Master Creative Intelligence and Technology

Unlabelled emotions: A responsive drawing interface for reflective expression

Name:	Miruna-Cristina Cotfas
Student ID:	s4469844
Date:	[21/08/2026]
1st supervisor:	Rob Saunders
2nd supervisor:	Tessa Verhoef

Master's Thesis in Creative Intelligence and Technology

Leiden Institute of Advanced Computer Science
Leiden University
Einsteinweg 55
2333 CC Leiden
The Netherlands

Contents

1 Introduction **3**
- 1.1 Research Question 4
- 1.2 Thesis Structure 4

2 Related Work **4**
- 2.1 Conceptual Foundations 4
- 2.2 Methodological Approach 9

3 Responsive Sketchpad Design & Implementation **10**
- 3.1 Design Principles 10
- 3.2 Mapping Exploration 11
- 3.3 System Architecture and Data Logging 16

4 Experiment Methodology **16**

5 Results **18**
- 5.1 Participants 18
- 5.2 Users' Perception of Video Stimuli 19
- 5.3 Drawing Movement Metrics 19
- 5.4 User Experience 20
- 5.5 Relations between individual context and drawing movement 22

6 Discussion **25**
- 6.1 Supporting Emotional Reflection 25
- 6.2 Interaction Roles 25
- 6.3 What Does This All Say About the Affective Mappings? 26
- 6.4 Limitations 26

7 Conclusion **27**

Appendices **28**

A Additional figures **28**

B Questionnaire & Sketchpad **28**

C Data Archive **28**

Abstract

Drawing behaviour has, in the realm of affective computing, frequently been approached as a source of categorising emotions. However, the subjectivity of both art and felt experiences raise questions about how suitable it is to only leverage such data for classification tasks. This thesis explores a different direction. It uses drawing as implicit input and instead of labelling emotions, it uses previously researched associations between the two to produce visual outputs fed as responses in an interactive sketchpad, creating a feedback loop hypothesised to support affective self-reflection.

The web-based prototype analyses stylus characteristics such as speed, pressure, direction variation, stroke length, spatial density and rhythm and turns them into real-time transformations applied to users' strokes. The system was evaluated by 16 participants in a within-subject experiment, where they interacted with both a neutral and the responsive drawing condition following exposure to contrasting emotional video stimuli. Metrics gathered from drawing were analysed both independently and next to survey answers to gauge emotional reflection and interaction perception.

The responsive sketchpad received higher ratings across all measured dimensions of emotional reflection, particularly attention to internal state and depth of reflection. Changes in drawing behaviour, such as longer pauses and more direction variation or reversals can be associated with increases in reflection ratings. However, the responsive condition was frequently perceived as distracting or unpredictable and was more successful at altering the trajectory of the drawing process than at revealing previously unnoticed emotions. The results suggest that the value of such a system may lie in creating moments where users stop and reconsider the interaction.

1 Introduction

There have undoubtedly been many situations in which I started drawing mindlessly on the corner of a notebook. Although I never gave the next stroke a second thought, I could feel how my emotional state at that point in time was reflecting onto the drawing unravelling at the tip of my ballpoint pen. Whether it's a decision, a hesitation, a repetition or an attempt to represent something, a simple line can hold many meanings.

How humans engage with technology has always been fascinating to me, especially the idea that interaction can extend beyond what a user tells the system and be shaped by how they behave in doing so. As more and more people nowadays turn towards digital devices for tasks such as writing or drawing, these patterns become observable; things that seem as ordinary as moving a pen across a surface can become data for various systems. By leveraging this data, many technological advancements can be, and have already been achieved in the field of affective computing, with the primary aim of successfully classifying emotional states [Yang and Qin, 2021]. Various metrics, ranging from speed and pressure to stroke length and tilt have been researched in past years and have re-shaped the way we understand affect as an embodied characteristic. Nevertheless, focusing only on predictions and classifications can prove risky when dealing with such subjective experiences, losing sight of the complexity of both art and emotion itself.

Moreover, much of this research is oriented towards clinical assessment. This is largely driven by EMOTHAW, undoubtedly the most prominent publicly available dataset that connects handwriting and drawing motion metrics to emotional states [Ayzeren et al., 2019]. Since this resource leverages the DASS (Depression Anxiety Stress Scales) set of questionnaires, it has shaped the direction of the field towards identifying these three affective states. In trying to create tools that could reveal certain mental health conditions, the target narrowed in on negative sentiments, leaving a significant part of the wheel of emotion untouched. This includes positive inner states, but also those that are difficult to name and fit into neatly predefined categories.

Another limitation identified in the field is that drawing data is commonly used for analysis after experiments have been finalised [Yang and Qin, 2021]. What remains less explored is how the knowledge that we have about these relationships can be used amidst the interaction itself. When these insights are used so that a system can respond in real time, we create a feedback loop, where not only does the drawing behaviour influence the system, but the responses in turn influence how the user interprets, and thus continues the drawing.

Consequently, to me, the question remains: what more can a system actually do with this data? Prior research informs this study by setting foundations, uncovering gaps and highlighting interesting insights. This project shifts the perspective towards interactive interfaces, aiming to understanding how a system can respond to drawing behaviour in real time, so as to help the user reflect on their own emotions. A working prototype has been developed for such a responsive sketchpad, which deliberately steers clear of printing out exact labels of a participant's emotional state. Instead of informing them of what affective state they are experiencing, the system aims to make these characteristics understandable implicitly through visual metamorphoses of their own strokes. Participants are ultimately invited to interpret these visual reactions as they feel resonates most.

Additionally, to evaluate the impact of this product, a comparative research method is proposed. The responsive sketchpad will be analysed against a neutral one. By doing this, examining the change introduced by the responsiveness, as well as its limitations or improvement points, becomes possible.

1.1 Research Question

This project is centred around two main research questions:

RQ1: *How can a responsive sketchpad, compared with a neutral one, support users in reflecting on their current emotional state?*

RQ2: *How do users perceive the role of each sketchpad in the drawing process, and what aspects of the interaction shape these perceptions?*

To investigate these questions, stylus data gathered from the drawing experiments will be analysed, as well as qualitative feedback from the questionnaire. Finally, relationships between the two will be established in order to better understand the extent to which an individual's thought process or context influences the results. All in all, this work is meant to uncover whether an affective computing system can successfully turn towards more open and participatory experiences, touching upon possibilities for self-reflection through digital expression. The secondary analytical aim is to get a better understanding of how insights from prior literature, explored in Section 2, are supported or challenged by these findings.

1.2 Thesis Structure

This paper is further structured as follows:

Chapter 2 *reviews prior work that has been done in the field of connecting drawings to emotions*

Chapter 3 *explains the implementation of the responsive sketchpad system*

Chapter 4 *describes the experiment set-up and methodology*

Chapter 5 *explores the results*

Chapter 6 *develops a discussion on the findings*

Chapter 7 *provides subsequent conclusions and identifies possibilities for future work*

2 Related Work

Previous research surrounding the intersection of (digital) drawing and emotions focuses on two different analysis directions: stylus movement characteristics or the final image, most often its content or composition. Both are relevant to this study and will be further explored in this section, alongside other relevant themes.

2.1 Conceptual Foundations

First, we will take a look at online collection, which is preferred for its ability to capture dynamic data such as pen position, status, timestamps, pressure and tilt, which are considered "behavioural biometrics". As was touched upon already, the focus of prior work is strongly oriented towards negative emotions, mental health afflictions and clinical applications. The main reason for this, in my view, is the available data in the field. [Ayzeren et al., 2019] reviewed all datasets that contain (soft-)biometric data, emphasising those with connections to handwriting. They find that EMOTHAW, which contains both handwriting and drawing samples, was the only publicly available dataset with emotion labels back in 2019. To this day, it remains the most relevant dataset, supplying an end-to-end methodology, from tasks to gather the data, to questionnaires, feature extraction and supervised classification. [Likforman-Sulem et al., 2017] situate their dataset in line with standard tests such as House-Tree-Person (HTP) or clock-drawing tests, which aim to find deviations from a healthy mental state. By specifically narrowing the focus to the three key emotions: depression, anxiety and stress allowed them to leverage the DASS questionnaires, introduced in [Parkitny and McAuley, 2010], and turn these three categories in prediction targets. This created a gap in both concept and methodology in the field, as positive or non-clinical states are significantly less explored, especially in process-based stylus analyses.

[Yang and Qin, 2021] provides a thorough overview of works that chose this route, the majority of which have opted to look at the correlations between metrics and emotional states as a classification problem. Fine details about the implementation of the algorithms are less relevant to the discussion, but results definitely provide insights of utmost importance. While earlier studies achieved accuracies ranging between 59% and 74%, more extensive feature sets allowed [Nolazco-Flores et al., 2022] to bring the number all the way up to 82% for recognising depression, and almost 75% for anxiety or stress. [Rahman and Halim, 2023], as another example, fused temporal and stroke-based measures with frequency-domain features, horizontal and vertical pen velocity. The classification then increased further to 89% for depression, 80.03% for anxiety and 75.39% for stress. The strongest results come from [Khan et al., 2024], with 92.64% for depression, 83.22% for anxiety and 79.41% for stress. This is achieved using an attention-based transformer model. All these findings support the hypothesis that there are predictive associations that can be made between stylus behaviour and labels for depression, anxiety or stress gathered from the DASS set of surveys, but they do not confidently provide any correlations between individual features and the emotions covered.

[Santosa et al., 2026] strengthens this claim. Throughout their study, they gave participants no clear instructions as to what emotion to depict, but asked them to complete a digital diary study over the course of two weeks. Every day, they had to create a drawing of their choice that reflected their inner state, alongside reporting their valence and arousal. When combining qualitative feedback with both pen data and content analysis, they found no significant association between expression patterns and valence or arousal. For this project, this finding was a turning point. Even when classification seems to achieve high accuracy on controlled datasets, we might still be missing crucial aspects of art or emotions. Structured as a review paper in the field of psychology and affect, [Lin et al., 2022] also uncovered that when understanding correlations between drawing and mental health or behavioural traits, results are rarely stable across findings. They combined comparable indicators across 17 studies and found that most reported associations only appeared in one paper. This seems to be a trend in the field and only goes to show that emotion and art are both much too complex to be encapsulated wholly in very exact studies, and achieving high classification accuracy does not automatically mean that relationships are universal.

Given the subjectivity of the fields at the intersection of which this project stands, context and individual uniqueness become shaping factors. What does science have to say about how nuanced and personal felt experiences can be? To begin with, the main factor that makes it difficult to point out the clear emotion of an interactor is the complexity feelings hold. Humans themselves often struggle to obtain clarity on their own feelings and might even experience contradictory ones simultaneously. [Larsen et al., 2001] strengthens the claim that positive and negative activation are not always mutually exclusive poles, but they can hold space in someone's soul at the same time. [Berrios et al., 2015] also provides strong evidence that it is possible to experience mixed emotions, which discard previously believed narratives that such occurrences are simply errors in measurement or reporting. While we don't fully understand how or why this happens, their sole existence is proof enough that feelings form a much more detailed mesh than we sometimes realise.

Furthermore, emotion can change from one split second to another. Affective episodes are defined by multiple phases: the start, the rise, the peak, the maintenance and the recovery. These can be shortened or prolonged due to various factors from one's immediate context, which ultimately contribute to its perceived intensity. [Verduyn et al., 2009] highlights this dynamism that characterises the intensity trajectory of an emotion, concluding that the latter should not be described only by its peak or average strength. Specific emotions can also have different temporal signatures, such as steeper onsets, gradual builds or numbers of peaks. This paints a complex picture, one that leads to the realisation that attempting to point out someone's emotion at a given moment in time will rarely be fruitful, as changes can occur frequently.

Emotion is, at first glance, something that attains only to the soul. [Nummenmaa et al., 2014] however discovered that it exceeds internal boundaries and is embodied in patterns of bodily sensation or reflected in actions that can be observed. [Castellano et al., 2007] showed that even when humans do the same common gesture, their movement dynamics can still lead to detecting emotional differences. It becomes apparent that an action itself is not the sole indicator of emotion, but also how it unfolds. From this perspective, we can project the same understanding onto handwriting or drawing, which are embodied motor behaviours.

Speaking your feelings out loud is thought to bring clarity. Words can help people categorise and regulate their emotions, but works such as [Lieberman et al., 2007] uncovered that labelling affect can actually also alter an emotional response. What happens is that psychological or neural reactivity is reduced by using certain control systems that dwindle the activity in the amygdala and other related regions. It remains uncertain if language is the one that plays the biggest part in this reaction, but even if symbolic categorisation were more generally sound, the conclusion still stands: assigning an emotion to a label can shift someone's experience from emotional engagement towards a more rational state.

Thus, this project considers a less mediated form of expression that does not ask users to translate their emotions verbally, or in pre-defined symbols: free-form line drawings. From [Yang and Qin, 2021], we learn that prior work has looked at both digital writing and drawing tasks. The latter proved to bring more success to detecting anxiety and stress when combined with the first. For depression, drawing as the sole task led to the best results. This is a testament to the fact that opting for drawing tasks provides a better window into a user's emotional state, allowing them to be expressive in a flexible environment.

How open should drawing be to achieve best results? Many papers opted for a rather structured approach, prompting test subjects to draw specific shapes, copy them from a given example or complete already established tests such as HTP. In the latter, participants are asked to draw a house, a tree and a person, and features like their size, missing elements and other details are thought to reveal indicators of various emotions. Although it has been widely used in psychological research on connecting drawing to emotion, especially in children audiences, [Lin et al., 2022] debunks the idea that such tests can show stable relationships between drawings and mental health. Using several deep neural networks, they detected poor inferences on mental health diagnoses from HTP drawing samples. In trying to find a balance between rule and creative freedom, I turned to the idea of line drawings as a contained but still flexible task.

[Stamatopoulou, 2007] argues that lines are enough to communicate artistic meaning and embodied feelings, even if they don't represent something literal in their final form. In a more recent work, [Lin et al., 2024] reinforce this focus and show that when prompted to represent certain emotions through lines, motion features contributed more strongly than static ones. Many of these influential characteristics were frequency-based, which goes hand in hand with how [Stamatopoulou, 2007] takes a closer look at rhythm and treats it as no single variable but a connection between repetitions, regularity, temporal patterns and duration of the process. This calls for closer attention to the fact that in line drawings, emotional expression is not only communicated through the shape of a line, but also through the whole rhythm of motion. All in all, both papers amplify the idea that while static characteristics are relevant, combining how a line looks and the rhythmic movement of its creation captures a wider range of affective expression.

Movement is only one of the layers that can make affect perceptible in drawings. Previous work has also focused on visual features, and the most well researched link is, without a doubt, colour. Papers such as [Nijdam, 2008] and [Weng et al., 2024] uncover that anger is most often depicted using red, happiness is associated with yellow and black for fear. These correlations intuitively make sense, although they might be slightly stereotypical understandings of the affect a colour can carry. Some of these associations are very familiar to us, and we can turn to various art forms to understand them, one being the renowned animation "Inside Out", where each character represents an emotion and is given a main colour which mostly falls in line with the cases mentioned beforehand. Nevertheless, even from just these two papers, we already uncover differences in results: for [Nijdam, 2008], blue connects to calmness, while for [Weng et al., 2024], it means sadness.

By studying drawings created for six different emotions, [Damiano et al., 2023] showed that both computers and human observers identified associations such as green = disgust, based on the percentage of pixels filled with the specific colour. When comparing line drawings with colour drawings directly, the work shows that the conventional colour and emotional links are stronger than those between line and emotion, due to the fact that we are primed to learn them, like an "internal colour template".

[Weng et al., 2024] broadens the perspective by introducing more abstract colour associations, that look at hue, saturation, brightness and fill instead of only the colour itself. In addition to the familiar links detected in other works on the topic, they show that happiness is defined by high saturation and brightness, while fear is the opposite. The same goes for how much of the canvas is filled, with the highest percentages connecting to signs of fear or more negative states. Framing their findings on a valence-arousal model, they help provide a more complex palette of understanding how certain aesthetic choices made in drawings reflect a user's emotion.

The studies described above focus on complementary sources from which affective information can be extracted when looking at drawing, both as a process and as a result. Studies that focus on dynamic pen characteristics indicate that correlations exist between these and the drawer's emotional state. This conclusion is primarily based on each feature's contribution to classification tasks, rather than on consistent and direct correlations. However, some links could be identified or hypothesised based on research insights. Visual content oriented studies analysed colour choices and dimensions, placement and overall composition. Both interweave in our understanding of the connections between drawing and emotions and thus, will play a consistent part in the approach described in Section 3. Table 1, which can be found below, provides an encompassing overview of the relations that were most stably noted across referenced papers.

Table 1: Previously observed correlations between drawing features and affective state

Pen characteristic/ movement	Negative Activation	Positive Activation
Fast movement	x [1]	
Speed variation	x [2]	
Direction variation	x [3]	
Heightened pressure	x [4]	
Long strokes		x [5]
Upward strokes		x [6]
Dense coverage	x [7]	
Smooth lines		x [8]
Low azimuth, high altitude	x [9]	
High brightness, high saturation		x [10]
Dark palette	x [11]	
Greater colour-filled areas	x [12]	
High number of strokes		x [13]
Angular, sharp lines	x [14]	
Verticality	x [15]	
Curves		x [16]
Red	x [17]	
Yellow		x [18]

[1] Based on [Stamatopoulou, 2007] [Lin et al., 2024] [Santosa et al., 2026].
[2] Based on [Santosa et al., 2026] [Khan et al., 2024] [Lin et al., 2024].
[3] Based on [Lin et al., 2024].
[4] Based on [Yang and Qin, 2021] [Lin et al., 2024].
[5] Based on [Damiano et al., 2023].
[6] Based on [Lin et al., 2024].
[7] Based on [Damiano et al., 2023].
[8] Based on [Lin et al., 2024].
[9] Based on [Rahman and Halim, 2023] [Han et al., 2019].
[10] Based on [Weng et al., 2024].
[11] Based on [Weng et al., 2024].
[12] Based on [Weng et al., 2024] [Damiano et al., 2023].
[13] Based on [Santosa et al., 2026].
[14] Based on [Damiano et al., 2023].
[15] Based on [Damiano et al., 2023] [Rahman and Halim, 2023].
[16] Based on [Stamatopoulou, 2007].
[17] Based on [Damiano et al., 2023] [Weng et al., 2024].
[18] Based on [Damiano et al., 2023] [Weng et al., 2024].

In thinking about how these relationships should best be used to shape the interaction, the core idea of the study was also guided by [Ge et al., 2024], where the difference between explicit and implicit user feedback is noted. While the prior requires users to actively report their preferences, thoughts, evaluations, etc., the latter infers this feedback from the way in which the interaction unfolds. This allows for the flow to be kept intact, avoiding disrupting elements. Implicit feedback is often gathered in interactive systems from clicks, pauses or time spent on various steps. This project began from the curiosity of whether emotion could constitute a form of implicit input, gathered through the process of drawing using lines.

[Ball, 2002] provided a great introduction into artwork as an active process of change, especially in relation to emotions. It finds art art as something that should not be interpreted after its creation, but treated as something that takes part in an interaction by passing along feelings and structuring someone's behaviour. With a strong focus on art therapy, it opens up the discussion to understand art as a window of self-reflection. Ball stands for the idea that we should analyse artworks by also keeping in mind the small moments of the interaction where emotional patterns shift, before asking anyone to verbalise their feelings or make sense of them for themselves.

Interpreting art as a dialogue has long been backed within various fields of research. In Computational Creativity, some systems, defined as co-creative, are designed to collaborate with the user throughout the creation process. This, in turn, makes us understand art as a process rather than looking only at the product. A shining example in this sense is [Davis et al., 2015], a web-based drawing application, centred on the trade-off between a user's drawing and the system's response in return. In this case, the system waits until the user is done with their turn and then adds its own line onto the canvas. In correlation with Ball's perspective, explored above, this project takes interactive drawing one step further and decides to alter the look and feel of the user's lines as they are being created, before they are set on the sketchpad.

A good testament to the impact of real-time affective feedback is [Wang and Chen, 2020]. They observe brainwaves to understand if the subject is focused, relaxed, calm or anxious and represents these back to them through animation, colour palettes and sound. The paper concluded that participants felt the visual and audio feedback helped them recognise their emotional state or the changes within it. This directly inspired the set-up presented in this paper.

2.2 Methodological Approach

Following the established ideation of this project and the overall conceptual direction, the methodology aimed to preserve expressive freedom while allowing for a clear comparison between neutral and responsive drawing interfaces. Learnings regarding how to study this through experiments can be found in prior work as well. Between the papers discussed in the previous subsection, several recurring choices can be found for the methodology. The studies that aim to predict emotions first have to collect behavioural data and before doing so, many use external elicitation materials such as images, sounds or clips. The drawing tasks afterwards vary, and usually fall under one of these two approaches: relying on pre-defined tasks to capture felt emotions or providing the emotions outright and requiring participants to express them through free drawing.

From all papers, the one closest to the chosen methodology for this thesis project is [Han et al., 2019]. It aims to uncover whether a user's emotional state can be recognised from features of their digital writing and drawing. To achieve this, they set out three tasks for their participants: re-writing words, free doodling and writing simple, yet repetitive characters. Situated in a broader human-computer interaction framework, they envision systems that can grasp such implicit feedback without requiring self-reporting. Their aim however remains classification-oriented, with emotional states being organised in four quadrants, based on the high to low ranges of both valence and arousal. Regardless of the differences, the research does remain a handy reference for building the foundation of this study. It demonstrates that emotionally charged stimuli, qualitative reporting and tablet-based drawing can be combined in one work successfully.

Affective state, in their case, is induced through film clips selected from the Emotional Movie Database. According to [Gross and Levenson, 1995], such materials are a strong method to induce affective and arousal changes in subjects, but this approach does not come without risk. While many datasets with labelled movie clips exist, these snippets can be intense without succeeding in projecting a single, very specific emotion onto the watchers. With this caution in mind, the current study uses video clips to create an affective context in broader terms, aligning with the core goal to not attempt to name the user's emotion directly. In another more recent database, EmoStim, described in [Somarathna et al., 2023], we are presented with a publicly available inventory of movie clips, rated by 1567 participants along a 1-5 Likert scale. These labels cover fourteen discrete emotion categories extracted from the Differential Emotion Scale: fear, anxiety, anger, shame, warm-heartedness, joy, sadness, satisfaction, surprise, love, guilt, disgust, contempt, and calm.

The work presented in this paper adopts the strategies explored above, but slightly changes what is meant to be achieved in each. To ensure there is a good trade-off between structure and freedom within the drawing tasks, the experiment asks user to draw using only lines. Moreover, to refrain from labelling emotions but still be able to somehow compare results with literature-driven insights, video stimuli are used to establish a broad affective context. Last but not least, the stylus dynamics are used in the experiment itself rather than post-hoc. All of these considerations make up important aspects of the responsive sketchpad and provide guidelines for how the study should be structured. The remaining unsolved variable is how the design of the sketchpad itself can achieve all the goals that have been set out.

3 Responsive Sketchpad Design & Implementation

In this section, the creation process of the proposed responsive sketchpad will be explored. Chapter 2 provided a look into existing literature and summarised the correspondences that can be backed by theory. These have been integrated in the interface by mapping them onto visual responses shown to the users in real-time. The following subsections will describe the design process from the drawing board, to the integration of these relations and the technical implementation. The latter will include an exploration of both the visual interface that the users see and work with, and the data collection logic that happens behind the scenes.

3.1 Design Principles

The overarching principle behind this interactive interface was that the system should at no point attempt to tell the user how they are feeling or ask them to name their emotions. Instead, by changing the visual representation of their strokes, it takes a step back from the classification focus and moves towards a self-reflective, expressive environment. If the system were to point out that a user's stroke indicates a specific emotion, the risk is not only being inaccurate, but also re-shaping their own internal process towards reflecting on the provided "diagnosis" instead of their own emotional compass.

Secondly, a balance should be achieved between ambiguity and interpretability. The goal is to create visuals that feel matched to a user's inner state without claiming that they are equivalents of specific feelings. The idea behind this reasoning is to give participants a signal to be more aware of what within them prompts them to draw as they do and what meaning the responses might carry. They can, and are encouraged to, interpret these correspondences for themselves. The responses constitute visual or material changes that can suggest general affect activations, such as weight, saturation, roughness or floating.

Last but not least, the third goal is achieving a sense of dialogue between the user and the sketchpad, a collaboration of sorts. The sketchpad should not feel overwhelming or hindering and users should find their own agency in the process in order to keep closely in touch with their emotions. Since the sketchpad is designed to alter the appearance, movement and material behaviour of the strokes as they are still being created, there is a fine line between understated and exaggerated. In case the responses are too subtle, reflection might fail because a user doesn't register them. In the opposite scenario, they may shift the interactor's attention away from their own emotional state towards trying to understand the behaviour of the system instead. Thus, the design decided upon is that effects will show as the user is moving the pen across the tablet surface; the most disruptive ones, those based on animations, will immediately settle upon release, while softer ones, such as hue changes, will remain as a memory of the response. I do not assume that these approaches are automatically beneficial. The aim remains to investigate if a trade-off between control and surprise or reflection and curiosity can be achieved through these decisions.

3.2 Mapping Exploration

In order for the responsive sketchpad to work as conceptualised, it required a mapping layer, which is driven by insights uncovered from prior research on associating pen movement, colour characteristics and affective activation, as well as my personal understanding of them. The sketchpad doesn't decide on whether the user experiences negative or positive activations, but it instead transforms the visual and material qualities of their drawing in a way that supports their self-reflection.

As was previously explored, negative and positive emotions can be experienced simultaneously. Because of this, although the distinction made between positive and negative activation based on background research remains relevant to the design, the grey area of the interplay between both remained a question that the mapping system needed to address. The identified solution was separating two conceptual layers: a local and a global one. The first holds both colouristic and kinetic mappings, some with directly inferred properties from Table 1 and others based on HCI oriented reasoning around expressive motion and ambient feedback.

The second does not focus on one stroke at a time only, but on the way the over-time progression of the drawing process reflects meaning. Based on the idea that the rhythm of the process can be more important than the visual form or individual characteristics, this layer is meant to look at the session until that point and establish activation as a whole, posing curiosities such as how active, continuous or intensive it appears to be. This is all generalised under the umbrella term of "pacing" and affects the general atmosphere of the sketchpad. This separation was a way to consider the complexity of features such as speed, which are known in research to mirror negative activation but might also simply mean that a user is engaged and excited about the interaction. Thus, a fast stroke itself is mapped onto a different visual response, a local one, while its impact on the overall rhythm leads to a global one. Put together, they ensure that the sketchpad reacts to the moments that are affectively relevant without confining them to a single line of judgement.

Table 2 provides the summary of mappings used in the mechanics of the sketchpad. Columns 1 and 2 are directly taken from Table 1. Thus, insights from prior work in the field are linked based on their activation family: negative or positive. The ones most applicable to basic drawing movements are kept as input, while the more abstract ones are used as a visual output, which allows for correspondences to be as grounded in literature as possible.

Table 2: Mappings between stylus movement and sketchpad response informed by prior work

Input: Pen dynamic features	**Output: Research derived qualities**	**Additional visual design translations**
Fast movement	Direction variation	Intensified vibration, pulse and recoil effects
Speed variation	Angular, rough strokes	Vibration-like effect
Heightened pressure	Low brightness, Red hue/ dark palette, Greater colour-filled areas	Dripping effect
Long strokes and upper-oriented strokes	High brightness, Yellow hue, Smooth strokes	Halo effect, Floating effect
Dense coverage	Low brightness, roughness	Near-black hue, Break-free effect
Verticality	Low brightness	Shadow echo

Column 3 describes the animations layered with the stroke alterations described in Column 2, to add dynamism and heightened interactivity to the platform. These effects don't add new meanings, but are simply behavioural translations of the prior work derived outputs. By giving the sketchpad a more "alive" nature, these moments of noticeable feedback are meant to support affective reflection during the interaction. A more detailed description of the idea behind each of them can be seen in Table 3.

Table 3: Stroke behavioural effects

Effect	**What pen characteristicsdoes it translate?**	**Expressive qualities**
Vibration	Speed and direction variation	Instability
Halo	Brightness	Openness
Floating	Long, smooth, upper strokes	Flow
Break-free	Dense coverage	Resistance to accumulation
Recoil	Pressure and fast movement	Push-back
Dripping	Intensified pressure	Material saturation

Tables 2 and 3 cover the concept behind the mappings that stand as pillars of this research. The following formulas describe the implementation logic. Beginning from the basics, whenever a participant draws, the system logs a sequence of pointer samples containing the following details: position on the canvas, both horizontal and vertical; timestamps; pressure; tilt; pen status and estimated area of contact with the screen surface; the point in the stroke creation when the sample is recorded (start, middle, finish).

The simplified representation of a stroke focuses only on the core values that play into the effect calculations:

$$s_i = (x_i, y_i, t_i, p_i) \tag{1}$$

where x_i and y_i are the sample position, t_i is the timestamp, and p_i is the pressure value.

This allows the system to analyse the drawing process based on temporal sequences of movement, spatial relations, pressure and pauses. The stroke that is actively being drawn is continuously evaluated, reshaped and shown to the user, with the end of this cycle being represented by the lift of the pen off the screen. Using these raw sequences, pen features used as input can be calculated:

- fast movement →mean speed across consecutive stroke samples
- speed variation →standard deviation of local speed values
- heightened pressure →mean and standard deviation of pressure
- long strokes →stroke path length
- verticality →relative vertical extent of the stroke's bounding box
- upward trend →upward start-to-end displacement relative to stroke path length

These are not interpreted in isolation, but rather compared to a baseline that is established to be tailored to the user's natural drawing rhythm. At the start of a session, before either of the drawing conditions takes place, the system takes 25 seconds to calibrate. In the responsive sketchpad, every stroke x is evaluated compared to the calibration levels b_x using relative deviation. Depending on whether the specific feature is higher or lower than the unique baseline, responses are triggered.

One restriction is applied on how many effects can show at a time in order to keep the canvas from being too overstimulating. Stroke features are converted into response scores, which represent the strength of a possible visual response. Each possible effect has a score S_i, a threshold θ_i, and a priority weight q_i. There is a separate score for each effect, and those that exceed their thresholds become candidates for the trigger set. Three response categories are selected:

$$T = top_3\{e_i \mid S_i > \theta_i\} \tag{2}$$

ranked by:

$$S_i \cdot q_i \tag{3}$$

where T is the set built of final active triggers for a stroke

As mentioned, motion-based transformations are transient. The strokes themselves, however, remain visible, but in order to avoid certain effects connected to overlapping strokes or canvas fill to trigger erroneously, they have a set lifespan, controlled by their "age":

$$age = t_{now} - t_{created} \tag{4}$$

Between 45–135 seconds after their creation, they gradually fade away until they reach 25% opacity. Due to the colour choices made, that was found to be just enough for the memory of the interaction to remain while visually opening the canvas back up.

Throughout the following equations, unless another range is stated explicitly:

$$clamp(x) = \min(1, \max(0, x)), \tag{5}$$

The output qualities presented in Table 2 are calculated as follows:

1. Direction variation:

$$\text{directionChangeRate} = \frac{\text{directionReversals}}{\max(1, N)} \tag{6}$$

where N is the number of samples in the stroke

2. Stroke angularity:

$$\text{angularity} = clamp\left(\frac{\text{curvatureMean}}{1.1} + 2.2 \cdot \text{directionChangeRate}\right) \tag{7}$$

where

$$\text{curvatureMean} = mean(\Delta\theta_i) \tag{8}$$

3. Dense coverage:

$$\text{density} = \frac{\sum_{c \in C} \sum_{o \in O} hit(c, o)}{\max(1, |C|)} \tag{9}$$

where C is the set of current stroke samples and O is the set of samples from recent previous strokes; an overlap score is extracted from the density value:

$$S_{\text{overlap}} = clamp\left(\frac{\text{density}}{5.5}\right) \tag{10}$$

and a hit is defined as:

$$hit(c, o) = \begin{cases} 1, & dist(c, o) < r \\ 0, & dist(c, o) \geq r \end{cases} \tag{11}$$

4. Smoothness:

$$\text{smoothVerticalMotion} = clamp\,(\text{verticality} \cdot \text{directness} \cdot (1 - \text{angularity}) \cdot 1.4) \tag{12}$$

5. Directness:

$$directness = clamp\left(\frac{dist(s_1, s_N)}{\max(1, L)}\right) \tag{13}$$

where s_1/ s_N are the first/ last samples of the stroke and L is the total stroke path length

The dynamic effects presented in Table 3 are calculated as follows:

1. Vibration:

$$S_{\text{vibration}} = clamp\left(0.36\Delta^+_{\text{speed}} + 0.28\Delta^+_{\text{speedVariation}} + 0.30 \cdot \text{directionChangeRate} \cdot 3\right) \tag{14}$$

$$\text{vibrationAmount} = 6.4 \cdot S_{\text{vibration}} \tag{15}$$

2. Halo:

$$S_{halo} = clamp\left(0.32\Delta^+_{length} + 0.32 \cdot directness + 0.20(1 - angularity) + 0.18 \cdot upwardTrend\right) \tag{16}$$

$$\text{haloAmount} = 8.5 \cdot S_{\text{halo}}. \tag{17}$$

3. Floating:

$$S_{pressure} = clamp\Bigg(0.45\Delta^{+}_{pressure} + 0.30\Delta^{+}_{pressureVariation} + 0.28\,clamp\left(\frac{meanPressure - 0.50}{0.32}\right) + 0.10\,clamp\left(\frac{density}{5}\right)\Bigg) \tag{18}$$

$$\text{floatAmount} = 4.2 \cdot S_{\text{smoothFlow}} \tag{19}$$

4. Break-free:

$$overlapBreak = (\text{noisyOverlap} \in T) \vee (\text{roughnessAmount} > 7) \tag{20}$$

$$breakFreeScale = \max\left(0,\ 1 - \frac{age}{2600}\right) \tag{21}$$

5. Recoil:

$$\text{forceImpact} = clamp\left(0.55\Delta^{+}_{\text{pressure}} + 0.45\Delta^{+}_{\text{speed}} + 0.22\Delta^{+}_{\text{speedVariation}}\right) \tag{22}$$

$$S_{\text{recoil}} = clamp\left(\text{forceImpact} \cdot (0.45 + 0.55\Delta^{+}_{\text{speed}})\right) \tag{23}$$

$$\text{recoilAmount} = 8 \cdot S_{\text{recoil}} \tag{24}$$

6. Dripping:

$$S_{\text{pressure}} = clamp(0.45\Delta^{+}_{\text{pressure}} + 0.30\Delta^{+}_{\text{pressureVariation}} + 0.28 \cdot clamp\left(\frac{\text{meanPressure} - 0.50}{0.32}\right) + 0.10 \cdot clamp\left(\frac{\text{density}}{5}\right)) \tag{25}$$

$$\text{dripAmount} = 18 \cdot S_{\text{pressure}} \tag{26}$$

The constants mentioned in the formulas that are presented above have been settled on throughout the design iterations and their role is to attempt balancing how visible each reaction can be while still being visually restrained. Some effects, as you may have noticed, combine multiple stroke features. To determine how much control each has over the final effect, numerical weights are added. It is important to note, however, that these values come from fine-tuning over iterations and are not backed by scientific literature and have not formally been evaluated before the experiment was run.

3.3 System Architecture and Data Logging

The responsive sketchpad was implemented as a web-based application, coded using React and TypeScript. The system can thus run smoothly in a browser, with all the real-time events required, and does not call for any additional software to be installed by the users. Its main modules separate study-state control, stroke analysis, rendering, stimulus selection and data logging.

The latter is essential to the project for it to remain an instrument of research and not only an interactive experience. Since the plan was to have users interact with it from their own devices, safely storing all pen-related data recorded during each interaction session was crucial. On top of this data, the aspect that users can access the application remotely also called for a way to gather insights into device diagnostics. Consequently, a small set of such information is recoded, most importantly pointer types used in the session and available pressure range. The storage system used to achieve this overarching aim was Firebase Firestore.

Each time a user completes the experiment, a unique session identifier is generated. Individual marks are also distinguished within each session by using stroke IDs, and a pseudonymous participant ID is retained to link all elements of the study, which will be further detailed in Section 4. As a complete session results in many values, storing it as a single file proved to not only be fragile, but also to exceed the available limits of the platform. Thus, a workaround needed to be found. Thus, the recorded events are serialised and metadata is uploaded in a session-level document, while sample-level records is saved in two sub-collections: jsonChunks and eventChunks. The first divided the full JSON representation in smaller batches, while the second already formats relevant data, to make inspection easier.

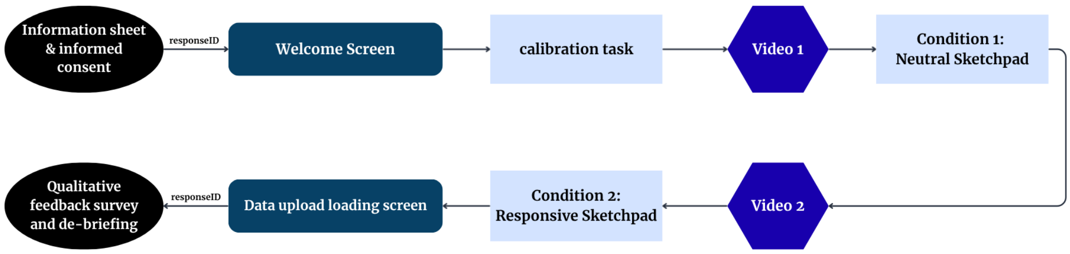
The participant ID was an important consideration within the architecture of the sketchpad and its integration in the experiment. To later link the app with the surveys, a pseudonymous response identifier is received from Qualtrics, by adding it as a query parameter to the URL that re-directs the user to the web application. Deployed using GitHub and Vercel, the app extracts this value and logs it in the database, thus allowing connections to be made between the surveys and the experiment data. Nevertheless, the email which connects this randomly generated string to a person is stored only in Qualtrics and never gets sent to Firebase.

4 Experiment Methodology

This section focuses on the full protocol, starting with the recruitment of users and following the pipeline all the way to the final survey design. Firstly, this study is a within-subject experiment, as all users interact with both the neutral and the responsive sketchpad. The first is comprised of a common canvas, similar to any other digital drawing tool. Although it records the drawing movement, it does not interfere with the visuals at all throughout the task duration, which the responsive sketchpad does through mappings explored in Section 3. The order of the conditions is always the same: the neutral one first, and the dynamic one second. The comparison between the two aims to examine how they differ in the way they support emotional reflection.

Participants were recruited online, using WhatsApp Groups targeted at university students or teachers and Reddit forums for survey swapping or digital art enthusiasts. The eligibility criteria were having access to a pen-based tablet to complete the experiment on, regardless of its brand or model and speaking English fluently in order to complete the survey and be able to give informed consent. Participants are not required to have any drawing experience, and more varied demographics are welcome in order to provide a more complete outlook on correlations. The interface highlights that artistic skill or aesthetic quality are not important and participants are repeatedly encouraged to just draw whatever comes to mind. The flow of the experiment is described in detail in Figure 1.

Figure 1: Experiment Setup



When launching the experiment, participants received an information sheet and were required to provide informed consent through Qualtrics. The survey data and interaction records were linked through the pseudonymous participant ID described in Section 3.4. Afterwards, broad instructions are presented in the welcome screen, refraining from detailed explanations regarding the scope of research, so as to avoid biasing users before the task starts. Then, the calibration needs to be completed, which takes 25 seconds, during which the user's typical drawing behaviour baseline is established. Then, the two drawing conditions follow.

Before each of them, a video stimulus is shown to the user, who is invited to immerse in the overall atmosphere of these film clips, and can only move to the next step after the video finishes playing. These two videos have been extracted from the EmoStim database, proposed by [Somarathna et al., 2023], and have been chosen based on the following three rationales: they should represent a generally positive and negative activation; they should not exceed 2 minutes; they should be ethically appropiate and contain no graphic or traumatic content. The final selected film clips come in the form of a snippet from "The Pianist" for the negative condition (a higher average score for negative affect: $M = 2.78$ and a low average score for positive/calm affect: $M = 1.44$ and a montage of Planet Earth videos with relaxing background music (high average score for positive/calm affect: $M = 3.50$ and a low average score for negative affect: $M = 1.41$). To understand drawing behaviour on both sketchpads in a more thorough manner, these two videos are assigned in a changing order, so the same affective context doesn't always precede the same drawing task.

Both the neutral and the responsive sketchpad interactions are timed, and a maximum of three minutes is offered per task. Users can stop the process and move on to the next step whenever they feel comfortable within this time limit. After the web application part of the study is completed, participants have to wait for the data to be successfully uploaded and are then redirected once more towards Qualtrics, along with their unique participant ID.

The questionnaire is meant to collect qualitative insights into the interaction experience, alongside demographic characteristics such as age range, tech-savviness, dominant hand and digital drawing experience. The main focus of this survey is to evaluate both sketchpad conditions by answering a set of identical, directly comparable questions. These revolve around how much users noticed their emotional state, whether they felt supported in self-reflection, how reactions influenced their drawing decisions and how they perceived each sketchpad. Additionally, for the responsive sketchpads, a couple extra questions are posed regarding the effects and what they meant to each participant. The survey also assesses the participant's understanding of the video stimuli they were exposed to in order to see if the labelled affective context fits the perceived one. Questions are both open ended and fixed, in forms of multiple choice or five-point Likert scale.

At the end of the process, a debriefing is presented to participants, where the more in-depth research goal is presented and users are again made aware of their right to withdraw their data within the first month. After this is completed, all data is safely stored for analysis.

5 Results

This section serves as a deeper dive in the data gathered from the interaction experiment and survey study. Analysed both independently and in correlation, these findings aid the understanding of recurring themes, notable moments and meaningful patterns across the drawing experiences. Additionally, results will be weighed against prior literature knowledge in order to examine where they converge, or which existing claims about drawing features, visual qualities or affective expression can be challenged.

5.1 Participants

A total of 16 unique participants completed the study, out of which 12 identify as female and 4 as male. The largest age group represented was 25-34, characterising 8 of the participants (50%). 5 participants were aged 18-24 (31.3%), 2 others 35-44 (12.5%) and one was aged 45 or older (6.3%). All of the interactors described themselves as right-handed.

The survey captured the frequency with which testers draw on a normal basis. 6 participants reported doing so a few times per year (37.5%) and another 6 a few times per month (37.5%), marking the most common trends across the sample. 2 participants reported drawing a few times per week (12.5%). 2 respondents positioned themselves at the two extremes: 1 indicated that they drew almost daily (6.3%) and 1 described this habit as almost never occurring (6.3%). Taking this a step further, to professional experience or formal training, 11 participants reported having none (68.75%). 2 participants indicated that they had some experience (13.3%) and 3 are trained professionals (20.0%).

Another important consideration in the demographic analysis was the users' familiarity with tasks such as drawing or even writing on a digital device, as this can influence the confidence in movement across such a surface. The survey results uncover varying levels for this metric, which was measured on a scale from 0 to 4. The most common rating was 4, reported by 5 participants (31.3%), followed by 3, indicated by 4 others (25.05%). 3 participants selected ratings of 2 (18.8%) and 1 (18.8%). Lastly, 1 participant selected 0 (6.3%). Looking across the entire sample, the mean rating stands at 2.56 out of 4, which suggests a moderate familiarity.

Although the study was designed mainly for stylus-based interactions, 6 participants ignored the indication and completed the experiment using different input modalities (touch or mouse). Such situations have been anticipated, and fallback mechanisms have been implemented to ensure that the study can run smoothly even when finer details cannot be tracked. In these situations, the system substitutes default values: 0.5 for pressure and 0 for tilt. These cases will be evaluated cautiously and included only in the analysis of relevant, input-independent features, and filtered out for others. This allows the following analysis to preserve as much of the interaction data as possible, as it can still contribute to a fuller understanding of the interaction experience. Out of the remaining 10 participants, 4 used Samsung tablets and 6 used iPads.

5.2 Users' Perception of Video Stimuli

To examine how the two video stimuli impacted the user's emotional charge and how they were overall interpreted, they were asked to describe their perception of each video. Table 4 provides an overview of the recurring themes identified in the open-ended responses.

Table 4: Affective themes identified in participants' responses

Stimulus	**Affective Theme**	**Key words used in survey responses**
Negative	Anxiety and tension	apprehension, fearful, worried, agitated
	Sadness and vulnerability	sad, devastated, powerless, lonely
	Heightened arousal	shocked, dynamic, loud, energetic
	Reflective responses	fragility, loss, the darker aspects of humanity
Positive	Calmness and relaxation	calming, peaceful, relaxed, content
	Happiness and gratitude	happy, grateful, appreciation of nature
	Immersion	nostalgic, dreamy, hopeful, desire to travel and experience nature
	Reflective responses	sadness concerning environmental deterioration
	Mixed responses	a bit bored, puzzling, nothing in particular

Generally, the responses imply that the movie clips elicited contrasting affective experiences, falling under the two intended umbrella affective states: positive and negative. However, the variation in the provided descriptions highlights that the videos touched upon broad enough content to still allow unique interpretations and even spark more existential reflective experiences in some cases.

5.3 Drawing Movement Metrics

Throughout the interaction part of the experiment, 17 sessions were collected. One participant completed the interaction study twice, and both sessions were examined for the purposes of this section, as it focuses on drawing behaviour independently. For participant-level analyses, only the user's first completed session was retained.

A total of 2397 strokes were logged across the three drawing tasks: calibration, neutral sketchpad and responsive sketchpad. Table 5 summarises general drawing behaviour across the last two and focuses on a selected set of variables, which were chosen for the interpretability of drawing extent, interaction pacing, mark characteristics and spatial placement. Figure 5 in the Appendix shows these differences plotted across the two sketchpad conditions for each individual session.

While within the responsive sketchpad, users generally drew fewer strokes, these were more extended. The rhythm of drawing is visibly different between the two conditions, and the responsive interface produced more variation in speed and pauses before new strokes. More shape-related measures showed a similar tendency, but less uniformly. Local density was lower for the responsive sketchpad, meaning that interactors did not place marks in the same areas as much as in the neutral one.

To understand how the video stimuli impacted the behaviours described above, analysis was run on the movement qualities recorded after each video was shown to the participants.

Table 5: Descriptive movement summaries for the neutral and responsive sketchpads

Metric	Neutral M (SD)	Responsive M (SD)	Mean difference	Higher in responsive
Number of strokes	60.82 (59.68)	50.12 (36.75)	-10.71	8
Stroke length (CSS px)	219.48 (143.59)	241.90 (175.11)	22.42	9
Pause before stroke (ms)	1018.36 (586.72)	1483.07 (721.59)	464.71	14
Total path length (CSS px)	9523.13 (5075.05)	11076.09 (8111.75)	1552.96	7
Speed (CSS px/ms)	0.44 (0.28)	0.51 (0.25)	0.07	14
Speed variation (CSS px/ms)	0.27 (0.16)	0.33 (0.20)	0.06	12
Local density (hits/sample)	13.69 (17.01)	6.24 (4.91)	-7.44	6
Bounding area (CSS px^2)	15427.59 (21407.98)	14782.79 (20148.26)	-644.80	9
Pressure	0.12 (0.14)	0.13 (0.15)	0.01	9
Direction reversals (count/stroke)	1.26 (1.30)	1.83 (2.21)	0.57	12
Angularity (0-1)	0.30 (0.11)	0.37 (0.17)	0.07	11
Curvature (rad)	0.27 (0.11)	0.29 (0.12)	0.02	11
Straightness (0-1)	0.73 (0.18)	0.69 (0.20)	-0.04	8
Verticality (0-1)	0.49 (0.15)	0.53 (0.07)	0.04	12
Smooth vertical motion (0-1)	0.35 (0.18)	0.33 (0.17)	-0.02	7
Upward trend (0-1)	0.12 (0.12)	0.09 (0.06)	-0.03	10

The selected stroke features which contribute to the response mechanism were standardised across the dataset using z-scored, and averaged into a composite score where multiple metrics are used in its definition. The mean standardised scores for both stimuli and their difference can be found in Table 6. The biggest differences can be observed for verticality, angular lines and direction variation. Almost no difference occurred for fast movement or long strokes.

Table 6: How participants drew after each video stimulus

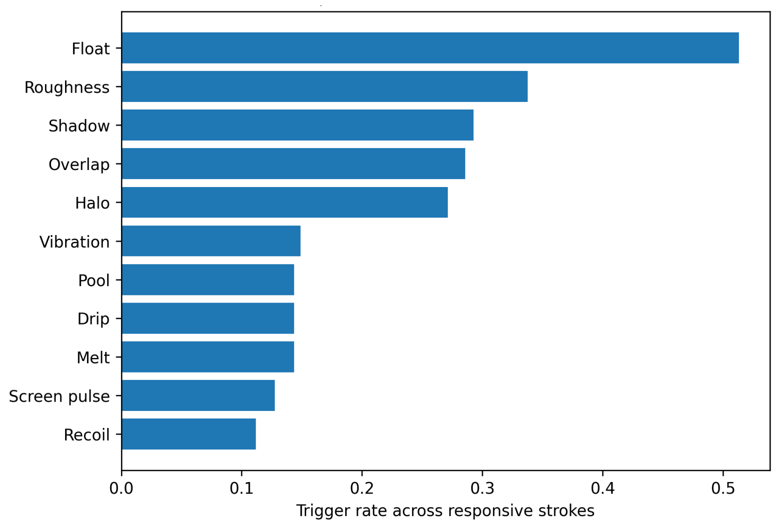
Drawing Feature	Positive M (SD)	Negative M (SD)	Neg. – Pos.
Fast movement	-0.09 (0.56)	-0.10 (0.44)	-0.01
Speed variation	-0.13 (0.58)	-0.07 (0.45)	0.06
Direction variation	-0.09 (0.36)	0.20 (0.62)	0.29
Heightened pressure	-0.41 (0.74)	-0.29 (0.80)	0.11
Dense coverage	-0.29 (0.21)	-0.15 (0.46)	0.14
Angular/sharp lines	-0.12 (0.41)	0.11 (0.58)	0.23
Verticality	-0.20 (0.33)	0.15 (0.41)	0.35
Long strokes	0.11 (0.39)	0.15 (0.54)	0.04
Upward strokes	-0.03 (0.22)	-0.11 (0.32)	-0.08
Smooth lines	-0.01 (0.36)	-0.16 (0.49)	-0.14
Curves	0.04 (0.18)	-0.06 (0.22)	-0.09
High number of strokes	-0.05 (0.96)	0.05 (1.07)	0.10

How does this all reflect in the visual responses provided by the responsive sketchpad? Based on the tracked data, an overal effect activation rate of 0.96 was measured, meaning that almost all strokes produced some sort of interactive mapping. Furthermore, in each session, 7.92 effects were triggered on average. Based on these results, it can be confidently stated that the responsive sketchpad was highly active, and exposed participants to many forms of visual feedback during the interaction. The distribution of these can be seen in Figure 2.

5.4 User Experience

When looking at the qualitative feedback independently, the main aim is to better understand how the responsive sketchpad, in comparison to the neutral one, supported emotional

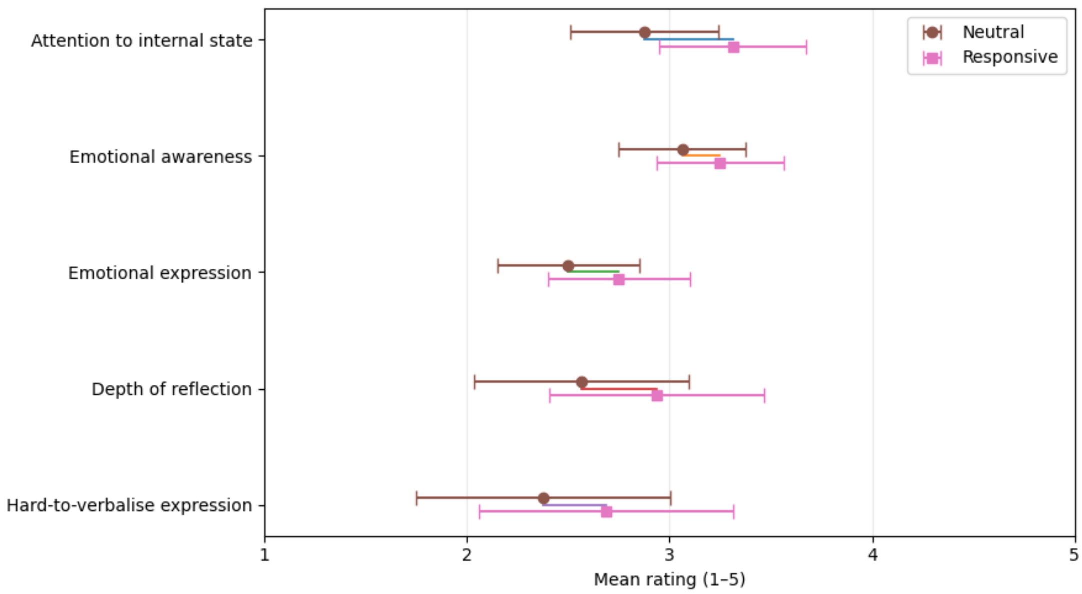
Figure 2: Detailed overview of triggered visual effects



self-reflection, how the users perceived the interaction and what interesting patterns can be observed across the open responses.

In the questionnaire, after drawing on each sketchpad, the users were asked to quantify their emotional self-reflection on a scale from 1 to 5. The questions looked at this through various lenses, and inquired about how much attention participants paid to their internal state, whether drawing increased their emotional awareness, how well they were able to express their feelings through drawing, how deeply they reflected on those feelings, and whether the sketchpad helped them express something that would have been difficult to put into words. A comparison between the two sets of results can be observed in Figure 3. Across all five questions, the responsive sketchpad scored higher, with the largest differences identified for attention and depth of reflection. Of these two measures, depth of reflection showed greater variability across participants.

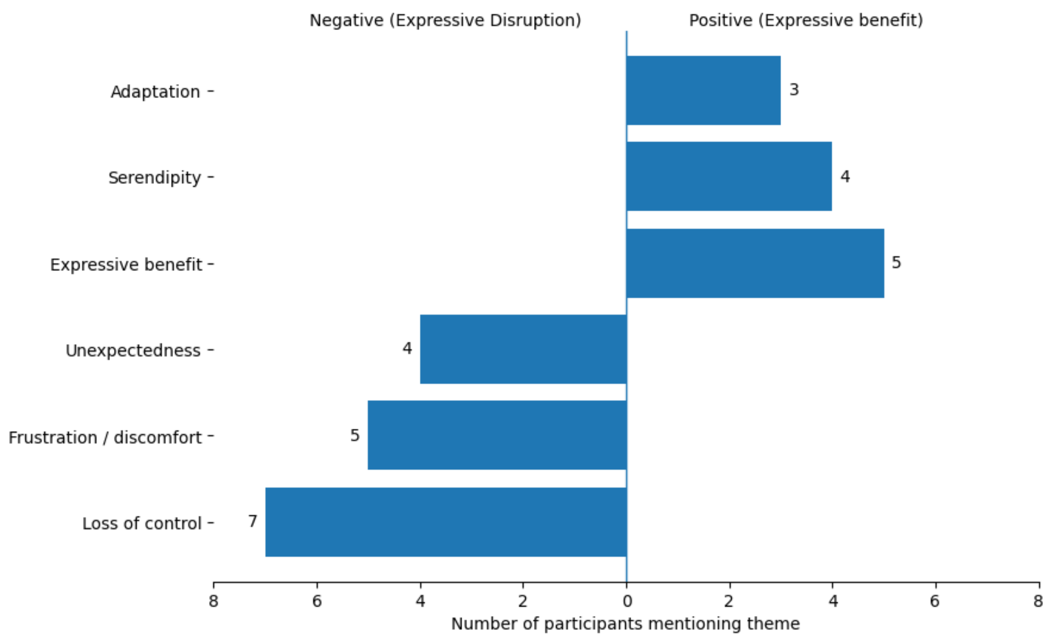
Figure 3: Affective reflection during interaction



Moreover, regarding the responsive sketchpad, one could more confidently state that it affected the trajectory of the drawing process than that it revealed previously unnoticed emotions in users. Responses for the latter were more mixed, with $M = 2.63$, while the prior achieved a higher rating of $M = 3.50$.

Users were also asked to provide their own interpretation of the second drawing condition, focusing on mismatches between felt emotions and visual responses, and especially meaningful moments that stood out to them. The responses were mixed and suggest that there is currently a tension between the need for agency and the new expressive opportunities that the sketchpad introduces in the interaction. A more detailed analysis of recurrent themes within answers can be seen in Figure 4.

Figure 4: Expressive experiences when interacting with the responsive sketchpad



This finding is consistent with responses that directly relate to the role that users assigned to each of the sketchpads. The first was most frequently described as controllable, namely by 11 participants (68.75%). In the open descriptions, this drawing condition was said to be nice, familiar and an overall smooth experience. The responsive sketchpad, however, was primarily characterised by the word "distracting", occurring in 9 responses (56.25%), accompanied by two different interpretations. Some found the visual transformations applied to their strokes to be frustrating and unpredictable, while others found this trade-off between agency and responsiveness to have generative outcomes and reflect the emotions felt after watching the video more accurately. In direct comparison, the second sketchpad was slightly more often described as collaborative than the neutral one.

5.5 Relations between individual context and drawing movement

This part of the paper focuses on describing the most visible links found between drawing metrics and qualitative feedback for the 16 users. Firstly, drawing experience was examined as a factor that could influence pen movements. Participants with some or professional experience tended to draw longer strokes on the responsive sketchpad in comparison to those with none. Experienced participants appeared to use less direct strokes, which implies a more experimental approach, more exploration of the canvas overall. However, this change was also visible in the first drawing condition, so it reflects more on drawing style than it does on differences caused by the two sketchpads. The rate at which visual responses were activated during the interaction was very similar across the groups.

It was previously established that the responsive sketchpad was not primarily experienced as a system that revealed hidden emotions, but rather as one that could interrupt, redirect, or complicate the drawing process. This can bring interesting insights to center stage, especially ones related to co-creation or reflection-in-action. Relevant links in this sense, as well as with self-reflection ratings, are explored in Table 7.

Table 7: Most relevant correlations identified between user answers and stylus metrics, with a focus on reflection and reconsideration. Increases are calculated as responsive minus neutral.

Survey measure	Stylus / system metric	Spearman ρ	p
Reflection rating increase	Direction reversals increase	.63	.008
Reflection rating increase	Any-effect activation rate	.49	.053
Pause/reconsider/change rating	Direction reversals increase	.42	.108
Pause/reconsider/change rating	Recoil activation rate	.36	.168

The clearest association found between the individual ratings of users and their subsequent movement data appeared between the increase in the reflection-related ratings and a higher number of direction reversals. A weaker link was found between the same survey measure and the activity of the second sketchpad interface overall, which is calculated as the proportion of responsive strokes that triggered at least one visual effect. When it comes to the experience of pausing/ reconsidering, contrary to expectations, the relation with literal pause duration was weak. Participants who gave high ratings for pausing, reconsidering, or changing direction did nevertheless show a larger average increase in pause duration than those who gave lower ratings (+693.28 ms compared with +266.18 ms).

Shifting the focus towards sessions where participants reported thinking the responsive sketchpad was distracting or uncontrollable, Table 8 establishes the correlation between these responses and the movement metrics. The strongest pattern can be found between sessions in which users expressed uncertainty or randomness in visual responses and increases in direction reversals and direction-change rate. Links with angularity are uncovered as well. When the sketchpad was defined as distracting, higher speed variation and vibration activation rate were found in the stylus data. Last but not least, loss of control can be connected to lower directness in stroke paths.

Table 8: Most relevant correlations identified between user answers and stylus metrics, with a focus on lack of control and disruption. Values show means for sessions in which the participant report was present or absent. Differences are calculated as present minus absent.

Report	Interaction metric	Present	Absent	Diff.	p
Uncertainty/randomness	Direction reversals	1.41	-0.48	+1.89	.031
Uncertainty/randomness	Angularity	0.13	-0.01	+0.13	.031
Uncertainty/randomness	Direction change	0.04	0.01	+0.03	.042
Distracting	Speed variation	0.13	-0.02	+0.15	.031
Distracting	Vibration rate	0.22	0.07	+0.15	.057
Loss of control	Directness	-0.10	0.00	-0.10	.091

In Figure 2, an overview of effect occurrences is provided. In the questionnaire, users were asked to name any visual response that stood out to them. The correlation between the amount of times an effect appeared on screen and participants' recall of it can be found in Table 9. This allows us to understand if recollection is tied directly to exposure, or if some responses simply leave a mark. The results suggest that the latter is true for some, such as vibration and roughness. which proved to be more salient than their frequency alone would suggest. Even though they both appeared proportionally less than other effects during the sessions where they were triggered, 5 and respectively 4 users pointed them out as especially noticeable. Drip occurred in fewer sessions and also had a very low rank when it did, and took third place in users' memory nevertheless. The opposite pattern can be seen for float, which was definitely the most frequent and appeared in all sessions, but was only mentioned by two users. Overall, the table suggests that participants remembered effects not only because they occurred often, but because they changed the felt quality of drawing in noticeable ways.

Table 9: Relationship between effect exposure and reported salience

Effect	**Sessions where triggered**	**Median rank (1-7)**	**Users who pointed it out**
Vibration	10	7	5
Roughness	16	4	4
Drip	9	7	3
Float	16	2	2
Pulse	15	6.5	1
Halo	16	4	1
Overlap	16	3.5	0
Shadow	16	5	0
Recoil	11	7	0

Last but not least, participants were invited to describe the level to which their emotional state oscillated during the experiment. Metrics gathered from the participants whose answers fell either under "stable" or "changing" as broad categories were compared, and can be found in Table 10. A changing emotional state would, at first glance, be associated to a more interrupted drawing flow. Contrary to this expectation, results show that this user group did not have higher pause durations or speed variations. However, the effect diversity that they were exposed to was slightly higher.

Table 10: Session-level differences by self-reported emotional stability. Differences are calculated as changing minus stable.

Metric	Stable M (SD)	Changing M (SD)	Difference	p
Pause change	907.61 (585.61)	446.64 (782.19)	-460.96	.374
Speed-variation change	0.18 (0.24)	0.00 (0.13)	-0.18	.304
Effect diversity	7.25 (0.96)	7.90 (1.20)	+0.65	.303

This analysis was done at session level (n=16) and stroke features were aggregated. Spearman correlations where used to identify associations between survey responses and drawing metrics that changed together, in the case of Table 7. Mann–Whitney U tests were used for Table 8 and Table 10 when the analysis compared two independent groups of sessions, with small and uneven populations, created from survey responses. The statistical analyses were meant to identify descriptive patterns that indicate possible links between interaction behaviour and user experience, but cannot be considered generalisable due to the small sample size.

6 Discussion

This study primarily aimed to investigate what happens when a responsive sketchpad interface does not ask for explicit emotion-related input, nor does it attempt to classify implicit input in well-defined categories. Instead, it makes traits of a user's own drawing behaviour perceptible to them through transforming strokes in real-time throughout the drawing process. What results point towards is that this shift in design rationale changed how participants engaged in the process of drawing: emotional reflection appears to have emerged less from the accuracy of an inferred emotional representation and more from the dialogue created between action and response.

6.1 Supporting Emotional Reflection

The first research question targeted affective self-reflection in users and how such an interactive system could support it. Results show that in comparison to a neutral sketchpad, the implemented prototype succeeded in heightening attention to one's internal state, overall emotional awareness, expression and depth of affective reflection. This effect was not large, but it was consistently present, meaning that the sketchpad made emotional experience more central during the drawing process. A personal hypothesis when setting out to do this research was that using such an interface would help uncover emotions that were previously not noticed. However, the results prove that the stronger claim is that such an interaction makes people pause, reconsider and adjust their drawing. Consequently, reflection can be understood as something that is not sparked by the system's ability to translate emotion into visual responses, but rather in the participant's response to the changing visual feedback.

From the movement data, we gather that when using the responsive condition, participants tended to pause longer before starting a new stroke. There were also more reversals in direction, heightened angularity and speed variation. Even though fewer strokes were created, they were longer. Altogether, this data indicates that there was less automatism in the interaction and more negotiation between the sketchpad and the drawer. The strongest association found between movement and feedback was the increase in direction reversals for higher reflection ratings, which highlights the concept that drawing trajectories tend to change alongside deeper pondering upon one's emotional state.

6.2 Interaction Roles

The second focus was slightly different: how have users actually perceived the interface's role in this dialogue? Results show that the neutral condition was described as predictable and controllable, which came as no surprise. The responsive one reportedly introduced both instability and surprise into the drawing dynamic, which was perceived as distracting by more than half of the participants, but sometimes also as collaborative. The change in role is central here, as feedback shows that what made the first sketchpad pleasant, or even boring to some, was the fact that it could be used as a transparent tool. The second sketchpad however became an active surface. This made some feel as though the drawing process was disrupted, whereas for others, it created moments of affective awareness.

To gather more context around the users' perception of the interface, we turn to movement data. Participants who would describe the interface as random showed bigger increases in direction reversals and rate of direction changes. Those who found the sketchpad distracting showed more speed variation and lost control can be associated with less direct stroke paths. Thus, those who experienced unpredictability also tended to produce a more unstable trajectory. This has to be understood within the feedback loop that is inherent to this experiment, where irregular movements could have caused less predictable visual reactions, which in turn could have made users switch directions. This could be seen as a complication, but it truly represents an important characteristic of the system, a play between input and output, a conversation between user and system.

6.3 What Does This All Say About the Affective Mappings?

Prior insights into literature are the backbone of the responsive sketchpad implementation. The binary category approach towards choosing the videos that precede the drawing tasks in the experiment allow this research to strengthen or challenge assumptions as to which movements attain to a negative or positive activation. Given that the perception of the videos aligns with the design intention, this analysis remains relevant. Only three of the 12 input pen movement characteristics did not match the expected dominant activation: fast movement, long strokes and high number of strokes. This suggests that the prior mappings were partially supported by the interaction data, but also discourage the interpretation of these are reliable translation of pen movement into specific emotions.

A very interesting additional layer comes in the form of open feedback. When asked to explain their understanding of the videos and how it made them feel, the majority followed the same thematic patterns. However, nuanced emotions were not uniform. The negative video projected feelings from anxiety to heightened arousal and reflections on how dark humanity can be. The positive video was mostly linked to calmness and relaxation, but also triggered sadness over environmental issues that we are facing.

Rather than undermining the project, this partial correspondence reinforces one of its original motivations. The literature review already identifies contradictions in attempts to connect individual drawing characteristics with emotional categories. The current findings likewise suggest that some affective tendencies may recur while others depend strongly on the participant, stimulus and interaction context. The results therefore support the decision to use such mappings as suggestive material transformations instead of diagnostic rules. Combined with the individuality of drawing style, this all strengthens the idea of not focusing on specific identifications. Reflective validity can be found in encouraging users to interpret for themselves and preserve ambiguity throughout the experience.

6.4 Limitations

Can responsiveness become too responsive? The short answer is yes.
Although support is found in the results for the main design direction, the data also shows an important weakness in the perceived aggressiveness of the interface. Effects were activated with a rate of .96, which translates to "almost every responsive stroke triggered some form of visual response". Approximately 7.92 effects were activated per session on average, which seems to have reduced their impact and the overall contrast between ordinary and meaningful responses. In the design principles explored earlier, a desired balance was described between responses that are engaging and overwhelming or distracting. Based on qualitative feedback, it becomes clear that reposes were triggered too often, so the thresholds were potentially set too low. Additionally, the salience results uncover that what matters most for a visual response to stand out is not frequency, but perceptual and material character. For future iterations, the system could benefit from prioritising fewer effects, but allowing them more

time and space to become memorable. Thresholds should be raised and the pacing index could be leveraged to better understand when the interface should stay quiet and when not.

In terms of methodological considerations, the main limitations are the sample size, the uneven distribution of gender and age and the heterogeneity of input devices. The order in which users interact with the two sketchpads is also fixed. This was intended, allowing the neutral sketchpad to set the scene of the interaction in a more familiar setting, before novelty is introduced, while also ensuring that the exposure to visual responses does not influence behaviour in the neutral condition, as this would have been more likely than the opposite scenario. However, this decision could introduce an order bias between the two conditions. The study should therefore be understood as evidence of a promising interaction mechanism rather than validation of a finished affective model.

7 Conclusion

This thesis started with a curiosity: what can affective computing offer beyond the strive to accurately classify emotional states? In searching for a perfect form of creative and free, yet potentially implicit input, drawing appeared to be a rich source of information. The way someone draws holds meaning, but translating it into stable emotion labels remains a hardship. Subjectivity and dependence on context heavily influence such insights, forming an ambiguity that I embraced and integrated in the interaction itself.

A responsive sketchpad was designed, which transforms a user's strokes in real-time, applying visual and material changes informed by hypothesised relations between movement and activations from prior work. This was compared against a neutral digital drawing environment, to clearly see what changes responsiveness brings in emotional self-reflection and user perception.

Regarding RQ1, the responsive sketchpad did support emotional reflection more than the neutral one. This mainly happened due to the users becoming more attentive towards their own ongoing behaviour while processing the responses that came from the interface.

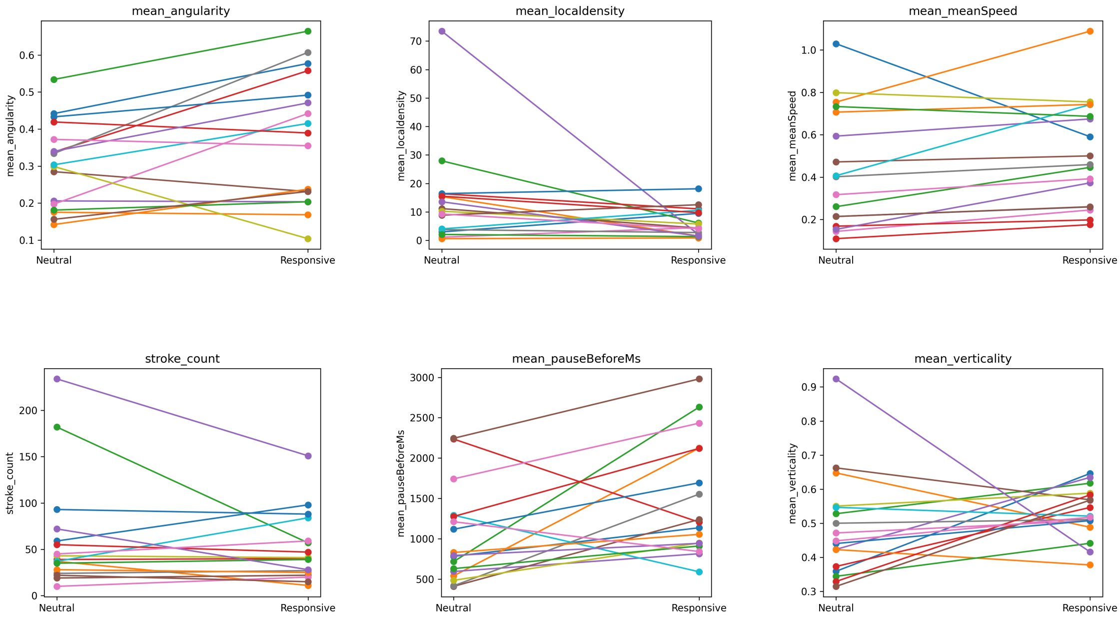
The answer to RQ2 is that participants perceived the responsive sketchpad as both collaborative and disruptive. It was meaningful enough to influence the participants' behaviour, but it exceeded the comfortable boundaries of agency for some users, which was perceived as intrusive. These differences, observed on a user to user basis, support the decision to not design a system that looks for general patterns and assumes they have the same meaning for each participant, but also underscore area of possible future improvements.

To conclude, this work suggests that uncertainty should not be seen as a problem to solve when it comes to affective computing or human-computer interaction applications. When experiences as personal and subjective to change as emotion become the focus, vagueness can become useful and resisting the temptation to work with definitive answers can turn into valuable insights. The responsive sketchpad prototyped for this project shows that turning behavioural information into a response offered to the person who ultimately produced it can be re-framed from diagnosis to dialogue, and personal interpretation can then guide the interaction further.

Appendices

A Additional figures

Figure 5: Comparison between movement behaviours across the two sketchpads, within the same participant session



B Questionnaire & Sketchpad

The entire study, comprised of both the questionnaires and the interface flow, used in this research project is available at the following link:
`https://leidenuniv.eu.qualtrics.com/jfe/form/SV_cUPtDMhyumPnwkm`

C Data Archive

The most relevant data gathered throughout this study can be found at: `https://docs.google.com/spreadsheets/d/1DG0pg68APMZBnAMfy93s0eQ2JKs45Ddo/edit?usp=sharing&ouid=102095298686742525537&rtpof=true&sd=true`

References

[Ayzeren et al., 2019] Ayzeren, Y. B., Erbilek, M., and Celebi, E. (2019). Emotional state prediction from online handwriting and signature biometrics. *IEEE Access*, 7:164759–164774.

[Ball, 2002] Ball, B. (2002). Moments of change in the art therapy process. *The Arts in Psychotherapy*, 29(2):79–92.

[Berrios et al., 2015] Berrios, R., Totterdell, P., and Kellett, S. (2015). Eliciting mixed emotions: A meta-analysis comparing models, types, and measures. *Frontiers in Psychology*, 6.

[Castellano et al., 2007] Castellano, G., Villalba, S. D., and Camurri, A. (2007). Recognising human emotions from body movement and gesture dynamics. In Paiva, A. C. R., Prada, R., and Picard, R. W., editors, *Affective Computing and Intelligent Interaction*, volume 4738 of *Lecture Notes in Computer Science*, pages 71–82. Springer.

[Damiano et al., 2023] Damiano, C., Gayen, P., Rezanejad, M., Banerjee, A., Banik, G., Patnaik, P., Wagemans, J., and Walther, D. B. (2023). Anger is red, sadness is blue: Emotion depictions in abstract visual art by artists and non-artists. *Journal of Vision*, 23(4).

[Davis et al., 2015] Davis, N., Hsiao, C.-P., Singh, K. Y., Li, L., Moningi, S., and Magerko, B. (2015). Drawing apprentice: An enactive co-creative agent for artistic collaboration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, C&C '15, page 185–186, New York, NY, USA. Association for Computing Machinery.

[Ge et al., 2024] Ge, Y., Liu, S., Fu, Z., Tan, J., Li, Z., Xu, S., Li, Y., Xian, Y., and Zhang, Y. (2024). A survey on trustworthy recommender systems. *ACM Transactions on Recommender Systems*, 3(2):1–68.

[Gross and Levenson, 1995] Gross, J. J. and Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion*, 9(1):87–108.

[Han et al., 2019] Han, J., Chernyshov, G., Zheng, D., Gao, P., Narumi, T., Wolf, K., and Kunze, K. (2019). Sentiment pen: Recognizing emotional context based on handwriting features. In *Proceedings of the 10th Augmented Human International Conference 2019*, pages 1–8. Association for Computing Machinery.

[Khan et al., 2024] Khan, Z. A., Xia, Y., Aurangzeb, K., Khaliq, F., Alam, M., Khan, J. A., and Anwar, M. S. (2024). Emotion detection from handwriting and drawing samples using an attention-based transformer model. *PeerJ Computer Science*, 10.

[Larsen et al., 2001] Larsen, Jr., J. T., McGraw, A. P., and Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, 81(4):684–696.

[Lieberman et al., 2007] Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., and Way, B. M. (2007). Putting feelings into words: Affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science*, 18(5):421–428.

[Likforman-Sulem et al., 2017] Likforman-Sulem, L., Esposito, A., Faundez-Zanuy, M., Clemencon, S., and Cordasco, G. (2017). EMOTHAW: A novel database for emotional state recognition from handwriting and drawing. *IEEE Transactions on Human-Machine Systems*, 47(2):273–284.

[Lin et al., 2024] Lin, Y., Yang, G., Ze, Y., Zhang, L., Xing, B., Liu, X., and Lyu, R. (2024). The impact of motion features of hand-drawn lines on emotional expression: An experimental study. *Computers & Graphics*, 119.

[Lin et al., 2022] Lin, Y., Zhang, N., Qu, Y., Li, T., Liu, J., and Song, Y. (2022). The House-Tree-Person test is not valid for the prediction of mental health: An empirical study using deep neural networks. *Acta Psychologica*, 230.

[Nijdam, 2008] Nijdam, N. A. (2008). Mapping emotion to color. Technical report 1, University of Twente. Sensitive Artificial Listener (SAL) Project.

[Nolazco-Flores et al., 2022] Nolazco-Flores, J. A., Faundez-Zanuy, M., Velázquez-Flores, O. A., Del-Valle-Soto, C., Cordasco, G., and Esposito, A. (2022). Mood state detection in handwritten tasks using PCA–mFCBF and automated machine learning. *Sensors*, 22(4).

[Nummenmaa et al., 2014] Nummenmaa, L., Glerean, E., Hari, R., and Hietanen, J. K. (2014). Bodily maps of emotions. *Proceedings of the National Academy of Sciences*, 111(2):646–651.

[Parkitny and McAuley, 2010] Parkitny, L. and McAuley, J. (2010). The depression anxiety stress scale (DASS). *Journal of Physiotherapy*, 56(3):204.

[Rahman and Halim, 2023] Rahman, A. U. and Halim, Z. (2023). Identifying dominant emotional state using handwriting and drawing samples by fusing features. *Applied Intelligence*, 53:2798–2814.

[Santosa et al., 2026] Santosa, S. S., Wan, Q., Yu, J., and Luo, Y. (2026). EmoFlow: From tracking to sense-making of emotions through creative drawing. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*.

[Somarathna et al., 2023] Somarathna, R., Vuilleumier, P., and Mohammadi, G. (2023). EmoStim: A database of emotional film clips with discrete and componential assessment. *IEEE Transactions on Affective Computing*, 15(3):1202–1212.

[Stamatopoulou, 2007] Stamatopoulou, D. (2007). Embodied imagination: Emotional expression perception of dynamic line-drawings created by artists. *Imagination, Cognition and Personality*, 26(3):213–248.

[Verduyn et al., 2009] Verduyn, P., Van Mechelen, I., Tuerlinckx, F., Meers, K., and Van Coillie, H. (2009). Intensity profiles of emotional experience over time. *Cognition & Emotion*, 23(7):1427–1443.

[Wang and Chen, 2020] Wang, C. and Chen, Y. (2020). Design of an interactive mind calligraphy system by affective computing and visualization techniques for real-time reflections of the writer's emotions. *Sensors*, 20(20).

[Weng et al., 2024] Weng, H. C., Huang, L. Y., Imcha, L., et al. (2024). Drawing as a window to emotion with insights from tech-transformed participant images. *Scientific Reports*, 14:11571.

[Yang and Qin, 2021] Yang, L. and Qin, S.-f. (2021). A review of emotion recognition methods from keystroke, mouse, and touchscreen dynamics. *IEEE Access*, 9:162197–162213.