

**AURIVA – ADAPTIVE LEARNING PLATFORM FOR
ENGLISH LANGUAGE DEVELOPMENT IN AUTISTIC
CHILDREN IN SRI LANKA**

INDIVIDUAL THESIS (DRAFT)

R. S. Liluksha

(IT22216700)

Individual component: Motor-Informed Adaptive Handwriting Learning
for Children with Autism Spectrum Disorder

Group ID: R26-IT-006

Report submitted in partial fulfilment of the requirements for the
B.Sc. (Hons) Degree in Information Technology Specialized in Information
Technology

Department of Information Technology
Sri Lanka Institute of Information Technology

Sri Lanka

October 2026

DECLARATION AND SUPERVISOR CERTIFICATION

Candidate declaration. I declare that this thesis is my own work and does not incorporate, without acknowledgement, material previously submitted for a degree or diploma at any university or institute of higher learning. To the best of my knowledge, it contains no material previously published or written by another person except where acknowledgement is made in the text. I grant the Sri Lanka Institute of Information Technology the non-exclusive right to reproduce and distribute this thesis, in whole or in part, in print, electronic, or other media, while retaining the right to use the content in future work.

Academic integrity and AI disclosure. I have disclosed all use of generative AI and other automated tools in accordance with the module policy. I remain responsible for the accuracy, originality, citations, code, data, analyses, and conclusions in this thesis.

Student name: R. S. Liluksha

Student ID: IT22216700

Signature: _____ Date: _____

Supervisor certification. The above candidate carried out this research for the bachelor's degree thesis under my supervision.

Role and name	Signature	Date
Supervisor: Ms. Manori Gamage		
Co-supervisor: Ms. Suriya Kumari		

ABSTRACT

Some children with Autism Spectrum Disorder (ASD) experience handwriting difficulties related to fine motor control and movement coordination, yet many digital handwriting applications judge only the final traced form and give teachers little evidence about how a letter was produced. This thesis presents a motor-informed adaptive handwriting learning component for Auriva, an adaptive English learning platform for autistic children in Sri Lanka. The component captures handwriting on a tablet as time-ordered coordinates with stroke boundaries, compares each attempt with a reference using direction-invariant Dynamic Time Warping (DTW), and derives interpretable measures of smoothness, pause control, direction control and speed variation. A weighted letter Motor Score drives a transparent workflow of three-attempt cycles with decreasing support, mastery decided from the unguided third attempt at a threshold of 70, at most three cycles per letter and day, exact-letter remediation and teacher-approved worksheets for additional home practice, followed by rule-based word learning. Teachers receive a report with exact mastery evidence and teacher-initiated Writing Check results. Exploratory clustering is kept separate from these decisions. In controlled data from seven learners (42 shape trajectories), K-Means, Ward clustering and Gaussian mixture modelling produced the same two-cluster partition (silhouette 0.412) across 100 initialisations, although leave-one-participant-out agreement was lower (mean ARI 0.643). A descriptive letter motor-pattern model, built from 140 third-attempt observations aggregated to participant level, showed weak-to-moderate structure. The research pipeline reproduced the deployed feature values exactly, and all 7,609 automated tests passed. The results support transparent, motor-informed handwriting practice in which clusters remain descriptive rather than diagnostic. Teacher usability results are still being compiled.

Keywords: Autism spectrum disorder, online handwriting analysis, dynamic time warping, adaptive learning, interpretable machine learning

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Ms. Manori Gamage, for her guidance, encouragement and constructive feedback throughout this research. I also thank my co-supervisor, Ms. Suriya Kumari, for her support.

I am grateful to the principal, teachers and staff of the participating special-education school for making the controlled data collection possible, and I thank the children for their time and effort.

I thank the other members of the Auriva team for their collaboration in building the integrated platform, and the Sri Lanka Institute of Information Technology for the facilities and academic support provided. Finally, I thank my family and friends for their patience and encouragement.

TABLE OF CONTENTS

DECLARATION AND SUPERVISOR CERTIFICATION	i
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	x
LIST OF TABLES	xi
LIST OF ABBREVIATIONS	xii
1 INTRODUCTION	1
1.1 Overall project context	1
1.2 Individual component context	2
1.3 Problem statement	3
1.4 Research gap	4
1.5 Aim and objectives	6
1.5.1 Research Aim	6
1.5.2 Research Objectives	6
1.6 Research questions or hypotheses	7
1.7 Scope limitations and assumptions	9
1.7.1 Scope	9
1.7.2 Limitations	10
1.7.3 Assumptions	11
1.8 Individual contribution	11
1.9 Thesis structure	12
2 LITERATURE REVIEW AND RELATED WORK	14
2.1 Review method and scope	14

2.2 Domain and theoretical foundations	15
2.2.1 Handwriting as a perceptual-motor skill	15
2.2.2 Handwriting and motor characteristics in children with ASD	16
2.2.3 Motor learning, feedback and scaffolding	17
2.3 Related research	18
2.3.1 Online handwriting acquisition and process measures	18
2.3.2 Automated handwriting assessment using machine learning.....	18
2.3.3 Trajectory comparison using Dynamic Time Warping.....	19
2.3.4 Tablet-based handwriting learning.....	20
2.3.5 Unsupervised learning and cluster validation	21
2.3.6 Interpretability in educational AI	21
2.3.7 Summary of related research.....	22
2.4 Existing systems products and techniques	23
2.5 Gap synthesis and conceptual framework.....	25
3 RESEARCH METHODOLOGY	28
3.1 Research design.....	28
3.2 Data participants or study materials.....	29
3.2.1 Pilot participants	29
3.2.2 Study materials and equipment	30
3.2.3 Teacher participants for usability evaluation	31
3.3 Experimental or study procedure	31
3.3.1 Pilot data-collection procedure	31
3.3.2 Teacher usability evaluation procedure	32
3.4 Measures and evaluation metrics	32
3.4.1 Handwriting and motor features	32
3.4.2 Scoring measures	34

3.4.3 Clustering evaluation metrics.....	35
3.4.4 Software and usability metrics	35
3.5 Data analysis	36
3.5.1 Trajectory preprocessing and DTW	36
3.5.2 Score calibration.....	36
3.5.3 Exploratory clustering.....	37
3.6 Ethics privacy security and responsible AI.....	38
3.6.1 Ethical approval and consent	38
3.6.2 Privacy and data protection.....	39
3.6.3 Security	39
3.6.4 Responsible AI.....	40
3.7 Reproducibility and research artefacts	40
4 PRODUCT AND SOLUTION DESIGN.....	42
4.1 Requirements.....	42
4.2 Component architecture within the complete solution.....	44
4.3 Design alternatives and decisions	47
4.4 Detailed design.....	48
4.4.1 Handwriting learning workflow	48
4.4.2 Letter template system	49
4.4.3 Three-attempt practice cycle and mastery rule	50
4.4.4 Attempt processing sequence.....	51
4.4.5 Word learning.....	52
4.4.6 Teacher-facing design	53
4.4.7 Motor-pattern analysis design	54
4.5 Standards and professional practice	55
5 IMPLEMENTATION INTEGRATION AND TESTING.....	57

5.1 Development environment and technology stack	57
5.2 Component implementation	57
5.2.1 Trajectory capture	58
5.2.2 Initial assessment	58
5.2.3 Feature extraction and scoring	59
5.2.4 Practice sequencing and mastery	59
5.2.5 Remediation and word learning	60
5.2.6 Teacher Report, Writing Check and worksheets.....	60
5.2.7 ML service	61
5.2.8 Data-collection mode and data integrity	62
5.3 Integration with the other components.....	62
5.4 Verification and software testing	62
5.4.1 Testing strategy	62
5.4.2 Automated test status	63
5.4.3 Functional test cases.....	64
5.5 Deployment and operation	65
6 RESULTS AND VALIDATION	66
6.1 Research results.....	66
6.1.1 Trajectory representation and pipeline verification (RQ1)	66
6.1.2 Scoring behaviour (RQ1)	66
6.1.3 Initial-shape exploratory analysis (Dataset A, RQ3)	67
6.1.4 Letter motor-pattern model (Dataset B).....	70
6.2 Product and technical results.....	71
6.3 User or stakeholder validation	72
6.4 Integration results.....	72
6.5 Summary of findings.....	72

7 DISCUSSION	74
7.1 Interpretation of findings.....	74
7.2 Comparison with literature baselines and existing solutions	75
7.3 Research contribution.....	76
7.4 Product contribution.....	76
7.5 Limitations and threats to validity.....	76
7.6 Implications.....	78
8 COMMERCIALISATION AND SUSTAINABILITY	79
8.1 Value proposition and target users	79
8.2 Market and competitive position.....	80
8.2.1 Target market in Sri Lanka	80
8.2.2 Competitive position	80
8.3 Business or sustainability model.....	81
8.3.1 Revenue model.....	81
8.3.2 Cost structure and resources.....	81
8.3.3 Expected adoption barriers.....	82
8.4 Scalability intellectual property and compliance	82
8.5 Contribution to the Sustainable Development Goals	83
9 CONCLUSION AND FUTURE DIRECTIONS	84
9.1 Conclusion	84
9.2 Achievement of objectives.....	84
9.3 Individual research and product contributions	85
9.4 Future directions.....	85
REFERENCES.....	87
GLOSSARY.....	92
APPENDICES	93

Appendix A Research instruments.....	93
A.1 Controlled data-collection protocol checklist	93
A.2 Teacher usability evaluation tasks.....	93
A.3 Evaluation measures.....	93
Appendix B Academic integrity and AI use disclosure	94

LIST OF FIGURES

Figure 2.1: Conceptual framework of the motor-informed handwriting learning component	27
Figure 3.1: Research design stages (Stage 6 - Evaluation in progress)	29
Figure 4.1: High-level architecture of Auriva showing the handwriting component and its interfaces	45
Figure 4.2: Handwriting learning workflow, including remediation, Additional Home Practice and the Teacher Report.....	49
Figure 4.3: Three-attempt practice cycle, mastery rule, remediation and escalation for one letter and case	51
Figure 4.4: Processing sequence for a letter attempt	52
Figure 4.5: Letter motor-pattern analysis: model fitting, runtime use and separation from operational decisions	55
Figure 5.1: Letter-practice screen (Attempt 1, highest support)	58
Figure 5.2: Initial assessment summary	59
Figure 5.3: Word-writing activity	60
Figure 5.4: Teacher Report	61
Figure 6.1: Dataset A exploratory analysis pipeline	67

LIST OF TABLES

Table 2.1: Comparison of representative related research with the proposed component	22
Table 2.2: Capability comparison of existing solutions and the proposed component	25
Table 2.3: Mapping of identified gaps to the proposed component.....	26
Table 3.1: Handwriting and motor features derived for each attempt	33
Table 4.1: Functional requirements of the handwriting component	42
Table 4.2: Non-functional requirements of the handwriting component.....	43
Table 4.3: Auriva components and their relationship to the handwriting component.....	46
Table 4.4: External interfaces of the handwriting component	46
Table 4.5: Main design decisions and alternatives considered	47
Table 4.6: Word-learning activities.....	52
Table 5.1: Technology stack of the handwriting component.....	57
Table 5.2: Automated test results, run of 30 August 2026	63
Table 5.3: Representative functional test cases.....	64
Table 6.1: Candidate clustering solutions for Dataset A.....	68
Table 6.2: Stability of the selected two-cluster solution (Dataset A)	68
Table 6.3: Robust-scaled participant-level features for Dataset A	69
Table 6.4: Exploratory associations between Dataset A features and teacher ratings.....	70
Table 6.5: Letter motor-pattern model validation from the saved model metadata (Dataset B).....	70
Table 6.6: Implementation status of the main functions	71
Table 6.7: Summary of findings by research question.....	72
Table 8.1: Buyers, users and beneficiaries	79
Table 8.2: Main cost areas for deployment and operation	81
Table 9.1: Achievement of objectives.....	84
Table A.1: Controlled data-collection protocol checklist	93
Table A.2: Teacher usability evaluation tasks	93
Table B.1: Use of generative AI and automated tools	94

LIST OF ABBREVIATIONS

Abbreviation	Meaning
AI	Artificial Intelligence
API	Application Programming Interface
ARI	Adjusted Rand Index
ASD	Autism Spectrum Disorder
CH	Caliński–Harabasz index
CORS	Cross-Origin Resource Sharing
CV	Coefficient of Variation
DBI	Davies–Bouldin Index
dp	Density-independent pixel
DTW	Dynamic Time Warping
FR	Functional Requirement
GMM	Gaussian Mixture Model
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
JSON	JavaScript Object Notation
JWT	JSON Web Token
LOPO	Leave-One-Participant-Out
ML	Machine Learning
NFR	Non-Functional Requirement
ORM	Object-Relational Mapping
OWASP	Open Worldwide Application Security Project
PDF	Portable Document Format
PDPA	Personal Data Protection Act
REST	Representational State Transfer

Abbreviation	Meaning
RQ	Research Question
SD	Standard Deviation
SDG	Sustainable Development Goal
SUS	System Usability Scale
TLS	Transport Layer Security
WCAG	Web Content Accessibility Guidelines

1 INTRODUCTION

1.1 Overall project context

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition associated with differences in communication, behaviour, learning and motor functioning. These differences can affect how children participate in educational activities, particularly when tasks require sustained attention, repeated practice, fine motor coordination and clear instructional support. In this context, digital learning technologies can provide structured activities, consistent guidance and progress information that may assist teachers when supporting learners with different educational needs [1], [2].

Auriva is an integrated learning platform developed to support selected learning and developmental activities for children with ASD. The platform combines several learning components within a shared system, with each component addressing a different area of the learner experience. Rather than replacing the role of the teacher, Auriva is designed to provide structured learning activities and additional evidence that can support teacher observation, monitoring and decision-making.

Within the overall Auriva platform, the handwriting component focuses on digital handwriting practice and the analysis of the writing process. Handwriting is an important part of early educational participation and requires the coordination of visual, cognitive and motor processes. For some children with ASD, difficulties related to fine motor control and visual-motor coordination can make handwriting activities more challenging [1], [3], [4]. The handwriting component therefore contributes to the wider Auriva system by providing structured handwriting activities for the learner and process-level progress information for the teacher.

1.2 Individual component context

The individual component presented in this thesis focuses on **motor-informed adaptive handwriting learning** for children with ASD. The component was developed from the view that handwriting performance should not be judged only by the final visual appearance of a letter. The way in which the learner produces the writing, including the movement path, timing, pauses, smoothness and changes in direction, can provide additional information about handwriting performance that is not visible from the final written form alone.

During handwriting activities, the system captures online trajectory data containing spatial, temporal and stroke-level information. These data are used to derive interpretable measures of handwriting performance. Dynamic Time Warping (DTW) is used to compare the learner's trajectory with a reference trajectory while allowing differences in writing speed and sequence length. Additional motor-related measures describe characteristics such as smoothness, speed variation, pause behaviour, direction control and completion of the writing task. These measures are used within the component to support scoring, progress monitoring and later research analysis.

The learning workflow is organised as a sequence of initial assessment, letter practice, remediation, word-level activities and teacher monitoring. Letter practice is structured around three attempts with decreasing levels of support, where the least-supported third attempt provides the operational evidence for mastery. Learners who do not meet the mastery criterion can continue through bounded practice cycles, with targeted warm-up activities introduced when repeated difficulty is detected. The same general structure is implemented separately for lowercase and uppercase letters, followed by word-level learning and word-writing activities.

Teacher support forms an important part of the component. The system provides progress information, handwriting evidence, Writing Check information, letter and word progress, and additional-practice support. Where repeated difficulty is identified, the system can recommend a worksheet that the teacher can generate for offline practice and later review through the system. Exploratory machine-learning

analysis is also included to examine relative motor-pattern structure in the collected handwriting data. However, these clustering results are kept separate from child-facing mastery, remediation and progression decisions. Operational progression remains controlled by explicit rules so that the basis of each learning decision can be understood and reviewed by the teacher.

Overall, the component combines handwriting trajectory capture, interpretable motor analysis, transparent practice progression and teacher-facing evidence within one integrated workflow. The aim is not to diagnose handwriting disorders or ASD-related severity, but to provide structured educational support and clearer evidence about how the learner performs during handwriting activities.

1.3 Problem statement

Handwriting is not only a visual task; it depends on the coordinated execution of movements over time. A learner may produce a letter that appears acceptable at the end of the task while still showing irregular movement, repeated corrections, pauses, inconsistent speed or poor control during writing. For some children with ASD, difficulties related to fine motor control, visual-motor integration and motor planning can make handwriting activities more challenging and can affect letter formation and writing consistency [1], [3], [4].

Many digital handwriting applications support children through tracing, repetition and visual guidance. However, these systems commonly emphasise whether a letter was completed or how closely the final trace follows a predefined path. This provides limited information about the handwriting process itself and may not give teachers enough evidence to understand how the learner performed the movement or why additional support may be required. Research on online handwriting analysis shows that trajectory, timing and movement characteristics can provide useful process-level information, but these measures are often studied for assessment, screening or diagnosis rather than being integrated into day-to-day handwriting practice [5]–[8].

A further issue is the transparency of personalised learning decisions. When a system changes activities or progression based on automated analysis, teachers should be able to understand the evidence behind that decision. This is especially important in a child-facing educational setting, where an unexplained model output should not determine whether a learner is considered to have mastered a skill or requires remediation [9].

Therefore, the problem addressed by this research is the limited integration of handwriting-process analysis with transparent, structured learning support and teacher-facing evidence for children with ASD. There is a need for a handwriting-learning approach that can capture and interpret the writing process, use that evidence to support structured practice and remediation, and present the resulting information in a form that teachers can understand and review.

1.4 Research gap

Previous studies have established that handwriting difficulties in some children with ASD are associated with factors such as fine motor control, visual-motor integration, movement consistency and letter formation [1], [3], [4], [10]–[13]. Research using digital tablets and digitising devices has further shown that process-level characteristics such as trajectory, timing, velocity, smoothness and pauses can provide information about handwriting performance that cannot be obtained from the final written form alone [14]–[18]. However, much of this work has focused on understanding, assessing or identifying handwriting difficulties rather than integrating these measures into an ongoing handwriting-learning process.

A second group of studies has applied machine-learning techniques to handwriting data, particularly for the screening or identification of handwriting disorders such as dysgraphia [6]–[8]. These approaches demonstrate the value of digital handwriting features, but they are generally designed for classification or assessment rather than for supporting day-to-day handwriting practice. In addition, many of these approaches rely on labelled datasets or richer stylus information such as pressure and

tilt, which may not always be available in a standard tablet-based learning environment.

Digital handwriting-learning applications, on the other hand, commonly provide tracing, visual guidance, demonstrations and repeated practice [5], [19]. Although these approaches can support letter learning, they generally place greater emphasis on task completion or the final traced form than on providing teachers with interpretable information about the movement process. As a result, there remains limited connection between the handwriting evidence captured during practice and the support, remediation and progress information provided to the learner and teacher.

Another important consideration is the transparency of adaptive educational systems. When automated analysis is used to influence a learner's activities, teachers should be able to understand the basis of the recommendation or progression decision [9]. This is particularly important when the available machine-learning evidence is exploratory and based on a small participant sample, where model outputs should not be treated as validated educational or clinical classifications.

The research gap addressed by this component is therefore the limited integration of online handwriting trajectory analysis, interpretable motor features, transparent adaptive practice and teacher-facing progress evidence within one handwriting-learning workflow for children with ASD.

The proposed component addresses this gap by combining:

- digital trajectory capture and motor-feature extraction;
- DTW-based trajectory comparison;
- transparent rule-based letter and word practice;
- bounded remediation following repeated difficulty;
- exploratory motor-pattern analysis separated from child-facing progression;
- and
- teacher-facing evidence and additional practice support.

This combination is intended to provide a more interpretable connection between how the child writes, how the system responds, and what the teacher can observe.

1.5 Aim and objectives

1.5.1 Research Aim

The main aim of this research is:

To design, implement and evaluate an interpretable motor-informed handwriting learning component for children with Autism Spectrum Disorder that captures digital handwriting trajectories, analyses relevant handwriting and motor characteristics, provides structured adaptive practice, and presents understandable progress evidence to teachers.

1.5.2 Research Objectives

The following objectives were defined to achieve the research aim.

O1 - Handwriting trajectory acquisition

To design and implement a digital handwriting capture mechanism that records spatial, temporal and stroke-level information during initial assessment, letter practice and word-writing activities.

O2 - Handwriting and motor-feature analysis

To derive interpretable handwriting and motor-related measures from the captured trajectories, including trajectory similarity using Dynamic Time Warping and selected measures related to smoothness, speed variation, pauses and directional behaviour.

O3 - Structured adaptive handwriting practice

To design and implement a transparent handwriting-practice workflow for lowercase letters, uppercase letters and words using progressively reduced support, explicit mastery criteria, bounded practice cycles and targeted remediation when repeated difficulty is observed.

O4 - Exploratory motor-pattern analysis

To investigate whether selected handwriting motor features reveal meaningful and sufficiently stable relative motor-pattern structure within the collected pilot data using exploratory unsupervised-learning techniques.

O5 - Teacher-facing monitoring and additional support

To provide teachers with understandable handwriting-progress information, mastery evidence, motor-performance summaries, Writing Check information and additional-practice support, including worksheet recommendations for learners who demonstrate repeated difficulty.

O6 - Evaluation of the developed component

To evaluate the developed handwriting component through software verification, integration testing, analysis of the collected handwriting data and usability evaluation with teachers and users involved in the learning workflow.

These objectives separate the research contribution, mainly O2 and O4, from the software implementation, mainly O1, O3 and O5, while O6 provides evaluation evidence for both.

1.6 Research questions or hypotheses

The following research questions guide the individual component and are aligned with the research objectives defined in Section 1.5.

RQ1 - How can digital handwriting trajectory data be transformed into interpretable measures of handwriting and motor performance?

This question focuses on the capture and analysis of spatial, temporal and stroke-level handwriting data. It considers how trajectory similarity and selected motor-related characteristics, including smoothness, speed variation, pauses and directional behaviour, can be derived in a form that is understandable and useful within the handwriting-learning component.

RQ2 - How can handwriting-performance evidence be incorporated into a transparent rule-based workflow to support letter and word practice, progression and remediation?

This question examines how the analysed handwriting evidence can be used within the operational learning workflow while keeping progression decisions understandable. It includes the use of progressively reduced support, explicit mastery criteria, bounded practice cycles and targeted remediation when repeated difficulty is observed.

RQ3 - Can selected handwriting motor features reveal meaningful and sufficiently stable relative motor-pattern structure within the collected pilot data?

This question investigates the exploratory use of unsupervised-learning techniques to identify relative structure within the collected handwriting data. The purpose is descriptive rather than diagnostic. Any identified clusters are treated as exploratory motor-pattern groupings and are not interpreted as ASD severity levels, clinical categories, measures of intelligence or direct indicators of educational ability.

RQ4 - How usable and understandable is the developed handwriting component for teachers and users who support children during handwriting practice and review their progress?

This question considers the usability of the implemented workflow, the clarity of teacher-facing progress information and recommendations, and the extent to which users can understand and complete the main handwriting-related tasks within the system.

1.7 Scope limitations and assumptions

1.7.1 Scope

This research focuses on the design, implementation and evaluation of a digital handwriting-learning component for children with ASD. The component operates as part of the wider Auriva platform and is intended to support structured handwriting practice under the guidance of teachers or responsible adults.

The scope of the component includes an initial pre-writing assessment using six basic shapes, followed by lowercase and uppercase letter practice, remediation activities, word-level learning activities and word writing. During handwriting tasks, the system captures online trajectory information consisting primarily of spatial coordinates, timing and stroke-related information. These data are used to derive selected handwriting and motor-related measures, including trajectory similarity, movement smoothness, speed variation, pause behaviour and directional characteristics.

Letter practice is organised using three attempts with decreasing levels of support. The least-supported third attempt is used as the operational evidence for mastery, while repeated difficulty can lead to additional practice cycles and targeted remediation. The component also provides word-learning activities and word-writing tasks after the relevant letter-learning stages.

The teacher-facing scope includes learner progress information, letter mastery evidence, handwriting-performance summaries, Writing Check information, word-practice progress and recommendations for additional practice. Where repeated difficulty is identified, the system can recommend supplementary worksheet-based practice that the teacher can generate, provide to the learner offline, upload after completion and review within the system.

The research also includes exploratory analysis of handwriting motor features using unsupervised-learning techniques. This analysis is intended to investigate relative patterns within the collected research data and is kept separate from the operational

rules that determine learner mastery, remediation and progression. The component therefore functions as an educational support tool rather than a diagnostic or clinical assessment system.

1.7.2 Limitations

Several limitations should be considered when interpreting the findings of this research.

First, the handwriting datasets available for the research are relatively small and were collected within a limited research setting. The clustering results and other quantitative analyses should therefore be interpreted as exploratory findings rather than evidence that the identified patterns generalise to the wider population of children with ASD.

Second, the handwriting capture mechanism is based mainly on spatial coordinates, timing and stroke information. Advanced stylus characteristics such as pen pressure, tilt, pen angle and grip are not available in the current implementation. Consequently, the analysis represents only the motor characteristics that can be derived from the available trajectory data.

Third, several operational decisions used by the prototype, including scoring weights and mastery thresholds, were established as engineering choices for the current research implementation. Although these values were examined during development and testing, they have not been clinically validated and should not be interpreted as universal standards for evaluating handwriting ability.

Fourth, the exploratory machine-learning analysis does not provide a clinical diagnosis, determine ASD severity, measure intelligence or classify a learner's educational ability. The clusters describe relative structure within the selected handwriting features and collected sample only.

Finally, the current research evaluates the functioning, usability and short-term behaviour of the developed component but does not establish long-term

improvement in handwriting ability. Demonstrating sustained learning outcomes would require a larger longitudinal study with repeated measurements over an extended period and an appropriate experimental design.

1.7.3 Assumptions

The component assumes that handwriting activities are completed using a compatible touch-enabled device capable of recording the required trajectory information and that a teacher or responsible adult can provide supervision when required.

It is assumed that the handwriting trajectory recorded by the system represents the learner's genuine attempt and that suitable reference trajectories are available for the handwriting activities that require trajectory comparison. The system also assumes that the learner profile used during a session correctly represents the child performing the activities.

For research analysis, multiple handwriting attempts produced by the same learner are treated as repeated observations from that participant rather than as independent participants. Where comparisons are made across learners or machine-learning stability is evaluated, the participant structure of the data is therefore preserved.

1.8 Individual contribution

This research represents the individual contribution of the author within the wider Auriva project and includes both a research contribution and a software-development contribution. The work focuses specifically on the design and implementation of a motor-informed handwriting-learning component and its integration with the shared Auriva platform.

From a research perspective, the individual contribution includes the investigation of online handwriting trajectories as a source of interpretable handwriting and motor information. This involved the use of Dynamic Time Warping for trajectory comparison, the derivation and analysis of selected motor-related features, the development of an interpretable scoring approach, and exploratory clustering to examine relative motor-pattern structure within the collected pilot data. The clustering analysis is used only for research and descriptive purposes and is kept separate from child-facing mastery and progression decisions.

From a system-development perspective, the individual component includes the initial six-shape handwriting assessment, lowercase and uppercase letter practice, the three-attempt practice structure, mastery handling, repeated practice cycles, targeted remediation, word-level activities and word-writing functionality. The work also includes teacher-facing features such as progress reporting, exact mastery evidence, Writing Check information, motor-performance summaries and worksheet-based additional practice support.

The individual contribution also includes integration of the handwriting component with the common Auriva application architecture. The mobile application communicates with the shared Node.js/Express backend, handwriting-related records are stored in PostgreSQL through the backend data layer, and machine-learning functionality is accessed through a FastAPI service via the backend. The mobile application does not communicate directly with the machine-learning service.

A further contribution of the work is the separation between operational learning decisions and exploratory machine-learning analysis. Mastery, remediation and progression are controlled through explicit and testable rules, while clustering is used only to provide descriptive information about patterns within the research data. This design was adopted to maintain transparency and to avoid allowing an exploratory model trained on a small pilot sample to make unsupported educational decisions.

1.9 Thesis structure

The rest of this thesis is organised as follows.

Chapter 2 - Literature Review and Related Work reviews previous research on handwriting difficulties in children with ASD, digital handwriting analysis, motor features, Dynamic Time Warping, adaptive learning, clustering and teacher-facing support. It also compares existing systems and identifies the research gap.

Chapter 3 - Research Methodology explains the research design, participants, data collection, handwriting features, evaluation methods, data analysis, ethical considerations and usability testing procedures.

Chapter 4 - Product and Solution Design describes the system requirements, architecture, main design decisions and the detailed workflow of the handwriting component.

Chapter 5 - Implementation, Integration and Testing explains how the component was developed, integrated with the Auriva platform and tested.

Chapter 6 - Results and Validation presents the findings from the handwriting analysis, machine-learning experiments, software testing and user evaluation.

Chapter 7 - Discussion discusses the results in relation to the research questions and previous studies, and explains the main contributions, limitations and implications of the work.

Chapter 8 - Commercialisation and Sustainability discusses the possible use of the system in practice, target users, market potential, scalability, costs and sustainability considerations.

Chapter 9 - Conclusion and Future Directions summarises the research, reviews the achievement of the objectives and presents possible improvements and future research directions.

2 LITERATURE REVIEW AND RELATED WORK

2.1 Review method and scope

This literature review examines research relevant to the development of a motor-informed handwriting-learning component for children with ASD. The review was conducted as a structured narrative review rather than a full systematic review. Its purpose was to establish the theoretical and technical foundations of the study, compare relevant research approaches and existing solutions, and identify the research gap addressed by the developed component.

Literature was mainly identified through **Google Scholar and IEEE Xplore**. Priority was given to peer-reviewed journal articles, conference papers, systematic reviews and review studies published within approximately the last twenty years, with greater emphasis placed on recent research from 2015 onwards. Older studies were considered only when they described a foundational concept or method that remained directly relevant to the study.

Search terms were organised around the main themes of the research. These included combinations of terms such as *autism spectrum disorder*, *children*, *handwriting*, *fine motor skills*, *visual-motor integration*, *online handwriting*, *digital handwriting*, *tablet handwriting*, *handwriting trajectory*, *kinematics*, *Dynamic Time Warping*, *handwriting assessment*, *adaptive learning*, *motor learning*, *clustering*, *machine learning*, and *explainable artificial intelligence*. References cited in relevant recent review papers were also examined when they contributed directly to the research questions.

Studies were selected when they contributed to at least one of the following areas: understanding handwriting and motor characteristics in children with ASD; digital capture and quantitative analysis of handwriting processes; computational techniques relevant to trajectory comparison and motor-pattern analysis; or digital handwriting-learning and teacher-support approaches. Studies that focused only on unrelated handwriting recognition or general machine-learning applications were not included unless their methods were directly relevant to the techniques used in this research.

The remainder of this chapter first discusses handwriting as a perceptual-motor activity and the motor characteristics reported in children with ASD. It then reviews online handwriting acquisition, handwriting features, machine-learning approaches, trajectory comparison, digital handwriting-learning methods and interpretability. Finally, related systems are compared, and the findings are synthesised to identify the research gap addressed by the Auriva Handwriting component.

2.2 Domain and theoretical foundations

2.2.1 Handwriting as a perceptual-motor skill

Handwriting is a complex skill that requires the coordinated involvement of motor, perceptual and cognitive processes. Producing a written letter requires the learner to control hand and finger movements while using visual information to guide the position, direction and size of the strokes. Recent research also describes handwriting as a sensorimotor activity involving visual-perceptual processing, motor coordination, memory and linguistic processes rather than simply the reproduction of a visual shape. Babushkin et al. [20], for example, emphasise that handwriting legibility is influenced by both hand and stylus kinematics, demonstrating that the movement used to produce writing is an important part of handwriting performance.

The distinction between the **handwriting product** and the **handwriting process** is particularly important when evaluating handwriting ability. The product refers to the final written output, such as its legibility, size or spatial arrangement, whereas the process represents how that output was produced over time. Digital handwriting systems make it possible to examine characteristics such as movement duration, speed, pauses and other kinematic or temporospatial measures. Chang and Yu [21] demonstrated this distinction through a computerised handwriting evaluation involving primary-school children, where handwriting performance, motor control, speed and automation, pauses and other temporal characteristics were represented as separate dimensions of handwriting behaviour.

Quantitative handwriting research also indicates that individual aspects of handwriting control develop differently during childhood. Watanabe et al. [22], in a study involving children aged four to twelve years, found age-related changes in

accuracy, writing pressure and movement time, suggesting that handwriting development cannot be represented by a single performance measure. More recent reviews similarly highlight the growing use of spatiotemporal and motor-related measures to understand handwriting difficulties and note that technology-based assessment can capture aspects of writing that are difficult to observe reliably from the final written product alone.

These findings support the approach taken in this research. Rather than evaluating handwriting only from the completed letter, the Auriva handwriting component records the writing trajectory and derives selected process-level measures. This allows characteristics of the movement used to produce the handwriting to be considered alongside the resulting written form. The approach does not attempt to measure every aspect of handwriting motor control; instead, it focuses on trajectory and timing information that can be captured reliably using the available tablet-based interaction.

2.2.2 Handwriting and motor characteristics in children with ASD

Handwriting difficulties in children with ASD are frequently associated with differences in fine motor control, visual-motor coordination and movement consistency. A recent review of handwriting research in ASD reported that handwriting performance varies considerably between individuals and highlighted fine motor and graphomotor abilities as important factors influencing handwriting quality [12]. Recent neurological research has also continued to identify motor-coordination impairment as an important associated characteristic in autistic children [23].

Digital handwriting studies provide further evidence that these difficulties can be observed in the writing process itself. Rosenblum et al. found significant differences between children with ASD and typically developing children in both handwriting process and product measures, suggesting that examining both aspects provides a more complete understanding of handwriting performance [24]. Johnson et al. similarly showed that handwriting size, consistency and movement characteristics can vary depending on visual guidance and task conditions [25].

Motor differences in ASD are also not limited to handwriting alone. Talkar et al. identified differences in fine motor coordination using handwriting and other time-series signals, supporting the wider relationship between motor coordination and purposeful hand movements in ASD [26]. These findings support analysing multiple handwriting-process features rather than relying only on whether the final letter appears visually correct.

2.2.3 Motor learning, feedback and scaffolding

Handwriting learning is influenced not only by repeated practice, but also by the type and amount of support provided during practice. Recent small-scale intervention studies with autistic learners indicate that structured handwriting programmes, combining guided practice with repeated teacher support, can help develop handwriting-related skills [27], [28]. In particular, Ch'ng and Ahmad reported improved visual-motor and letter-formation scores for one autistic student after a six-week structured handwriting programme in which the teacher used visual cues and verbal guidance [28].

However, support should not remain constant throughout learning. Motor-learning principles suggest that learners need opportunities to perform with reduced assistance so that performance reflects their own control rather than continuous dependence on guidance. This is relevant to handwriting systems because strong visual guides may improve immediate performance but do not necessarily represent independent writing ability. Johnson et al. also showed that visual writing guides can influence handwriting performance in children with ASD, highlighting the importance of considering the level of guidance when interpreting writing behaviour [25].

These findings support the use of **progressively reduced support** in this research. The Auriva handwriting workflow therefore provides greater guidance during the earlier attempts and evaluates mastery using the least-supported attempt. Additional practice and remediation are provided when repeated difficulty is observed, while progression decisions remain based on explicit and understandable rules.

2.3 Related research

2.3.1 Online handwriting acquisition and process measures

Digital handwriting can be analysed from either the final written product or the writing process. Online handwriting systems capture the movement while writing takes place, allowing temporal and spatial information such as stroke trajectory, duration, speed and pauses to be analysed. Recent computerised handwriting studies show that this process-level information can provide a more detailed view of handwriting performance than visual inspection of the final written form alone [20], [21]. Chang and Yu, for example, evaluated both handwriting product and process measures in primary-school children and identified distinct dimensions related to motor control, speed, automation and halting behaviour [21].

Recent work has also examined richer kinematic information captured from digital handwriting devices. Babushkin et al. analysed stylus and hand-motion features and showed that kinematic information can contribute to the assessment of handwriting legibility [20]. Such studies demonstrate the value of recording handwriting as a time-ordered sequence rather than treating it only as a static image.

These findings support the data-capture approach used in this research. The Auriva handwriting component records spatial coordinates, timing and stroke information during writing, which are then used to derive measures such as trajectory similarity, smoothness, speed variation, pauses and directional behaviour. The current implementation does not depend on advanced stylus measurements such as pressure or tilt, allowing the analysis to operate with the trajectory information available from the tablet interaction.

2.3.2 Automated handwriting assessment using machine learning

Machine-learning methods have increasingly been applied to digital handwriting data to support automated assessment of handwriting quality and handwriting difficulties. Recent studies commonly use kinematic, temporal and spatial features captured from tablets or digital pens and train models to distinguish between handwriting patterns or predict assessment outcomes. For example, Kunhoth et al. analysed on-surface and in-air handwriting characteristics and evaluated several classifiers for dysgraphia

screening, showing that a reduced set of online handwriting features could still provide useful classification performance [29].

More recent work has also explored deep-learning approaches. Babushkin et al. used temporal convolutional networks with stylus and hand-kinematic features to evaluate handwriting legibility and reported that the addition of hand-motion information improved model performance. Their study also used explainability techniques to identify influential features such as pressure variability, pen orientation and hand-speed characteristics [20].

Although these approaches demonstrate the potential of machine learning for handwriting assessment, they generally rely on labelled outcomes, larger datasets, or richer sensing information than is available in the present study. The current research therefore uses machine learning more conservatively, applying exploratory unsupervised clustering to selected motor features rather than attempting diagnostic classification or automated prediction of learner ability. This is more appropriate for the small pilot dataset and the absence of clinically validated class labels.

2.3.3 Trajectory comparison using Dynamic Time Warping

Dynamic Time Warping (DTW) is widely used to compare time-series or trajectory data when two sequences may differ in length or execution speed. Instead of comparing corresponding points directly, DTW finds an alignment that minimises the cumulative distance between the two sequences. This makes it suitable for handwriting, where the same letter can be written at different speeds while following a similar overall trajectory [30], [31].

Recent trajectory-recognition studies continue to use DTW because it can handle temporal stretching and local timing differences while preserving sequence similarity. Tang et al. proposed a structured DTW method for recognising gestures from continuous hand trajectories, illustrating that DTW-based comparison remains applicable to continuous trajectory data, although that work concerned hand gestures rather than handwritten letters [32]. Di Brina et al. also showed that DTW can be used to study handwriting-form variability in children with poor handwriting, supporting its relevance for temporal handwriting data [31].

DTW also has limitations. Its output depends on how trajectories are normalised and represented, and large differences in starting point, writing direction or stroke structure can increase the calculated distance even when the final shape appears similar. For this reason, DTW should not be treated as a complete measure of handwriting quality by itself. In this research, it is used as one measure of trajectory similarity and is considered together with other movement-related features such as smoothness and speed variation.

2.3.4 Tablet-based handwriting learning

Tablet-based handwriting systems have become increasingly common because they allow children to practise writing while also capturing information about how the writing is produced. Bonneton-Botté et al. evaluated a stylus-based digital notebook in kindergarten classrooms and found that tablet-supported handwriting practice could benefit some learners, although the effect depended on the child's initial handwriting level [33]. This suggests that digital handwriting support should not be assumed to work equally well for every learner and that individual differences should be considered.

The type of digital interaction also influences learning. Mayer et al. compared handwriting with pencil, handwriting with a tablet stylus and typing on a virtual keyboard in kindergarten children. Their results showed that handwriting-based conditions supported literacy learning differently from keyboard-based input, while also highlighting that writing on a touchscreen changes the motor demands because of the lower surface friction [34].

Tablet systems can also provide immediate feedback and personalised practice. Ecalle et al. demonstrated that multisensory tablet-based handwriting training affected both literacy outcomes and handwriting kinematics in young children [35]. However, these findings also indicate that the value of tablets lies not simply in replacing paper, but in how guidance, feedback and recorded handwriting information are used within the learning activity.

These studies support the use of a tablet as the interaction medium in this research. In Auriva, the tablet is used not only to present handwriting activities but also to capture

trajectory and timing information that can support structured practice and teacher monitoring.

2.3.5 Unsupervised learning and cluster validation

Unsupervised learning is useful when validated class labels are not available and the aim is to explore whether meaningful structure exists within a dataset. Clustering methods group observations according to similarity in their feature values, but different algorithms make different assumptions about cluster shape and distribution. K-Means is widely used because of its simplicity and interpretability, while hierarchical clustering, Gaussian Mixture Models (GMM) and density-based methods such as DBSCAN provide alternative ways of examining data structure [36], [37].

Because clustering will always produce some form of grouping, the resulting clusters should not automatically be treated as meaningful. Internal validation measures such as the Silhouette coefficient, Davies–Bouldin Index and Calinski–Harabasz Index are commonly used to examine cluster compactness and separation [38]. However, recent research has also shown that these measures can favour different solutions and should not be interpreted as proof that a discovered cluster represents a real or clinically meaningful group [38].

Cluster stability provides an additional form of evidence by examining whether similar groupings are obtained when the data or participants are changed. This is particularly important for small datasets, where the apparent structure may be strongly influenced by individual observations [39]. Therefore, this research compares multiple clustering approaches and considers both internal validity and stability before interpreting any motor-pattern grouping. The resulting clusters are treated as exploratory patterns rather than predefined learner categories.

2.3.6 Interpretability in educational AI

Interpretability is important in educational AI because teachers need to understand how automated systems reach conclusions that may influence learning activities or student support. Khosravi et al. noted that educational AI has specific explainability

requirements because different stakeholders, including teachers and learners, need explanations that are understandable and relevant to their role [9].

Recent research also shows that transparency alone is not enough; the information presented by an intelligent system should be meaningful and actionable for teachers. Ley et al. found that teacher-facing learning analytics are more useful when system models are presented in language that supports educational decision-making rather than only exposing technical model outputs [40].

These considerations influenced the design of the Auriva handwriting component. Child-facing mastery, remediation and progression are controlled using explicit rules that can be explained to teachers, while the exploratory clustering model is kept separate from these operational decisions. The machine-learning output is therefore used as descriptive motor-pattern information rather than as an automatic decision about a learner’s ability or progression.

2.3.7 Summary of related research

Table 2.1 compares representative studies with the proposed component.

Table 2.1: Comparison of representative related research with the proposed component

Study	Population / data	Main technique	Primary purpose	Limitation relative to this work
Fuentes et al. [3]	Children with ASD and controls; paper-based handwriting	Standardized handwriting assessment	Characterise handwriting difficulties	No digital trajectory capture or adaptive practice
Johnson et al. [4]; Grace et al. [10]	Children with ASD; digitiser-based handwriting	Kinematic and temporal analysis	Explain handwriting difficulty mechanisms	Analyse handwriting process and motor characteristics
Asselborn et al. [6]; Mekyska et al. [7]; Drótár and Dobeš [8]	Children with handwriting difficulties; tablet/digitiser data	Supervised machine learning	Screening or classification	Requires labelled data; mainly diagnostic rather than instructional
Digital handwriting intervention studies [34], [35]	Children using tablet-based handwriting activities	Guided digital practice and feedback	Support handwriting learning	Limited motor-feature analysis and teacher-facing monitoring

Study	Population / data	Main technique	Primary purpose	Limitation relative to this work
Explainable educational AI studies [9], [40]	Educational AI environments	Explainable/interpretable AI	Improve transparency of automated decisions	General educational context; not specific to ASD handwriting
Proposed component	Children with ASD; tablet handwriting trajectories	DTW, motor features, rule-based progression and exploratory clustering	Support structured practice, remediation and teacher monitoring	Small-scale evaluation; exploratory ML not used for autonomous decisions

Overall, previous studies provide strong evidence for handwriting assessment, motor analysis and digital practice, but there is limited integration of these elements within a transparent adaptive learning workflow for children with ASD.

2.4 Existing systems products and techniques

Several existing digital products provide learning support for children with ASD or handwriting practice for young learners. The systems most relevant to this research include **Otsimo**, **MITA**, **AutiSpark**, **Writing Wizard** and **LetterSchool**, as these represent either autism-focused learning platforms or digital handwriting-practice applications.

Otsimo is designed as a special-education learning platform for children with autism and includes gamified activities covering language, cognitive and early learning skills. It also provides activities for drawing lines and letters, where children follow visual guidance and practise basic finger control. However, its handwriting activities are mainly focused on learning and tracing letter forms rather than analysing detailed handwriting-process characteristics such as trajectory smoothness, pauses or movement consistency.

MITA (Mental Imagery Therapy for Autism) is an adaptive application developed specifically for children with ASD. It provides visual and language-based cognitive exercises, automatically adjusts task difficulty according to learner performance and allows parents or therapists to monitor progress. However, its primary purpose is language and cognitive development rather than handwriting learning, and it does not provide handwriting-specific trajectory analysis or motor-based writing feedback.

AutiSpark is another autism-focused educational platform that provides games and worksheets for areas such as matching, sequencing, spelling, colours, numbers and early literacy. It also includes tracing activities for letters, numbers and patterns. Although these activities support early writing practice, the platform primarily provides guided educational activities and does not report detailed motor characteristics of the handwriting process or connect such measures to teacher-facing handwriting progression.

Writing Wizard and **LetterSchool** represent handwriting-focused applications rather than autism-specific learning platforms. They are designed around guided tracing and repeated practice of letters, numbers and words, typically using visual demonstrations and interactive feedback. These approaches are useful for teaching letter formation and encouraging repeated practice; however, they mainly emphasise guided tracing and task completion rather than analysing the learner's underlying movement characteristics or providing detailed motor evidence for teachers.

Overall, the reviewed products demonstrate two main approaches: autism-focused applications that provide broad adaptive educational activities, and handwriting-focused applications that provide guided tracing and letter-formation practice. There remains limited integration of these approaches with detailed online handwriting-process analysis, transparent progression and teacher-facing motor evidence.

Table 2.2 summarises the capabilities of these categories of solution as described in their published material, compared with the proposed component. Where a capability is not described publicly it is marked "Not reported" rather than assumed absent.

Table 2.2: Capability comparison of existing solutions and the proposed component

Capability	Otsimo	MITA	AntiSpark	Writing Wizard	LetterSchool	Proposed Component
Designed specifically for ASD learning	✓	✓	✓			✓
Letter / handwriting practice	✓		✓	✓	✓	✓
Guided tracing / visual support	✓		✓	✓	✓	✓
Adaptive learning or difficulty adjustment	✓	✓	Limited	Limited	Limited	✓
Online handwriting trajectory capture				Limited / not reported	Limited / not reported	✓
Motor-feature analysis						✓
Transparent mastery / remediation workflow						✓
Teacher-facing handwriting evidence	Limited	Progress monitoring	Limited	Limited	Limited	✓
Exploratory motor-pattern analysis						✓
Additional offline practice support			Worksheets available			✓

2.5 Gap synthesis and conceptual framework

The literature reviewed in Sections 2.2-2.4 shows that several parts of the handwriting problem have already been studied separately. Research on ASD has identified differences in fine motor control, visual-motor coordination and handwriting execution, while digital handwriting studies have demonstrated the value of analysing trajectory, timing and other process-level measures. Other studies have applied machine learning to handwriting assessment, and tablet-based applications have shown the usefulness of guided digital practice. However, these approaches are often developed for assessment, screening, diagnosis or general handwriting practice rather than as one integrated educational workflow.

A second gap concerns how handwriting analysis is connected to learning support. Existing handwriting applications commonly provide tracing, repetition and visual feedback, while research systems can provide detailed quantitative analysis. There is

less evidence of systems that combine online handwriting capture, interpretable motor-related measures, transparent progression and remediation, and teacher-facing evidence in a single workflow designed for children with ASD. In addition, exploratory machine-learning findings should not automatically control educational decisions when the available dataset is small or lacks validated clinical labels.

The conceptual approach used in this research therefore separates operational learning support from exploratory motor-pattern analysis. Handwriting trajectories are first captured and transformed into interpretable features. These features support transparent rule-based scoring, practice progression and remediation, while selected features are also analysed separately using exploratory clustering. Teacher-facing reporting then presents progress and motor-related evidence without allowing the clustering model to determine mastery or progression.

Table 2.3: Mapping of identified gaps to the proposed component

Identified gap	Evidence	Response in this component	Related objective
Practice tools emphasize the final shape more than the movement process.	[5], [19]	Online x–y–time–stroke capture with DTW and motor features for every attempt	O1, O2
Process analysis is often used for assessment rather than day-to-day learning support.	[6]–[8]	Motor evidence used inside a structured practice workflow	O2, O3
Automated personalisation may be difficult for teachers to interpret.	[9]	Rule-based mastery and remediation are used; ML remains descriptive only.	O3, O4
ASD learners vary widely, while small datasets limit supervised modelling.	[3], [10]	Exploratory clustering with stability checks is used for motor-pattern analysis, not decisions.	O4
Teachers lack clear process-level evidence of learner progress.	[1], [9]	Teacher report, mastery evidence, Writing Check, and worksheet support improve visibility.	O5

The conceptual framework that results from this synthesis is shown in Figure 2.1. It links how the child writes (trajectory capture and motor features), how the system responds (transparent rules), and what the teacher can observe (teacher-facing

evidence). Exploratory clustering is attached to the motor evidence as a descriptive layer, without a path into operational decisions.

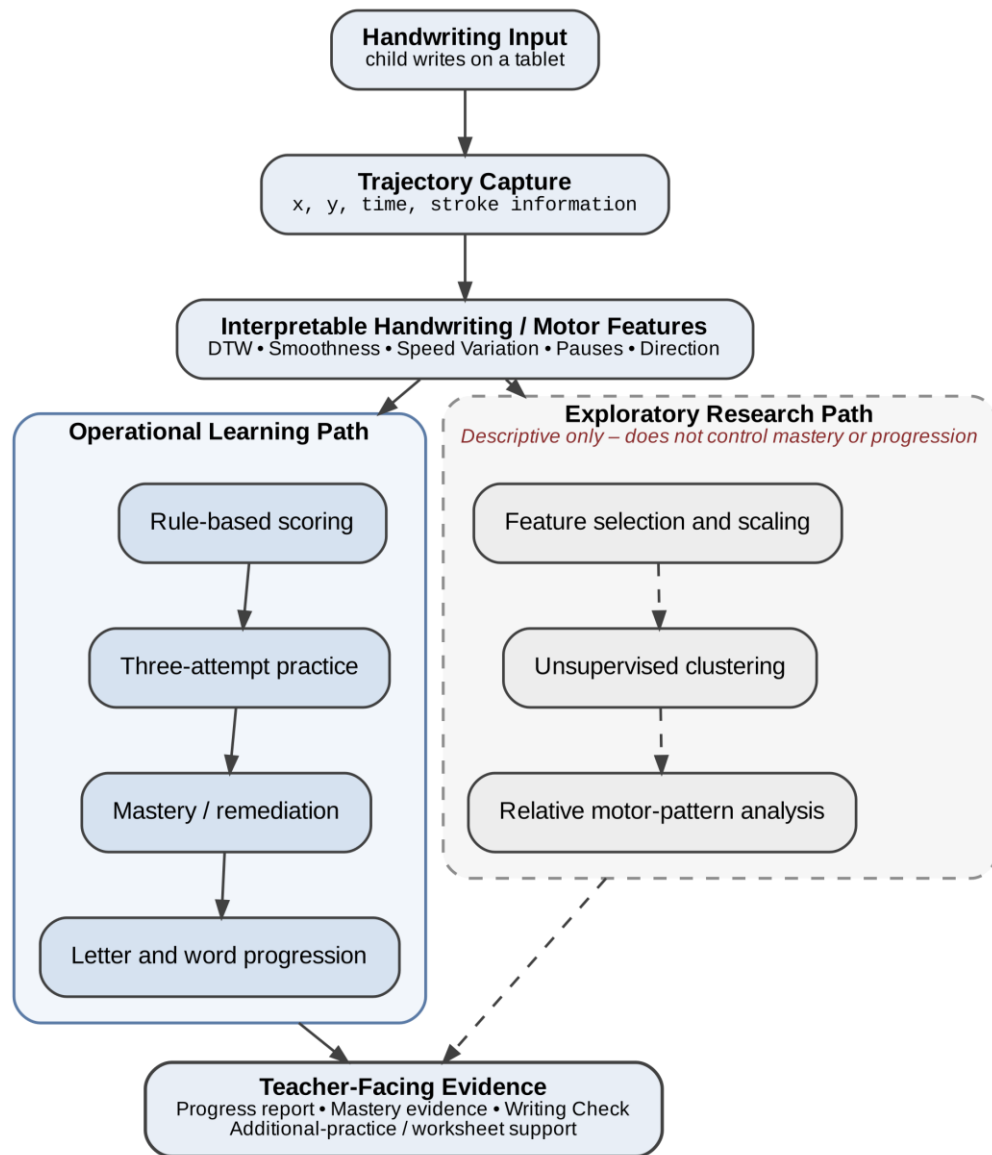


Figure 2.1: Conceptual framework of the motor-informed handwriting learning component

3 RESEARCH METHODOLOGY

3.1 Research design

This research follows a design-oriented mixed-methods approach. A software artefact, the handwriting learning component, is designed and developed to address a practical problem, while evidence about the artefact and the data it produces is gathered through quantitative analysis and user evaluation. This approach fits the dual nature of the work, which includes both a research contribution and a product development contribution.

The research was carried out in six overlapping stages, as shown in Figure 3.1. The problem and research gap were first identified through the literature review and discussions with the supervisor and project team. Requirements were then refined, and the component was developed iteratively. Pilot handwriting data were collected during a school visit and analysed to examine handwriting features, scoring behaviour and exploratory motor-pattern structure. Findings from this analysis informed further refinement of the component, including changes to scoring, progression logic and data-handling mechanisms. The final stage includes software verification, integration testing and usability evaluation with users of the system.

The research questions are addressed using different forms of evidence. RQ1 is examined through handwriting feature extraction and trajectory analysis, RQ2 through the design and verification of the rule-based learning workflow, RQ3 through exploratory clustering and stability analysis, and RQ4 through usability testing and user feedback.

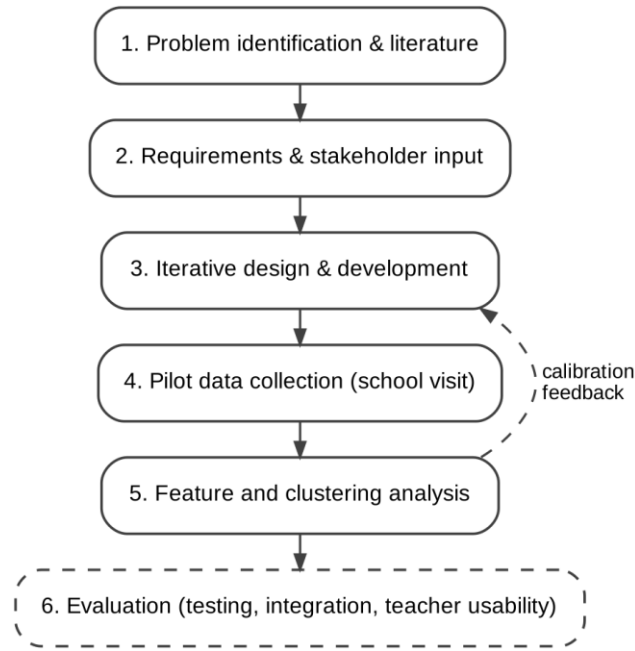


Figure 3.1: Research design stages (Stage 6 - Evaluation in progress)

The research questions are addressed using different forms of evidence. RQ1 is examined through feature extraction, DTW analysis and score calibration; RQ2 through the design and verification of the rule-based practice workflow; RQ3 through exploratory clustering and stability analysis; and RQ4 through usability evaluation.

3.2 Data participants or study materials

3.2.1 Pilot participants

Pilot handwriting data were collected from children receiving special-education support at a partner school during the dedicated data-collection sessions. Participants were identified through the school, and participation followed the consent procedures described in Section 3.6. Pseudonymous participant identifiers were used for research analysis, and the system did not collect diagnostic or clinical information about the children.

Two separate datasets were derived from the pilot collection because they were used for different analyses:

Dataset A - Initial-shape dataset: Contains trajectories from the six pre-writing shapes completed by 7 participants. This dataset was used to explore differences in shape-level handwriting and motor features between participants.

Dataset B - Letter dataset: Contains letter trajectories from 7 participants, with each participant completing 20 reference letters using three attempts per letter under the fixed collection protocol. The least-supported Attempt 3 records were used for the exploratory letter motor-pattern analysis.

The two datasets are analysed separately because the shape assessment and letter-writing tasks represent different forms of handwriting evidence and address different parts of the research.

3.2.2 Study materials and equipment

Data collection was conducted using the Auriva mobile application on a Samsung Galaxy Tab S7+ (SM-T970) running Android. The same tablet was used for all controlled data-collection sessions in landscape orientation to minimise device-related variation. Handwriting input was entered using a stylus on the touchscreen, with a fixed 791×427 dp drawing canvas. The application recorded spatial coordinates, timestamps and stroke boundaries for each handwriting trajectory. Stylus pressure was not recorded.

The controlled assessment consisted of six pre-writing tasks: horizontal line, vertical line, full circle, half circle, zigzag and curve-wave. These represented straight, curved and more complex movement patterns. The letter-data collection additionally used 20 reference letters, with each participant completing three attempts per letter. The same task definitions and reference trajectories used by the handwriting component were used during controlled data collection to maintain consistency between the research data and the implemented system.

The recorded trajectory data provided the basis for calculating handwriting characteristics such as trajectory distance, duration, speed variation, smoothness, pauses, directional changes and DTW-based similarity during subsequent analysis.

3.2.3 Teacher participants for usability evaluation

The usability evaluation involved teachers who regularly support children during learning and handwriting activities at the participating special-education setting. Their experience working directly with the target learners made them suitable for evaluating the practical usefulness of the handwriting component, particularly the progress reports, mastery evidence, Writing Check information, recommendations and worksheet-support workflow.

Teachers participated in task-based evaluation sessions using the implemented system and provided feedback on the usability, clarity and usefulness of the child-facing and teacher-facing features. Their responses were used as stakeholder evidence for evaluating whether the component could be understood and practically used in an educational setting.

3.3 Experimental or study procedure

3.3.1 Pilot data-collection procedure

A dedicated data-collection mode was used to ensure that all participants completed the same controlled protocol. In this mode, adaptive gating was disabled so that task progression was not altered by the learner's score, and collection records were marked separately from normal-practice data using a collection-session identifier.

Each participant completed the session under supervision in the participating special-education setting. The procedure was as follows:

- A pseudonymous learner profile was selected and the collection session was started.
- The participant completed the six pre-writing tasks in a fixed order: horizontal line, vertical line, full circle, half circle, zigzag and curve-wave.
- The participant then completed 20 reference letters, with three attempts per letter using decreasing levels of support.
- Under the controlled collection protocol, Attempt 3 retained a faint visual guide. This differs from normal practice, where Attempt 3 is fully unguided.

- Each completed attempt was submitted to the backend and stored with its trajectory and derived handwriting information. A raw-payload safety log was also retained to reduce the risk of data loss if a database write failed.
- After data collection, stored records were checked using database verification procedures before being included in the research analysis.

Across the seven participants, the controlled protocol produced 42 initial-shape trajectories and 420 repeated letter attempts. For the controlled letter-level motor-pattern analysis, one eligible Attempt-3 observation per participant–letter pair was retained, resulting in 140 observations.

3.3.2 Teacher usability evaluation procedure

The teacher usability evaluation was conducted as a task-based study followed by a questionnaire survey. Each teacher completed representative tasks using the handwriting component, including reviewing a learner’s initial assessment summary, interpreting letter mastery evidence, examining the Writing Check information, generating a worksheet from a recommendation, and uploading a completed worksheet.

During the task-based evaluation, task completion, difficulties and required assistance were observed. After completing the tasks, teachers completed the System Usability Scale (SUS) [41] and additional questions on the clarity, usefulness and understandability of the progress information and recommendations. The evaluation procedure and questionnaire are provided in Appendix A.

3.4 Measures and evaluation metrics

3.4.1 Handwriting and motor features

The handwriting component derives a set of spatial, temporal and movement-related features from each recorded trajectory. The mobile application captures the raw x- and y-coordinate sequence, timestamps and stroke boundaries. Selected measures required by the letter-writing pipeline, including DTW-based trajectory comparison and smoothness-related inputs, are computed on the client, while the backend

validates and normalises the submitted values, derives additional motor measures, computes the authoritative Motor Score and applies the mastery policy.

The extracted features include trajectory distance, duration, average speed, speed variation, smoothness, pause behaviour, directional changes, DTW distance, stroke count and completion-related information. These measures were selected because they provide complementary information about how the handwriting movement was produced rather than relying only on the final written shape. The raw trajectories and derived features are retained for scoring, progress monitoring and research analysis.

The features used in the current implementation are summarised in Table 3.1.

Table 3.1: Handwriting and motor features derived for each attempt

Feature	Description	Interpretation	Computed in
Path distance	Total spatial length of the recorded handwriting trajectory	Indicates the amount of movement used to complete the task; unusually long paths may reflect over-tracing or wandering	Client; retained and validated by backend
Duration	Time between the beginning and end of the handwriting attempt	Overall time taken to complete the writing task	Client; backend can derive a fallback from timestamps
Average speed	Total trajectory distance divided by writing duration	General movement pace during handwriting	Client; backend can recompute as a fallback
Speed variation (speed CV)	Coefficient of variation of instantaneous segment speeds	Higher values indicate less consistent movement speed	Backend
Smoothness	Mean angular change between successive trajectory segments, converted to a normalised smoothness score	Higher scores indicate smoother movement with fewer abrupt directional changes	Client; normalised/used by backend
Pauses	Number of consecutive captured points separated by more than 300 ms	Indicates interruptions or hesitation during writing	Client; retained by backend
Direction reversals	Score derived from the proportion of directional reversals greater than 90°	Lower scores indicate more frequent sharp reversals or back-and-forth corrections	Backend
DTW distance	Normalised Dynamic Time Warping distance between the learner trajectory and the reference trajectory	Lower values indicate greater trajectory similarity to the reference	Client for letter comparison; retained/normalised by backend

Feature	Description	Interpretation	Computed in
Stroke count	Number of separate captured strokes	Describes how many stylus contacts were used to produce the writing	Captured on client; validated/derived by backend

3.4.2 Scoring measures

Two separate scoring approaches are used in the handwriting component and are treated as different score domains.

For letter practice, the backend computes a Motor Score on a 0–100 scale using five normalised components:

$$M = 0.35A + 0.25S + 0.15Pc + 0.15R + 0.10V$$

where A represents accuracy, S smoothness, Pc pause control, R direction control, and V speed control. Accuracy is derived from the available accuracy measure and, for letter records where a direct accuracy value is unavailable, a normalised DTW-based value may be used. The score is null-aware; if a component is unavailable, its weight is excluded and the remaining weights are renormalised. The coefficients were selected as pilot engineering weights and were not statistically learned or clinically validated.

A separate scoring method is used for the six initial pre-writing shapes. The current shape score combines trajectory similarity and movement smoothness as:

$$\text{Shape Score} = 0.80 (\text{DTW Score}) + 0.20(\text{Smoothness Score})$$

The shape score is used only within the initial-assessment workflow and is not directly compared with the letter Motor Score. Keeping these two scoring domains separate avoids treating scores produced from different tasks and feature combinations as equivalent.

For operational letter practice, mastery is evaluated using the least-supported third attempt and the current mastery threshold. The scoring weights and threshold values are prototype settings and should therefore be interpreted as engineering parameters rather than validated clinical measures.

3.4.3 Clustering evaluation metrics

Because the clustering task does not have validated ground-truth labels, conventional classification accuracy is not used. Instead, clustering performance is evaluated using the following measures:

- Silhouette Coefficient: ranges from -1 to 1 . Higher values indicate better separation between clusters [42].
- Davies–Bouldin Index (DBI): lower values indicate better-separated and more compact clusters [43].
- Caliński–Harabasz Index (CH): higher values indicate better-defined cluster structure [44].
- Adjusted Rand Index (ARI): used to measure the similarity between cluster partitions and to evaluate clustering stability [45].
- Leave-One-Participant-Out (LOPO) analysis: the clustering process is repeated while excluding one participant at a time to examine how stable the resulting grouping remains when the participant set changes.

Repeated K-Means initialisations are also used to check whether the selected partition remains consistent across different starting conditions.

3.4.4 Software and usability metrics

Software quality was evaluated using automated unit, regression, parity and error-path testing across the mobile application, backend and ML service. Functional test cases were mapped to the system requirements defined in Section 4.1, while manual end-to-end testing was used to verify communication between the mobile application, backend, database and ML service on the target tablet environment.

Usability was evaluated using task completion, observed difficulties, assistance required during tasks, System Usability Scale (SUS) scores, and teacher responses regarding the clarity, usefulness and understandability of the system. Open-ended feedback was also considered to identify usability issues and possible improvements to the handwriting and teacher-facing workflows.

3.5 Data analysis

3.5.1 Trajectory preprocessing and DTW

The letter-template pipeline runs on the mobile client. Before comparison, learner and reference trajectories are normalised into a common coordinate space so that differences in canvas size do not distort the comparison. Each letter is defined using a shared set of ordered waypoints in `LETTER_PATHS`, which is also used to generate the visual guide, tracer and DTW reference. This keeps the displayed letter and the reference used for comparison consistent. Multi-stroke letters are handled by matching learner strokes with the most appropriate reference strokes before the DTW distance is calculated.

DTW is used to compare the learner trajectory with the reference while allowing differences in writing speed and sequence length. During development, the comparison was also made direction-invariant by evaluating both the original and reversed learner trajectory and retaining the smaller DTW distance. This reduces unnecessary penalties when a similar path is produced in the opposite drawing direction. The controlled research analysis used the same direction-invariant approach for the relevant trajectory comparisons.

For letters containing multiple strokes, stroke matching is performed before DTW calculation so that the result is less dependent on the exact order in which the learner completed the strokes. All reference letters used in the controlled protocol were checked against the implemented templates to maintain consistency between the practice interface and the trajectory-analysis process.

3.5.2 Score calibration

During development, multiple scoring approaches existed across different parts of the handwriting pipeline. Analysis of pilot attempts showed that some earlier formulas produced compressed scores, with visually different attempts receiving similarly high values. To improve consistency, letter scoring was consolidated around the backend Motor Score described in Section 3.4.2, while the initial-shape assessment continues to use its separate shape-specific scoring method.

The current Motor Score combines accuracy, smoothness, pause control, direction control and speed control using fixed prototype weights. These weights were selected during development rather than learned statistically from the participant data. The current operational mastery threshold is also treated as an engineering parameter rather than a clinically validated value.

Dynamic or family-specific performance thresholds were explored during development but are not enabled in the current operational workflow. Letter mastery therefore remains based on the fixed mastery policy and the least-supported third attempt. Further calibration using a larger and more diverse handwriting dataset would be required before introducing dynamically changing thresholds or statistically optimised scoring weights.

3.5.3 Exploratory clustering

Clustering analysis was conducted in Python using scikit-learn [46]. Repeated observations from the same learner were not treated as independent participants, and participant-level stability was considered during evaluation.

For Dataset A, the six initial-shape tasks from seven participants were aggregated into participant-level representations. Six features were retained: straight-task trajectory distance, curved-task trajectory distance, complex-task trajectory distance, straight-task accuracy, curved-task accuracy, and complex-task DTW distance. Complex-task DTW was transformed using $\log(1+x)$, and the features were normalised using RobustScaler. K-Means, Ward hierarchical clustering, and Gaussian Mixture Models (GMM) were evaluated for $k = 2$ and $k = 3$. Cluster structure was compared using the Silhouette Coefficient, Davies–Bouldin Index, Calinski–Harabasz Index and Adjusted Rand Index. Stability was additionally examined across repeated K-Means initialisations and leave-one-participant-out analysis.

For Dataset B, one eligible Attempt-3 observation was retained for each participant-letter pair, producing 140 letter-level observations from seven participants and 20 reference letters. The letter motor-pattern analysis uses three features: smoothness score, DTW distance and speed coefficient of variation. The 140 Attempt-3

observations were then aggregated into seven participant-level feature vectors, one for each participant, and the StandardScaler and the fixed two-cluster K-Means model were fitted to these seven vectors. The fitting unit is therefore the participant, not the individual letter attempt. A reference-range check is also used to avoid forcing a cluster assignment when new evidence falls outside the range represented by the model data.

The clustering outputs are interpreted only as relative motor-pattern groupings within the available research data. They are not treated as diagnostic categories, ASD severity levels or measures of learner ability, and they do not control mastery, progression, activity sequencing or child-facing support.

3.6 Ethics privacy security and responsible AI

3.6.1 Ethical approval and consent

Ethical clearance for the study was obtained from the Sri Lanka Institute of Information Technology (SLIIT). Permission to conduct the data-collection activities was also obtained from the administration of the participating special-education school before the study commenced. A formal request describing the planned data-collection procedure was submitted to the school and approved in advance

For the child participants, consent was provided by the school under its institutional care arrangements on behalf of guardians. Separate written consent from individual parents or guardians was not obtained, and this is acknowledged as a limitation of the consent procedure. Participation remained voluntary at the child level; children were free to decline or stop participating at any time and were not encouraged to continue once they disengaged.

For the teacher usability evaluation, teachers were provided with a consent statement explaining that participation was voluntary, that the evaluation concerned the application rather than the teacher, that names were not recorded on the survey form, and that participants could stop at any time. Teachers indicated their agreement before taking part in the usability tasks.

3.6.2 Privacy and data protection

The handwriting component follows a data-minimisation approach by collecting only information required for handwriting analysis, progress monitoring and research. Stored information includes handwriting trajectories, derived motor features, scores and learning-progress records. The system does not record video, audio or photographs of children during handwriting activities.

Research data are pseudonymised before analysis and reporting. Participant records are re-coded using study-specific identifiers, and identifying information is not included in the analysed research dataset.

Completed paper worksheets may be uploaded by teachers as images for later review. Because these images contain the learner's handwriting and may potentially include identifying information, they are treated as personal data and access is restricted to authorised users. The privacy requirements of the component therefore include pseudonymous research identifiers, data minimisation and restricted access to worksheet images.

The handling of personal data is designed with reference to Sri Lanka's Personal Data Protection Act, No. 9 of 2022 [47]. The system is treated as a research prototype, and stronger production-level controls would be required before deployment with larger populations.

3.6.3 Security

The handwriting component uses JSON Web Token (JWT) authentication [48] together with role-based and learner-ownership controls. Teacher-facing records are returned only to authenticated users who are authorised to access the relevant learner. These controls reduce the risk of unauthorised access to handwriting, progress and worksheet information.

The current implementation is a research prototype and is not yet suitable for production deployment. A security review identified several remaining limitations. The mobile application currently communicates with the backend through a local HTTP endpoint rather than HTTPS/TLS, and the backend address is configured directly in the frontend. The current development configuration also includes

credentials and secrets that require strengthening before deployment. In addition, the FastAPI ML service does not currently enforce its own authentication, and the CORS configuration can fall back to a permissive setting.

These limitations are treated as deployment requirements rather than completed security controls. Production deployment would require HTTPS/TLS communication, secure and rotated secrets, restricted CORS configuration, stronger administrative credentials, and authenticated communication between the backend and the ML service. Security considerations were also reviewed against common risks such as broken access control and security misconfiguration identified in the OWASP Top 10 [49].

3.6.4 Responsible AI

Responsible AI principles are incorporated into the handwriting component to ensure that automated analysis remains transparent and supportive of teacher decision-making. The system is designed so that educational decisions can be understood and reviewed rather than being controlled by an opaque machine-learning model.

Three safeguards are applied:

1. Mastery, progression, remediation and word unlocking are controlled by explicit rule-based logic. The K-Means clustering model does not determine these operational learning decisions.
2. Motor-pattern information is presented to teachers as descriptive evidence only. Cluster outputs are not shown as diagnostic labels or fixed classifications of the learner.
3. The system does not diagnose ASD, estimate ASD severity or predict intelligence. Final educational interpretation and decisions remain with the teacher.

3.7 Reproducibility and research artefacts

The following measures were used to improve reproducibility and traceability of the handwriting component:

- **Version control.** Source code for the mobile application, backend and ML service is maintained in the project GitHub repository. This provides a record of implementation changes and supports consistent development across the three system components.
- **Data provenance.** Research collection records include collection-mode and session identifiers, allowing controlled research data to be distinguished from normal application use and traced back to the relevant collection session.
- **Data integrity.** Additional safeguards were introduced after development testing identified a schema-related data-loss risk. The backend includes checks for the expected database structure, while a raw-payload safety log retains collection requests so that records can be recovered if a database write fails.
- **Verification scripts.** Database records and collection results are checked using repeatable scripts rather than relying only on manual inspection. This supports consistent verification of the collected datasets.
- **Model artefacts.** The exploratory clustering models and their preprocessing information are stored as versioned artefacts. The letter motor-pattern implementation includes saved model metadata and model files used by the FastAPI ML service. Model loading and prediction behaviour are also checked to ensure that saved artefacts reproduce the expected cluster assignments.
- **Data availability.** Raw handwriting data are not publicly released because the dataset concerns children receiving special-education support. Research records used for analysis are pseudonymised, and de-identified feature-level information may be provided for academic examination where permitted by the study conditions.

4 PRODUCT AND SOLUTION DESIGN

4.1 Requirements

The requirements of the handwriting component were derived from the research objectives in Section 1.5, the literature reviewed in Chapter 2, supervisor feedback and the needs of teachers who support learners during handwriting activities. Functional requirements describe what the component does for the learner and the teacher, and they are listed in Table 4.1. Non-functional requirements describe the qualities that the component should have, and they are listed in Table 4.2. Some non-functional requirements describe the target for a production deployment rather than the current research prototype. Where this is the case, the gap between the requirement and the current implementation is reported in Chapters 5 and 6.

Table 4.1: Functional requirements of the handwriting component

ID	Requirement	Objective
FR1	The system shall capture each handwriting attempt as time-ordered x and y coordinates with timestamps and stroke boundaries during the initial assessment, letter practice, word writing and Writing Check.	O1
FR2	The system shall provide an initial assessment of six pre-writing shapes (horizontal line, vertical line, full circle, half circle, zigzag and curve-wave), calculate a shape score and create an initial summary and baseline for the learner.	O1, O2
FR3	The system shall derive the trajectory and motor measures listed in Table 3.1 and compare learner trajectories with reference trajectories using direction-invariant DTW.	O2
FR4	The backend shall calculate a letter Motor Score from 0 to 100 using the weighted, null-aware formula in Section 3.4.2, and this score shall be the authoritative input to the letter mastery policy.	O2
FR5	The system shall provide lowercase and uppercase letter practice and store lowercase and uppercase mastery separately.	O3
FR6	Each practice cycle shall contain three attempts with decreasing support. Attempt 3 shall be the mastery attempt, and a letter shall be mastered when the Attempt-3 Motor Score is at least 70.	O3
FR7	The system shall allow at most three cycles, and therefore at most nine attempts, per letter, case and day, using Asia/Colombo dates.	O3
FR8	A failed Cycle 1 shall lead to Cycle 2. A failed Cycle 2 shall lead to exact-letter remediation of up to two warm-up activities before Cycle 3.	O3
FR9	A failed Cycle 3 shall move the learner forward with the letter left unmastered and shall create an Additional Home Practice candidate.	O3, O5
FR10	A technically incomplete Attempt 3 shall be retried without consuming the cycle.	O3

ID	Requirement	Objective
FR11	The system shall provide word Activities A–E for each word. A word shall be complete only when all five activities are marked correct or good.	O3
FR12	Word writing shall pass when the word-writing score is at least 50 and the required letter-region coverage is met. Spacing and size feedback shall be advisory only.	O3
FR13	Word groups A, B and C shall be unlocked initially. Group D shall unlock after all catalogue words for A, B and C are complete, and groups E–Z shall then unlock in sequence.	O3
FR14	The system shall keep the exact Attempt-3 record that established mastery as mastery evidence.	O5
FR15	The system shall provide a Teacher Report covering the initial assessment, motor performance, letter mastery, word progress, Writing Check, motor-pattern information, worksheets and recommendations.	O5
FR16	The system shall allow a teacher to start a Writing Check, let the learner complete the handwriting input and return the result to the teacher for review, with teacher ownership enforced.	O5
FR17	The teacher action "Generate Worksheet" shall produce an A4 PDF worksheet, and the teacher shall be able to upload a photo of the completed worksheet for later review. Worksheet completion shall not change mastery or scoring.	O5
FR18	The backend shall obtain a descriptive motor-pattern summary from the ML service. This summary shall not affect mastery, progression, remediation or word unlocking.	O4
FR19	The system shall provide a data-collection mode with a fixed protocol and disabled gating, and shall mark collected records with a collection-session identifier.	O1, O4

Table 4.2: Non-functional requirements of the handwriting component

ID	Category	Requirement
NFR1	Usability (learner)	Large writing area, simple instructions, consistent screen layout and immediate, non-punitive feedback.
NFR2	Usability (teacher)	Progress information shall be understandable without technical knowledge, and every recommendation shall show the reason for it.
NFR3	Performance	Feedback for an attempt should be returned quickly enough not to interrupt practice.
NFR4	Reliability	Collected attempts shall not be lost silently, and repeated submission or finalisation of the same record shall not create duplicate results.
NFR5	Security	API access shall require JWT authentication with role and ownership checks. A production deployment shall also use HTTPS/TLS, managed secrets, restricted CORS and authenticated backend-to-ML communication.
NFR6	Privacy	Data minimisation, pseudonymous research identifiers and restricted access to worksheet images.

ID	Category	Requirement
NFR7	Interpretability	Operational learning decisions shall be rule-based and explainable, and ML output shall be presented as descriptive information only.
NFR8	Maintainability	Letter shapes shall have one definition (LETTER_PATHS), letter mastery shall use one authoritative scoring policy, and each layer shall be covered by automated tests.
NFR9	Compatibility	Letter templates and scoring shall not be distorted by differences in canvas size or screen aspect ratio.
NFR10	Accessibility	Screen design should follow WCAG 2.1 principles where they apply, such as sufficient contrast and not relying on colour alone [50].

4.2 Component architecture within the complete solution

Auriva is an adaptive learning platform for English language development in autistic children in Sri Lanka. All of its components share one three-tier architecture, shown in Figure 4.1. The React Native / Expo mobile application is the client. It communicates with a Node.js / Express backend through REST requests made with Axios, and each request carries a JSON Web Token (JWT). The backend contains the business logic, including the authoritative letter Motor Score, the mastery policy, word unlocking and worksheet plans, and it accesses a PostgreSQL database through the Sequelize ORM. Machine-learning functions run in a separate Python FastAPI service that only the backend calls. The mobile application never communicates with FastAPI directly. In the research deployment used in this study, the mobile application reaches the backend through a local-network HTTP endpoint. A production deployment would require HTTPS/TLS (Section 3.6.3).

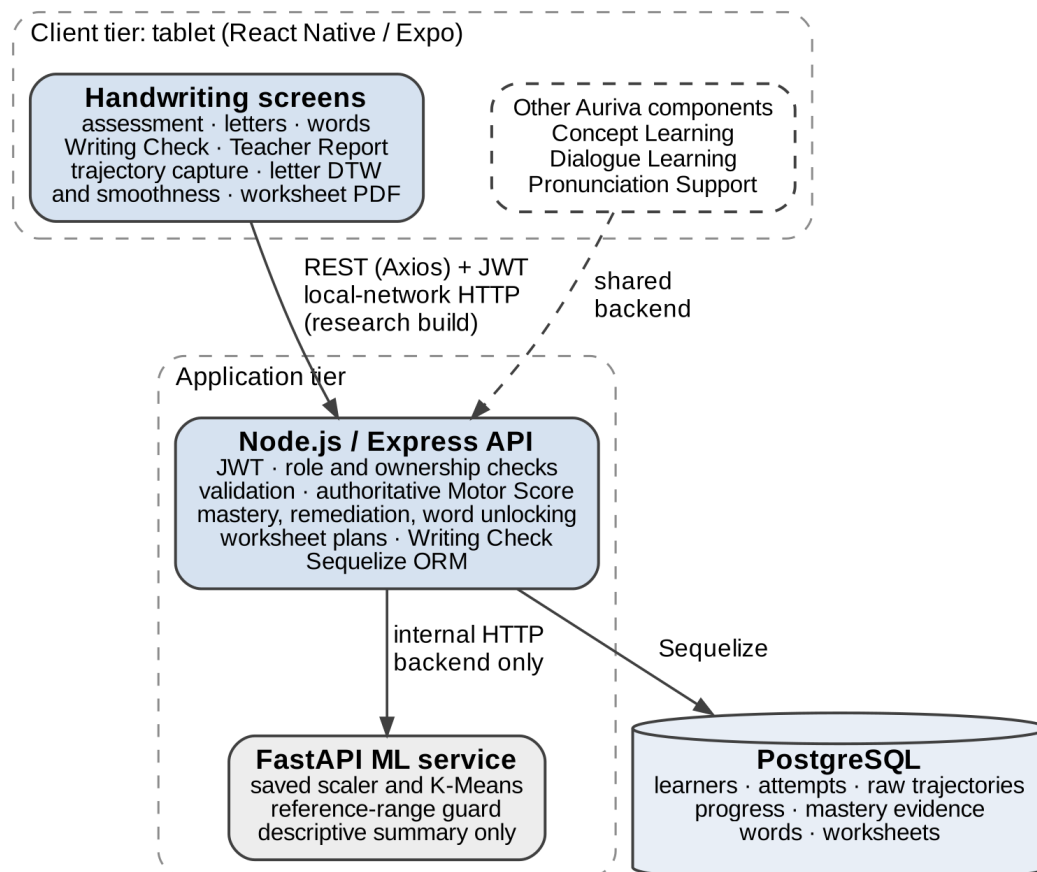


Figure 4.1: High-level architecture of Auriva showing the handwriting component and its interfaces

The handwriting component spans all three tiers. On the client it includes the handwriting screens, trajectory capture, the LETTER_PATHS template pipeline and the calculation of letter DTW and smoothness inputs. On the backend it includes handwriting routes, controllers, scoring utilities, sequencing rules, Writing Check endpoints and worksheet planning. It also includes the handwriting-related tables in the shared database and the letter motor-pattern model in the ML service. Table 4.3 summarises the four Auriva components and how they relate to the handwriting component.

Table 4.3: Auriva components and their relationship to the handwriting component

Component	Owner (student ID)	Purpose	Relationship with the handwriting component
-----------	-----------------------	---------	---

Component	Owner (student ID)	Purpose	Relationship with the handwriting component
Motor-informed adaptive handwriting learning (this thesis)	IT22216700	Handwriting assessment, adaptive letter and word practice, and teacher-facing handwriting evidence	Not applicable
Adaptive Concept Learning Module	IT22084590	Foundational English vocabulary taught through a three-tier, mastery-based concept framework with sensory-preference personalisation	Shares the Auriva mobile application and backend
Adaptive Dialogue Learning and Progression Tracker	IT22166838	Functional spoken English dialogue taught across three levels aligned with the Grade 1–3 English syllabus	Shares the Auriva mobile application and backend
Pronunciation Support Component	IT22266378	Word-level and phoneme-level pronunciation practice with automated speech analysis and teacher analytics	Shares the Auriva mobile application and backend

The main external interfaces of the handwriting component are summarised in Table 4.4.

Table 4.4: External interfaces of the handwriting component

Interface	From → To	Protocol	Main content	Security in the research build
Attempt submission	App → Backend	REST (JSON) over local-network HTTP	Learner, activity, letter, case, attempt number, raw trajectory and client-computed letter features	JWT bearer token; ownership checks
Teacher Report retrieval	App → Backend	REST (JSON)	Assessment summary, motor performance, mastery, word progress, Writing Check, motor patterns, worksheets and recommendations	JWT; role and ownership checks
Writing Check	App ↔ Backend	REST (JSON)	Start, progress and complete requests for a teacher-initiated check	JWT; teacher ownership enforced
Worksheets	App ↔ Backend	REST (JSON); image upload	Worksheet plan from the backend; photo of the completed worksheet to the backend	JWT; restricted access to images
Data persistence	Backend → PostgreSQL	Sequelize ORM	Attempts, trajectories, features, scores, progress and evidence	Database credentials in backend configuration

Interface	From → To	Protocol	Main content	Security in the research build
Motor-pattern analysis	Backend → FastAPI	Internal HTTP (JSON)	Feature values in; cluster and reference-range summary out	Not reachable from the app; no service-level authentication in the current build

4.3 Design alternatives and decisions

The main design decisions and the alternatives considered are summarised in Table 4.5.

Table 4.5: Main design decisions and alternatives considered

Decision	Alternatives considered	Choice	Justification
Mobile framework	Native Android or iOS; Flutter	React Native / Expo	Shared by the Auriva team; one codebase for Android and iOS tablets; JavaScript on both client and backend
Backend	Python Django or Flask; serverless functions	Node.js / Express with Sequelize	Shared Auriva backend; same language as the client; established middleware for JWT and request validation
Database	Document database such as MongoDB	PostgreSQL	Learners, attempts, cycles, mastery evidence and words are strongly related records that suit a relational model with integrity constraints
ML hosting	Python inside the Node process; ML on the device	Separate FastAPI service called by the backend	Uses the Python data-science libraries; keeps ML away from the client; can be updated or disabled without changing practice logic
Trajectory comparison	Point-wise Euclidean distance; image-based classifier	Direction-invariant DTW [30]	Tolerates differences in writing speed and sequence length; needs no labelled training data; easy to explain as a distance
Location of letter DTW	Server-side computation	Client-side, using the same LETTER_PAT HS templates as the display	Keeps the reference used for scoring identical to the guide shown to the child. The trade-off is that the backend cannot independently recompute letter DTW (Section 7.5)
Progression logic	ML-driven personalisation; dynamic thresholds per learner or letter family	Fixed rule-based Attempt-3 threshold of 70	Transparent to teachers and safer with small pilot data. Dynamic thresholds were developed but are disabled in the operational system
Support across attempts	Same support for every attempt	Three attempts with decreasing support	Reduced guidance in the assessed attempt reflects the learner's own control (Section 2.2.3) [51], [52]

Decision	Alternatives considered	Choice	Justification
Exploratory clustering	Single algorithm	K-Means compared with Ward hierarchical clustering and GMM	Agreement across algorithms and initialisations is stronger evidence of structure than one algorithm alone [36], [38]

Dynamic or family-specific performance thresholds were explored during development. Logic that adjusts the target threshold from recent handwriting evidence was implemented, but it is disabled in the current operational workflow. There are three reasons for this. First, the pilot dataset is too small to show that changing thresholds would be fair and educationally valid. Second, the current Motor Score constants still show some compression (Section 3.5.2). Third, a fixed threshold is easier for teachers to interpret. Letter mastery is therefore decided only by the fixed threshold of 70 applied to Attempt 3, and dynamic thresholds are treated as future work (Section 9.4).

4.4 Detailed design

4.4.1 Handwriting learning workflow

The learner workflow is shown in Figure 4.2. The learner first completes the six-shape initial assessment, which produces a shape score for each task, an initial summary and a baseline that appear in the Teacher Report. The learner then practises letters in three-attempt cycles. Lowercase practice and uppercase practice are separate parts of the practice flow, and mastery is stored separately for each case. Letter practice is followed by word learning. Throughout the workflow, the evidence produced by each activity is collected in the Teacher Report. The teacher can also start a Writing Check at any time and review its result.

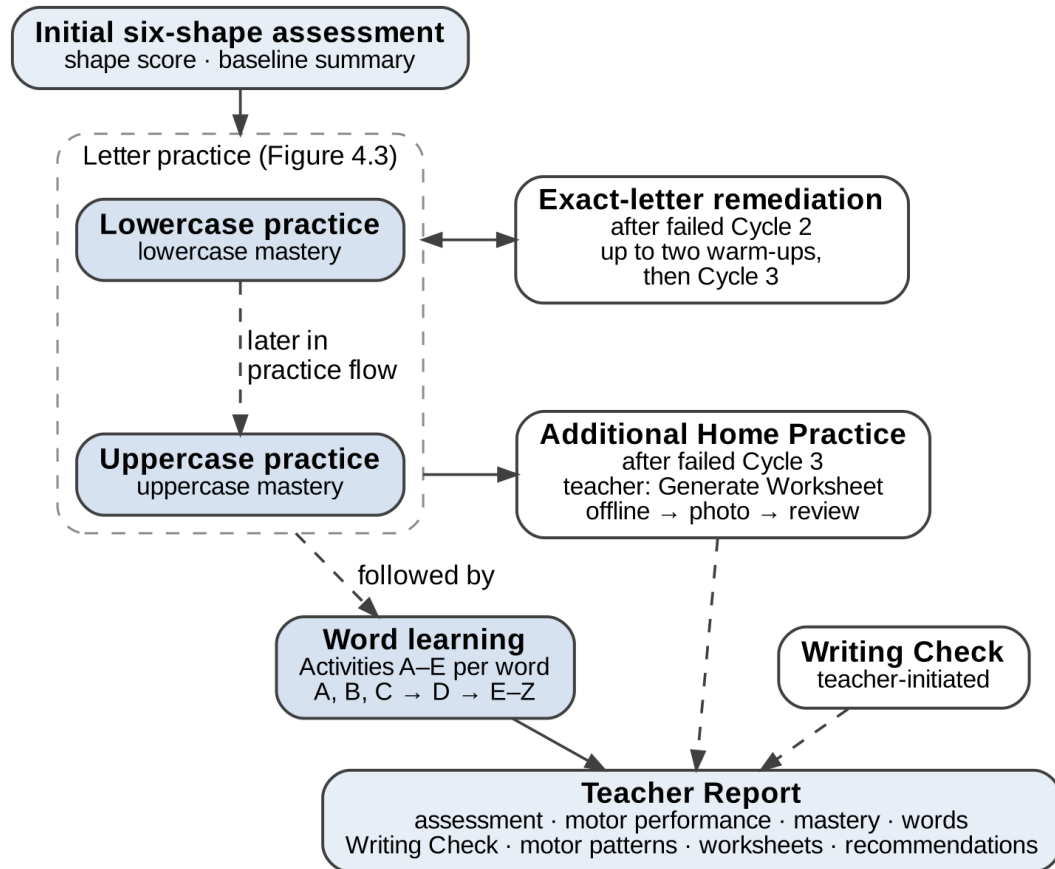


Figure 4.2: Handwriting learning workflow, including remediation, Additional Home Practice and the Teacher Report

4.4.2 Letter template system

Every letter is defined once in LETTER_PATHS as a list of strokes, each made of ordered waypoints in a normalised coordinate space. This single definition is used in three places: the ghost guide shown behind the writing area, the animated tracer that demonstrates stroke order, and the reference trajectory used for DTW. Before this design was introduced, the visual guide and the scoring reference were defined separately and could disagree, so a child could follow the guide accurately and still receive a low score. The redesign also corrected aspect-ratio distortion on wide tablet screens, grouped straight-line letters as ANGULAR_LETTERS so that their corners are kept, and added multi-stroke support in which learner strokes are matched to reference strokes before DTW is calculated. The 20 reference letters used in the controlled protocol were checked against the implemented templates.

4.4.3 Three-attempt practice cycle and mastery rule

The practice cycle for one letter and case is shown in Figure 4.3. Attempt 1 gives the most visual guidance, Attempt 2 gives reduced guidance, and Attempt 3 is unguided in normal practice. Only Attempt 3 is assessed for mastery. If Attempt 3 is technically incomplete, it can be retried without using up the cycle. If the Attempt-3 Motor Score is at least 70, the letter is marked as mastered for that case, and the exact Attempt-3 record is kept as mastery evidence.

If the cycle fails, the next step depends on which cycle it was. After a failed Cycle 1, the learner starts Cycle 2. After a failed Cycle 2, the learner completes exact-letter remediation of up to two warm-up activities and then starts Cycle 3. After a failed Cycle 3, the learner moves on with the letter left unmastered, and an Additional Home Practice candidate is created for the teacher. A learner can therefore complete at most three cycles, or nine attempts, per letter, case and day. Days are counted using Asia/Colombo dates.

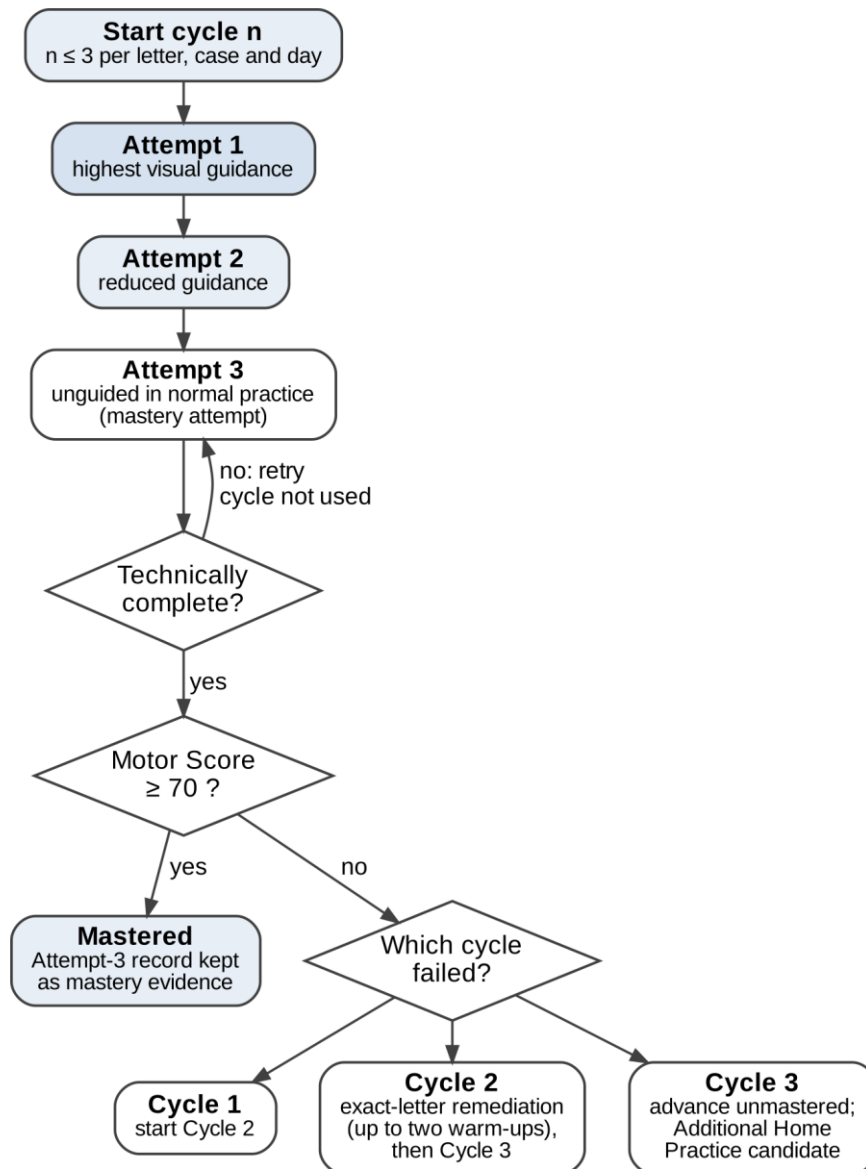


Figure 4.3: Three-attempt practice cycle, mastery rule, remediation and escalation for one letter and case

4.4.4 Attempt processing sequence

The processing of a single letter attempt is shown in Figure 4.4. Trajectory capture and the calculation of letter DTW and smoothness inputs take place on the client, because the letter-template pipeline is located there. The backend validates and normalises the submitted values, derives the remaining motor measures, calculates the authoritative Motor Score and applies the mastery policy. Keeping these decisions on the backend means that the rules are applied in the same way for every learner and can be audited from the stored records, which include the raw trajectory.

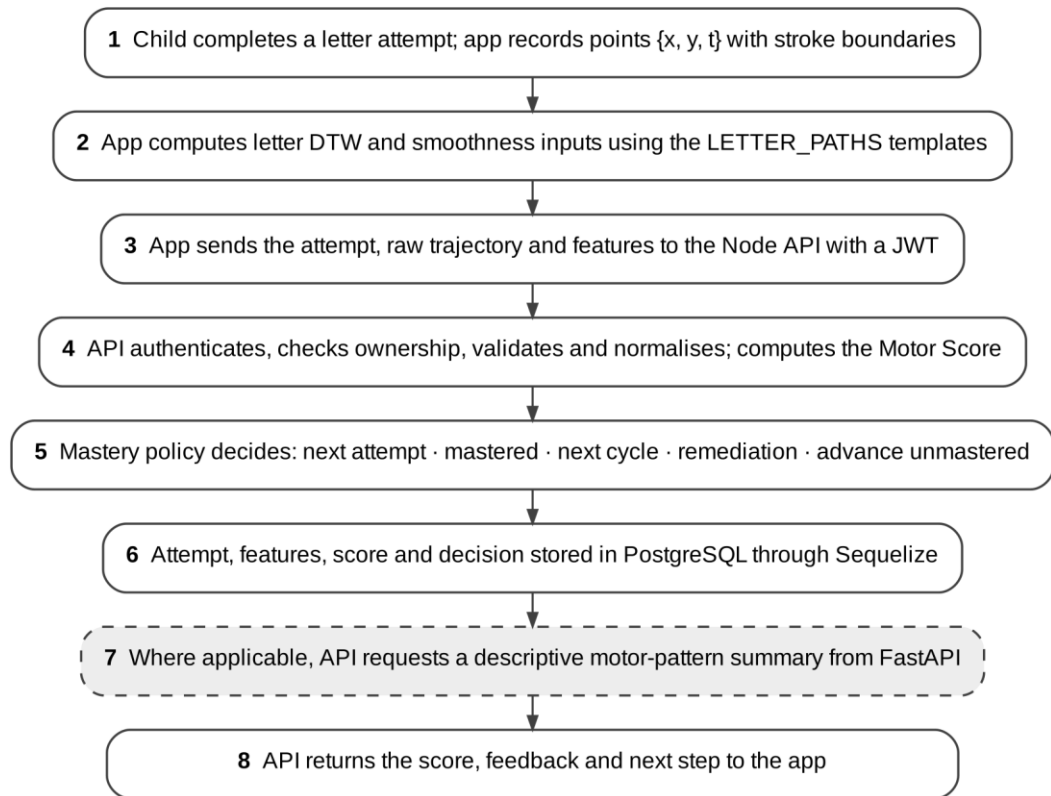


Figure 4.4: Processing sequence for a letter attempt

4.4.5 Word learning

Word learning uses a catalogue of 154 unique words grouped by initial letter. For each word, the learner completes the five activities shown in Table 4.6. A word is complete only when all five activities are marked correct or good. Word groups A, B and C are available from the start. Group D unlocks after all catalogue words in groups A, B and C are complete, and groups E to Z then unlock one after another as each previous group is completed.

Table 4.6: Word-learning activities

Activity	Description	Evidence recorded
A – First letter	Learner identifies the first letter of the word	Correct or not
B – Picture	Learner matches the word with its picture	Correct or not
C – Missing letter	Learner completes the word by choosing the missing letter	Correct or not
D – Shuffled word construction	Learner builds the word from shuffled letters	Correct or not

Activity	Description	Evidence recorded
E – Word writing	Learner writes the whole word on the canvas	Word-writing score, letter-region coverage, advisory spacing and size feedback

Word writing uses its own score domain, which is separate from the letter Motor Score and the initial-shape score. A word-writing attempt passes when its score is at least 50 and the required letter-region coverage is met. Spacing and size feedback is shown to support the learner and the teacher, but it is advisory and is not the main pass criterion.

4.4.6 Teacher-facing design

The Teacher Report brings the handwriting evidence for a learner together in one place. It contains the initial assessment, motor performance, letter mastery, word progress, Writing Check, motor-pattern information, worksheet information and recommendations. Mastery evidence links to the exact Attempt-3 record that established mastery, so the teacher can see what the learner actually wrote.

Writing Check is a teacher-initiated handwriting-pattern check. The teacher starts a Writing Check, the learner completes the handwriting input, and the result is returned to the teacher for review. The check uses dedicated backend start, progress and complete endpoints, and teacher ownership is enforced, so a teacher can only run and view checks for their own learners.

The worksheet workflow connects digital practice with paper practice in the classroom. Three failed same-day cycles for a letter can create an Additional Home Practice candidate. Persistent difficulty is also tracked separately over a longer pattern of attempts. The teacher's "Generate Worksheet" action is the approval step. The backend stores the worksheet plan, and the mobile client renders it as an A4 PDF that can be saved, shared or printed. After the learner completes the worksheet offline, the teacher uploads a photo, and the worksheet is stored in the learner's history for later review. Completing a worksheet does not change mastery or scoring, and the system does not score worksheet images automatically. Judgement of the paper work stays with the teacher.

4.4.7 Motor-pattern analysis design

The design of the letter motor-pattern analysis is shown in Figure 4.5. It has two stages. In the model-fitting stage, the 140 eligible Attempt-3 observations from controlled Dataset B supply the letter-level evidence, which is aggregated into seven participant-level vectors of smoothness score, DTW distance and speed coefficient of variation. A `StandardScaler` and a two-cluster K-Means model are fitted to these seven vectors and saved as model artefacts. In the runtime stage, the backend sends a learner's Attempt-3 letter evidence for the three features to the FastAPI service. The service applies the saved scaler and centroids, uses a reference-range guard so that evidence outside the range represented by the model is not forced into a cluster, and returns a descriptive summary for the Teacher Report or research view.

The figure also shows that no path leads from the model to mastery, progression, remediation or word unlocking. These operational decisions are made only by the rules in Sections 4.4.3 and 4.4.5. Dataset A clustering (Section 3.5.3) is a separate research analysis and is not part of the learner workflow at all.

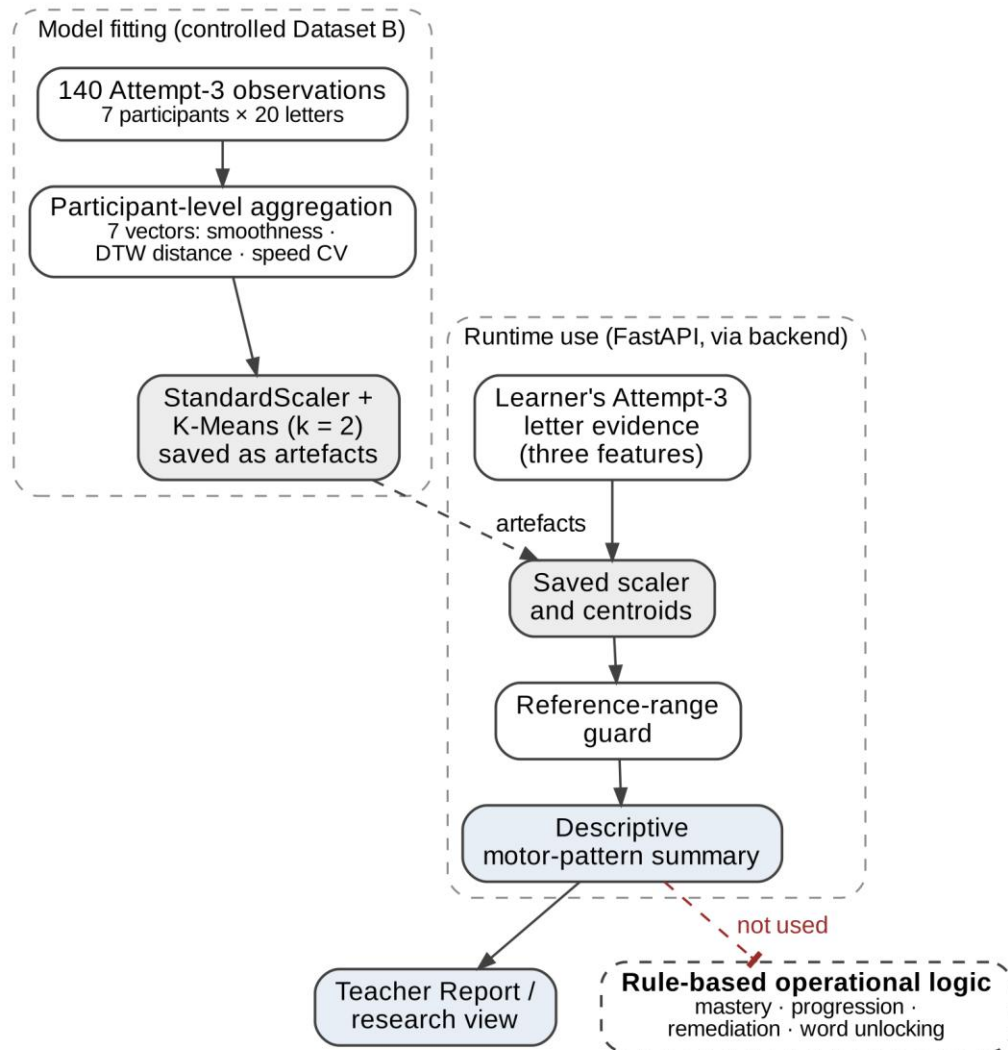


Figure 4.5: Letter motor-pattern analysis: model fitting, runtime use and separation from operational decisions

4.5 Standards and professional practice

The design and development follow recognised standards and practices where they apply to a student research prototype:

- **Software quality.** Quality attributes were considered using the ISO/IEC 25010 product quality model [53], particularly functional suitability, reliability, usability, security and maintainability. These attributes correspond to the non-functional requirements in Table 4.2.

- **Usability.** Usability evaluation is framed using ISO 9241-11, which defines usability in terms of effectiveness, efficiency and satisfaction [54], and it uses the System Usability Scale (SUS) [41].
- **Accessibility.** WCAG 2.1 principles [50], such as sufficient contrast and not relying on colour alone, were used as design guidance. A formal accessibility evaluation was not carried out.
- **Security.** Token-based authentication follows the JWT standard [48], and the OWASP Top 10 [49] was used as a checklist of common risks. The review identified transport, secret-management, CORS and ML-service authentication gaps in the research deployment (Section 3.6.3).
- **Data protection.** Personal data handling was designed with reference to Sri Lanka's Personal Data Protection Act [47].
- **Engineering practice.** The component was developed iteratively with version control on GitHub, automated tests at each layer (including scoring-parity and idempotent-finalisation tests), script-based database verification, a single definition of letter templates, a single authoritative policy for letter mastery, and documented audits when defects were found.

5 IMPLEMENTATION INTEGRATION AND TESTING

5.1 Development environment and technology stack

The technologies used in the handwriting component are summarised in Table 5.1. Most of them are shared with the rest of the Auriva platform, which simplifies integration.

Table 5.1: Technology stack of the handwriting component

Layer	Technology	Use in this component
Mobile client	React Native with Expo (JavaScript)	Handwriting canvas, initial assessment, letter practice, word learning, Writing Check and Teacher Report screens; worksheet PDF rendering
Client–server communication	Axios; JWT	Authenticated REST calls to the backend; local-network HTTP in the research build
Backend	Node.js with Express	Handwriting routes and controllers, scoring utilities, mastery and sequencing rules, word unlocking, Writing Check endpoints and worksheet plans
Data access	Sequelize ORM	Models and queries for learners, attempts, progress, words, evidence and worksheets
Database	PostgreSQL	Persistent storage for handwriting records
ML service	Python with FastAPI; scikit-learn [46]	Loading the saved scaler and K-Means artefacts and returning descriptive motor-pattern summaries
Research analysis	Python 3.13 with NumPy 2.2.1, pandas 2.2.3, SciPy 1.17.1 and scikit-learn 1.6.0	Feature reconstruction, clustering and stability analysis for Datasets A and B
Study device	Samsung Galaxy Tab S7+ (SM-T970), Android, stylus input	Controlled data collection in landscape orientation
Version control	Git and GitHub	Source history for the mobile application, backend and ML service

5.2 Component implementation

This section describes how the main functions of the component were implemented. Screenshots of the implemented screens are shown in Figures 5.1 to 5.4.

5.2.1 Trajectory capture

The handwriting canvas (for example in LetterWritingScreen.js) records touch input as a sequence of points with x and y coordinates and timestamps, and it keeps the

stroke boundaries separately. A new stroke begins whenever the stylus or finger is lifted and placed down again. In the controlled study, all sessions used the same Samsung Galaxy Tab S7+ in landscape orientation with a fixed 791×427 dp writing canvas and stylus input. Pressure, tilt, grip and stylus angle are not recorded. The raw trajectory of each attempt is sent to the backend and retained for research and audit purposes.

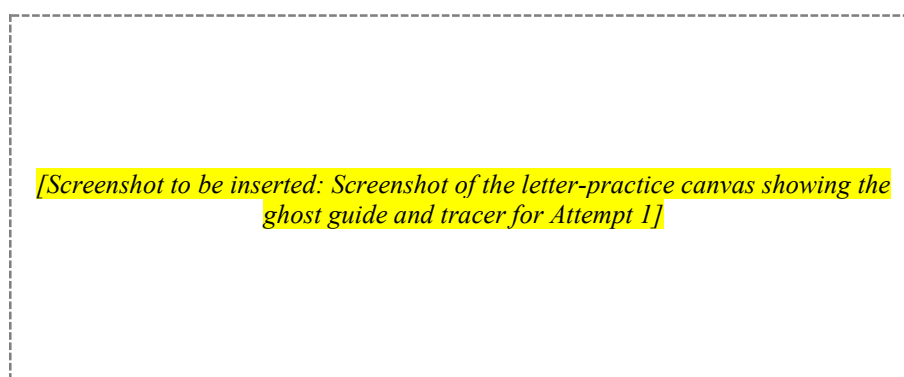


Figure 5.1: Letter-practice screen (Attempt 1, highest support)

5.2.2 Initial assessment

The initial assessment presents the six pre-writing tasks in a fixed order: horizontal line, vertical line, full circle, half circle, zigzag and curve-wave. For each task, the component calculates a shape score using the initial-shape formula in Section 3.4.2:

$$\text{Shape Score} = 0.80 (\text{DTW Score}) + 0.20 (\text{Smoothness Score})$$

The shape score belongs to the initial-assessment domain only and is not compared with the letter Motor Score. The results form an initial summary and a motor baseline, which are shown in the Teacher Report.

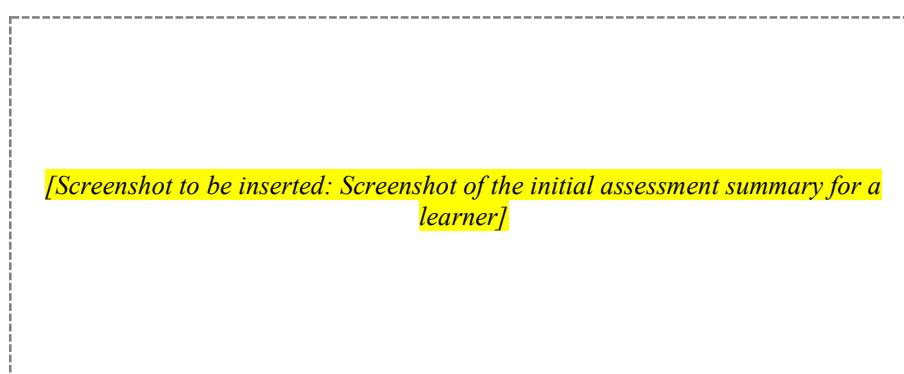


Figure 5.2: Initial assessment summary

5.2.3 Feature extraction and scoring

For letters, the client normalises each attempt into the common coordinate space of the LETTER_PATHS templates and compares it with the reference using DTW. Both the original and the reversed learner trajectory are compared, and the smaller distance is kept. For multi-stroke letters, strokes are matched before DTW is calculated. The client also calculates the smoothness score from the turning angles between successive trajectory segments and counts pauses, which are consecutive captured points separated by more than 300 ms.

On the backend, the handwriting controller (handwritingController.js) validates and normalises the submitted values and derives the remaining measures in Table 3.1, including speed CV and the direction score based on reversals of more than 90°. The scoring utilities then calculate the letter Motor Score:

$$M = 0.35A + 0.25S + 0.15Pc + 0.15R + 0.10V$$

The calculation is null-aware. If a component is not available for an attempt, its weight is removed and the remaining weights are renormalised. This backend score is the only score used for letter mastery. The initial-shape score, the word-writing score, descriptive feedback measures and research-analysis measures are separate. Some older scoring utilities remain in the codebase, but the mastery policy does not use them.

5.2.4 Practice sequencing and mastery

The practice rules are implemented in the backend sequencing module (adaptiveSequencing.js). For each letter and case, the module determines the next attempt within a cycle, applies the Attempt-3 mastery threshold of 70, counts cycles per day using Asia/Colombo dates, schedules exact-letter remediation after a failed Cycle 2, and moves the learner on with the letter unmastered after a failed Cycle 3, creating an Additional Home Practice candidate. A technically incomplete Attempt 3 is retried without using up the cycle. Lowercase and uppercase progress are stored separately. When mastery is reached, the identifier of the Attempt-3 record is stored with the mastery entry so that the evidence can be retrieved later.

The threshold utilities (thresholdUtils.js) and related learner-model fields developed for dynamic thresholds remain in the codebase. However, the active mastery check uses the fixed threshold of 70, and the dynamic-threshold logic is not connected to the operational mastery decision (Section 4.3).

5.2.5 Remediation and word learning

Remediation consists of up to two warm-up activities for the exact letter that failed Cycle 2, after which the learner starts Cycle 3. Word Activities A–E are implemented as separate screens that share the same submission and progress mechanisms. Word unlocking follows the group policy in Section 4.4.5, using the 154-word catalogue. Word writing reuses the trajectory capture and comparison pipeline at word level. A word-writing attempt passes when its score is at least 50 and the required letter-region coverage is met. Spacing and size are calculated and shown as advisory feedback.

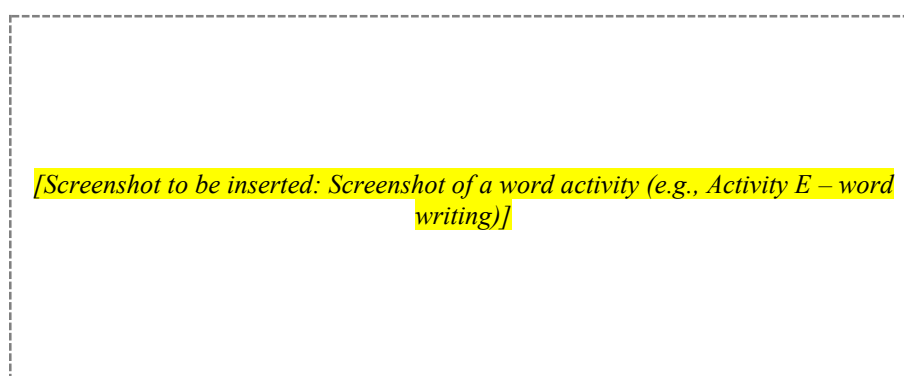


Figure 5.3: Word-writing activity

5.2.6 Teacher Report, Writing Check and worksheets

The Teacher Report screen retrieves one consolidated report from the backend containing the initial assessment, motor performance, letter mastery, word progress, Writing Check, motor-pattern information, worksheet information and recommendations.

Writing Check uses dedicated backend endpoints to start a check, record its progress and complete it. The teacher starts the check, the learner completes the handwriting input, and the result is returned to the teacher for review. Teacher ownership is checked on these endpoints.

For worksheets, the backend selects and stores the worksheet plan for the letter or letters behind an Additional Home Practice candidate. When the teacher chooses "Generate Worksheet", the mobile client renders the plan as an A4 PDF (generateWorksheetPdf, using Print.printToFileAsync), and the device share sheet is used to save, share or print it. After offline completion, the teacher uploads a photo of the worksheet through the backend, and it is stored in the learner's worksheet history for review. Completing a worksheet does not change mastery or scoring, and uploaded images are not analysed automatically.

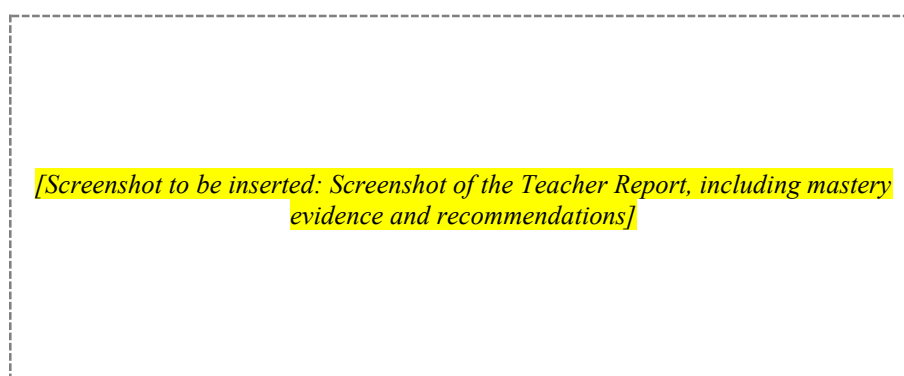


Figure 5.4: Teacher Report

5.2.7 ML service

The FastAPI service loads the saved scaler, K-Means model and model metadata when it starts. It accepts the three letter features from the backend, applies the saved scaler and centroids, checks the reference-range guard and returns a cluster summary. The backend attaches this summary to teacher-facing and research views only. Automated tests check service start-up, model loading and the stability of the saved cluster assignments.

5.2.8 Data-collection mode and data integrity

The data-collection mode reuses the normal screens with a fixed protocol and with adaptive gating disabled. Records are marked as collection data and linked to a collection-session identifier. In this mode, Attempt 3 keeps a faint guide, which differs from normal practice (Section 3.3.1).

While preparing for data collection, a schema-drift audit found that automatic schema alteration through Sequelize had silently dropped a column. Automatic alteration was replaced with controlled schema changes. Two further safeguards

were added: a raw-payload safety log (`collectionSafetyLog.js`) that keeps a copy of incoming collection payloads, and a start-up diagnostic query that confirms the expected columns exist before the backend serves requests. These safeguards protected the controlled data collected in the study.

5.3 Integration with the other components

The handwriting component is integrated with the other Auriva components, namely the Adaptive Concept Learning Module, the Adaptive Dialogue Learning and Progression Tracker and the Pronunciation Support Component (Table 4.3), through the shared mobile application and backend. Handwriting routes are part of the shared Express application and use the shared JWT authentication. The design document of the Dialogue component describes using learners' mastery profiles from the concept, speaking and writing modules when it generates Level 2 dialogue activities. That cross-component use belongs to the Dialogue component and was not evaluated in this thesis.

The integration with the ML service uses a small JSON contract: the backend sends the three letter features and receives a cluster identifier with a descriptive summary. Because no operational decision uses the ML output, letter practice and word learning do not depend on the ML service. The behaviour of the Teacher Report when the ML service is unavailable was not separately evaluated.

5.4 Verification and software testing

5.4.1 Testing strategy

Verification combines automated testing with manual checking on the target tablet:

- Automated unit tests for scoring, feature handling, touch mapping, mastery cycles and React Native screens.
- Automated backend route tests for request handling, validation, authorisation, ownership and error paths. These tests use mocked dependencies and do not run against a real database.
- Parity tests that check that equivalent scoring logic gives the same results in different parts of the system.

- ML service tests for start-up, model loading, input validation, output format and clustering stability.
- Regression tests added when defects were fixed, such as DTW direction sensitivity and the schema-drift issue.
- Manual end-to-end and device checks on the tablet across the app, backend, database and ML service. There is currently no automated real-device end-to-end test framework.
- Teacher usability evaluation (Section 3.3.2).

5.4.2 Automated test status

Table 5.2 reports the latest verified automated test run, completed on 30 August 2026. The tests cover scoring parity, touch mapping, assessment, navigation, mastery cycles, word activities, reporting, authorisation, ownership, finalisation and idempotency, ML start-up and clustering stability. Because the backend suite is mostly unit-level with mocked dependencies, these results show the extent of automated unit, regression, parity and error-path checking. They do not show database-backed integration or end-to-end behaviour, which were checked manually.

Table 5.2: Automated test results, run of 30 August 2026

Layer	Test suites	Tests	Result
Mobile client (React Native / Expo)	142	4,150	All passed
Backend (Node.js / Express)	139	3,209	All passed
ML service (FastAPI / Python)	Not applicable	250 tests and 33 subtests	All passed
Total ordinary tests	–	7,609	All passed

5.4.3 Functional test cases

Functional test cases were designed from the requirements in Section 4.1. Representative cases are listed in Table 5.3. The results of running these cases manually on the final build are still being confirmed, so they are not reported in this version.

Table 5.3: Representative functional test cases

ID	Requirement	Scenario	Expected result
----	-------------	----------	-----------------

ID	Requirement	Scenario	Expected result
TC01	FR1	Learner draws a two-stroke letter	Coordinates and timestamps stored with two stroke boundaries
TC02	FR2	Learner completes all six assessment tasks	Shape scores, initial summary and baseline appear in the Teacher Report
TC03	FR3	Correct letter shape drawn in the reverse direction	Direction-invariant DTW does not penalise the reversal
TC04	FR4	One Motor Score component is unavailable	Remaining weights are renormalised and a valid score is returned
TC05	FR6	Attempt-3 Motor Score of exactly 70	Letter marked as mastered; Attempt-3 record kept as evidence
TC06	FR6	Attempts 1 and 2 score 90 but Attempt 3 scores 60	Letter not mastered; next step follows the cycle rules
TC07	FR8	Cycle 2 fails	Up to two exact-letter warm-up activities, then Cycle 3
TC08	FR9	Cycle 3 fails	Learner advances with the letter unmastered; Additional Home Practice candidate created
TC09	FR7	Learner attempts a fourth cycle for the same letter and case on the same day	No fourth cycle is started
TC10	FR10	Attempt 3 is technically incomplete	Attempt 3 can be retried without using up the cycle
TC11	FR5	Learner masters lowercase "a" only	Uppercase "A" remains unmastered; progress shown separately
TC12	FR12	Word written with a score of 50 or more but missing a required letter region	Word writing not passed
TC13	FR13	All catalogue words in groups A, B and C completed	Group D unlocks
TC14	FR16	A teacher opens another teacher's Writing Check	Request refused by the ownership check
TC15	FR17	Teacher generates a worksheet and uploads a completed photo	A4 PDF produced; photo stored in history; mastery unchanged
TC16	FR18	Motor-pattern summary returned for a learner	Summary shown in the teacher view; mastery and progression unchanged
TC17	NFR4	The same attempt is finalised twice	Only one result is recorded
TC18	NFR5	Request without a valid JWT	Request rejected with an authentication error

5.5 Deployment and operation

In the research deployment used for this study, the mobile application ran on the Samsung Galaxy Tab S7+ and reached the Node.js backend through a local-network HTTP endpoint configured in the frontend. The backend accessed the PostgreSQL database through Sequelize and called the FastAPI service internally. When the backend starts, the diagnostic query checks the database schema before requests are served, and during data collection the safety log provides an extra recovery path.

This setup is suitable for supervised research use but not for production. A production deployment would require HTTPS/TLS, a configurable backend endpoint, managed and rotated secrets, stronger administrative credentials, restricted CORS, authenticated calls between the backend and the ML service, and a defined backup and restore procedure (Sections 3.6.3 and 8.4).

6 RESULTS AND VALIDATION

6.1 Research results

This section reports the research results for RQ1 and RQ3. The Dataset A results come from the controlled research analysis of the six-shape assessment. The Dataset B results come from the saved letter motor-pattern model and its metadata, and they are reported as implementation and model-validation evidence rather than as a separate research experiment.

6.1.1 Trajectory representation and pipeline verification (RQ1)

The controlled data showed that time-ordered coordinates, timestamps and stroke boundaries captured on a standard tablet are enough to derive the interpretable measures in Table 3.1 without pressure or tilt information. The research pipeline was checked against the deployed implementation. The participant-level feature representation rebuilt from the raw trajectory export showed a maximum numerical difference of 0.0 from the values produced by the deployed implementation, and all 14 controlled complex-shape DTW values were reproduced exactly. This indicates that the features used in the research analysis are the same as those calculated by the system.

Two development findings were important for the validity of these measures. First, DTW was sensitive to drawing direction, so a correctly shaped stroke drawn in reverse could receive a large distance. Comparing both the original and the reversed learner trajectory and keeping the smaller distance removed this problem (Section 3.5.1). Second, separate definitions of the visual guide and the scoring reference produced inconsistent scores. This was resolved by the single LETTER_PATHS definition (Section 4.4.2).

6.1.2 Scoring behaviour (RQ1)

Analysis of pilot attempts during development showed that some earlier scoring formulas compressed scores, so that visibly different attempts received similarly high values. Letter scoring was therefore consolidated around the backend Motor Score in Section 3.4.2, while the initial-shape assessment kept its own score. The current Motor Score coefficients are prototype values chosen by engineering priority, and a

later source audit found that they still show some compression. Further calibration with a larger dataset is therefore needed before the scores or the mastery threshold can be treated as validated (Section 7.5).

6.1.3 Initial-shape exploratory analysis (Dataset A, RQ3)

The Dataset A analysis pipeline is shown in Figure 6.1. The 42 controlled shape trajectories from seven participants were aggregated so that each participant formed one analytical unit ($P = 7$). Six features were used: straight-, curved- and complex-task trajectory distance, straight- and curved-task accuracy, and complex-task DTW distance. Complex-task DTW was transformed with $\log(1 + x)$, and all features were scaled with RobustScaler.

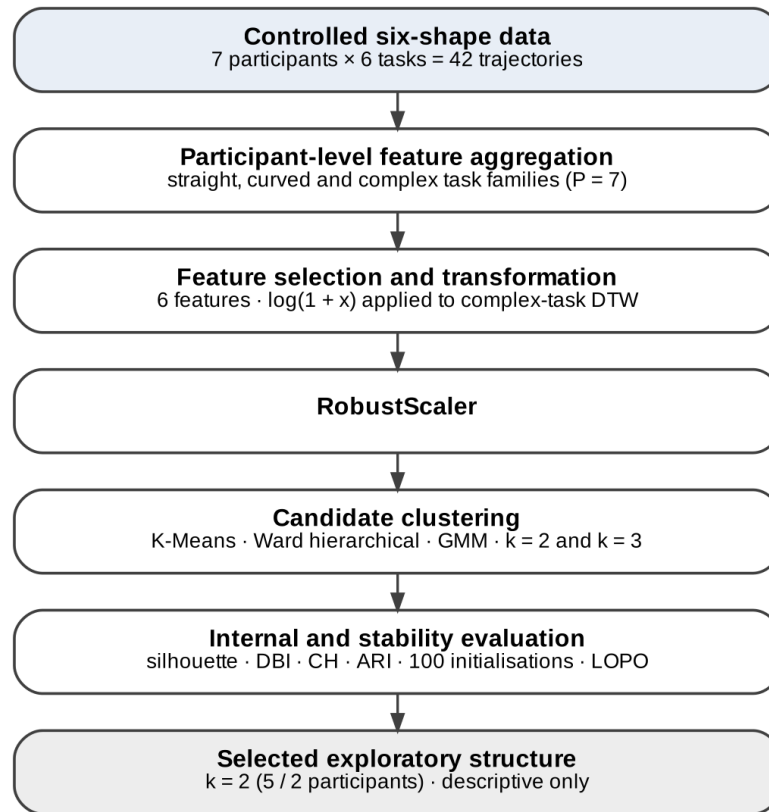


Figure 6.1: Dataset A exploratory analysis pipeline

The candidate clustering solutions are compared in Table 6.1. At $k = 2$, K-Means, Ward hierarchical clustering and GMM each produced the same 5/2 participant partition, with a silhouette coefficient of 0.412, a DBI of 0.739 and a CH score of 7.001. At $k = 3$, the solutions differed between algorithms and had lower silhouette

values. The two-cluster solution was therefore selected as the most consistent exploratory structure.

Table 6.1: Candidate clustering solutions for Dataset A

Method	k	Cluster sizes	Silhouette	DBI	CH
K-Means, Ward and GMM (identical partition)	2	5 / 2	0.412	0.739	7.001
K-Means and Ward	3	2 / 4 / 1	0.267	0.617	5.542
GMM	3	2 / 2 / 3	0.233	0.948	5.046

The stability results for the selected solution are summarised in Table 6.2. All 100 K-Means initialisations reproduced the same partition, and the pairwise ARI between the three algorithms was 1.0. The leave-one-participant-out (LOPO) analysis was more sensitive. The seven fold agreements were -0.174, 1.000, 0.348, 1.000, 1.000, 0.324 and 1.000, giving a mean ARI of 0.643 and a median of 1.000. Agreement between held-out participants and the reference partition was 6 out of 7 (85.71%), compared with 5 out of 7 (71.43%) for assigning every participant to the majority cluster. This value is agreement with a reference partition, not supervised classification accuracy.

Table 6.2: Stability of the selected two-cluster solution (Dataset A)

Check	Result
Pairwise ARI between K-Means, Ward and GMM	1.0
100 K-Means initialisations	All reproduced the same partition (ARI = 1.0)
LOPO ARI	Mean 0.643; median 1.000
Held-out reference-partition agreement	6/7 (85.71%), compared with 5/7 (71.43%) for majority assignment

The selected partition contained cluster C1 (RS01–RS04 and RS06) and cluster C2 (RS05 and RS07). The robust-scaled feature values are shown in Table 6.3. Within this sample, C2 had mainly greater straight-task trajectory distance, lower straight-task accuracy and greater complex-task DTW distance. These are descriptive differences within seven participants. The cluster labels are nominal and do not indicate ASD severity, ability or a motor subtype.

Table 6.3: Robust-scaled participant-level features for Dataset A

Participant (cluster)	Straight distance	Curved distance	Complex distance	Straight accuracy	Curved accuracy	Complex DTW
RS01 (C1)	-0.04	1.24	-0.08	-0.29	-0.67	0.12
RS02 (C1)	-0.30	-0.28	0.38	0.00	0.22	-0.20
RS03 (C1)	0.11	0.00	1.21	0.29	0.67	0.00
RS04 (C1)	-0.36	-0.55	-0.59	0.14	0.00	-0.46
RS05 (C2)	1.56	-0.13	0.00	-1.29	-0.44	1.22
RS06 (C1)	0.00	1.76	-0.33	0.57	1.33	-1.19
RS07 (C2)	1.55	0.35	1.97	-2.57	-0.67	2.33

Two further checks were carried out in the research analysis. First, robustness was tested by applying small identity-preserving perturbations (rotation of $\pm 2^\circ$, scaling of 0.97–1.03, temporal scaling of 0.95–1.05, coordinate jitter of $\sigma = 0.05$ dp, and combinations of these) to the real training trajectories within the LOPO design, while held-out participants were left unchanged. Mean held-out agreement ranged from 0.857 to 0.871 across the individual perturbation conditions, and the median stayed at 0.857 in every condition. Across 50 seeds of combined perturbation, mean agreement was 0.866 (SD 0.061) and the median was 0.857, with 94% of runs matching or exceeding the result obtained with real data only. The perturbations therefore preserved the partition but did not improve median held-out agreement.

Second, the features were compared with teacher ratings. Teachers rated each participant after the controlled session and before the computerised analysis, and these ratings were not used in clustering or model fitting. Spearman rank correlations were used only as an external descriptive comparison (Table 6.4). The strongest association was between curved-task accuracy and the teacher's curve rating ($\rho = 0.706$). Complex-task DTW showed the expected negative direction, because a larger DTW distance means a less similar trajectory. None of the associations was statistically significant before or after Holm correction. With seven participants, these values are exploratory effect estimates and not evidence of criterion validity.

Table 6.4: Exploratory associations between Dataset A features and teacher ratings

Comparison	Spearman ρ	Raw p	Holm-adjusted p
------------	-----------------	-------	-----------------

Comparison	Spearman ρ	Raw p	Holm-adjusted p
Straight accuracy vs. straight rating	0.401	0.373	0.701
Curved accuracy vs. curve rating	0.706	0.077	0.230
Complex DTW vs. complex rating	-0.418	0.350	0.701

6.1.4 Letter motor-pattern model (Dataset B)

The deployed practice pathway produced 420 letter-attempt records in the controlled collection (7 participants $\times$ 20 reference letters $\times$ 3 attempts). One eligible Attempt-3 record was kept for each participant–letter pair, giving 140 controlled letter-level observations. These observations provided the smoothness, DTW and speed-variability evidence for the letter motor-pattern model.

The K-Means model was not fitted to 140 independent points. The 140 observations were aggregated into seven participant-level vectors, and the scaler and two-cluster model were fitted to those vectors. The saved scaler reports `n_samples_seen_ = 7`, which confirms this. The validity and stability values recorded in the saved model metadata are shown in Table 6.5.

Table 6.5: Letter motor-pattern model validation from the saved model metadata (Dataset B)

Item	Value
Underlying letter evidence	140 Attempt-3 observations (7 participants $\times$ 20 letters)
Fitting unit	7 participant-level vectors (scaler <code>n_samples_seen_ = 7</code>)
Features	Smoothness score, DTW distance, speed CV
Scaling and algorithm	StandardScaler; K-Means, $k = 2$
Silhouette coefficient	0.2563
Davies–Bouldin index	1.1504
Caliński–Harabasz index	3.4936
LOPO ARI	Mean 0.8413; median 1.0; 6 of 7 runs identical to the reference partition

These values show a weak-to-moderate two-group structure. The silhouette coefficient is positive but modest, and the DBI is above 1. The partition was generally stable to leave-one-participant-out perturbation, with six of the seven runs

reproducing it exactly. Because only seven fitting units were available, the model is used only to give teachers descriptive motor-pattern information. The model was also fitted to collection-protocol Attempt-3 evidence, in which a faint guide was still shown, whereas Attempt 3 is unguided in normal practice. Its behaviour on normal-practice attempts has therefore not been validated (Section 3.5.3).

6.2 Product and technical results

The implementation status of the main functions is summarised in Table 6.6. The automated test results are reported in Table 5.2: all 7,609 ordinary tests passed in the run of 30 August 2026. Response times for attempt scoring were not measured in this study, so NFR3 has not yet been verified quantitatively.

Table 6.6: Implementation status of the main functions

Function	Status
Initial six-shape assessment with shape score and baseline	Implemented
Trajectory capture and motor-feature derivation (Table 3.1)	Implemented
Weighted, null-aware letter Motor Score	Implemented; prototype weights; further calibration needed
Dynamic or family-specific thresholds	Developed; disabled in the operational system
Lowercase and uppercase three-attempt cycles with Attempt-3 mastery	Implemented
Exact-letter remediation and Additional Home Practice candidates	Implemented
Word Activities A–E, word-writing scoring and group unlocking	Implemented
Teacher Report and exact mastery evidence	Implemented
Writing Check	Implemented
Worksheet generation (A4 PDF), photo upload and review	Implemented
Descriptive letter motor-pattern summaries through FastAPI	Implemented
Data-collection mode and data-integrity safeguards	Implemented
JWT authentication with role and ownership checks	Implemented
HTTPS/TLS, managed secrets, restricted CORS, ML-service authentication	Not implemented; production requirements
Automated real-device end-to-end tests	Not implemented

6.3 User or stakeholder validation

The teacher usability evaluation followed the task-based procedure in Section 3.3.2. The evaluation materials recorded task completion, help required, ease ratings, SUS responses, ratings of clarity and usefulness, and open feedback. The results are still being compiled and are not reported in this version. Children also used the handwriting component in practice sessions. Those sessions are not reported as usability evidence, because they were not part of the structured usability evaluation, and they are kept separate from the controlled research data.

6.4 Integration results

The integration between the mobile application, backend, database and ML service was checked mainly through manual end-to-end use on the study tablet. The controlled data collection itself depended on this integration: all 42 shape trajectories and 420 letter-attempt records were captured on the tablet, processed by the backend and stored in the shared database with their collection-session identifiers. The automated backend tests use mocked dependencies and therefore do not demonstrate database-backed integration. The results of the manual functional test cases in Table 5.3 are still being confirmed.

6.5 Summary of findings

Table 6.7 summarises the findings for each research question.

Table 6.7: Summary of findings by research question

Research question	Finding	Status
RQ1 – Interpretable measures from trajectory data	Trajectories captured on a standard tablet support the measures in Table 3.1; the research pipeline reproduced the deployed feature values exactly; DTW direction sensitivity was resolved; scoring was consolidated but still needs calibration	Addressed, with calibration as an open issue
RQ2 – Transparent rule-based workflow	Three-attempt cycles with Attempt-3 mastery, bounded cycles, exact-letter remediation, Additional Home Practice and word unlocking are implemented and covered by automated tests	Addressed at implementation level; manual test results pending

Research question	Finding	Status
RQ3 – Stable relative motor-pattern structure	Dataset A: the same 5/2 partition across three algorithms and 100 initialisations, with weaker LOPO stability (mean ARI 0.643). Dataset B: weak-to-moderate separation at participant level (silhouette 0.2563, LOPO mean ARI 0.8413)	Partly addressed; exploratory only
RQ4 – Usability and understandability for teachers	Evaluation carried out; results being compiled	Pending

7 DISCUSSION

7.1 Interpretation of findings

For RQ1, the results show that handwriting process data captured on a standard tablet can be turned into measures that are meaningful for practice. The exact reproduction of the deployed feature values by the research pipeline shows that the measures reported in this thesis are the same ones the system calculates. The development work also showed that the validity of such measures depends on details that are easy to overlook. If the reference shown to the child differs from the reference used for scoring, if the aspect ratio is distorted, or if DTW penalises a correct stroke drawn in reverse, the score no longer reflects how well the child wrote. In the same way, a score that places most attempts near the top of the scale cannot support a meaningful mastery threshold. The remaining compression in the current Motor Score coefficients is therefore an open issue rather than a solved one. The location of computation also matters. Because letter DTW is calculated on the client, the backend cannot verify it independently, which limits auditability.

For RQ2, the implemented workflow shows that motor evidence can drive structured practice through a small number of explicit rules. These are decreasing support over three attempts, one assessed attempt with a threshold of 70, a limit of three cycles per letter, case and day, exact-letter remediation after a failed second cycle, and an Additional Home Practice candidate after a failed third cycle. Each decision can be explained to a teacher in one sentence, which was a design goal. The rules are also covered by automated tests. However, the manual functional test results and the teacher evaluation are still needed to show that the workflow behaves as intended on the final build and is understood by teachers.

For RQ3, the two datasets give different but consistent pictures. In Dataset A, three algorithms and 100 initialisations produced the same two-cluster partition, which shows internal consistency in the participant-level representation. The leave-one-participant-out analysis was more sensitive (mean ARI 0.643), however, so the clusters should not be read as stable motor phenotypes. In Dataset B, the participant-level letter model showed only weak-to-moderate separation but fairly stable

assignments when participants were left out. These results agree with the variability in handwriting among children with ASD reported in the literature [12], [24]: learners do not fall into sharply separated groups, but some relative differences in accuracy, trajectory similarity and speed consistency can be seen within the sample. This supports the decision to use clustering descriptively and to keep it away from decisions about the child.

RQ4 cannot yet be answered, because the results of the teacher usability evaluation are still being compiled (Section 6.3).

7.2 Comparison with literature baselines and existing solutions

Diagnostic and screening studies such as those of Asselborn et al. [6], Mekyska et al. [7], Drotár and Dobeš [8] and Kunhoth et al. [29] train supervised models on labelled data and report classification accuracy. This component does not classify learners, so its results are not directly comparable with theirs, and it reports cluster validity and stability instead. Its use of process measures is closer to the computerised assessment work of Rosenblum et al. [15], [24] and Chang and Yu [21], but the measures are used inside a learning workflow rather than in a one-off assessment. The component also works without pressure or tilt, which makes it simpler to run on general-purpose tablets but loses information that richer sensing can provide [20].

The tablet-learning literature also helps interpret the design. Bonneton-Botté et al. found that the benefit of a tablet handwriting application depended on the child's initial handwriting level [33], which supports starting from an initial assessment rather than giving every learner the same practice. Mayer et al. noted that writing with a stylus on a touchscreen places different motor demands on the writer because of the lower surface friction [34], which is one reason why tablet scores should not be treated as equivalent to pencil-and-paper handwriting.

Compared with the products reviewed in Section 2.4 (Table 2.2), the autism-focused platforms provide broad adaptive learning activities and the handwriting applications provide guided tracing. None of the reviewed products was found to combine online trajectory capture, motor-feature analysis, a transparent mastery and remediation

workflow, and teacher-facing motor evidence. A direct empirical comparison with these products was outside the scope of this work.

7.3 Research contribution

The research contributions of this component are:

- an interpretable set of trajectory and motor features, derived from tablet trajectories without pressure or tilt, together with a weighted, null-aware letter Motor Score and an honest account of its calibration limits;
- a consistent DTW reference framework in which the visual guide, the tracer and the scoring reference share one letter definition, with stroke matching and direction-invariant comparison;
- an exploratory, stability-checked analysis of relative motor-pattern structure in controlled data from children with ASD, evaluated across three algorithms, repeated initialisations, participant removal, trajectory perturbation and teacher-referenced comparison; and
- a design principle, shown in a working system, that keeps exploratory machine learning separate from decisions that affect the child.

7.4 Product contribution

The product contribution is a working handwriting learning component integrated into the Auriva platform. It provides the initial six-shape assessment, lowercase and uppercase letter practice with Attempt-3 mastery, exact-letter remediation, word learning across a 154-word catalogue, a Teacher Report with exact mastery evidence, a teacher-initiated Writing Check, a worksheet workflow for additional home practice and a descriptive motor-pattern service. It is supported by data-collection tooling, data-integrity safeguards and 7,609 automated tests across the three layers.

7.5 Limitations and threats to validity

- **Sample size and setting.** The controlled data came from seven learners with ASD aged 10–15 years in one special-education setting. The clustering and correlation results may not generalise beyond this sample.

- **Participant-level models.** Both clustering analyses were fitted to only seven participant-level units, which limits the reliability of the validity and stability indices.
- **Construct validity of scores.** The Motor Score weights and the mastery threshold of 70 are engineering choices, not clinically validated values, and the current coefficients still show some compression. A score of 70 means that the system's criterion was met, not that the handwriting is age-appropriate.
- **Training and operational mismatch.** The letter model was fitted to collection-protocol Attempt-3 evidence with a faint guide, whereas Attempt 3 is unguided in normal practice, and the source audit found a substantial DTW distribution shift between the two conditions.
- **Input conditions.** The controlled data were collected with a stylus on one tablet model, and pressure, tilt and grip were not recorded. Finger input and other devices were not evaluated.
- **Location of computation.** Letter DTW and smoothness inputs are calculated on the client, so the backend cannot recompute them independently.
- **Security of the research deployment.** The tested build uses local-network HTTP, development secrets, a permissive CORS fallback and an unauthenticated ML service, so it is suitable only for supervised research use.
- **Test depth.** The automated backend tests are mainly unit-level with mocked dependencies, there is no automated real-device end-to-end testing, and the manual functional test results and response-time measurements are not yet reported.
- **Consent procedure.** Consent for the child participants was given by the school under its institutional care arrangements, and separate written consent from each parent or guardian was not obtained (Section 3.6.1).
- **Transfer to paper.** Improvement on the tablet may not transfer to handwriting on paper. The worksheet workflow partly addresses this, but worksheets are not scored automatically.
- **Evaluation stage.** The teacher usability results are still being compiled, and long-term learning outcomes were not measured.

- **Researcher involvement.** The developer of the system also carried out the data collection and analysis, which may introduce bias. The fixed protocol, the separation of teacher ratings from model fitting and the script-based verification reduce this risk but do not remove it.

7.6 Implications

For practice, the work suggests that teachers can be given process-level evidence about handwriting without being asked to interpret raw data or opaque model outputs, provided that the evidence is linked to understandable rules and to the learner's actual attempts. For researchers and developers, the project shows that small engineering inconsistencies, such as mismatched templates, distorted scaling and direction-sensitive DTW, can invalidate handwriting measures, so these checks should be part of any study design. For the responsible use of AI in special education, the work shows how machine learning can add descriptive value while decisions about the child stay transparent and under the teacher's control [9], [40].

8 COMMERCIALISATION AND SUSTAINABILITY

8.1 Value proposition and target users

The handwriting component is part of Auriva, an adaptive learning platform for English language development in autistic children in Sri Lanka. Within the platform, it offers structured handwriting practice for learners and process-level handwriting evidence for teachers, using general-purpose tablets rather than specialist equipment. Its value proposition can be summarised as follows: teachers can see how a child writes, not only what the child writes, and they can see why the next practice step was chosen.

The Auriva project identified two main customer personas. The first is the autistic learner, estimated to be between 5 and 15 years old, who may have limited verbal or non-verbal communication, may be learning English as a second language, may find abstract or unsupported instruction difficult and may have fine motor difficulties with handwriting. This learner needs structured, visual and repetitive content, clear feedback, consistent pacing and a non-verbal way of taking part when needed. The second is the special educator, estimated to be between 21 and 60 years old, who lacks data-driven tools for tracking individual progress, finds it difficult to identify specific learning gaps manually and needs ready-to-use, curriculum-aligned content. This teacher wants to see each child's scores and trends at a glance, identify which children need additional support and deliver structured lessons without technical expertise.

A distinction is made between the buyer, who approves and pays for adoption, and the users, who work with the system every day. Table 8.1 summarises this distinction for the handwriting component.

Table 8.1: Buyers, users and beneficiaries

Stakeholder	Role	Value received
Management of special-education schools and units, and owners of learning or therapy centres	Buyer and approver	Structured, documented handwriting support and evidence of learner progress for reporting

Stakeholder	Role	Value received
Special-education teachers	Primary user	Teacher Report, exact mastery evidence, reasons for recommendations, Writing Check and ready-made worksheets
Learners with ASD	Beneficiary and user	Structured practice with decreasing support and non-punitive feedback
Parents and guardians	Secondary beneficiary	Worksheets for additional home practice under the teacher's guidance

8.2 Market and competitive position

8.2.1 Target market in Sri Lanka

The initial target market is special-education provision in Sri Lanka. According to the Ministry of Education's Annual School Census, 7,574 students were enrolled in special-education units within government schools in 2024. In the same year, 27 government-assisted special schools enrolled 1,956 students and employed 468 teaching staff [55]. Not every student in these settings has ASD or needs handwriting support, so these figures describe the reachable institutional market rather than the number of users. They also show that the market is relatively small and concentrated, which suits institutional licensing better than individual consumer sales.

8.2.2 Competitive position

The products reviewed in Section 2.4 fall into two groups (Table 2.2): autism-focused learning platforms such as Otsimo, MITA and AutiSpark, which offer broad adaptive learning activities, and handwriting applications such as Writing Wizard and LetterSchool, which offer guided tracing and letter-formation practice. The handwriting component is positioned between these groups:

- Compared with autism-focused platforms, it provides handwriting-specific trajectory analysis, an explicit mastery rule and teacher-facing handwriting evidence.
- Compared with handwriting applications, it analyses the movement process rather than only completion of the trace, and it links difficulty to remediation and paper worksheets.
- Within Auriva, it can be adopted together with the concept, dialogue and pronunciation components as one platform rather than as a separate product.

8.3 Business or sustainability model

8.3.1 Revenue model

The Auriva project proposes an institutional revenue model with three entry levels:

- a freemium trial, allowing a school or centre to evaluate the platform before committing to it;
- Tier I, for up to 30 students, aimed at small special-education units; and
- Tier II, for 30 to 120 students, aimed at mid-scale schools and larger special-education units.

Linking tiers to the number of students reflects the fact that hosting, storage and support costs grow with the number of learners and handwriting attempts. Prices for the tiers have not yet been set or validated with customers, and no return-on-investment estimate is claimed. For government special-education units, which may have limited budgets, adoption could also be supported through education authorities or sponsored programmes.

8.3.2 Cost structure and resources

The main cost areas for deploying and operating the component are described in Table 8.2. Validated cost estimates are not yet available, so the costs are described qualitatively.

Table 8.2: Main cost areas for deployment and operation

Cost area	Description
Cloud hosting	Compute for the backend and the ML service, which grows with the number of active learners
Database and storage	Managed PostgreSQL, storage of raw trajectories and worksheet images, and backups
Devices	Tablets for schools that do not already have suitable devices
Maintenance	Bug fixes, dependency and security updates, and recalibration of scores and models
Teacher onboarding and support	Training sessions, user guides and help-desk support
Security and compliance	Production hardening (Section 5.5), data-protection documentation and agreements with institutions

8.3.3 Expected adoption barriers

The following barriers are expected in deployment. They are based on the design and context of the component and are not findings measured in the teacher evaluation:

- limited availability of suitable tablets and reliable connectivity in some schools;
- limited budgets in government special-education units;
- teacher time needed for training and for reviewing reports;
- language needs, because the current letter templates cover English letters only; and
- concerns from parents and schools about collecting children's data.

These barriers can be reduced through the freemium trial and student-based tiers, short teacher onboarding, a planned extension to Sinhala and Tamil templates, and a clear data-protection position (Section 8.4). The component can technically accept touch input as well as stylus input, which could make it usable on more devices. However, only stylus input was evaluated in the controlled study, so finger input should be treated as a possible accessibility option that still needs evaluation.

8.4 Scalability intellectual property and compliance

Scalability. The backend and ML service can be scaled out because they do not hold session state, and the database can be provided as a managed service. New letters or scripts can be added by extending LETTER_PATHS, which offers a route to Sinhala and Tamil, although each new script would need new templates and validation.

Maintenance and support. A production service would need a maintenance plan covering security updates, dependency upgrades, monitoring, backups, periodic recalibration of scores and retraining of the motor-pattern model as more data become available.

Intellectual property. The source code, the letter template definitions, the scoring and sequencing rules and the trained model artefacts are the main intellectual assets of the project. Ownership and commercialisation arrangements would need to follow SLIIT and project policy.

Compliance and data protection. Commercial use would require compliance with the Personal Data Protection Act [47], including a lawful basis for processing children's data, data-processing agreements with institutions, defined retention periods and access controls. The security gaps identified in the research deployment (Section 3.6.3) would need to be closed first. The product would also need to be described clearly as an educational support tool and not as a diagnostic or medical device.

8.5 Contribution to the Sustainable Development Goals

The component contributes most directly to SDG 4, Quality Education [56], and in particular to target 4.5, which calls for equal access to education for persons with disabilities, and target 4.a, which calls for disability-sensitive learning environments. By making individualised handwriting practice and evidence available on general-purpose tablets, the component could reduce the dependence on scarce specialist time. It also contributes indirectly to SDG 10, Reduced Inequalities, by supporting the inclusion of learners with additional needs. Possible negative effects include widening a digital divide where schools lack devices and relying too much on screen-based practice. The paper worksheet workflow and the use of general-purpose tablets are intended to reduce these risks.

9 CONCLUSION AND FUTURE DIRECTIONS

9.1 Conclusion

This thesis addressed the limited integration of handwriting-process analysis with transparent, structured learning support and teacher-facing evidence for children with ASD. It has shown that such an approach can be built on a general-purpose tablet with an interpretable design. The handwriting process is captured as online trajectories and analysed with direction-invariant DTW and motor features such as smoothness, pause control, direction control and speed variation. A weighted letter Motor Score drives a transparent three-attempt practice workflow with a fixed Attempt-3 mastery threshold, bounded daily cycles, exact-letter remediation, Additional Home Practice and rule-based word unlocking. Teachers see this evidence in a Teacher Report that includes exact mastery evidence, Writing Check results and worksheet support.

The research analysis showed that the deployed feature pipeline could be reproduced exactly and that the controlled six-shape data contained a two-cluster structure that was consistent across three algorithms and 100 initialisations but sensitive to removing individual participants. The letter motor-pattern model showed weak-to-moderate structure at participant level. Both results are exploratory and descriptive. Machine learning therefore adds information for teachers and researchers but is deliberately kept out of decisions about the child's progression. The usability of the component for teachers remains to be confirmed once the evaluation results have been compiled.

9.2 Achievement of objectives

The achievement of each objective is summarised in Table 9.1.

Table 9.1: Achievement of objectives

Objective	Evidence	Status
O1 – Handwriting trajectory acquisition	Trajectory capture implementation; 42 shape trajectories and 420 letter-attempt records collected (Sections 5.2.1, 5.2.8, 6.4)	Achieved

Objective	Evidence	Status
O2 – Handwriting and motor-feature analysis	DTW and motor features; exact reproduction of the deployed features (Sections 3.4, 5.2.3, 6.1.1–6.1.2)	Achieved; score calibration still needed
O3 – Structured adaptive handwriting practice	Letter cycles, remediation, word learning and unlocking, with automated tests (Sections 4.4, 5.2.4–5.2.5, 5.4)	Achieved at implementation level; manual test results pending
O4 – Exploratory motor-pattern analysis	Dataset A and Dataset B clustering and stability analysis (Sections 6.1.3–6.1.4)	Partly achieved (exploratory)
O5 – Teacher-facing monitoring and support	Teacher Report, mastery evidence, Writing Check and worksheets (Sections 4.4.6, 5.2.6); teacher evaluation (Section 6.3)	Implemented; teacher validation pending
O6 – Evaluation of the developed component	Automated testing, research analysis and usability evaluation (Sections 5.4, 6.1–6.4)	Partly achieved; usability and manual test results pending

9.3 Individual research and product contributions

The individual contributions of this work are:

- the motor-informed feature and scoring framework, together with an honest account of its calibration limits;
- the shared letter template and direction-invariant DTW reference system;
- the rule-based three-attempt practice workflow, with mastery evidence, remediation, Additional Home Practice and word learning;
- the Teacher Report, Writing Check and worksheet workflow;
- the exploratory Dataset A and Dataset B motor-pattern analyses, with the letter model served through FastAPI and kept separate from operational decisions; and
- the data-collection mode and data-integrity safeguards that made the controlled study possible.

9.4 Future directions

- **Larger and longitudinal studies.** Collect data from more learners, schools and sessions to test whether practice with the component improves handwriting over time and on paper, compared with a fixed practice sequence.

- **Calibrated scores and thresholds.** Use larger datasets and teacher or occupational-therapist ratings to calibrate the Motor Score weights and the mastery threshold, and only then re-evaluate dynamic thresholds.
- **Server-side verification.** Move or duplicate the letter-template and DTW pipeline on the backend so that submitted features can be checked independently.
- **Production hardening.** Deploy over HTTPS/TLS with managed secrets, restricted CORS and authenticated ML-service calls, and add database-backed integration tests and automated device end-to-end tests.
- **Model retraining.** Retrain the motor-pattern model on unguided normal-practice Attempt-3 data, and evaluate it on more participants.
- **Input options.** Evaluate finger input as an accessibility option, and add stylus pressure and tilt where the hardware allows it.
- **Worksheet analysis.** Investigate assisted analysis of uploaded worksheet images, with the teacher remaining the final decision-maker.
- **Local scripts.** Extend the template system to Sinhala and Tamil letters.

REFERENCES

- [1] A. Kushki, T. Chau, and E. Anagnostou, "Handwriting difficulties in children with autism spectrum disorders: A scoping review," *J. Autism Dev. Disord.*, vol. 41, no. 12, pp. 1706–1716, 2011, doi: 10.1007/s10803-011-1206-0.
- [2] L. Chmiliar, "Improving learning outcomes: The iPad and preschool children with disabilities," *Front. Psychol.*, vol. 8, Art. no. 660, 2017, doi: 10.3389/fpsyg.2017.00660.
- [3] C. T. Fuentes, S. H. Mostofsky, and A. J. Bastian, "Children with autism show specific handwriting impairments," *Neurology*, vol. 73, no. 19, pp. 1532–1537, 2009, doi: 10.1212/WNL.0b013e3181c0d48c.
- [4] B. P. Johnson et al., "Understanding macrographia in children with autism spectrum disorders," *Res. Dev. Disabil.*, vol. 34, no. 9, pp. 2917–2926, 2013, doi: 10.1016/j.ridd.2013.06.003.
- [5] M. M. Patchan and C. S. Puranik, "Using tablet computers to teach preschool children to write letters: Exploring the impact of extrinsic and intrinsic feedback," *Comput. Educ.*, vol. 102, pp. 128–137, 2016.
- [6] T. Asselborn et al., "Automated human-level diagnosis of dysgraphia using a consumer tablet," *npj Digit. Med.*, vol. 1, Art. no. 42, 2018, doi: 10.1038/s41746-018-0049-x.
- [7] J. Mekyska et al., "Identification and rating of developmental dysgraphia by handwriting analysis," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 2, pp. 235–248, 2017, doi: 10.1109/THMS.2016.2586605.
- [8] P. Drotár and M. Dobeš, "Dysgraphia detection through machine learning," *Sci. Rep.*, vol. 10, Art. no. 21541, 2020.
- [9] H. Khosravi et al., "Explainable artificial intelligence in education," *Comput. Educ.: Artif. Intell.*, vol. 3, Art. no. 100074, 2022, doi: 10.1016/j.caeai.2022.100074.
- [10] N. Grace, P. G. Enticott, B. P. Johnson, and N. J. Rinehart, "Do handwriting difficulties correlate with core symptomology, motor proficiency and attentional behaviours?," *J. Autism Dev. Disord.*, vol. 47, no. 4, pp. 1006–1017, 2017, doi: 10.1007/s10803-016-3019-7.
- [11] S. Rosenblum, H. Amit Ben-Simhon, S. Meyer, and E. Gal, "Predictors of handwriting performance among children with autism spectrum disorder," *Res. Autism Spectr. Disord.*, vol. 60, pp. 16–24, 2019, doi: 10.1016/j.rasd.2019.01.002.

- [12] H. C. Handle, M. Feldin, and A. Pilacinski, "Handwriting in autism spectrum disorder: A literature review," *NeuroSci*, vol. 3, no. 4, pp. 558–565, 2022, doi: 10.3390/neurosci3040040.
- [13] N. van den Bos and S. Rosenblum, "The underlying mechanisms of handwriting of individuals with autism spectrum disorder: A scoping review," *J. Occup. Ther. Sch. Early Interv.*, vol. 16, no. 4, pp. 556–576, 2023, doi: 10.1080/19411243.2022.2129900.
- [14] R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 63–84, 2000, doi: 10.1109/34.824821.
- [15] S. Rosenblum, S. Parush, and P. L. Weiss, "Computerized temporal handwriting characteristics of proficient and non-proficient handwriters," *Am. J. Occup. Ther.*, vol. 57, no. 2, pp. 129–138, 2003, doi: 10.5014/ajot.57.2.129.
- [16] J. Danna, V. Paz-Villagrán, and J.-L. Velay, "Signal-to-noise velocity peaks difference: A new method for evaluating the handwriting movement fluency in children with dysgraphia," *Res. Dev. Disabil.*, vol. 34, no. 12, pp. 4375–4384, 2013, doi: 10.1016/j.ridd.2013.09.012.
- [17] N. Hogan and D. Sternad, "Sensitivity of smoothness measures to movement duration, amplitude, and arrests," *J. Mot. Behav.*, vol. 41, no. 6, pp. 529–534, 2009, doi: 10.3200/35-09-004-RC.
- [18] S. Balasubramanian, A. Melendez-Calderon, and E. Burdet, "A robust and sensitive metric for quantifying movement smoothness," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 8, pp. 2126–2136, 2012, doi: 10.1109/TBME.2011.2179545.
- [19] Common Sense Media, "LetterSchool – Learn to write," app review. [Online]. Available: <https://www.commonsensemedia.org/app-reviews/letterschool> (accessed Oct. 7, 2026).
- [20] V. Babushkin, H. Alsuradi, M. O. Al-Khalil, and M. Eid, "Analyzing handwriting legibility through hand kinematics," *Front. Artif. Intell.*, vol. 8, Art. no. 1426455, 2025, doi: 10.3389/frai.2025.1426455.
- [21] S.-H. Chang and N.-Y. Yu, "Computerized handwriting evaluation and statistical reports for children in the age of primary school," *Sci. Rep.*, vol. 12, Art. no. 15675, 2022.
- [22] Y. Watanabe, T. Ohtoshi, T. Takiguchi, A. Ishikawa, and S. Takada, "Quantitative evaluation of handwriting skills during childhood," *Kobe J. Med. Sci.*, vol. 66, no. 2, pp. E49–E55, 2020.
- [23] M. Cao, T. Gu, X. Li, and J. Jing, "Socioeconomic status and motor coordination function among children with autism," *Front. Psychiatry*, vol. 16, Art. no. 1619918, 2025, doi: 10.3389/fpsy.2025.1619918.

- [24] S. Rosenblum, H. A. Ben-Simhon, and E. Gal, "Unique handwriting performance characteristics of children with high-functioning autism spectrum disorder," *Res. Autism Spectr. Disord.*, vol. 23, pp. 235–244, 2016, doi: 10.1016/j.rasd.2015.11.004.
- [25] B. P. Johnson, J. G. Phillips, N. Papadopoulos, J. Fielding, B. Tonge, and N. J. Rinehart, "Do children with autism and Asperger's disorder have difficulty controlling handwriting size? A kinematic evaluation," *Res. Autism Spectr. Disord.*, vol. 11, pp. 20–26, 2015.
- [26] T. Talkar et al., "Assessment of speech and fine motor coordination in children with autism spectrum disorder," *IEEE Access*, vol. 8, pp. 127535–127545, 2020, doi: 10.1109/ACCESS.2020.3007348.
- [27] H. L. Ch'ng and A. C. Ahmad, "The evaluation of 'Efficient Writing' on autistic student in intervention center," in *Proc. Int. Conf. Special Education*, vol. 4, 2021. [Online]. Available: <https://publication.seameosen.edu.my/index.php/icse/article/view/274>
- [28] H. L. Ch'ng and A. C. Ahmad, "How to improve autistic students' handwriting," in *Proc. Int. Conf. Special Education*, vol. 4, 2021. [Online]. Available: <https://publication.seameosen.edu.my/index.php/icse/article/view/260>
- [29] J. Kunhoth, S. Al-Maadeed, M. Saleh, and Y. Akbari, "Exploration and analysis of on-surface and in-air handwriting attributes to improve dysgraphia disorder diagnosis in children based on machine learning methods," *Biomed. Signal Process. Control*, vol. 83, Art. no. 104715, 2023.
- [30] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 26, no. 1, pp. 43–49, 1978, doi: 10.1109/TASSP.1978.1163055.
- [31] C. Di Brina, R. Niels, A. Overvelde, G. Levi, and W. Hulstijn, "Dynamic time warping: A new method in the study of poor handwriting," *Hum. Mov. Sci.*, vol. 27, no. 2, pp. 242–255, 2008.
- [32] J. Tang, H. Cheng, Y. Zhao, and H. Guo, "Structured dynamic time warping for continuous hand trajectory gesture recognition," *Pattern Recognit.*, vol. 80, pp. 21–31, 2018, doi: 10.1016/j.patcog.2018.02.011.
- [33] N. Bonneton-Botté et al., "Can tablet apps support the learning of handwriting? An investigation of learning outcomes in kindergarten classroom," *Comput. Educ.*, vol. 151, Art. no. 103831, 2020.
- [34] C. Mayer, S. Wallner, N. Budde-Spengler, S. Braunert, P. A. Arndt, and M. Kiefer, "Literacy training of kindergarten children with pencil, keyboard or tablet stylus: The influence of the writing tool on reading and writing

performance at the letter and word level," *Front. Psychol.*, vol. 10, Art. no. 3054, 2020, doi: 10.3389/fpsyg.2019.03054.

- [35] J. Ecalles, A. Boisson, H. Labat, R. Versace, and A. Magnan, "Spatial sonification of letters on tablets to stimulate literacy skills and handwriting in 5 y-o children: A pilot study," *Hum. Mov. Sci.*, vol. 79, Art. no. 102844, 2021, doi: 10.1016/j.humov.2021.102844.
- [36] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci.*, vol. 622, pp. 178–210, 2023.
- [37] A. E. Ezugwu et al., "A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects," *Eng. Appl. Artif. Intell.*, vol. 110, Art. no. 104743, 2022.
- [38] M. Gagolewski, M. Bartoszek, and A. Cena, "Are cluster validity measures (in)valid?," *Inf. Sci.*, vol. 581, pp. 620–636, 2021.
- [39] U. von Luxburg, "Clustering stability: An overview," *Found. Trends Mach. Learn.*, vol. 2, no. 3, pp. 235–274, 2010.
- [40] T. Ley et al., "Towards a partnership of teachers and intelligent learning technology: A systematic literature review of model-based learning analytics," *J. Comput. Assist. Learn.*, vol. 39, no. 5, 2023, doi: 10.1111/jcal.12844.
- [41] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds. London, U.K.: Taylor & Francis, 1996, pp. 189–194.
- [42] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, 1987, doi: 10.1016/0377-0427(87)90125-7.
- [43] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, 1979, doi: 10.1109/TPAMI.1979.4766909.
- [44] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Stat.*, vol. 3, no. 1, pp. 1–27, 1974, doi: 10.1080/03610927408827101.
- [45] L. Hubert and P. Arabie, "Comparing partitions," *J. Classif.*, vol. 2, no. 1, pp. 193–218, 1985, doi: 10.1007/BF01908075.
- [46] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

- [47] Parliament of the Democratic Socialist Republic of Sri Lanka, Personal Data Protection Act, No. 9 of 2022, 2022.
- [48] M. Jones, J. Bradley, and N. Sakimura, "JSON Web Token (JWT)," IETF RFC 7519, May 2015.
- [49] OWASP Foundation, "OWASP Top 10:2021," 2021. [Online]. Available: <https://owasp.org/Top10/> (accessed Oct. 7, 2026).
- [50] World Wide Web Consortium (W3C), "Web Content Accessibility Guidelines (WCAG) 2.1," W3C Recommendation, Jun. 2018. [Online]. Available: <https://www.w3.org/TR/WCAG21/>
- [51] A. W. Salmoni, R. A. Schmidt, and C. B. Walter, "Knowledge of results and motor learning: A review and critical reappraisal," *Psychol. Bull.*, vol. 95, no. 3, pp. 355–386, 1984, doi: 10.1037/0033-2909.95.3.355.
- [52] D. Wood, J. S. Bruner, and G. Ross, "The role of tutoring in problem solving," *J. Child Psychol. Psychiatry*, vol. 17, no. 2, pp. 89–100, 1976, doi: 10.1111/j.1469-7610.1976.tb00381.x.
- [53] Systems and Software Engineering—Systems and Software Quality Requirements and Evaluation (SQuaRE)—System and Software Quality Models, ISO/IEC 25010:2011, 2011.
- [54] Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts, ISO 9241-11:2018, 2018.
- [55] Statistics Branch, Ministry of Education, Sri Lanka, "Annual School Census of Sri Lanka – Summary Report 2024," Jan. 2025. [Online]. Available: https://www.statistics.gov.lk/Resource/en/Education/School_Census/Annual_School_Census_SummaryReport_2024.pdf
- [56] United Nations General Assembly, "Transforming our world: The 2030 Agenda for Sustainable Development," A/RES/70/1, Oct. 2015.

GLOSSARY

Additional Home Practice candidate – A record created when a learner fails three same-day cycles for a letter, which the teacher can act on by generating a worksheet.

Attempt 3 – The least-supported attempt in a practice cycle and the only attempt used to decide letter mastery.

Collection mode – An operating mode with a fixed protocol and disabled gating, used to collect controlled research data.

Ghost guide – A faint outline of the target letter shown behind the writing area.

Mastery evidence – The stored Attempt-3 record that established mastery of a letter.

Motor Score – The weighted 0–100 letter score combining accuracy, smoothness, pause control, direction control and speed control.

Online handwriting – Handwriting captured as a time-ordered sequence of positions while it is being written.

Reference-range guard – A check that stops the motor-pattern model from assigning a cluster to evidence outside the range represented by its training data.

Remediation – Up to two exact-letter warm-up activities given after a failed second cycle.

Tracer – An animation that demonstrates the stroke order and direction of a letter.

Writing Check – A teacher-initiated handwriting-pattern check whose result is returned to the teacher for review.

APPENDICES

Appendix A Research instruments

A.1 Controlled data-collection protocol checklist

Table A.1: Controlled data-collection protocol checklist

Step	Action	Responsible
1	Confirm that SLIIT ethical clearance and the participating school's approval of the data-collection request are in place; participation of each child is arranged by the school under its institutional care arrangements	Researcher and school
2	Select the learner's pseudonymous profile and start a collection session (collection-session identifier assigned)	Researcher
3	Confirm that the backend start-up schema check passed and that the collection safety log is active	Researcher
4	Learner completes the six pre-writing tasks in order: horizontal line, vertical line, full circle, half circle, zigzag and curve-wave	Learner, supervised
5	Learner completes the 20 reference letters with three attempts each (a faint guide is retained on Attempt 3 under this protocol)	Learner, supervised
6	Stop if the learner declines or disengages; do not encourage the learner to continue	Teacher and researcher
7	After the session, check stored records with the verification scripts before analysis	Researcher

A.2 Teacher usability evaluation tasks

Table A.2: Teacher usability evaluation tasks

Task	Description
T1	Review a learner's initial assessment summary
T2	Interpret the letter mastery evidence for a learner
T3	Examine the Writing Check information
T4	Generate a worksheet from a recommendation
T5	Upload a completed worksheet

A.3 Evaluation measures

For each task, the evaluation recorded whether the task was completed, whether help was required and how easy the teacher found it. After the tasks, teachers completed the System Usability Scale [41] and questions on the clarity and usefulness of the progress information and recommendations, and they gave open feedback.

Appendix B Academic integrity and AI use disclosure

Table B.1: Use of generative AI and automated tools

Tool	Purpose	Sections or artefacts affected	Verification performed
Claude (Anthropic)	Structuring the thesis against the template and marking rubric; drafting text from the author's project notes, implementation details and results; drawing diagrams; literature search support	Initial draft of Chapters 2–9; revision of Chapters 4–9; front matter; citation checking; Figures 2.1, 3.1, 4.1–4.5 and 6.1	[To be completed by the author: technical statements checked against the implementation; references checked against original sources]
ChatGPT (OpenAI)	Thesis structuring; academic wording; interpretation of source and security audit findings; survey design; technical review of the draft	[list the sections and artefacts affected]	[To be completed by the author]