



A study of children emotion and their performance while handwriting Arabic characters using a haptic device

Jezia Zakraoui¹ · Moutaz Saleh¹  · Somaya Al-Maadeed¹ · Jihad Mohamad AlJa'am¹

Received: 18 April 2022 / Accepted: 19 July 2022 / Published online: 5 August 2022
© The Author(s) 2022

Abstract

Emotion and focus of children during handwriting are essential for language learning. Handwriting for young children is challenging because it needs deep motivation and willingness to complete the task regardless of its difficulty. Recently, emerged haptic guidance systems have a good potential to offer children a better sense of engagement to keep their interests awake. Yet, handwriting in 3D is more challenging for children due to many reasons including demotivation, out of focus, and visuomotor coordination difficulties. In this paper, we study the effectiveness of a haptic device in analyzing schoolchildren emotion, attentiveness and handwriting performance of Arabic letters. We conducted the experiments for a period of four weeks with an immersive environment where the subjects practiced writing in VR environment using a haptic device-controlled stylus. We assessed the childrens' emotions to get insights into their engagement during such hard learning environment. We found that our approach improved the participants' fine-motor skills and handwriting quality. However, our analysis revealed that such task was effective on detecting emotions (angry/neutral) only as a negative/positive contributor of performance. Overall, the obtained findings can well contribute to the understanding of the relationship between students' emotions and other variables in an attempt to support the development of adaptive learning technologies.

Keywords Emotion · Haptic · Handwriting · Learning skills · Arabic vocabulary

1 Introduction

The study of children emotions when learning a new task is highly important as they influence their engagements and performances during the learning process. Recent studies (Anagnostopoulou et al., 2021; Baltrusaitis et al., 2016; Karumbaiah et al.,

✉ Moutaz Saleh
moutaz.saleh@qu.edu.qa

¹ Department of Computer Science and Engineering, Qatar University, 2713 Doha, Qatar

2022) attempt to improve the quality of learning techniques by continuously detecting the development and change in students' emotional experiences in order to better support individual students and maximize their engagement and learning gains (Anagnostopoulou et al., 2021). However, the task of learning to write vocabulary in any language is a challenging task especially for young children and those with low handwriting skills and moderate memory, which leads to repeated failure and may result in negative emotions. Some studies (Baltrusaitis et al., 2016; Bowers & Hayle, 2020) used haptic devices to help children learn the language independently and efficiently. While others, for example (Grundmann et al., 2021) monitored the emotional status of each student and provided information about student's mood to the teacher in real time when e-learning. Some studies focused mainly on features such as pen pressures as indicators of enjoyment and frustration (Anagnostopoulou et al., 2021); gaze direction (Khan, 2020), facial expression (Young-Seok Kim et al., 2013), pose estimation (Guneysu Ozgur et al., 2020) and many other techniques (Ashwin & Guddeti, 2020; Dewan et al., 2019). However, few research works have been carried out on how children with low handwriting skills react to different virtual reality environment (VR) writing tasks (Baltrusaitis et al., 2016; Savov et al., 2018; Zhao et al., 2018).

To the best of our knowledge, this is the first study exploring child emotion and attention while handwriting Arabic in VR. However, learning to write vocabulary in Arabic is a challenging task especially for young children and those with low handwriting skills and moderate memory. Some Latin native children learning Arabic as a second language mimic Latin letters when they write in Arabic. They start writing from left to right while in Arabic words should be written from right to left. Moreover, Arabic letters have substantially different shapes depending on whether it will be connecting with a preceding and or a succeeding letter, thus all independent letters have conditional forms. These forms occasionally give confusion to new learners especially for young children. Essentially, the technology of haptic devices can help children learn the language independently and efficiently. In fact, they can start practicing handwriting letters with higher motivation. In addition, they can learn at home and repeat the task several times. Psychologists hypothesized that children with writing difficulties will be cognitively empowered while interacting in a VR during the learning process. VR offers positive reinforcement and a better sense of engagement to keep children's interests awake.

In view of the belief that children's emotions and attentiveness may have impact on the handwriting performance, in this research, we aim to explore the link between participants' emotions and their attentiveness and performance, if any, when learning to write Arabic letters. Moreover, there is reason to believe that emotions and level of attention may have a positive impact on performance and might further mediate the link between emotions and learning performance. This assumption was demonstrated in several research within the field of learning where it reveals that learning activities are always accompanied by emotions (Pekrun, 2006; Pekrun et al., 2007; Schrader & Kalyuga, 2020; Schrader & Nett, 2018; Schutz & Pekrun, 2007).

Even though these associations are not guided by any explicit predictions due to the inconsistent findings of existing research (Karumbaiah et al., 2022; Schrader &

Kalyuga, 2020), our study contributes to an improved understanding of the impact of students' emotion and attentiveness on their handwriting performance. However, the relationship between emotion, attentiveness and learning outcome appears to be more complex and not as straightforward as is generally assumed (Karumbaiah et al., 2022).

First, student emotional experience and its contribution to the learning process is not well-investigated (Schrader & Kalyuga, 2020). Only few (Schrader & Nett, 2018; Schutz & Pekrun, 2007) research have shown the importance of emotions in classroom, as they can both positively and negatively influence students' engagement (Schrader & Kalyuga, 2020) and their well-being (Hill et al., 2019) during learning, as a result, their performance. However, a recent review (Karumbaiah et al., 2022) found out that the affective states of students (confusion, frustration, and boredom) in VR learning have mixed relationships with their outcomes. Some of these researchers in the latter review pointed out that there is no significant difference in students' academic achievement and retention in various affective states for instance in social study learning (Karumbaiah et al., 2022). While some studies suggest that positive emotions promote self-regulation and motivation may not improve learning, some negative emotions may promote learning by triggering motivation to learn better (Worsley & Blikstein, 2015).

Second, student attentiveness has been proven to improve the learning performance (Fredricks & McColskey, 2012), (Au et al., 2016). However, a lack of student attentiveness and active engagement with learning activities has become a significant challenge that negatively influences learning outcome (Creelman, 2021; Derakhshandeh et al., 2021). Many studies have been performed to determine the attentiveness of students in a learning setting (Adam et al., 2009; Kainat et al., 2022; Lipp & Neumann, 2004), however all these studies focus more on determining the attention states rather than studying the effective relationship of attentiveness on learning performance. Only few (Creelman, 2021; Fredricks & McColskey, 2012) studies such relationship. For instance, (Creelman, 2021) found out that inattentive student accounts for failing to finish work, failing to show close attention to details.

To the best of our knowledge, there is no research on the concrete impact of participants' changes in emotional and attentional trends, at least not concerning Arabic handwriting performance. Our study contributes to an improved understanding of the impact of students' emotion and attentiveness on their handwriting performance in a VR environment. Our findings add value to a range of existing research works and confirm that emotional changes have no direct relation with writing performance, but only, affected by overall emotion. Besides, performance was not affected by attentional change, but only, by eye blinks at the beginning of the experiment only. The results are important because they have many implications and may help researchers to improve the conceptualization and the measurement of student emotion and attention changes, effectively. More precisely, our study focuses on the impact of a haptic device-based learning technology on students' emotions and attentiveness on their writing performance of Arabic letters. Furthermore, the continuous detection of emotion states through emotion detection models (Anagnostopoulou et al., 2021; K V & Bahel, 2021; Mukhopadhyay et al., 2020) and through eye blink detection model (Soukupová & Cech, 2016) as objective metrics, try to

cope with some of the methodological problems such as self-reporting (Schrader & Kalyuga, 2020).

We use a haptic device with a stylus-like tool that allows the children to physically touch the pen, feel the force and movement and write in a virtual space within a harsh environment to assess their motivation and focus while also fine-tuning their motor skills. Since, handwriting is generally difficult for young children even when writing in a 2D space, writing in a 3D or VR environment is more challenging for children due to many reasons including arm fatigue and visuomotor coordination issues like problems in line following, poor fine motor movement, etc. All that may cause out of focus, boredom and frustration. However, these factors when resolved can contribute to mastering the proper writing and strengthening the understanding of the language vocabulary. To this point, we aim to explore real-time data (e.g., temporal, and spatial), which provides objective measurements to assess the effect of a new haptic device on the learning skills such as enhancing memorization of vocabulary, endurance, focus, and engagement. We investigate whether using a haptic device to practice handwriting can be an effective technique to motivate the children for handwriting. Further, we examine if it resolves the difficulties in the writing process. Based on guided writing mode that follows the letter trajectories, students are able to evaluate and correct their writing errors and, as a result, improve their writing skills.

In summary there are two main research questions for our study:

1. First, to what extent participant's emotions (either at begin or at any stage) contribute to better handwriting performance.
2. Second, is there any effect of attention level (either at begin or at any stage) on handwriting performance.

The remainder of this paper is organized as follows. Section 2 presents our proposed approach and experiments settings to study the effectiveness of using a haptic device in analyzing schoolchildren emotions, attentiveness and handwriting performance of Arabic letters. Section 3 discusses the results and evaluations. Section 4 presents the key observations and implications, while Sect. 5 concludes the paper, highlights the limitations and futures future works.

2 The approach

2.1 Participants

We selected fifty-two children from primary schools in Doha, Qatar. These children were able to hold a pen and manipulate it appropriately. They are from different ethnicity and of an age between 5 and 11 years ($n=52$, $M=7.2$, $SD=1.2$) as shown in Table 1. We explained to them the objective of this study, discussed the details of the experiments, answered their questions, and gave them a training session to write anything they like with the haptic stylus in the VR environment and spend

Table 1 Dataset details

School	Girls	Boys	Native	Non-Native	Age Group
DIS School	5	10	15	0	5–11
Al Bayan School	25	0	25	0	5–9
Al Riffa School	12	0	12	0	5–11
Total	42	10	52	0	5–11

time using it to the point where the children felt relaxed, mastering the pen manipulation before starting the experiments.

2.2 Equipment

We use the Touch™ haptic device from 3D-Systems company (<https://www.3dsystems.com/haptics-devices/touch>). The haptic device is connected to a laptop with control algorithms allowing guided writing (i.e., with partial and full guidance) and free writing. We recorded all the experiments for further analysis. Initially, we obtained the Institutional Review Board (IRB) approval to test the haptic device with children in schools in Qatar. After obtaining informed consent from the parent, we gave instructions to the children about the task of handwriting the three letters in 3D using both modes. We observed and assisted when required. To evaluate the system, we asked the participant to write the letters before training and after training in free hand using a usual pen to analyze the change.

2.3 Scenarios

We design two scenarios for the experiments with tracing three letters in 3D using the haptic device as follows:

1. Write three independent letters in guided mode.
 - Show an image of an animal object.
 - Show an animal name with one missing letter.
 - Child writes the missing letter using full/partial mode.
2. Write three independent letters in free mode.
 - Show an image of an animal object.
 - Show an animal name with one missing letter.
 - Child writes the missing letter using record mode.

2.4 Haptic task procedure

The haptic task procedure is given in Algorithm 1. Each participant completed three sessions: a training (~8 Min), a pre-test session (~8 Min), and a post-test (~8 Min). Each participant spent 40 min on average for this study. The pre- and post-tests

consisted of three Letter tasks to observe whether there were any differences in performance.

Algorithm 1 The haptic task procedure.

Data: Pre-test, character traces
Result: 3D traces, questions
Initialization
while *not at end of the experiment* **do**
 exercise in guided writing mode
 take a break
 exercise in free writing mode
 if *score is above average* **then**
 go to Post-test task
 else
 take a break
 exercise in free writing mode
 end
Post-test
take questionnaire
end

In the task session, we ask the participant to repeat the letter tracing when we observe that the letter shape diverts largely from the true trajectory, which means task score, is very low. For participants with severe writing difficulty, we repeat the sessions, and a break of five minutes is added between two consecutive sessions. If a participant showed no knowledge than he/she repeats the tasks more times and we provide path cues to augment visual guidance for the participant. Thus, the participant got more opportunities and hints to practice and improve relevant fine motor skills in 3D. Note that we played short stories with animal characters to augment visual modality for all participants. At the end of the experiment, we ask the participant verbally a short survey and explained each question to the participants and encouraged them to express their feelings as such feedback will help make the system better.

2.5 Datasets

We collected datasets as depicted in Table 2. It comprises 520 handwritten Arabic characters (in both modes) and 52 video recordings for 3 different views, namely front, side and webcam view. For the task of character recognition, we retained only

Table 2 Details of collected handwriting dataset

	No. of children	Samples	Right handed	Left handed
Boys	10	100	9	1
Girls	42	420	39	3
Total	52	520	48	4

156 handwritten Arabic characters for the analysis since many handwriting attempts were excluded due to being either too noisy or incomplete.

Thereafter, we cleaned the noise around the characters. Most of the noise occur around the start and the end of the character trace. Some noise resulted because of hand movement when the child was trying to grip the pen and finding the start position in 3D. Figure 1 shows cleaned character samples from our dataset.

2.6 Metrics

To assess emotion, focus and handwriting performance of the participants, we use both objective and subjective metrics described in Table 3. For instance, we identify facial landmarks: eyebrows, eyes, nose and mouth as the critical features for objective emotions prediction. Whereas subjectively, we use the observations from the teachers on the participants' behavior (Schrader & Kalyuga, 2020). We can explore the results with the children questionnaire too and derive to some extent their motivation and engagement. Note the questionnaire reliability was 0.56 using Cronbach alpha (Wessa, 2021) at the time of development. The questions, given in Table 5, were verbally asked, and the answers were noted.

3 Results and analysis

In following section, we report our results and analyze them using Minitab software. The means and standard deviations of the measured variables are shown in Table 4 with p -value=0.002 for Group A and p -value=0.000 for Group B, so we rejected the null hypothesis that all means are equals. Note that Eye blink:B, Eye blink:M and Eye blink:E are the values of eye blinks extracted from the video sequences at begin, midway and the end of the recorded handwriting task.

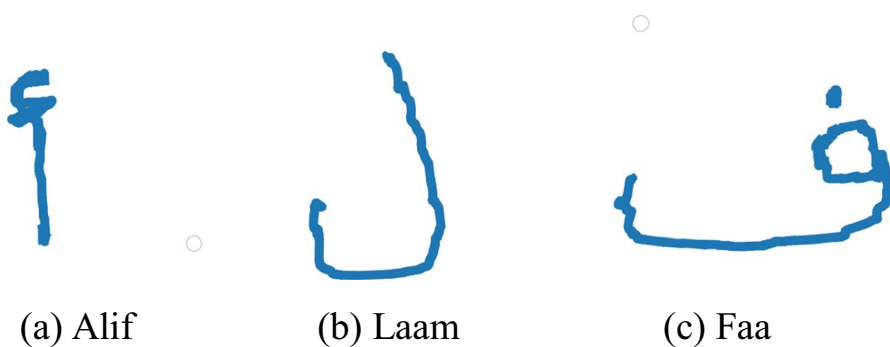


Fig. 1 Samples of cleaned Arabic characters. (a) Alif, (b) Laam, (c) Faa

Table 3 Objective and subjective metrics

Metric	Type	Description
Eye blink detection	Objective	Number of eye blinks detected by identifying the EAR (eye aspect ratio) where a low EAR is eye closed and high EAR is eye opened. An eye kept open denotes a focused state whereas multiple eye blinks denote a disturbed state or non-attentive state. Attentiveness is thus inferred with number of eye-blinks identified from videos of the participants (Soukupová & Cech, 2016)
Emotion detection	Objective	Facial landmarks and facial expressions and their emotions. Several states of emotion were determined using a deep learning model which was trained to identify emotions based on the emotions of the people (Goodfellow et al., 2015)
Dynamic Time Warping (DTW) (Salvador & Chan, 2007)	Objective	Distance between the actual written letters and the reference letter, which is taken as a factor contributing to performance. This distance is used as an error score for writing performances. For a given letter, a lower error score indicates a closer grapheme to the reference letter
Task performance	Objective	DTW error score divided by the total effective time the participant used to complete a task
Character recognition	Objective	PHOG (Bosch et al., 2007) features extraction and three classifiers
Teacher observation	Subjective	How many attempts they did? Did they stop and how many times? Did they try to finish their writing following the template as precise and as fast as possible to engage with the activity?
Questionnaire	Subjective	Verbal questions using Very much (5), A little, Neutral, Not much, Not at all (1) Likert scale

Table 4 Mean and standard deviations for eye blinks, attentiveness and writing performance

Metric	Group A M(SD)	Group B M(SD)
Eye blink:B	27.41 (23.31)	42.10 (23.90)
Eye blink:M	44.00 (7.91)	42.90 (9.83)
Eye blink:E	12.72 (5.70)	12.95 (9.54)
Total blinks	125.1 (350.0)	181.5 (331.2)
Attentiveness	97.384 (4.234)	96.28 (5.91)
Performance	0.1254 (0.0711)	0.1370 (0.0579)

Table 5 Participants' feedback from the questionnaire

Questions	Group A Avg1 (SD)	Group B Avg2 (SD)
1. Do you like writing Arabic using this haptic device?	4.73 (1.01)	4.80 (0.89)
2. Do you enjoy the activity?	3.63 (1.45)	4.00 (1.37)
3. Do you like to repeat the activity?	2.73 (2.01)	3.00 (2.05)
4. Do you feel any pain or tiredness?	2.70 (1.70)	3.30 (1.62)

3.1 A. Participants' feedback

There were four questions included in the questionnaire asking the participants' feedback on their experience with handwriting task using the haptic device. As shown in Table 5, the survey results are based on two age groups A (5–7) and B (8–11). The four questions were scored on a five-point scale (where Very much (5), A little, Neutral, Not much, Not at all (1). In the following discussion, Avg1 and Avg2 stand for the average score obtained by the two groups, respectively. First, participants in both groups liked the handwriting task (Avg1 = 4.73, Avg2 = 4.80) and they enjoyed the activity (Avg1 = 3.63, Avg2 = 4.00). It indicated the positive feedback of the participants for the tasks. Second, group B could repeat the activity. However, we found that due to the gripper size, it took longer for the younger participants to perform grip adjustments; some did not manipulate the Haptic Gripper very well in the shorter time (Note that they performed guided mode only) especially when they tried to apply large finger pressure. In addition, the participants felt tired and had some pain (Avg1 = 2.70, Avg2 = 3.30). However, we still can see that the free mode was more difficult for the participants. We found that the participants' answers were to some extent consistent with their handwriting performances.

3.2 B. Participants' emotions and focus

Emotions and eye-blinks are an indication of intrinsic cognition through extrinsic responses (Hömke et al., 2018), (Neumann et al., 2004). Emotion and eye-blinks can be a determination of an individual's focus and attention (Lipp & Neumann, 2004), (Adam et al., 2009). This enables the use of them as performance metrics for attention and focus in this experiment. The methodology used for this purpose was

automated through deep learning using a CNN model for emotion detection (Goodfellow et al., 2015) and an eye blink detection model for eye blink detection (Soukupová & Cech, 2016). First, each frame of the video captured during the experiment is analyzed. Then, the eye-blink is captured in frames correlating with that of eyes closed in each frame. Eyes are detected with an EAR (eye aspect ratio measurement). Each frame of eye blink was captured and their count was accumulated and noted. The attentiveness through this model was measured by the following equation (Patil et al., 2021) and the obtained results by age group are illustrated in Figs. 2 and 3:

$$\text{Attentivness} = \frac{\text{TotalFrames} - \text{BlinkFrames}}{\text{TotalFrames}} \quad (1)$$

Further, in Fig. 4 we showed the changes in eye blinks over time, where begin corresponds to (Eye blinks:B), midway (Eye blinks:M) and end (Eye blinks:E) of the handwriting task. Table 6 reports the p -values for the significance of eye blinks while handwriting task. It is clear that the effects of the haptic handwriting was significant to align with p -value of 0.0000 and 0.001 in middle and end of the experiment, respectively.

For the emotion detection process seven emotional states, namingly, happy, sad, fearful, angry, surprised, neutral and disgusted were identified using state of art Facial Expression Recognition (FER) dataset trained model (Goodfellow et al., 2015) which used haar features to identify the face and detect emotion. The captured videos of each subject were analyzed using this model and emotions of each frame were noted. Eventually, the mode of all emotions were taken for an

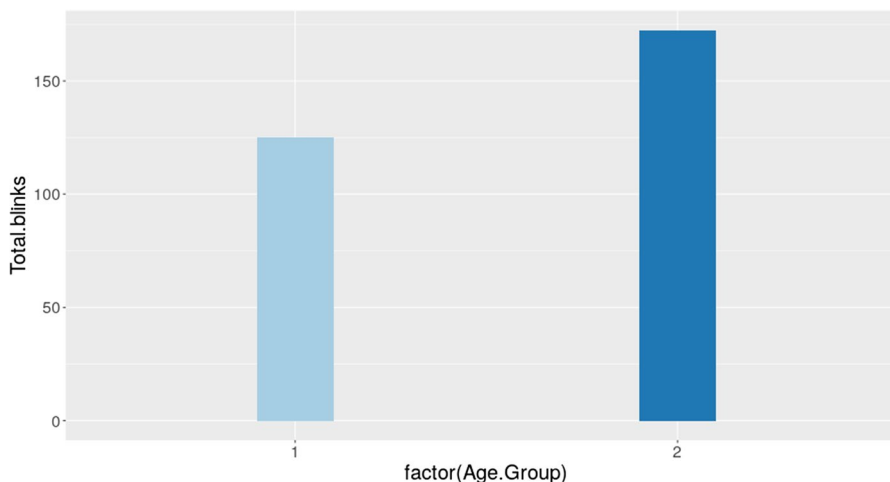


Fig. 2 Total blink rate by age group

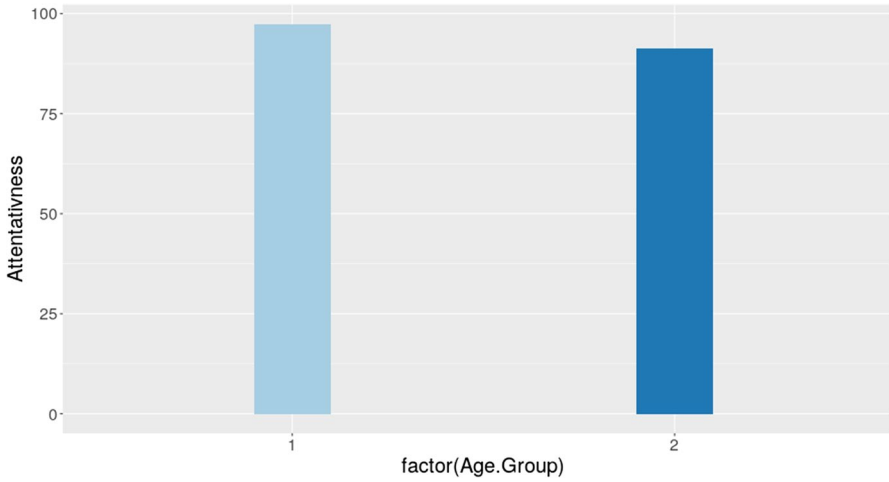


Fig. 3 Attentiveness by age group

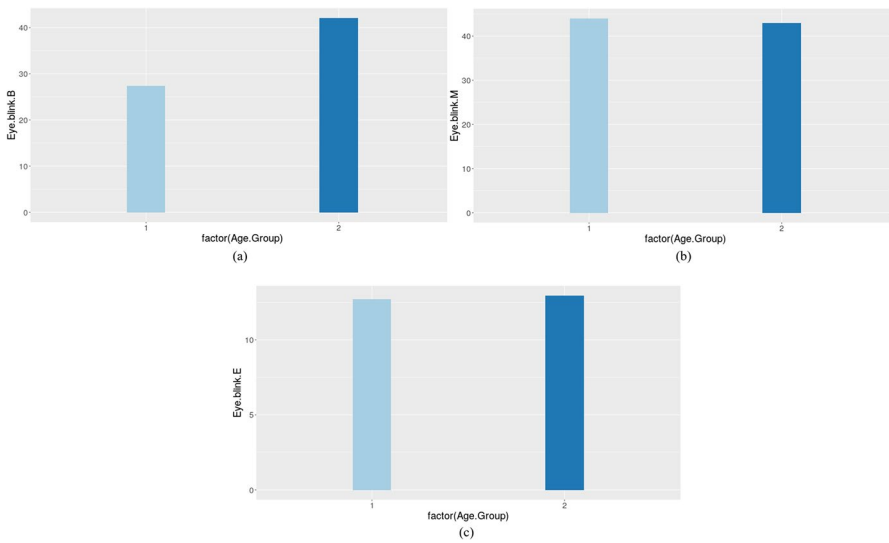


Fig. 4 Blink rate by age group in begin (a), mid (b) and end (c) of handwriting task

Table 6 *P*-value for blinks over time during the task

	Begin	Midway	End	Total blinks
<i>p</i> -value	0.405	0.000*	0.001*	0.844

* $p < 0.05$

overall emotional state description. The experiment was divided into the three zones; beginning, middle and end of the experiment with each zone's emotional state was captured at the end. This discriminated the emotion at each state of the experiment, and thus leading to a conclusion of the overall state of the participant throughout the experiment. As the deep learning models' accuracy is 63.2%, the results were further verified for random sets to validate the data and produce an observed analysis of the emotion. Figure 5 depicts the measured emotion from the video analysis performed during each scenario of the experiment. In cases where the emotions are unreliable, human observation is performed to decide if the models result correlates with the emotion predicted from the deep learning model used as in (Dahmane & Meunier, 2011).

Further, we showed the changes in participants' emotions over time (begin, midway, end, overall) during the handwriting task. Table 7 reports the *p*-values for the significance of emotional responses while doing handwriting. It is clear that the effects of the haptic handwriting was significant to modulate sadness with *p*-value of 0.034 and 0.014, midway and at the end, respectively.

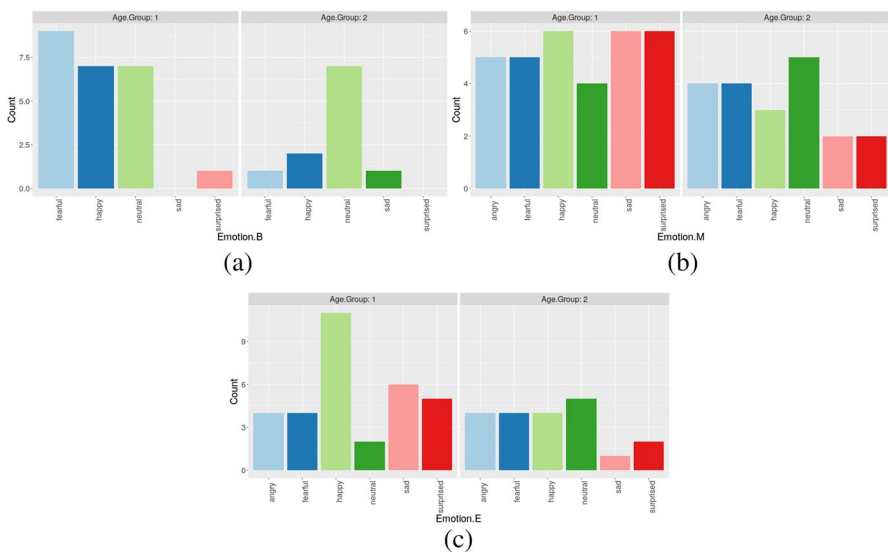
**Fig. 5** Emotions by age group in begin (a), mid (b) and end (c) of handwriting task

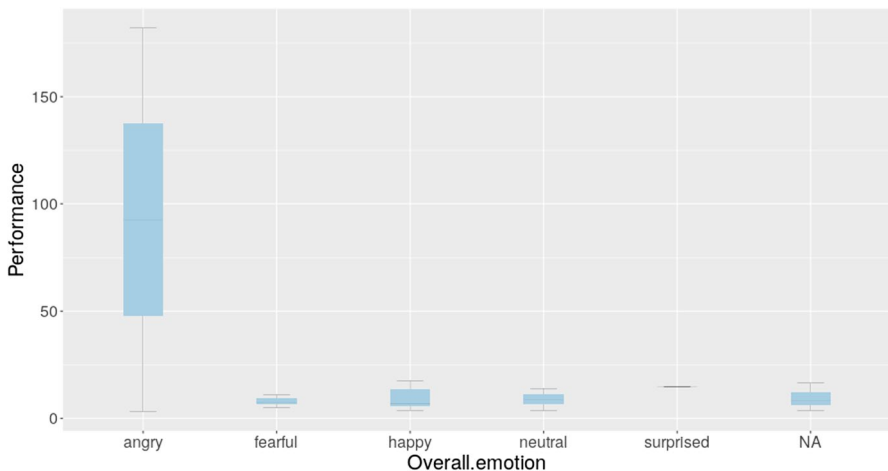
Table 7 *P*-value for emotions over time during the task

	Begin	Midway	End	Overall emotion
Fearful	0.605	0.108	0.226	0.334
Happy	0.278	0.263	0.870	0.143
Neutral	0.222	0.125	0.764	0.068
Sad	0.090	0.034*	0.014*	0.181
Surprised	0.726	0.191	0.727	0.180

* $p < 0.05$

3.3 C. Participants' performance

In this section, we first report the effect of changes of emotions and attentiveness during the process on the handwriting performance. Second, we focus on the effectiveness of our handwriting activity in fine motoring the child to write Arabic letters. We used Analysis of variance (ANOVA) to compare the differences across the means of the two groups. The results report a significant difference between participants' eye blinks at begin and their performance in the practice part of Arabic letters ($F = 74.36$, $p = 0.001$), indicating a positive learning gain in handwriting. While no significant effects of emotion changes on the performance were reported either midway or at the end of the task, a significant effect of overall emotion was reported with a p -value of 0.0014. Furthermore, as reflected by the mean of overall emotions, the performance was found to be affected by angry emotions with a mean of 92.6, as shown in Fig. 6. No significant impact of observed emotion on resulting performance was reported (p -value = 0.366). The performance was measured by the following equation and the obtained results by age group are illustrated in Figs. 8 and 12:

**Fig. 6** Performance versus overall emotion

$$\text{Performance} = \frac{\text{DTW score}}{\text{Completion time}} \quad (2)$$

Note that we were firstly interested in letter tracing accuracy only. When considering the completion time as a cost function (Williams et al., 2016), the performance as such allowed us to account for the speed-accuracy trade-offs that participants would make when performing the handwriting task. Thus, we could determine if better tracing accuracy was a result of increased handwriting skill or simply a slower movement. A higher performance indicated efficient and overall good task performance.

Regarding the pre/post-tests, it was scored on a five-point scale (Very much (5), A little, Neutral, Not much, Not at all (1)) to measure the progress in letters learnt during the sessions. To analyze if there was overall learning in writing letters for all children in all sessions, we did a paired t-test (since our data is normal distributed), which indicated that post-test scores were significantly higher than pre-test scores (p -value=0.015). However, ranking between 1 and 5 was not reflecting the improvement in writing performance of children who were initially unable to write at all. There were instances where the pre-test performance was not gradable (no letter was written) and the post-test performance was very low but comparably closer to actual writing. We clearly remark that some learning took place, but both performances received little rank. In order to reliably quantify the writing performance, by focusing on letter shape quality, we switched to DTW and demonstrated the results in Figs. 7 and 8. Moreover, the average DTW and performance scores in the training session for each participants are shown in Figs. 9 and 10 respectively.

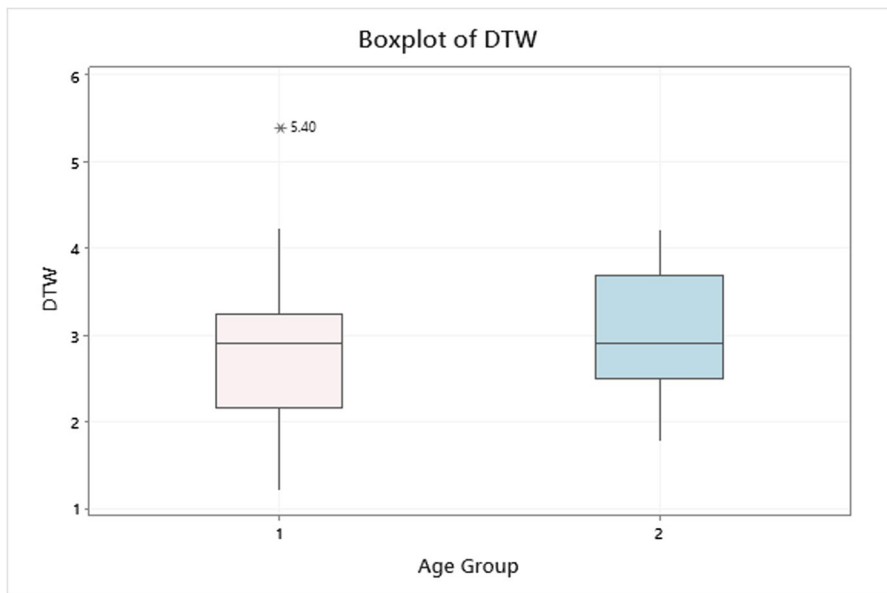


Fig. 7 Average DTW results by age group

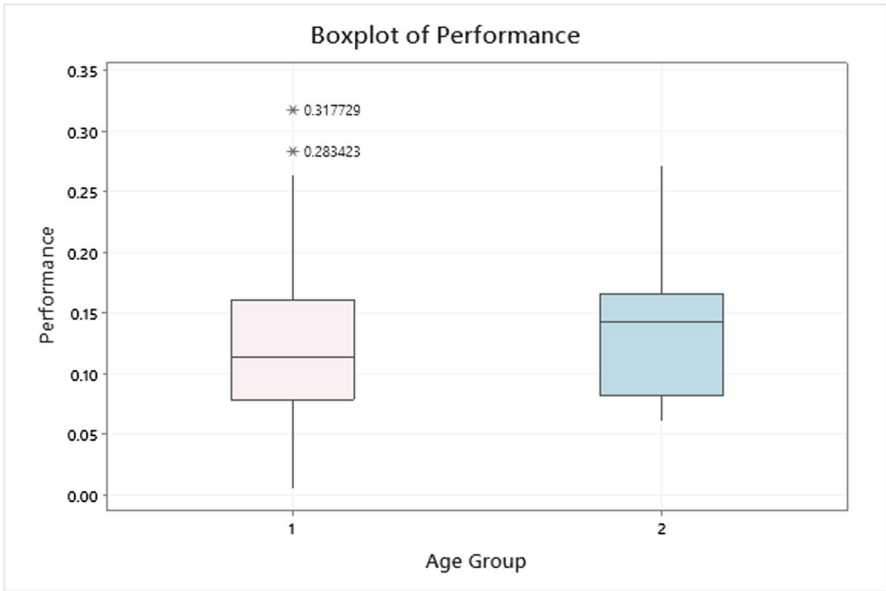


Fig. 8 Average performance by age group

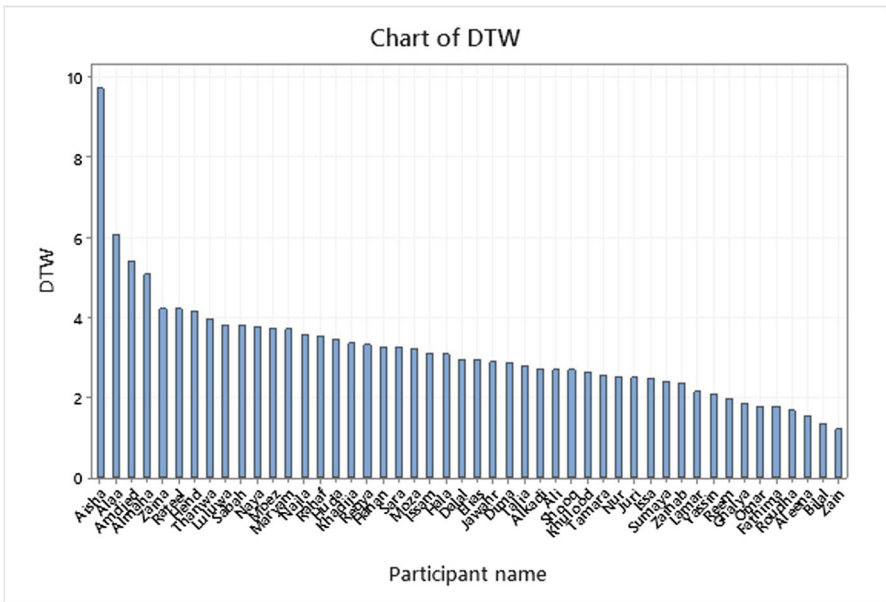


Fig. 9 Average DTW results for each participant

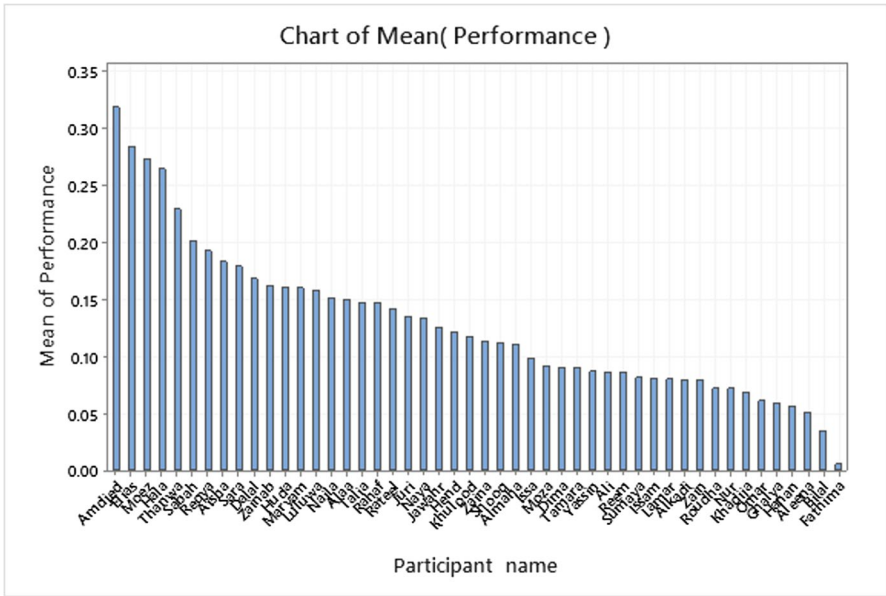


Fig. 10 Average performance results for each participant

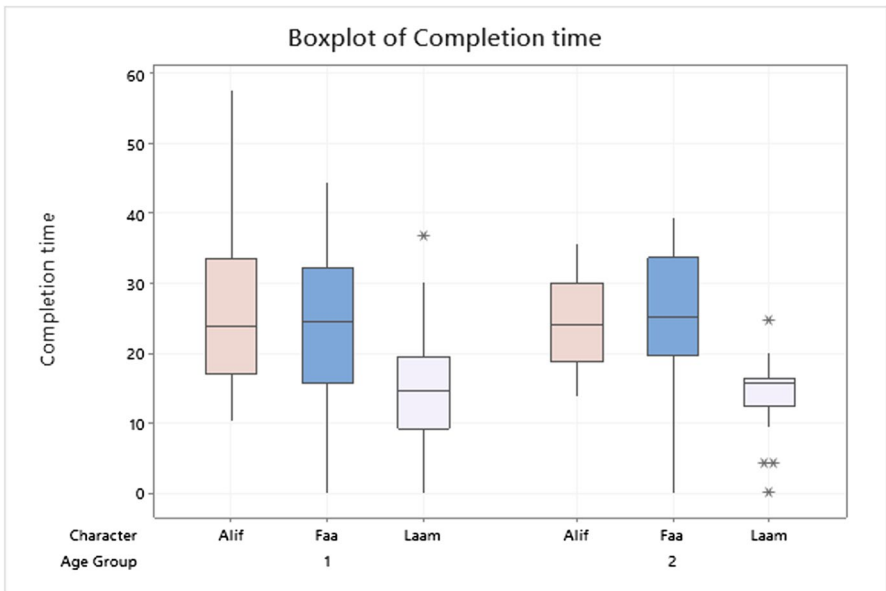


Fig. 11 Average completion time by age group and Arabic letter

Regarding the effective completion time, we can see that the average completion time for handwriting the letter laam was shorter for both groups, as shown in Fig. 11. The effect of age group is statistically not significant for both groups (p -value=0.114). However, completion time and post-test were significant with p -values of 0.001 and 0.008, respectively. It indicates that both groups improved their hand control of the haptic device to trace the letter correctly and improved their post-test as shown in Fig. 12. Figure 13 gives the correlation between all metrics, while Table 8 gives strong correlations only.

3.4 D. Participants' character recognition

For most computer vision applications, deep learning architectures have been the de facto choice (Ali et al., 2020; Balaha et al., 2020). Their effectiveness, however, is strongly reliant on access to labelled large-scale datasets. In this experiment, we are using a relatively small dataset with three classes only, thus machine-learning algorithms will be sufficiently enough. The dataset has 147 images and classified into three classes Alif (ا), Faa (ب) and Laam (ل), as shown in Table 10. The Alif has 51 samples while the Faa and Laam have 48 samples. In the experiment we used ten-fold cross-validation to estimate our proposed machine models skill.

To prepare the data for training, we used the PHOG Filter (Pyramid Histogram of Oriented Gradients) proposed by Bosch et al. (Bosch et al., 2007) for feature extraction. The PHOG stores information about how intensity gradients are oriented across an image. At each resolution level, it consists of a histogram of orientation gradients

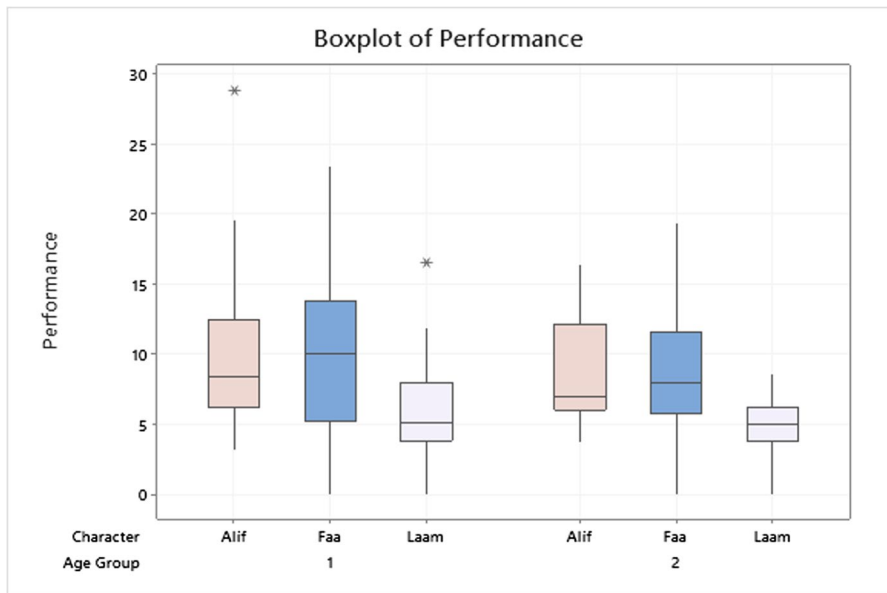


Fig. 12 Average Performance by age group and Arabic letter

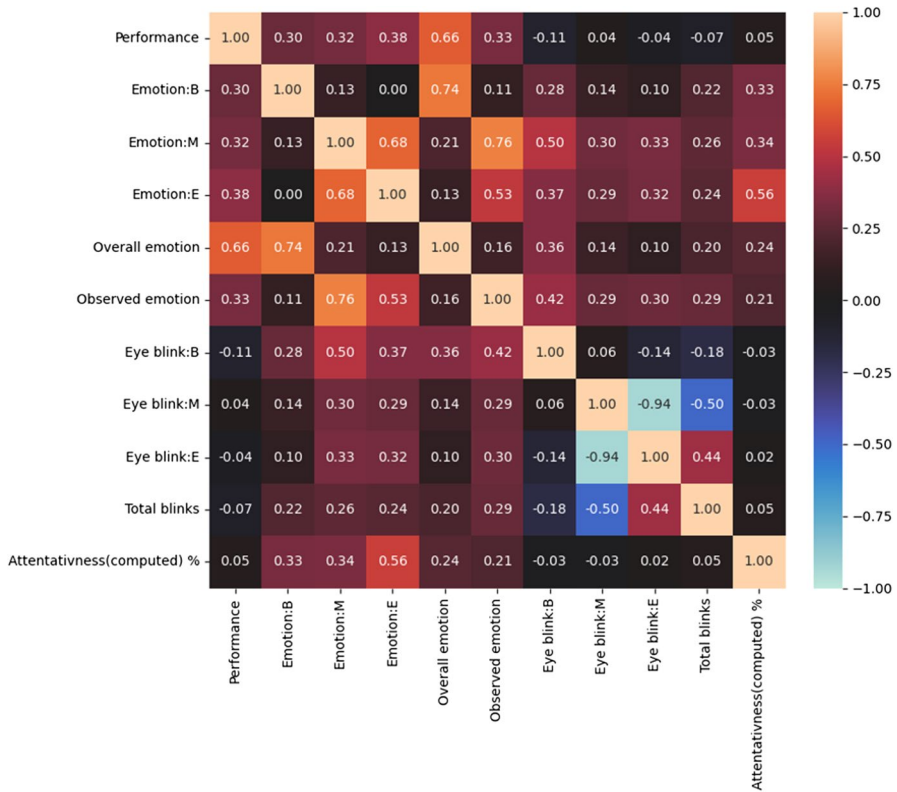


Fig. 13 Correlation matrix between all metrics

Table 8 Strong correlations between performance metrics

Metric 1	Metric 2	Correlation
Performance	Overall emotion	0.661238
Emotion:B	Overall emotion	0.740985
Emotion:M	Emotion:E	0.682690
Emotion:M	Observed emotion	0.763179
Emotion:E	Observed emotion	0.530217
Emotion:E	Attentiveness	0.556435
Eye blink:M	Eye blink:E	-0.936135

over each image sub-region. The distance between two PHOG image descriptors indicates how similar the images are in terms of shape and spatial layout.

Three classifiers were tested to classify images, namely SMO, Random Forest and Multilayer Perceptron. The Sequential Minimal Optimization (SMO) proposed (Platt, 1998) to tackle the quadratic programming (QP) problem that emerges during the training of Support Vector Machines. The Random Forest is made up of a huge number of individual decision trees that work together as an ensemble. Each tree in the random forest produces a class prediction, and the class with the most votes becomes the prediction of our model. The Multilayer Perceptron is a type of neural network. The strength of neural networks derives from their capability to learn how to effectively match the representation in the training data to the output variable to be predicted. Neural networks can learn any mapping function mathematically and have been demonstrated to be a universal approximation algorithm. The results on Table 9 show the performance of the three classifiers. We can see that the PHOG features help in gaining high accuracy with the three classifiers and the Multilayer Perceptron gained the best performance.

4 Discussion

4.1 A. Observations

Currently, studies on the effectiveness of learning technologies that also integrate haptic devices for educational purpose are not investigated for young students. Moreover, they involve some methodological issues such as small sample sizes (Schrader & Kalyuga, 2020). Therefore, in order to contribute to research on learning technologies with haptic devices, the purpose of this study was to determine the relationship between participants' emotion and attentiveness, and their resulting performance as well as the effectiveness of writing Arabic letters using such difficult VR environment. Overall, the results confirm that using the haptic device had a significantly positive effect on students' writing performance as measured at both pre-test and post-test. However, no significant effective change in emotion was detected. After a careful analysis, we can make the following key observations:

- **First**, results from this feasibility study indicated that the haptic task was acceptable to study participants. The participants did not face large difficulty in using the haptic device and the system. However, not all participants completed the whole session due to time constraints. While the training time for the hap-

Table 9 Results of using PHOG feature extraction for SMO, Random Forest and Multilayer Perceptron classification

classifier	Accuracy	Precision	Recall	F-Measure
SMO	96.599%	0.967	0.966	0.966
Random Forest	95.238%	0.953	0.952	0.952
Multilayer Perceptron	97.279%	0.973	0.973	0.973

tic device was quite limited, these results have to be interpreted with caution. According to (Pekrun, 2006), the emotional experience has an impact on learning outcomes in more extensive learning situations.

- **Second**, our results, as can be seen from Table 7, show no direct relation of emotion change with resulting performance. However, these findings were not in line with other research (Hill et al., 2019; Pekrun, 2006; Pekrun et al., 2007; Shute et al., 2015) that have presented theoretical and empirical evidence to support the impact of emotions on students' resulting performance. Our findings might be related to the emerging VR settings which seems to provide mixed results due to many reasons such as the experiment settings. For instance, the study (Karumbaiah et al., 2022) pointed out that there is no significant difference in students' academic achievement and retention in various emotional states in different VR learning, particularly in social study learning.
- **Third**, the analysis from Table 6 indicates that eye-blinks modulates participant's focus. For instance, statistically significant reduction of blinks in the middle and towards the end of the handwriting task, as shown by p -values < 0.05 . This finding partly confirms and builds on the common notions where after a while the student eye start to focus more and so the attention level. However, our results from Table 8 suggests that attentiveness has very weak effect on the resulting performance. These findings are not consistent with those of previous studies (Au et al., 2016; Fredricks & McColskey, 2012) which came in fact after a long history of studies investigating the impact of attentiveness in learning outcome. The latter studies show that attentiveness has a positive impact on performance and might further mediate the link between emotions and learning performance.
- **Fourth**, the correlation analysis from Table 7 indicates that the emotion of sadness affects the performance scores in both groups with a p -value of 0.014 and 0.034 for the two groups. This cannot be correlated with common notion that sadness indicates focus. However, this finding seems to be in line with common belief that some negative emotions may promote learning by motivating to learn better (Worsley & Blikstein, 2015). For instance, the participants who displayed sadness maybe also effected by other external factors such as reluctance to do the experiment after their school hours or the deep learning model has misinterpreted the emotion due to model bias.
- **Fifth**, regarding the children's performance in handwriting, we observed significant improvements for two young kids, especially who have had difficulty in tracing correct shape and orientation, see Arabic letter samples in Table 10, while most of participants performed good as their letter tracings were accurately recognized as demonstrated in Table 9.

To analyze if there was grasping in writing letters for all children in all sessions, we did an analysis comparing both groups which indicated that DTW error scores were significantly lower for group 1 than group 2 as shown in Fig. 7. Though no significant effect of age group variable was detected for both age groups ($p = 0.351$, $p = 0.356$). Figure 9 explores the results of DTW scores for all participants. It is observed that the errors were either due to the participant writing a segment in the

Table 10 Samples of collected handwriting (guided and free) with DTW error scores, cleaned versions, and pre- and post-tests for each character

Character	Pre-test	Guided writing	Free writing	Cleaned (free writing)	Post-test
Alif					
Laam					
Faa					

wrong direction, missing a segment or missing a directional change completely and continuing in the same direction. On the other hand, many participants start the letter tracing from left to right or down to up which causes large error score. Some missed the punctuation after finishing the letter tracing, while others lost the focus and had difficulties to control the grip of the pen.

We remark that maintaining a precise visuo-motor coordination was challenging for most of the participants. The stability of their hands' movement had impact on the grip strength and grip precision where we observed that participants were moving their hands while writing which caused unnecessary noise. In particular, at begin and end of the trace, the participant who do not know the shapes were drawing without notice. However, we believe that the training session under the guided mode followed by the use of force feedback is responsible of the results (improvement in the engagement, handwriting legibility and speed) observed here and presents a possible enhancement compared to traditional methods. This assumption appears to be in line with the literature interested in the contribution of haptic devices on learning (Anagnostopoulou et al., 2021).

As shown in Fig. 13, attentiveness and performance were weakly positively correlated. Likewise, total blinks and performance were weakly positively correlated, whereas we did not find any significant correlations between total blinks and attentiveness. In addition, the correlation between the observed and the predicted emotion was also weak; where it clearly indicates no correlation with each other's. This

finding reveals that no association between the output of the emotion detection model and human observation. Although our observations suffered from the potential problem that no self-reported or additional measures were used to correlate the emotion, the strong correlation between the performance and the overall emotions does provide evidence that it might be reliable to detect emotional states using CV models rather than human observations. Nonetheless, our results report strong correlation between observed emotion and participants' emotion midway. Table 8 shows strong correlations between performance metrics only.

4.2 B. Implications

Implications for teaching: As discussed in the first section, emotions and attentiveness are often treated as contributors to learning performance. However, as demonstrated in this work, emotion/attentiveness-outcome relationships are not as straightforward as is generally assumed. The relationship is even more complex when considering emotion changes and eye blink changes, long interaction time, physical and mental fatigue, etc. Except for overall emotion, performance was affected by overall emotion (angry/neutral) as well as by eye blinks at the beginning of the experiment only. Hence, more careful attention must be paid in learning in making assumptions about emotion's role in student learning. Other factors, such as learning context (duration, mental fatigue, etc.), may also play an important role, and will be valuable to examine in such a VR task. Thus, an optimal learning and teaching style can be assigned to a particular student to perform optimally with. This is also helpful to teacher to adapt to individual students, so that they might learn successfully.

Implications for research: there is mixed findings (Karumbaiah et al., 2022; Schrader & Kalyuga, 2020) on the role of student affective states in learning. More research is needed to establish how emotion/attentiveness should be conceptualized to measure it effectively. First, current emotion detection methods may need to be improved to capture rare emotional states and resolve object occlusions more efficiently. Second, the null results for rare could be a limitation of the current approaches that fail to recognize the role of sparse but potentially important cases. Third, there is a possibility to measure a wider range of variables beyond emotion/attentiveness and handwriting outcomes such as engagement. Fourth, observing and asking students about their emotions during the VR tasks distract their learning, which might also influence their judgments as mentioned in (Schrader & Kalyuga, 2020). To avoid such distraction, we need to implement better indicators of discrete emotional and attentional states.

5 Conclusion

There is limited research on exploring handwriting using haptic device to improve the children focus and fine-tune motor skills. This study shows the feasibility of utilizing such task to explore the challenges and the solutions while producing

handwriting that resemble to a harsh learning environment. As expected, significant changes in emotions and attentiveness as indicators of focus over the study were reported. Likewise, no direct relation of emotion change with writing performance was found. Performance was affected by overall emotion (angry/neutral) as well as by eye blinks at the beginning of the experiment only.

Although the findings of the present study were supported by statistical significance, some limitations should be mentioned. First, due to the short-term nature of the situation in schools (COVID restrictions), we experimented in a single batch without the possibility for extension, and eventually repeated experimentation on same groups of participants. Second, the current study relied on manual video editing, human observations and emotion detection models to measure perceived and emotional responses rather than psycho-physiological measures. Video editing and human observations measure the subjective experience of participants' emotions and thus can be inconsistent, unreliable, and difficult to reproduce (Devillers et al., 2005). Another limitation is the use of CV models for eye blink and emotion detection (Goodfellow et al., 2015; Soukupová & Cech, 2016). Some participants covered their faces with masks, which had affected on their detecting emotions as key determining facial features were covered. To identify the emotional state in these as well as the eye blink count; they were deduced by human observation and were noted through feedback from the participant. An overall observed emotional state could only be determined from this approach. The approaches used for the emotional and eye blink detection can further be improved for diversity of faces and occlusions that may provide results that are more accurate. However, as we were separately observing the students' emotions, we did not consider emotional engagement. Furthermore, motivation was not assessed either, as the means used for this study did not account for the measurement of engagement and motivation. Other studies (Hamari et al., 2016; Wang et al., 2016) relied on lengthy questionnaire during the experiments, which did not seem to be practical for our case. Other attributes such as gender, gaze direction and head pose are expected to further modulate emotional responses. Finally, the current study did not consider the interaction with the system in real-time, which would largely influence the emotional responses. A multi modal interaction could be an interesting direction for future research. However, more attractive tasks and a wide variety of fine-motor tasks must be developed to retain participants' attention, especially in a long-term study. For Future works, we are planning to test this system on individuals with visuo-motor skills issues and/or handwriting learning disabilities to improve and retain fine motor skills.

Funding Open Access funding provided by the Qatar National Library. This work was made possible by NPRP grant #10-0205-170346 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

Data availability The datasets used during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adam, A. R., Mallan, K. M., & Lipp, O. V. (2009). The effect of emotional and attentional load on attentional startle modulation. *International Journal of Psychophysiology*, 74(3), 266–273. <https://doi.org/10.1016/j.ijpsycho.2009.09.011>
- Ali, A. A. A., Suresha, M., & Ahmed, H. A. M. (2020). A survey on Arabic handwritten character recognition. *SN Computer Science*, 1(3). <https://doi.org/10.1007/s42979-020-00168-1>
- Anagnostopoulou, D., Efthymiou, N., Papailiou, C., & Maragos, P. (2021). Engagement estimation during child robot interaction using deep convolutional networks focusing on ASD children. 2021 IEEE International Conference on Robotics and Automation (ICRA). <https://doi.org/10.1109/icra48506.2021.9561687>
- Ashwin, T. S., & Guddeti Ram, M. R. (2020). Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. *Education and Information Technologies*, 25(2), 1387–1415. <https://doi.org/10.1007/s10639-019-10004-6>
- Au, O. T., So, R., & Lee, L. (2016). Attentiveness and self-studying are keys to academic performance. *International Symposium on Educational Technology (ISET)*, 2016, 61–64.
- Balaha, H. M., Ali, H. A., Saraya, M., & Badawy, M. (2020). A new Arabic handwritten character recognition deep learning system (AHCR-DLS). *Neural Computing and Applications*, 33(11), 6325–6367. <https://doi.org/10.1007/s00521-020-05397-2>
- Baltrusaitis, T., Robinson, P., & Morency, L.-P. (2016). OpenFace: An open source facial behavior analysis toolkit. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). <https://doi.org/10.1109/wacv.2016.7477553>
- Bosch, A., Zisserman, A., & Munoz, X. (2007). Representing shape with a spatial pyramid kernel. Proceedings of the 6th ACM International Conference on Image and Video Retrieval - CIVR '07. <https://doi.org/10.1145/1282280.1282340>
- Bowers, L., & Hayle, R. (2020). Creative haptics: An evaluation of a haptic tool for non-sighted and visually impaired design students, studying at a distance. *British Journal of Visual Impairment*, 39(3), 214–230. <https://doi.org/10.1177/0264619620912771>
- Creelman, K. (2021). A literature review of understanding and supporting students with attention deficit hyperactivity disorder in the classroom. *Northwest Journal of Teacher Education*, 16(1), Article 3. <https://doi.org/10.15760/nwjte.2021.16.1.3>
- Dahmane, M., & Meunier, J. (2011, March 1). Emotion recognition using dynamic grid-based HoG features. IEEE Xplore. <https://doi.org/10.1109/FG.2011.5771368>
- Derakhshandeh, Z., Esmacili, B. & Patel, B. (2021). Student engagement and attentiveness in online learning systems. In T. Bastiaens (Ed.), Proceedings of EdMedia + Innovate Learning (pp. 45–58). United States: Association for the Advancement of Computing in Education (AACE). Retrieved July 5, 2022 from <https://www.learntechlib.org/primary/p/219638/>
- Devillers, L., Vidrascu, L., & Lamel, L. (2005). Challenges in real-life emotion annotation and machine learning based detection. *Neural Networks*, 18(4), 407–422. <https://doi.org/10.1016/j.neunet.2005.03.007>

- Dewan, M. A. A., Murshed, M., & Lin, F. (2019). Engagement detection in online learning: a review. *Smart Learning Environments*, 6(1). <https://doi.org/10.1186/s40561-018-0080-z>
- Fredricks, J.A. & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In: Christenson, S., Reschly, A., Wylie, C. (eds) *Handbook of Research on Student Engagement*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4614-2018-7_37
- Goodfellow, I. J., Erhan, D., Luc Carrier, P., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.-H., Zhou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shawe-Taylor, J., Milakov, M., Park, J., & Ionescu, R. (2015). Challenges in representation learning: A report on three machine learning contests. *Neural Networks*, 64, 59–63. <https://doi.org/10.1016/j.neunet.2014.09.005>
- Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PLoS ONE*, 16(4), e0249792. <https://doi.org/10.1371/journal.pone.0249792>
- Guneyso Ozgur, A., Özgür, A., Asselborn, T., Johal, W., Yadollahi, E., Bruno, B., Skweres, M., & Dillenbourg, P. (2020). Iterative design and evaluation of a tangible robot-assisted handwriting activity for special education. *Frontiers in Robotics and A, I*, 7. <https://doi.org/10.3389/frobt.2020.00029>
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54, 170–179. <https://doi.org/10.1016/j.chb.2015.07.045>
- Hill, J., Healey, R. L., West, H., & Déry, C. (2019). Pedagogic partnership in higher education: Encountering emotion in learning and enhancing student wellbeing. *Journal of Geography in Higher Education*, 1–19. <https://doi.org/10.1080/03098265.2019.1661366>
- Hömke, P., Holler, J., & Levinson, S. C. (2018). Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLoS ONE*, 13(12), e0208030. <https://doi.org/10.1371/journal.pone.0208030>
- K V, K., & Bahel, V. (2021). Transfer learning approach for analyzing attentiveness of students in an online classroom environment with emotion detection. <https://doi.org/10.20944/preprints202105.0303.v1>
- Kainat, Ali, S., Khawaja, F.I., Avaz, Y., & Saiid, M. (2022). A review on different approaches for assessing student attentiveness in classroom using behavioural elements. 2022 2nd International Conference on Artificial Intelligence (ICAI) (pp. 152–158)
- Karumbaiah, S., Baker, R., Tao, Y., & Liu, Z. (2022). How does students' affect in virtual learning relate to their outcomes? A systematic review challenging the positive-negative dichotomy. *LAK22: 12th International Learning Analytics and Knowledge Conference*. <https://doi.org/10.1145/3506860.3506863>
- Khan, F. (2020). Facial expression recognition using facial landmark detection and feature extraction via neural networks. ArXiv:1812.04510 [Cs]. <http://arxiv.org/abs/1812.04510>
- Young-Seok Kim, Collins, M., Bulmer, W., Sharma, S., & Mayrose, J. (2013). Haptics Assisted Training (HAT) System for children's handwriting. 2013 World Haptics Conference (WHC). <https://doi.org/10.1109/whc.2013.6548469>
- Lipp, O. V., & Neumann, D. L. (2004). Attentional blink reflex modulation in a continuous performance task is modality specific. *Psychophysiology*, 41(3), 417–425. <https://doi.org/10.1111/j.1469-8986.00165.x>
- Mukhopadhyay, M., Pal, S., Nayyar, A., Pramanik, P. K. D., Dasgupta, N., & Choudhury, P. (2020). Facial emotion detection to assess learner's state of mind in an online learning system. Proceedings of the 2020 5th International Conference on Intelligent Information Technology. <https://doi.org/10.1145/3385209.3385231>
- Neumann, D. L., Lipp, O. V., & McHugh, M. J. (2004). The effect of stimulus modality and task difficulty on attentional modulation of blink startle. *Psychophysiology*, 41(3), 407–416. <https://doi.org/10.1111/1469-8986.2004.00185.x>
- Patil, A., Singh, A., & Chauhan, N. (2021). Attentiveness monitoring and user record maintenance in virtual classrooms. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). <https://doi.org/10.1109/iciccs51141.2021.9432169>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions. In *Emotion in Education* (pp. 13–36). <https://doi.org/10.1016/b978-012372545-5/50003-4>

- Platt, J. (1998) Sequential minimal optimization: A fast algorithm for training support vector machines, advances in kernel methods. Support Vector Learning, MIT Press.
- Salvador, S., & Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11(5), 561–580. <https://doi.org/10.3233/ida-2007-11508>
- Savov, T., Terzieva, V., & Todorova, K. (2018). Computer vision and internet of things. Proceedings of the 19th International Conference on Computer Systems and Technologies - CompSysTech'18. <https://doi.org/10.1145/3274005.3274014>
- Schrader, C., & Kalyuga, S. (2020). Linking students' emotions to engagement and writing performance when learning Japanese letters with a pen-based tablet: An investigation based on individual pen pressure parameters. *International Journal of Human-Computer Studies*, 135, 102374. <https://doi.org/10.1016/j.ijhcs.2019.102374>
- Schrader, C., & Nett, U. (2018). The perception of control as a predictor of emotional trends during gameplay. *Learning and Instruction*, 54, 62–72. <https://doi.org/10.1016/j.learninstruc.2017.08.002>
- Schutz, P. A., & Pekrun, R. E. (2007). *Emotion in education*. Elsevier Academic Press.
- Shute, V. J., D'Mello, S., Baker, R., Cho, K., Bosch, N., Ocumpaugh, J., Ventura, M., & Almeda, V. (2015). Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. *Computers & Education*, 86, 224–235. <https://doi.org/10.1016/j.compedu.2015.08.001>
- Soukupová, T., & Cech, J. (2016). Real-Time Eye Blink Detection using Facial Landmarks. *Proceedings of the 21st Computer Vision Winter Workshop (CVWW 2016)*, pp 1–8.
- Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction*, 43, 16–26. <https://doi.org/10.1016/j.learninstruc.2016.01.008>
- Wessa, P. (2021). Cronbach alpha (v1.0.6) in Free Statistics Software (v1.2.1), Office for Research Development and Education. https://www.wessa.net/rwasp_cronbach.wasp/. Accessed 28 June 2022
- Williams, C. K., Tremblay, L., & Carnahan, H. (2016). It pays to go off-track: Practicing with error-augmenting haptic feedback facilitates learning of a curve-tracing task. *Frontiers in Psychology*, 7, <https://doi.org/10.3389/fpsyg.2016.02010>
- Worsley, M. & Blikstein, P. (2015). Using learning analytics to study cognitive disequilibrium in a complex learning environment. In Proceedings of the Fifth International Conference on Learning Analytics and Knowledge (pp. 426–427).
- Zhao, H., Zheng, Z., Swanson, A., Weitlauf, A., Warren, Z., & Sarkar, N. (2018). Design of a haptic-gripper virtual reality system (Hg) for analyzing fine motor behaviors in children with autism. *ACM Transactions on Accessible Computing*, 11(4), 1–21. <https://doi.org/10.1145/3231938>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.