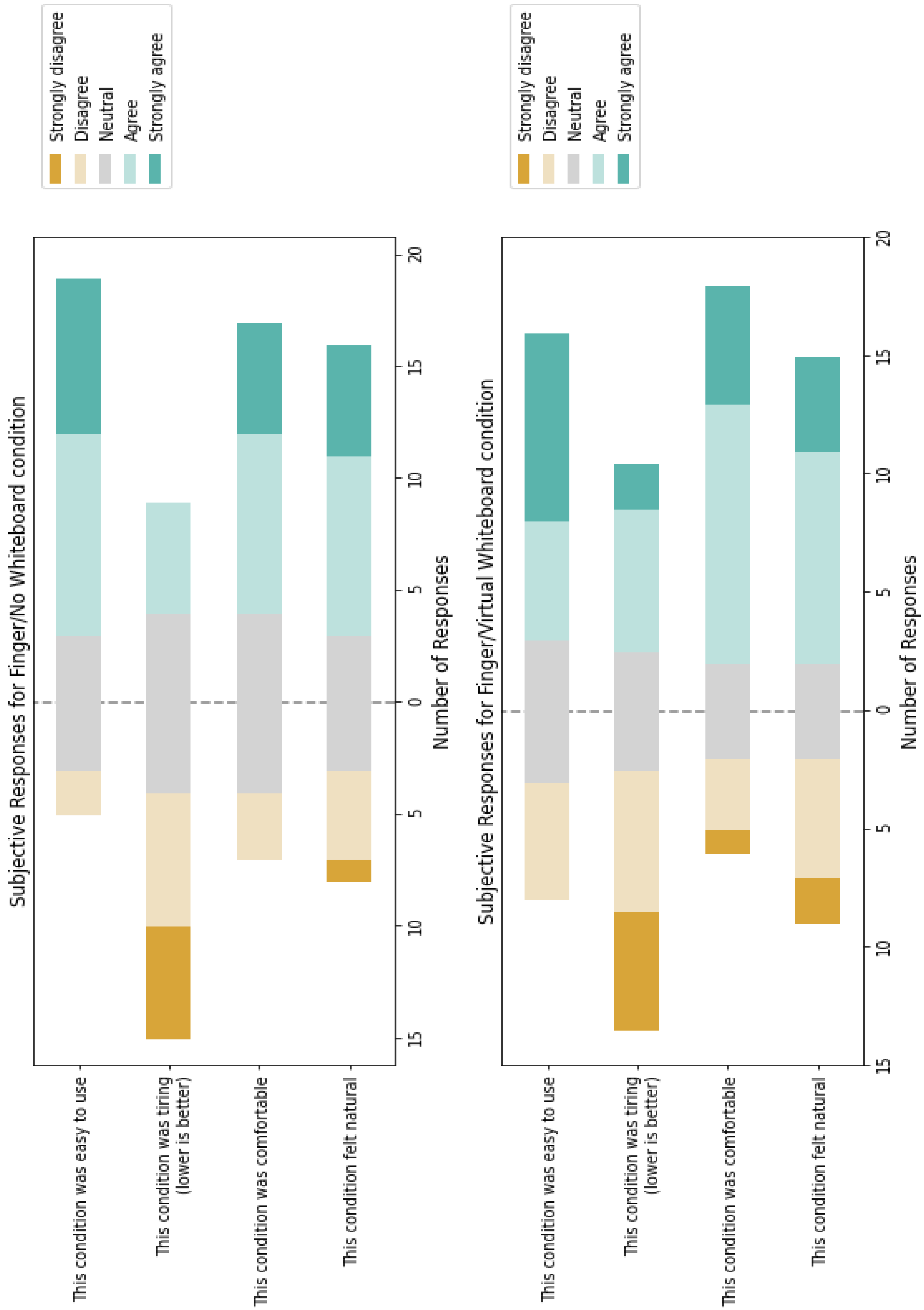


Figure 32: Subjective results for finger writing utensil in both conditions

easier even.". Lastly, some participants brought in comments about the lack of haptic feedback.

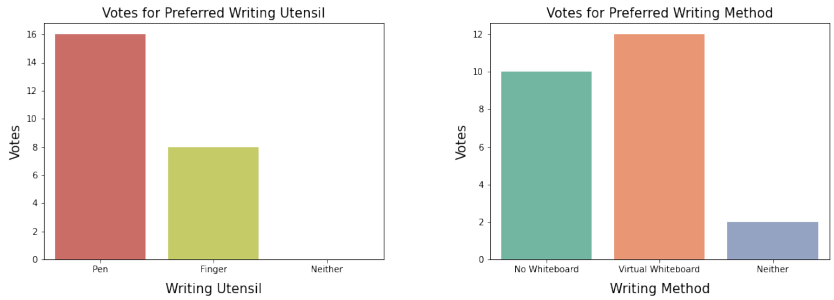
It seems that the simple implementation of the virtual whiteboard caused some drawbacks, as some participants expected it to work differently. The feedback can provide valuable insights into design considerations for future work. All of the comments can be found in our supplementary material.

13.2 preferred writing utensil and conditions

The post-study questionnaire asked participants for their favourite combination of writing utensils and writing methods at the end of the study. The results are shown in Figure 33, and the aggregated results are shown in Figure 34. From the results, we can see that the pen writing utensil was preferred over the finger, which is most likely due to it being closely related to the traditional method of writing. While the Virtual Whiteboard condition has more votes than the No Whiteboard condition, the combined results show that they are equal when the pen is the writing utensil.

Additionally, the feedback provides reasons as to why the pen is preferred over the finger. Some comments from participants who preferred the pen include **P3:** "The pen felt way more natural than the finger because I'm used to using pens to write in real life. The finger did feel like it had a bit more control/accuracy in free-hand, but it felt weird.". **P7:** "The pen felt more comfortable and natural. It felt to me like my writing was neater.". On the other hand, some preferred the finger, such as **P15:** "The finger felt more natural, and it allowed the benefit of not holding anything to write as that seemed more of a hassle." **P14:** "Felt less restrictive." (referring to the finger as the writing utensil). From these comments, we can conclude that the pen would be ideal as a straightforward option as it reflected the traditional method of writing. At the same time, the finger provided a lightweight alternative.

The number of votes for the two writing conditions is relatively close. Those who preferred having No Whiteboard, such as **P26** said, "I liked the free hand because I didn't feel constricted on where I had to write, as in, there were no set margins.". Note that free hand refers to no whiteboard. **P28** said, "The free hand made my writing look more natural and flowed better, which made me feel more confident



(a) (b)

Figure 33: Subjective results showing votes for the preferred writing utensil (left) and votes for the preferred writing method (right).

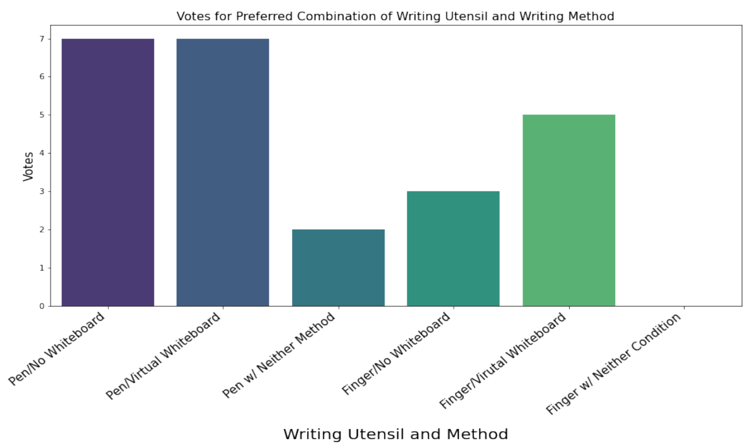


Figure 34: Aggregated results for most favoured writing conditions.

in my writing.". On the other hand, those who preferred the virtual whiteboard, such as **10:** "It made all the writing on a visible plane, which made it nicer to look at. I felt like I could go back and correct errors with much more ease than the free hand condition", and **P8:** "The whiteboard helped me keep my sentences in line with each other and made it feel more natural and close to writing on paper or writing on a whiteboard in real life." Based on the comments, participants who preferred the No Whiteboard methods preferred the freedom of being able to write in the 3D space without any boundary. Meanwhile, those who preferred the virtual whiteboard liked to use it as a form of guidance, which helped them write neatly. The close number of votes and comments suggests that both conditions have their pros and cons, which can encourage further investigation and provide design considerations.

14

DISCUSSION & LIMITATIONS

This chapter discusses the findings from the results, limitations, contributions, conclusions, and ideas for future work.

14.1 PREDICTING INTENT OF MIDAIR HANDWRITING

With an overall accuracy of ~85%, this showed promising results in predicting the intent of midair handwriting, which can serve as a starting point for this type of research. According to our aggregated results in Chapter 11, the metrics across all conditions appeared to be relatively close.

Using the same neural network architecture to train on data across all conditions, it was discovered that the different conditions did not have much of an impact. We investigated this further by conducting additional tests where the same models were tested on data from conditions they were not trained on. Our results from our additional tests showed that despite the models being tested on writing conditions they had never seen, the results still averaged within the same range of results across all metrics. We expected the models to perform worse in these additional tests since our visualizations from Chapter 12 showed that the presence of a virtual whiteboard could indeed affect the curvature of writing. We suspected that close results across all tests might be due to the small window size of 19-time steps. After plotting out the z-positions of a participant's writing utensil for a random sentence, we discovered that there was little change in the z-positions across 19-time steps. Compared to plotting across 300-time steps, there was much more change. Ideally, we would redo the data processing, model training and testing with varying window sizes. Unfortunately, due to time constraints, we could only gather results from a window size of 19 time steps. Thus, we conclude that our current methods can be refined based on our findings to achieve potentially better results.

Our findings comparing the two label selection methods, selecting the majority Boolean and selecting the last Boolean per sequence, showed that selecting the majority Boolean yielded better results overall. Additionally, from our confusion matrices, we found that our tests generally performed better in predicting true negatives. At the same time, there were varying outcomes when predicting true positives - some better and some worse. We cannot say with certainty as to why this is happening, but we hypothesize that this may be due to the higher number of false instances overall in our data. Further data cleanup should be implemented as it was later discovered that accidental writing was performed. These instances were included as part of our training and testing data.

While an accuracy of ~85% is relatively high, it would not be sufficient to implement in an actual application. Further research, training and testing should be conducted to create more accurate models. Looking at the metrics, the precision is overall lower than recall, which is less favourable in this application as higher precision would be more useful to detect the intent of midair handwriting. In an actual application, this can lead to frustration as ink may be placed during instances when the writer does not intend to write.

When grouped under different test cases, our results show very small differences, with **Test case 2** having the worst results. We believe this was due to the model being tested on sentences written by both new participants and sentences that fall under new sentence types. The models encountered only new participants or new sentence types compared to the other test cases. Even though the differences were small, this created a hypothesis that there were indeed patterns that the models picked up when training on data from different participants and sentences. If more data is gathered and the same tests are performed and result in larger differences, we can conclude with a more discerning conclusion.

Comparing our results from LSTM models to our preliminary tests, there is a clear distinction between positive results from using an LSTM neural network and classifiers. Additionally, making predictions using data in the form of sequences has proven to be more effective than predicting individual time steps. This means that valuable information can be retrieved from previous time steps during midair handwriting to make more accurate predictions. A limitation in our tests is finding which features are most important. It is possible that our neural network only saw one or a small number of features as important, which was not investigated as it was outside of the scope of our study. However, there are possible ways to identify the important features by using methods such as permutation importance and even feature importance specific for LSTM [55] networks.

While our LSTM neural network was able to train models to achieve relatively high accuracies, more options may yield better results. For example, transformers such as DistilBERT [54] can perform sequence classification. Although these transformers were originally intended for natural language processing, it is possible to re-purpose our sequenced data to perform classification using transformers. This act of using transformers for different classification tasks is also present in previous works, such as in cyber security and malware detection [44]. Alongside these are other options, such as attention models, which can be effective at identifying the important parts of our data and making accurate predictions.

Additionally, using the Matthews Correlation Coefficient to evaluate the models would provide better insights into the overall performance, as it is more effective at evaluating models with imbalanced classes. This was not known during the time of the study.

The last limitation is related to the data we gathered, as the sampling rate from Optitrack was lower than Unity's sampling rate, as mentioned in Section 10.2.1. Due to the lowered sampling rate, the objects tracked by Optitrack would come as duplicates of every other frame and require interpolation to smooth out the data, potentially affecting model training and testing. Had both systems been able to run at 90 Hz, this step would not have been needed. It was discovered in the later phase of the project that an upgraded software could resolve these issues, which must be ensured to guarantee the highest data quality.

14.1.1 *Various Writing Patterns*

All participants wrote in many different writing styles based on our results in Section 12.2. Although it is evident that the presence of a virtual whiteboard helped participants write with a flat characteristic, there were still degrees to the curvature we observed from different participant's handwriting. Some participants could still write somewhat flat without the virtual whiteboard, while others still wrote with a degree of curvature even when a virtual whiteboard was present. The same can be applied to the stair-casing pattern that we observed, which followed the same trend of being affected by the virtual whiteboard. These various writing patterns invite much investigation into possible causes. We could not investigate the causes in this study, but we believe much has to do with each individual's natural writing style. Perhaps these findings may be of more value if future research aims to correct the depth to make midair handwriting more legible. In addition to writing curvature, we observed many other characteristics, such as the range of writing, writing speeds, heights and depth. Additionally, the ergonomics of the VR headset and tools used in the experiments such as the wrist tracker and writing utensil could also factor into how people wrote differently. Lastly, each participant's experience writing in English can be another factor, as English may not be their primary language and could have been a possible influence on how each participant wrote. These characteristics may relate to how an individual writes, but whether they affected the data is unsure as they have not been investigated in this study.

A limitation of our participant study was that the new medium of writing in midair could have been an uncanny writing experience, which the participants might have had trouble adjusting to. In future studies, we encourage more time for participants to become adjusted to writing in midair. Additionally, using different technologies may be beneficial as this study served as many of the participant's first time using VR. We aim to provide a writing experience to be as comfortable as possible, and VR may not be suitable for some individuals. Perhaps a technology such as AR can help participants adjust more easily.

Another limitation regarding the many different writing patterns is the resulting legibility. We have not analyzed the legibility of each participant's handwriting across the different writing conditions. This can be important for future research especially when it comes to making midair handwriting presentable.

14.1.2 *Preferred Writing Methods*

Most participants agreed that all conditions were easy to use, which is a positive sign that many people can easily adapt a midair handwriting application, especially if it is their first time using this technology. Because of this, we can confirm that this transferable skill of traditional handwriting on paper and/or whiteboard can also be applied to midair. Across all conditions, many participants answered "Neutral" when asked if they found the technique to be comfortable. We strongly believe this was a drawback of midair handwriting since writing in this fashion required participants to keep their arms raised for a long time. Even though they were given breaks in between sentences and writing conditions, participants were still required to write many sentences within an hour. In the scope of our research, we aimed to focus on predicting the intent of midair handwriting for **sentences**. But we hypothesize that if this application were to be applied in a real-life scenario, it would be ideal for it to be used for writing short sentences, notes, annotations or even sketches [50]. Therefore, if another study were to be conducted, stimuli that encourage this type of quick writing would be more accurate for a real-life scenario and may even affect test results differently.

An interesting finding was that none of the participants selected "Strongly disagree" when asked if they found the pen conditions to be natural, which were most likely due to its close relation to traditional handwriting. In contrast, we received a small number of "Strongly disagree" when the writing utensil in question was the finger, which was further from the traditional method of writing, as stated in some of the comments. This would seem to be something obvious as we do not normally write with our fingers. While the pen writing utensil was more favoured than the finger, there was still much participant feedback on why they liked using their finger as the writing utensil. Most notable was the fact that the finger provided a lightweight means of writing, as the only weight applied to the participant was the finger tracker, which weighed 7 grams. Many participants said they liked the lightweight design and not worrying about holding onto an object to write. On the other hand, most participants preferred the pen because it closely resembles how people normally write daily. The pen gave the participants more control and a natural feeling as they wrote in midair.

Thus, here lies a trade-off when designing a midair handwriting system where one can utilize a writing utensil based on the traditional writing method or have a writing utensil that requires little effort. While the results show that one writing utensil was favoured over the other, the mixed feedback created room for investigating which writing utensils can be advantageous in different settings. It is possible that participants preferred the pen due to the design of our study and that participants may prefer the finger-writing utensil in a different type of study. Participants stated that they preferred having no virtual whiteboard because writing in the 3D space allowed them to write freely, although at the cost of the neatness of their handwriting. The participants ' writing space was restricted compared to writing with the virtual whiteboard, but they were provided with a way to write neater in the air. Hence, depending on the scenario, one writing method can be more advantageous than the other. If this application was applicable in real life, a user with a lot of space in their vicinity may prefer having all the space to write. While a user who aims to write in a presentable manner may prefer to write using a virtual whiteboard.

More considerations can be taken for the design of the application. As per the participant feedback, providing haptic feedback can accurately indicate when the ink is being placed in the air. The virtual whiteboard can be improved to accurately project the ink from the writing utensil's position and angle to the whiteboard and a cursor that visualizes where the ink would land. This would provide a more natural and easier writing experience with the virtual whiteboard. The simplicity of our virtual whiteboard also serves as a limitation in our study.

Another area of interest was to find a way to correct the deviations in the z-axis during midair handwriting, such as the curvatures seen in Section 12.2. This was an initial goal in our study where our idea was to make midair handwriting legible, hence why we recorded the stroke positions shown in appendix Table 40. However, we could not investigate this area due to scope and time limitations. Future research can use this data to investigate how to correct said stroke positions.

Lastly, we could not find links between the participant's writing patterns and preferences of writing conditions with details they answered in the pre-study questionnaire. These details included their past experiences with VR & writing on whiteboards. We did not include demographics in our questionnaires, asking participants about their English background, such as their English writing level and whether it was the first language they learned. Such information could also be valuable in analyzing writing styles and methods for each individual. There is potentially more to uncover in this area, which could encourage future research to investigate these writing patterns and preferences in midair handwriting.

14.2 additional use cases

The findings in our study are not limited to detecting the intention of midair handwriting. The applications of machine learning in our research can be extended to determine when a whiteboard should be instantiated after a writer lifts their writing utensil into the air. Additionally, our models have only been trained in writing short sentences. In future work, we can gather more data to train models that can detect any type of writing, such as point form notes and sketches. Our findings regarding writing patterns can also help future research in designing virtual whiteboards. It was evident that there were different levels of writing curvature when writing in the air, which can invite further investigation, as the various writing patterns can be potential areas of future research.

15 CONCLUSION

15.1 CONTRIBUTIONS

This chapter discusses the contributions of our study and our concluding remarks.

15.1.1 *Data Collection Study*

This thesis aimed to 1: Investigate how we can predict the intent of unconstrained midair handwriting and 2: Investigate whether the presence of a virtual whiteboard can impact how people write in midair, as well as any patterns that can be found. Since this type of data was not publicly available, we conducted a data-gathering study to gather this type of data. We provided the design of our study, the stimuli and how the study was conducted. From the study, we gathered both quantitative and qualitative data. The quantitative data was used to perform tests to answer the first question, and the qualitative data was used to answer the second.

15.1.1.1 *Data Cleanup and Processing*

Our published data consists of the raw data gathered from the experiments, as well as the data after cleanup and processing, prior to being used for machine learning. We took the steps necessary to remove outliers and any values that would disrupt the machine-learning process. We also standardized our data, ensuring that it is mean-centred with unit variance.

15.1.1.2 *Data Gathering Application*

The main application we developed was a prototype VR application that enabled any user to perform midair handwriting. This Unity application works alongside the Optitrack motion capture cameras in our lab to track our trackers, which we use for our writing utensils. We also included the design of our pen and finger trackers. The application can also be used on other setups, provided that they have the required hardware. This application can record midair handwriting data and save them as CSV files. In our study, we gathered data from 24 participants who wrote 20 sentences each, reaching a total of 480 sentences in our collection.

15.1.1.3 *Data Viewing Application*

We developed another application in Unity that can read the CSV files exported by our main application to recreate the scene for individual sentences. This was primarily used to take screenshots of different writing patterns from 6 different views, with an extra screenshot for the *Imagination* sentences that contained the original writing, which we used to serve as the ground truth. This totalled up to 2977 screenshots to compose 480 pictures, which outlines the various writing styles throughout our experiment.

15.1.2 *Predicting Intent of Midair Handwriting via Machine Learning*

Our main contribution was performing tests by predicting the intent of midair handwriting. We have conducted preliminary tests where we tested the feasibility of performing binary classification on individual time steps. Afterwards, we converted our data into sequences so our neural network can perform binary classification on data sequences rather than individual time steps. We created a neural network composed of LSTM layers and optimized the architecture and parameters through several testing iterations. We also conducted tests in various manners, training individual models for each condition and testing them on their respective & opposite conditions. We also divided the tests into different test cases, testing the model when exposed to different types of seen and unseen data. Based on our results, the different conditions have little impact on the accuracy and other metrics, including precision, recall, f1-score and AUC-ROC. It was also found that the test case when models were tested on new participants and new sentences had the lowest results out of all the tests, which stemmed a hypothesis that the difference can be more significant if more data is gathered. Despite our small amount of data, we achieved these findings through k-fold cross-validation and tests performed on individual sentences.

15.1.2.1 *Analysis of Midair Handwriting Patterns*

Our second biggest contribution was the discovery of different midair handwriting patterns that we found when we used our data-viewing application to recreate the sentences. We found many different types of patterns that can be present for different participants during midair handwriting, including the level of curvature, elevation and speed. Our results showed that while participants wrote in midair, they generally wrote in a sphere around them, where they would be the sphere's center. The text would curve when writing from left to right, while new lines of text also tend to follow the curvature of the sphere – usually coming closer to the participant in a stair-casing fashion. It was found that the presence of a virtual whiteboard tends to help writers keep their text flat despite not having any physical surface. However, this curvature and stair-casing varies in a spectrum across different participants as some participants still wrote with a level of curvature with the virtual whiteboard, while some participants managed to write relatively straight without the virtual whiteboard.

15.1.2.2 *Analysis of Subjective Feedback*

Another important aspect of our study was to determine what method of writing participants preferred while performing midair handwriting. We discovered that most participants preferred using the pen as the writing utensil as it closely resembled traditional handwriting. While fewer participants preferred using their fingers to write, it showed value as a lightweight method for midair handwriting, as many participants stated in the comments that they liked the idea of not having to hold onto a writing utensil. In terms of the virtual whiteboard, an equal number of participants preferred either writing with or without the virtual whiteboard. This was due to many reasons, primarily that having no virtual whiteboard provided the most freedom, while the virtual whiteboard helped participants keep their midair writing neat. There is no definitive answer as to which combination of writing utensils and methods is triumphant, but the feedback we received can provide design considerations depending on the need.

15.1.3 *Future Work*

Many improvements can be made in future work. A longer study with more participants and stimuli allows us to work with more data. We can improve our neural network to be more advanced or use alternatives mentioned in Section 14.1, which can potentially yield better results. Most importantly, after achieving better results, the model can be integrated with the Unity application to create a system that can be tested. Our virtual whiteboard should also be updated to project the ink to the whiteboard along the user's line of sight, potentially producing a more fluid virtual whiteboard experience. Afterwards, a study that further investigates the two writing utensils and two writing conditions can be conducted, as there are many strengths and weaknesses that can be compared.

15.1.4 *Supplementary Material*

Our supplementary material consists of spreadsheets for all of our machine-learning tests, as well as the confusion matrices and AUC-ROC graphs. We also include the 480 sentences in the form of 6 different views composed of 2977 screenshots. Additionally, it contains the original CSV files for the data recorded for the sentences and questionnaire answers and screen recordings that were recorded during the participant trials.

15.2 CONCLUSION

The results from our data gathering, processing and analysis show promise in predicting the intention of midair handwriting without constraints applied to where and how writing can be done, as long as they are within the midair handwriting environment. This handwriting method allows users to pick up a utensil and write using their movements and eye tracking. Our $\sim 85\%$ accuracy shows that this method of prediction is possible with room for improvement. We discovered many different patterns and preferences regarding midair handwriting, which opened new doors for investigation in the field of Human Computer Interaction. In summary, our work introduces new possibilities for supporting unconstrained midair handwriting in different ways, which has the potential to perform writing anywhere in an open space quickly and efficiently.

Part I

APPENDIX

You can put some informational part preamble text here. Illo principalmente su nos. Non message *occidental* anglo-romanic da. Debitas effortio simplificate sia se, auxiliar summarios da que, se avantiate publicationes via. Pan in terra summarios, capital interlingua se que. Al via multo esser specimen, campo responder que da. Le usate medical addresses pro, europa origine sanctificate nos se.

A

APPENDIX

A.1 COORDINATE MISALIGNMENT

Using the Optitrack CS-200 calibration square [8], we were able to set the Optitrack's coordinate system center of origin as close as possible to the Unity scene's center of origin. We also aligned the calibration square so that the direction of the positive x, y, and z axes from Optitrack's coordinate system was as close as possible to Unity's. The process of aligning these two coordinate systems was done by hand, and the center of origin was located in the real world, according to the Unity scene. To find the positive x, y and z directions in the Unity scene concerning the real world, we moved an HTC Vive remote from the center of the Unity scene while measuring the distance recorded in Unity. The recorded vector basis for the Unity scene is shown in Figure 8.

It was found that even after using the calibration square to set the center of origin and basis vectors of Optitrack's coordinate system, there was still a misalignment which grew more significant when moving further from the center of the 3D scene. This suggested that the problem could be due to both coordinate systems' rotation and/or scaling differences. We isolated the problem to a rotation difference by moving and measuring a tracked object 1 meter in the real world in different directions, which was reflected accurately in both the Unity and Optitrack systems. The solution to align both systems was to find the rotation matrix and translation vector that can be used on incoming Optitrack coordinates so they can be aligned to Unity.

Arun et al. [6] proposed a least-squares fitting method for two 3D point sets. Their solution was implemented in Python by nghia012 [56] on Github, and the script would return the optimal rigid transformation needed to align Optitrack coordinates to Unity coordinates. The first step was to gather data from both coordinate systems. We taped an Optitrack marker on top of a Vive remote to record these coordinates. Then, starting at the Unity scene's center, we recorded 1564 points by moving the Vive remote around in a sphere while extending and retracting the arm. This was done to record as many points as possible around the general writing area. After gathering the data, we used the script written by nghia012 [56] to perform mean-centering on both sets, followed by computing the covariance matrix and performing singular value decomposition (SVD) to isolate the left-singular vectors and right-singular vectors. The vectors represent the rotation vectors for both coordinate systems, and they were multiplied to obtain the rotation matrix needed to get from Optitrack to Unity.

Table 39: Prompts used to generate stimuli using ChatGPT, prior to modifications.

Prompt	Write me a sentence that begins with parenthesis
ChatGPT	(According to recent studies), regular exercise has been shown to improve cognitive function and reduce the risk of chronic diseases.
Prompt	Generate a short sentence that contains colons and brackets
ChatGPT	The recipe called for three main ingredients: flour, sugar, and butter (preferably unsalted).
Prompt	Generate a short sentence that contains a colon and brackets
ChatGPT	I have a message for you: (Remember to stay positive and keep pushing forward).
Prompt	Generate a short sentence that includes a colon and brackets
ChatGPT	The concert lineup was impressive: (rock bands, pop artists, and a surprise guest performer) will take the stage.
Prompt	Generate a few pangrams
ChatGPT	Certainly! Here are a few pangrams for you:

1. The quick brown fox jumps over the lazy dog.
2. Pack my box with five dozen liquor jugs.
3. Mr. Jock, TV quiz PhD, bags few lynx.
4. Crazy Fredericka bought many very exquisite opal jewels.
5. Jackdaws love my big sphinx of quartz.
6. The five boxing wizards jump quickly.
7. How quickly daft jumping zebras vex.
8. Waltz, bad nymph, for quick jigs vex.
9. Quick zephyrs blow, vexing daft Jim.
10. Bright vixens jump; dozy fowl quack.

I hope you find these pangrams interesting!

Prompt | Please generate some more

ChatGPT | Certainly! Here are a few more pangrams for you:

1. How vexingly quick daft zebras jump!
2. Jumpy veldt fox bids a quick phlegm zap.
3. Fix problem quickly with galvanized jets.
4. My jazzy elephant gave the big, lazy kangaroo four pizzas.
5. The wizard quickly jinxed the gnomes before they vaporized.
6. Quirky whims beggar exciting vixens.
7. Few black taxis drive up major roads on quiet hazy nights.
8. Jumping monkeys quickly seize the waxy bouquet.
9. A crazy gazelle jumps over a quivering fox.
10. Six big juicy steaks sizzled in a pan.

Data	Data Type	Used for Features
Frame	Integer	N
Timestamp	String	Y
Right handedness	Boolean	N
Head position	3D Vector (x, y, z)	Y
Writing utensil position	3D Vector (x, y, z)	Y
Wrist position	3D Vector (x, y, z)	Y
Gaze position	3D Vector (x, y, z)	Y
Alternate controller position	3D Vector(x, y, z)	N
Head rotation	Quaternion (w, x, y, z)	Y
Writing utensil rotation	Quaternion (w, x, y, z)	Y
Wrist rotation	Quaternion (w, x, y, z)	Y
Alternate controller rotation	Quaternion (w, x, y, z)	N
Left pupil position	Float	Y
Right pupil position	Float	Y
Left eye gaze origin	3D Vector (x, y, z)	N
Right eye gaze origin	3D Vector (x, y, z)	N
Left eye gaze direction	3D Vector (x, y, z)	N
Right eye gaze direction	3D Vector (x, y, z)	N
Left pupil diameter	Float	Y
Right pupil diameter	Float	Y
Left eye openness	Float	N
Right eye openness	Float	N
Left eye wideness	Float	N
Right eye wideness	Float	N
Left eye squeeze	Float	N
Right eye squeeze	Float	N
Ink Activated	Boolean	Y (Used for label)
Stroke number	Integer	N
Stroke positions	Array of 3D Vectors	N

Table 40: All the recorded data per sentence written by participants, including ones that were not used for data processing. Ink Activated is the only label.

Predicting Intent of Midair Handwriting using Logistic Regression

Condition: Pen/No Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
Same Participants New Sentence Types	0.67	0.67	0.67	0.66	0.65
New Participants New Sentence Types	0.66	0.66	0.66	0.66	0.64
Same Participants New Sentence Types	0.67	0.67	0.67	0.67	0.64

Table 41: Aggregated results from tests using logistic regression on data from Pen/No Whiteboard.

Predicting Intent of Midair Handwriting using Simple Neural Network

Condition: Pen/No Whiteboard			Label Selection Method: Majority		
Test Case	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
Same Participants New Sentence Types	0.79	0.74	0.74	0.73	0.88
New Participants New Sentence Types	0.78	0.74	0.71	0.72	0.87
Same Participants New Sentence Types	0.80	0.76	0.73	0.74	0.88

Table 42: Aggregated results from tests for models trained on data from Pen/No Whiteboard sentences.

BIBLIOGRAPHY

[1] Sylvia Ahern and Jackson Beatty. "Pupillary responses during information processing vary with Scholastic Aptitude Test scores." In: *Science* 205.4412 (1979), pp. 1289–1292.

[2] Emre Aksan, Fabrizio Pece, and Otmar Hilliges. "Deepwriting: Making digital ink editable via deep generative modeling." In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 2018, pp. 1–14.

[3] Denis Alamargot, Christophe Dansac, David Chesnet, and Michel Fayol. "Parallel processing before and after pauses: A combined analysis of graphomotor and eye movements during procedural text production." In: *Writing and Cognition*. Brill, 2007, pp. 11–29.

[4] Christoph Amma, Marcus Georgi, Tomt Lenz, and Fabian Winnen. "Kinemic wave: A mobile freehand gesture and text-entry system." In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2016, pp. 3639–3642.

[5] Christoph Amma, Marcus Georgi, and Tanja Sc hultz. "Airwriting: Hands-free mobile text input by spotting and continuous recognition of 3D-space handwriting with inertial sensors." In: *2012 16th International Symposium on Wearable Computers*. IEEE. 2012, pp. 52–59.

[6] K Somani Arun, Thomas S Huang, and Steven D Blostein. "Least-squares fitting of two 3-D point sets." In: *IEEE Transactions on pattern analysis and machine intelligence* 5 (1987), pp. 698–700.

[7] Thorsten Brants and Alex Franz. "Web 1T 5-gram Ver. 1." In: *LDC2006T13, Linguistic Data Consortium, Philadelphia* (2006).

[8] *Calibration Tools*. https://optitrack.com/accessories/calibration-tools/. Accessed: 2023-11-15.

[9] Leith KY Chan and Henry YK Lau. "MagicPad: the projection based 3D user interface." In: *International Journal on Interactive Design and Manufacturing (IJIDeM)* 6 (2012), pp. 75–81.

[10] *ChatGPT*. https://chat.openai.com/. Accessed: 2023-11-16.

[11] Mingyu Chen, Ghassan AlRegib, and Biing-Hwang Juang. "Air-writing recognition—Part I: Modeling and recognition of characters, words, and connecting motions." In: *IEEE Transactions on Human-Machine Systems* 46.3 (2015), pp. 403–413.

[12] Mingyu Chen, Ghassan AlRegib, and Biing-Hwang Juang. "Air-writing recognition—Part II: Detection and recognition of writing activity in continuous stream of motion data." In: *IEEE Transactions on Human-Machine Systems* 46.3 (2015), pp. 436–444.

[13] Ananya Choudhury and Kandarpa Kumar Sarma. "A Novel Approach for Gesture Spotting in an Assamese Gesture-Based Character Recognition System using a Unique Geometrical Feature Set." In: *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*. 2018, pp. 98–104. DOI: 10.1109/SPIN.2018.8474285.

[14] Çağla Çığ and Tevfik Metin Sezgin. "Gaze-based prediction of pen-based virtual interaction tasks." In: *International Journal of Human-Computer Studies* 73 (2015), pp. 91–106.

[15] *Clasification on imbalanced data*. https://www.tensorflow.org/tutorials/structured_data/imbalanced_data. Accessed: 2023-12-05.

[16] Andrew T Duchowski, Krzysztof Krejtz, Matias Volonte, Chris J Hughes, Marta Brescia-Zapata, and Pilar Orero. "3D gaze in virtual reality: vergence, calibration, event detection." In: *Procedia Computer Science* 207 (2022), pp. 1641–1648.

[17] *Eye and Facial Tracking SDK (Legacy)*. https://developer-express.vive.com/resources/vive-sense/eye-and-facial-tracking-sdk//. Accessed: 2023-11-15.

[18] Zhangjie Fu, Jiashuang Xu, Zhuangdi Zhu, Alex X. Liu, and Xingming Sun. "Writing in the Air with WiFi Signals for Virtual Reality Devices." In: *IEEE Transactions on Mobile Computing* 18.2 (2019), pp. 473–484. DOI: 10.1109/TMC.2018.2831709.

[19] Sabrina Gerth, Thomas Dolk, Annegret Klassert, Michael Fliesser, Martin H Fischer, Guido Nottbusch, and Julia Festman. "Adapting to the surface: A comparison of handwriting measures when writing on a tablet computer and on paper." In: *Human Movement Science* 48 (2016), pp. 62–73.

[20] Claire C Gordon, Cynthia L Blackwell, Bruce Bradtmiller, Joseph L Parham, Patricia Barrientos, Stephen P Paquette, Brian D Corner, Jeremy M Carson, Joseph C Venezia, Belva M Rockwell, et al. "2012 anthropometric survey of us army personnel: Methods and summary statistics." In: *Army Natick Soldier Research Development and Engineering Center MA, Tech. Rep* (2014).

[21] Eric Granholm, Robert F Asarnow, Andrew J Sarkin, and Karen L Dykes. "Pupillary responses index cognitive resource limitations." In: *Psychophysiology* 33.4 (1996), pp. 457–461.

[22] Zhengxin Guo, Fu Xiao, Biyun Sheng, Huan Fei, and Shui Yu. "WiReader: Adaptive air handwriting recognition based on commercial WiFi signal." In: *IEEE Internet of Things Journal* 7.10 (2020), pp. 10483–10494.

[23] Aakar Gupta, Cheng Ji, Hui-Shyong Yeo, Aaron Quigley, and Daniel Vogel. "Rotoswype: Word-gesture typing using a ring." In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. 2019, pp. 1–12.

[24] Douglas J Hacker, Matt C Keener, and John C Kircher. "TRAKTEXT: Investigating writing processes using eye-tracking technology." In: *Methodological Innovations* 10.2 (2017), p. 2059799116689574.

[25] Zhenyi He, Ruofei Du, and Ken Perlin. "Collabovr: A reconfigurable framework for creative collaboration in virtual reality." In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2020, pp. 542–554.

[26] *Htc Vive Pro Eye*. https://www.vive.com/sea/product/vive-pro-eye/overview/. Accessed: 2023-11-15.

[27] Fu-An Huang, Chung-Yen Su, and Tsai-Te Chu. "Kinect-based mid-air handwritten digit recognition using multiple segments and scaled coding." In: *2013 International Symposium on Intelligent Signal Processing and Communication Systems*. IEEE. 2013, pp. 694–697.

[28] Muhammad Zahid Iqbal and Abraham G Campbell. "From luxury to necessity: Progress of touchless interaction technology." In: *Technology in Society* 67 (2021), p. 101796.

[29] Muhammad Zahid Iqbal and Abraham Campbell. "The emerging need for touchless interaction technologies." In: *Interactions* 27.4 (2020), pp. 51–52.

[30] Marcel Adam Just and Patricia A Carpenter. "The intensity dimension of thought: Pupillometric indices of sentence processing." In: *Reading and language processing*. Psychology Press, 2013, pp. 182–211.

[31] Daniel Kahneman and Jackson Beatty. "Pupil diameter and load on memory." In: *Science* 154.3756 (1966), pp. 1583–1585.

[32] Çağla Çığ Karaman and Tevfik Metin Sezgin. "Gaze-based predictive user interfaces: Visualizing user intentions in the presence of uncertainty." In: *International Journal of Human-Computer Studies* 111 (2018), pp. 78–91.

[33] Florian Kern, Peter Kullmann, Elisabeth Ganal, Kristof Korwisi, René Stingl, Florian Niebling, and Marc Erich Latoschik. "Off-the-shelf stylus: Using xr devices for handwriting and sketching on physically aligned virtual surfaces." In: *Frontiers in Virtual Reality* 2 (2021), p. 684498.

[34] *Kinect*. https://learn.microsoft.com/en-us/windows/apps/design/devices/kinect-for-windows. Accessed: 2023-11-13.

[35] Michal Lech, Bozena Kostek, and Andrzej Czyzewski. "Virtual Whiteboard: A gesture-controlled pen-free tool emulating school whiteboard." In: *Intelligent Decision Technologies* 6.2 (2012), pp. 161–169.

[36] Anders Markussen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. "Vulture: a mid-air word-gesture keyboard." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014, pp. 1073–1082.

[37] *Meta Horizon Workrooms*. https://forwork.meta.com/horizon-workrooms/. Accessed: 2023-06-06.

[38] *Motive*. https://optitrack.com/software/motive/. Accessed: 2023-11-15.

[39] Ewa Niechwiej-Szwedo, David Gonzalez, Mina Nouredanesh, and James Tung. "Evaluation of the Leap Motion Controller during the performance of visually-guided upper limb movements." In: *PloS one* 13.3 (2018), e0193639.

[40] *Optitrack*. https://optitrack.com/cameras/prime-41/. Accessed: 2023-11-15.

[41] Markus Petrykowski, Philipp Berger, Patrick Hennig, and Christoph Meinel. "Digital collaboration with a whiteboard in virtual reality." In: *Proceedings of the Future Technologies Conference (FTC) 2018: Volume 1*. Springer. 2019, pp. 962–981.

[42] *Purdue Online Writing Lab*. https://ielts.idp.com/canada/prepare/article-punctuation. Accessed: 2023-11-16.

[43] *Purdue Online Writing Lab*. https://owl.purdue.edu/owl/. Accessed: 2023-11-16.

[44] Abir Rahali and Moulay A Akhloufi. "MalBERT: Using transformers for cybersecurity and malicious software detection." In: *arXiv preprint arXiv:2103.03806* (2021).

[45] Sai Deepika Regani, Chenshu Wu, Beibei Wang, Min Wu, and KJ Ray Liu. "mmWrite: passive handwriting tracking using a single millimeter-wave radio." In: *IEEE Internet of Things Journal* 8.17 (2021), pp. 13291–13305.

[46] Alexander Schick, Daniel Morlock, Christoph Amma, Tanja Schultz, and Rainer Stiefelhagen. "Vision-based handwriting recognition for unrestricted text input in mid-air." In: *Proceedings of the 14th ACM international conference on Multimodal interaction*. 2012, pp. 217–220.

[47] Garth Shoemaker, Leah Findlater, Jessica Q Dawson, and Kellogg S Booth. "Mid-air text input techniques for very large wall displays." In: *Graphics Interface*. 2009, pp. 231–238.

[48] Aradhana Kumari Singh, Lalit Kane, Abhirup Khanna, and Tanupriya Choudhury. "An Overview on State-of-Art in Mid-Air Writing and Recognition Systems." In: *Innovations in Cyber Physical Systems: Select Proceedings of ICICPS 2020* (2021), pp. 415–427.

[49] Trina D Spencer and Lydia Kruse. "Beery-Buktenica developmental test of visual-motor integration." In: *Encyclopedia of Autism Spectrum Disorders*. Springer, 2021, pp. 608–612.

[50] Arnold JWM Thomassen, Hans-Leo Teulings, and Lambert RB Schomaker. "Real-time processing of cursive writing and sketched graphics." In: *Human-Computer Interaction: Psychonomic Aspects*. Springer, 1988, pp. 334–352.

[51] *Ultraleap*. https://www.ultraleap.com/. Accessed: 2023-11-13.

[52] *Unity*. https://unity.com/. Accessed: 2023-11-15.

[53] Difeng Yu, Kaixuan Fan, Heng Zhang, Diego Monteiro, Wenge Xu, and Hai-Ning Liang. "PizzaText: Text entry for virtual reality systems using dual thumbsticks." In: *IEEE transactions on visualization and computer graphics* 24.11 (2018), pp. 2927–2935.

[54] *distilbert-base-uncased*. https://huggingface.co/distilbert-base-uncased. Accessed: 2023-12-12.

[55] *lstm-feature-importance*. https://www.kaggle.com/code/cdeotte/lstm-feature-importance. Accessed: 2023-12-12.

[56] nghiaho12. *rigid*$_{t}ransform_{3}D$. https://github.com/nghiaho12/rigid$_{t}ransform_{3}D$/
2021.