Evaluating a Multi-method Approach for User Emotional Dynamics Assessment in Software UX

Alina Khomenko*
akh@macpaw.com
MacPaw
Kyiv, Ukraine

Kateryna Filipchenko*
kate.f@macpaw.com
MacPaw
Kyiv, Ukraine

Daria Polishchuk
dariapo@macpaw.com
MacPaw
Kyiv, Ukraine

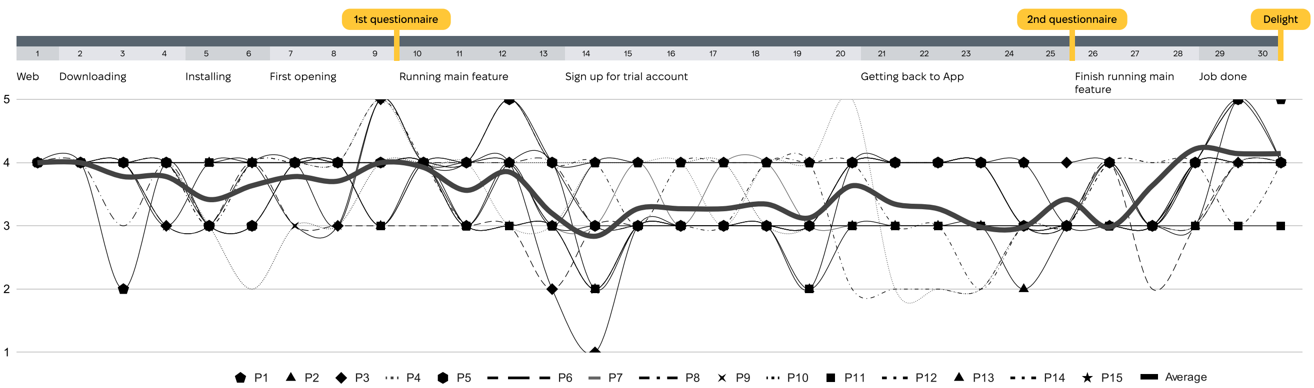


Figure 1: The Customer Journey Map (CJM) for the Utility App under study, each line corresponding to the observer's assessment of emotional dynamic changes of the user study participants, values from strong negative to strong positive intensity, one line per each participant, and average emotional dynamics.

ABSTRACT

User experience research plays a crucial role in software product development, focusing on user perceptions of the product and the emotions it invokes. However, many methods for measuring emotions still remain subjective and can lack sufficient accuracy and objectivity. We aim to address the subjectivity concern by proposing a multi-method user research approach, which could be applied in the context of interactions with software products and would be scalable and repeatable in remote user testing conditions. We combine self-reporting, behavioral observation analysis, direct user speech, and AI-powered facial expression analysis. We evaluate our method in two case studies with 15 participants, analyzing the emotional responses of users interacting with a Utility App and an App Marketplace, utilizing the Customer Journey Map Framework for deeper insights into emotional dynamics shifts. The analysis results indicate that, although AI analysis of emotions has limitations, the overall methodology partially correlates with observer analysis. Both methodologies are more effective in reporting emotional downs, while self-reported data tends to show emotional shifts more boldly.

CCS CONCEPTS

• **Human-centered computing** → **HCI design and evaluation methods**; *HCI theory, concepts and models*; **User studies**.

KEYWORDS

Emotional Response Analysis, UX Emotions, Remote User Testing, Customer Journey Map Framework

ACM Reference Format:
Alina Khomenko, Kateryna Filipchenko, and Daria Polishchuk. 2024. Evaluating a Multi-method Approach for User Emotional Dynamics Assessment in Software UX. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Emotions are central to User Experience (UX), significantly influencing human perceptions, choices, and actions. The subjective and complex nature of emotional responses, however, makes it difficult to quantify and measure these experiences accurately. This complexity is due to the unique and context-dependent nature of emotional reactions, compounded by various cognitive biases, including rationalization, memory recall, and the tendency to provide socially acceptable responses.

Our work focuses on evaluating emotions in software usage, recognizing their vital role in shaping decision-making and user

*Authors contributed equally to this research.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

interaction with products. We propose a multi-method approach in which we account for the limitations and applicability of different methods against our resources and objectives. For example, while biometric techniques offer precision, they also demand in-person participants, specific equipment, and software, which were beyond our scope due to budget and logistical constraints. We employ several emotion measurement techniques, namely self-reports, analysis of behavioral observations, direct user comments, and AI-based facial expression recognition. We integrate them into into a Customer Journey Map (CJM) for a holistic view of participants' emotional responses. Artificial Intelligence (AI) emerged as a more viable option, although it is not yet fully reliable in capturing and interpreting emotional feedback due to cultural, individual, and contextual variations. Hence, incorporating human judgment and analysis was essential, alongside direct user feedback and self-reports, to provide a comprehensive understanding of the emotional dynamics shifts.

We evaluate our proposed methodology in two user studies involving interaction with a desktop Utility App and an App Marketplace. In both studies, we aim to identify moments within the user experience during software use that trigger negative emotions in order to understand the underlying causes and develop strategies to address them in future product development iterations. At the same time, we aim to detect, highlight, and enhance interaction aspects that provoke positive emotions.

Both user studies highlight the individuality of emotional experiences, exemplifying diverse reactions among participants but also highlighting common patterns. Direct user comments analyzed through CJM emerged as the most precise and unbiased method. However, self-reports provided deeper insights at key points of the user journey. Interestingly, manual observations often coincided with AI-generated facial expression analyses.

Our findings suggest that while AI-based emotion analysis has its limits, it aligns to some extent with human observations. Both these approaches effectively identify negative emotional responses, whereas self-reports more vividly capture shifts in emotions.

2 RELATED WORK

In the domain of User Experience (UX), the measurement of emotional experiences encompasses a variety of methodologies, each providing distinctive perspectives on user interactions and their emotional reactions to products and services.

2.1 Self-reporting

Self-reporting techniques, such as surveys and questionnaires, capture subjective user feedback on emotional experiences. They represent a cost-effective and accurate alternative to more intrusive biometric methods while capturing nuanced emotional responses that might not be evident through behavioral observations alone. Today, various tools and frameworks are in use. In the social sciences, the Positive Affect Negative Affect Schedule (PANAS) scale [47] has been widely employed. The pleasure-arousal-dominance (PAD) emotional framework [31] has been implemented in several measurement instruments, including the Affect Grid [41] and the Self-Assessment Manikin (SAM) [5].

A notable advancement in self-reporting surveys use came with the adoption of Robert Plutchik's wheel of emotions [35] depicting the spectrum of emotions. A recent development in self-report visual scales is embodied by the use of expressive animated cartoons [27]. The Product Emotion Measurement Instrument [10] (PrEmo) is designed to assess 14 different emotions that are commonly evoked by product design. Despite the progress on survey-based measures of emotional response, the accuracy of self-reported emotions is generally enhanced when they directly reflect the emotions being felt at the moment [37]. Even in this case, though, there are concerns that not all individuals are aware of and capable of reporting on their momentary emotional states [24].

2.2 Biometric Methods

Biometric methods offer an objective lens through which to view users' emotional responses, utilizing data from physiological sources. These include cardiovascular measures (heart rate, blood pressure, total peripheral resistance, cardiac output, pre-ejection period, heart rate variability), electrodermal responding (skin conductance level or short-duration skin conductance responses), adrenaline levels, neural images, perspiration, tears, and muscle activity. Some evidence for autonomic nervous system specificity in differentiating discrete emotions has been reported [6, 12, 44]. The key advantage here is bypassing self-reporting subjectivity, although the invasiveness and the need for specialized equipment can be seen as downsides. Some studies promote the development of wearable consumer electronic devices for monitoring human emotions, such as emotional recognition using heart rate data from a wearable smart bracelet [43]. Biosensor measures of valence and arousal, calculated from Electroencephalography and Apple Watch, correlate with self-reported valence and arousal measured by the EmojiGrid[20], building a bridge between self-reporting and biometric methods.

Other options might be checking blood pressure, pupil dilation, or skin conductance[28]. Instead of measuring heart rate, cortisol levels could be tested through saliva samples[36]. Similarly, galvanic skin response used widely in pre-testing marketing campaigns and commercials, can be replaced by a voice pitch analysis[40].

2.3 Observing Behaviours

Kunin [22] proposed the use of facial expressions as scale descriptors to visually depict an individual's emotional state along a continuum from positive to negative rather than employing textual descriptions or numerical values.

The concept that individuals communicate emotions through facial expressions, as explored by Ekman [11] and Darwin [9], posits that emotions have an adaptive role in communication, leading to behaviors that inherently disclose one's emotional state to others.

Another set of theories that links emotional states to action dispositions [14, 25]. According to these theories, it should be possible to infer a person's emotional state from vocal characteristics, facial displays, and whole-body behaviors. Work by Paul and Friesen [34] focuses on the configurations of facial muscular movements, which have been shown to provide an accurate representation of the emotions felt by an individual. Paul and Friesen's "pictures of facial affect" have been developed to measure perceptions of emotions in others rather than as a self-report measure of an individual's feelings. Facial behaviors appear to reliably indicate the valence of a person's emotional state[39]. That builds a ground for developing qualitative research methods through interviews and observational studies, providing a rich tapestry of emotional insights. They allow researchers to delve into the 'why' behind user emotions, offering a depth of understanding that quantitative methods alone cannot achieve. This approach underscores the importance of a contextual and nuanced understanding of emotions, paving the way for more empathetic and user-centered product development. However, the approach has several drawbacks, including subjectivity bias and interpretation challenges.

2.4 AI-based Facial Emotion Recognition

With the advances in deep learning, the popularity of AI-based solutions for facial emotion recognition (FER) has soared in the last decade [8, 16]. Researchers and practitioners are actively exploring the applicability of this method of human emotional assessment, both in regards to the precision of the technique [21], and its fitness for use in diverse domains, such as tourist satisfaction when visiting sites [15], videoconference user experience improvement [3], or medical thermal imagery analysis [33]. AI-based FER is also finding its use in various UX studies, as the recent systematic studies [29, 46] show. Multi-modality in AI-based FER techniques can also refer to training machine learning (ML) models on diverse datasets comprising audio, video, and biometric information [30], which is out of the scope of this work.

2.5 Our Approach

While most of the approaches to emotional assessment are well-established on their own, the multi-modal methods field is still growing, especially when such studies incorporate AI/ML solutions. We situate our work within the growing research area of multi-method approaches to user emotions assessment in UX flow analyses, which combine the aforementioned techniques. Such combined approaches are reported in prior work, that combine biometric and self-reporting methods for webpage [49] and augmented reality apps [42] UX analysis, and emotional tension and desirability of software features [17], or combine UX Curve and self-reporting methods for mobile app UX analysis [13], evaluate combinations of biometric and AI-based FER methods for UX tasks in industry [26], report on technological product failure analysis via facial recognition and eWOM data combination [18].

While we recognize the potential and value of biometric-based methods, their dependence on specialized hardware limits their broad adoption both from the infrastructure setup and financial considerations. In our work, we focus on combining emotion assessment methods that are less infrastructure-dependent and, thus, more broadly applicable and scalable. We work with self-reporting, observational, and AI-based FER techniques, that were shown to combine naturally in the UX studies by recent research on multi-method UX studies limitations in the context of software use [23].

Given our specific focus on software use, we rely on the Customer Journey concept [45], previously successfully used in industry research [4, 48]. Our work closely relates to that of [2], where a Customer Journey Map is used to incorporate heterogeneous data within a user study, with the difference that we do not employ biometric-based methods.

3 PROBLEM AND METHODOLOGY

Defined as a complex psychological state, emotion typically emerges as an instinctive response to a stimulus, occurring spontaneously without conscious intent. This led us to the idea to assess the emotional trajectory of users as they interact with our two products (a Utility App and an App Marketplace) in a real-time setting. We aim to identify the emotional triggers impacting users positively and negatively at crucial points in their customer journey, creating a detailed map of the user emotional engagement. We employ a methodology combining low-moderated usability testing with several data collection methods, such as self-report questionnaires, observational analysis, direct communication, and AI-based facial emotion recognition.

3.1 Main Challenge

The principal challenge in quantifying users' emotional reactions during interaction with product interfaces lies in the subjective nature of emotions themselves. Emotional experiences are deeply personal and can vary significantly from one individual to another, influenced by many factors, including personal history, expectations, and the context of the interaction. This variability introduces a level of complexity that makes capturing a universally accurate representation of users' emotional states non-trivial. We hypothesize that by employing a multi-method approach we move towards a more objective emotion evaluation.

3.2 User Study Participants

In both of our studies, we engage 15 participants, selected for their alignment with the demographic characteristics of the existing user base of the Utility App and the App Marketplace, via an online platform dedicated to facilitating user studies. It was imperative that these individuals had no prior exposure to both products, ensuring that their feedback and reactions would be uninfluenced by previous experiences.

3.3 Defining the Customer Journey Map

We determine the most popular user flows for the Utility App and the App Marketplace using data from Google Analytics and our internally configured application analytics, based on the number of unique user sessions. Utilizing this data, we identify the pathways frequently navigated by newcomers and subsequently develop detailed Customer Journey Maps for each product. These maps are instrumental in foreseeing potential emotional shifts, enabling us to designate specific collection points for self-reported emotional feedback and tracking the observations during the user study course.

3.4 User Studies

Our objective is to analyze the emotional triggers impacting users at key junctures of their experience. To achieve this, we conduct all sessions remotely via 60-minute video calls, during which participants share their screens. Written tasks lead participants through each CJM phase, enhanced by two questionnaires designed to delve

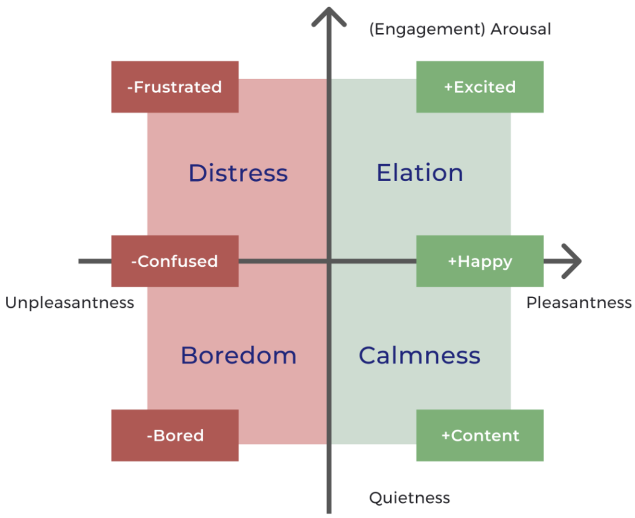
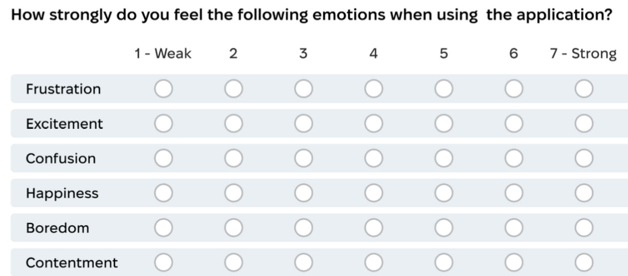


Figure 2: Emotional Spectrum, positive and negative emotions.



How strongly do you feel the following emotions when using the application?

	1 - Weak	2	3	4	5	6	7 - Strong
Frustration	○	○	○	○	○	○	○
Excitement	○	○	○	○	○	○	○
Confusion	○	○	○	○	○	○	○
Happiness	○	○	○	○	○	○	○
Boredom	○	○	○	○	○	○	○
Contentment	○	○	○	○	○	○	○

Figure 3: Self-reporting questionnaire, assessing participants' feelings of Frustration, Confusion, Boredom, Excitement, Happiness, and Contentment, 1 to 7 scale.

into the user's emotional experiences. Observers' cameras are intentionally deactivated to maintain the purity of user responses, minimizing any external influence on their genuine reactions.

3.5 Observation

The observing researchers follow and pinpoint the frequency and intensity of positive and negative sentiments expressed by users at each step of the customer journey, providing their measure of emotional response levels. Their evaluations include a 5-point scale assessment of both verbal and non-verbal (behavioral) expressions of sentiment, with ratings ranging from 1 (indicating strong negative emotions) to 5 (indicating strong positive emotions), based on the observer's subjective interpretation. The study also includes such metrics as the number of user comments with negative and positive sentiments in total and for each CJM step. This approach helped to gain an understanding of the reasons behind emotional ups and downs.

3.6 Self-reporting Questionnaires

Emotional Spectrum. The collection of self-reported data is designed to ascertain both the valence (positivity or negativity) and the intensity of users' emotional responses. To facilitate this, our questionnaire is structured around Russell's circumplex model of emotion [38]. The Emotional Spectrum tool is utilized at 2 pivotal points in the user journey to record a broad range of emotional experiences. Users are asked to assess their feelings of Frustration, Confusion, Boredom, Excitement, Happiness, and Contentment (Figure 2), providing a comprehensive insight into the emotional dimensions encountered during their interactions with the Utility App and the App Marketplace. At two key moments in the respective CJMs, participants are asked to rate the intensity of their emotions on a 1 to 7 scale, reflecting their experience (Figure 3).

Delight. In our study we also evaluate Customer Delight [1, 7, 19, 32]. Although there are various models of delight, they generally include a positive reaction to an unexpected experience. According to the Emotional Spectrum, delight is associated with high energy and pleasure (see Figure 4). For our assessment, participants provide responses to the statement "I felt delighted while using this app" on a 7-point scale, from "Strongly Disagree" to "Strongly Agree".

3.7 AI-Based Facial Expression Analysis

We aim to evaluate the efficacy of facial expression analysis tools by comparing their outcomes with traditional assessment methods, with the ultimate goal of streamlining the time required for such evaluations. To achieve a more nuanced understanding of the emotional spectrum encountered by users, we explored various AI-based tools tailored to our specific needs. The selected technology is capable of recognizing a range of emotions and their intensities, we employ it as third-party tool with expression measurement models for the voice, face, and language, with an open API. The emotional spectrum it analyzes is similar to that of our questionnaires, encompassing Confusion, Boredom, Excitement, and Happiness. We conduct this analysis with an emphasis on the intensity of these

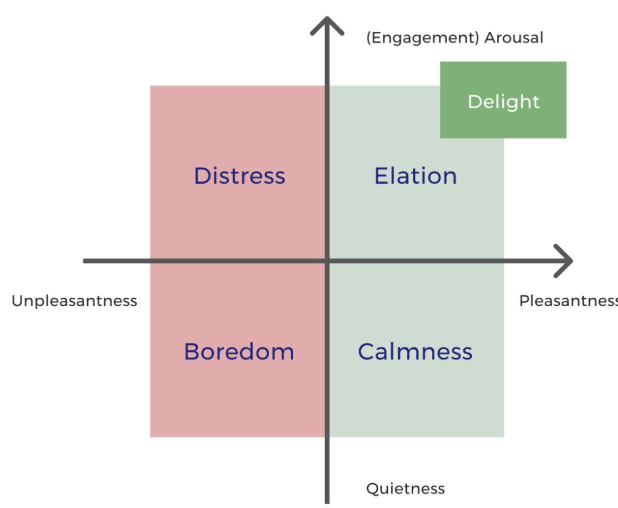


Figure 4: Delight in the Emotional Spectrum. Delight is associated with high energy and pleasure.

emotions, rated on a scale from 0.0 to 1.0, allowing us to delve deeply into the emotional nuances experienced by the users during product interface interactions.

To validate the efficiency of the AI-based approach for this particular study, we select a single user test recording for analysis, for each type of software product under study, the Utility app, and the App Marketplace. We have created an in-house dashboard that finds anomalies in data and helps to understand the user emotional dynamics at each CJM stage. The algorithms are calibrated on interviews where a person's face is located in the top left corner of the interview recording frame. Ensuring the participant is facing directly toward the monitor and front camera was crucial, as well as minimizing any other visual elements that could include people's faces. Also we document the time codes for each step and segment the recording, allowing the AI-based tool to exclusively analyze the participant's facial expressions during their interaction with the app interface. It is crucial to isolate user reactions from any potential bias introduced by their interaction with the research platform, such as when reading instructions.

4 USER STUDY: THE UTILITY APP

4.1 Participants

We recruit 15 participants (9 females, 6 males; average age 42.5; located in the USA; English native speakers) through an online research platform. The pre-screening process is designed to evenly distribute participants based on the existing segmentation of user profiles. This segmentation is determined by the type of computer they have and are using the most; the year of the device; previous experience with the apps from the same product category. Additionally, they had no prior experience using or exploring this particular Utility App. All studies are conducted remotely via video conference calls, requiring participants to have an active front camera so we could record and analyze their facial expressions. Participants receive compensation from the recruiting platform upon completing the 60-minute study.

4.2 Customer Journey Map

Using the data gained from product analytics, we design a Customer Journey Map of the most common flow of the user new to this Utility App prior to the user testing sessions. The defined CJM consists of 30 steps starting with the landing page of the Utility App (CJM Step 1), followed by the downloading and installation of the Utility App (CJM Steps 2-6), opening it for the first time (CJM Steps 7-9), running the main App's feature (CJM Steps 10-13), starting a free trial account (CJM Steps 14-23), returning to the App to complete the main task (CJM Steps 24-30).

The CJM serves as a foundational framework for our subsequent user testing. By mapping out the user journey, we highlight key interaction points within the app, which allows us to pinpoint stops where users experience significant emotional shifts.

4.3 Incorporating the Observer Data

Observers assign subjective ratings from 1 to 5 (where 1 is strong negative emotions, 3 is neutral, and 5 is an exceptional positive reaction) reflecting shifts in the participants' mood during the user test.

Average emotional change. By evaluating the average emotional level across respondents, we obtained a generalized picture, accounting for outlier cases, see Figure 1. The overall flow didn't cause significant emotional drops or peaks of excitement for the majority, indicating no major design flaws eliciting strong negative reactions or, conversely, exceptional excitement. However, some common patterns are observed. For instance, a drop when the user faces the necessity to sign up for a trial to continue the process (CJM Step 14, 2.93 out of 5); emotional rise when the task is completed at the end of the journey (CJM Step 29, 4.2 out of 5).

Emotional changes per user. Collecting data for each interviewee allowed us to understand that not every user experiences emotional shifts; some remain neutrally indifferent through the app's flow, as shown in the Figure 1. However, observations enabled us to discern what emotional changes are possible, under what circumstances they occurr (specific events within the app that trigger such emotions), and their intensity. We focus on peak points and drops to pinpoint exactly what aspects of the app caused these emotional changes. If a particular behavior or attitude is observed in even one participant out of our sample, this could likely be observed within the broader population, so we record that too.

Emotional drops generators. We observed that 6 steps on the CJM triggered emotional drops in some users: a UX issue during the installation process (Step 3); looking for the App to open it (Step 6); facing the need to sign up for a trial account (Step 14); facing the need to enter credit card information (Steps 19-20); difficulties with returning to the app after signing in (Steps 21-25); facing an unexpected pop-up (Step 27). Most of the mentioned emotional drops were caused by unexpected user interface behavior and/or the need for a commitment.

Positive emotions drivers. Four steps on the CJM lead to a positive emotional shift in some respondents: opening the app after the installation process (Step 9); a screen showing the preliminary results of the main feature running (Step 12); a success screen after signing up for a trial (Step 20); and a success screen after finishing the task (Step 29). It is worth mentioning, that all observed cases were linked with the feeling of accomplishment during the customer journey.

Analyzing emotional changes for each user revealed that certain stages (e.g., Step 20), could significantly drive positive emotions. However, due to the UX design flaws, these same stages could also lead to frustration among users who encounter these specific UX issues.

4.4 User Comments as Data

Each step on the CJM is enriched with user comments, both positive and negative. Direct quotes from users serve as a powerful data source, helping to determine the triggers of emotional state changes in subsequent analysis of emotions.

4.5 Incorporating Self-Reporting Data

To augment our methodology with quantitative data, we implement two self-reporting questionnaires at key points within the CJM. The Emotional Spectrum questionnaires are placed accordingly: a) after opening the app (Step 9); b) after creating a trial account (Step 25).

Additionally, we measure the Delight Rate at the end of the user journey to gauge overall satisfaction and emotional uplift (Step 30). These points on the CJM, identified through our hypotheses, are anticipated to trigger significant mood changes in users.

Table 1: First Questionnaire results. 1 to 7 scale, distribution of the number of respondents per each answer, % of total respondents

Emotion	1	2	3	4	5	6	7
Frustration	67%	13%	13%	0%	7%	0%	0%
Excitement	0%	0%	13%	20%	13%	40%	13%
Confusion	73%	13%	0%	7%	0%	7%	0%
Happiness	0%	0%	7%	27%	20%	33%	13%
Boredom	60%	20%	7%	13%	0%	0%	0%
Contentment	0%	0%	7%	40%	0%	33%	20%

First Questionnaire. Participants reported low levels of negative emotions (1 to 7 scale; average per sample): Frustration = 1.67, Confusion = 1.67, Boredom = 1.73; medium level of positive ones: Excitement = 5.2, Happiness = 5.2, Contentment = 5.2. However, respondents were more inclined to report the absence of negative emotions than assigning the highest positive rating to positive emotions. Overall, 10 respondents (67%) reported the lowest possible levels of Frustration, 11 respondents (73%) reported the lowest possible levels of Confusion, and 9 respondents (60%) reported Boredom; however, just 2 respondents (13%) reported the highest levels of Excitement and Happiness accordingly, and 3 respondents reported the highest levels of Contentment. See Table 1 for details.

Second Questionnaire. Fewer respondents reported low levels of negative emotions compared to the First Questionnaire: Frustration (4 respondents, -40%) and Confusion (7 respondents, -26%); 11 respondents reported the lowest level of Boredom (+13%). The number of respondents that reported the middle levels (answer 4 out of 7) of Excitement, Happiness, and Contentment rose: +27%, +13%, and +7%, respectively. See Table 2 for details.

On average, respondents reported higher levels of Frustration = 3.13 (+1.46) and Confusion = 2.47 (+0.8); a bit lower level of Boredom = 1.53 (-0.2); and lower positive emotions: Excitement = 4.6 (-0.6), Happiness = 4.47 (-0.73), Contentment = 4.8 (-0.4), as shown in Figure 5.

Table 2: Second Questionnaire results. 1 to 7 scale, distribution of the number of respondents per each answer, % of total respondents. Arrows show the change compared to the First Questionnaire (increased, decreased)

Emotion	1	2	3	4	5	6	7
Frustration	↓ 27%	20%	7%	↑ 27%	7%	7%	7%
Excitement	0%	7%	7%	↑ 47%	7%	↓ 27%	7%
Confusion	↓ 47%	7%	13%	20%	13%	0%	0%
Happiness	7%	7%	0%	↑ 40%	20%	20%	7%
Boredom	↑ 73%	7%	7%	13%	0%	0%	0%
Contentment	0%	7%	7%	↓ 33%	20%	20%	13%

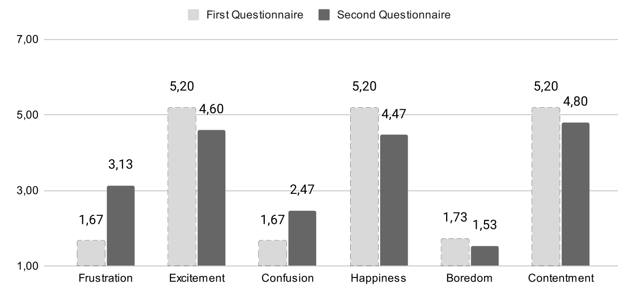


Figure 5: Dynamics of the average emotions rates, data from First and Second Questionnaires compared, 1 to 7 scale.

Delight Rate. 8 out of 15 respondents reported being delighted with the Utility App, opting for the highest ratings of 6 and 7, resulting in an average score of 5.07. Together with the observed and discussed earlier shift towards negative emotions, we conclude that certain interface behaviors are likely to trigger an emotional decline in the middle of the CJM, as shown in Figure 6.

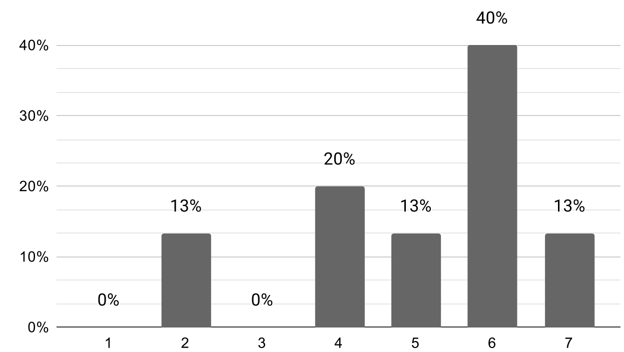


Figure 6: Delight rate. 1 to 7 scale, % of total respondents.

4.6 AI Facial Analysis Results

The AI-based FER analysis focused on Boredom, Confusion, Excitement, and Happiness. The emotional intensity scale spans from 0.0 (not observed) to 1.0 (intense). The general tendency goes in line with the data gathered by the observer and self-reported data: we see the drop in Excitement and Happiness that happened after 00h 20min (Figure 7). The timing corresponds to Step 15 of the CJM (the user faces the necessity to sign up for a trial and create an account to continue).

Happiness emerged as the leading emotion throughout the customer journey, showing a broad range from 0.09 to 0.91, with peak intensities at Steps 1, 9, 13, and 26. Similarly, Excitement peaked at the same steps of the CJM, however, with a lower intensity, ranging from 0.05 to 0.62. The analysis revealed that Boredom exhibited a higher intensity range (0.09 - 0.61, peaking at 0.99) compared to

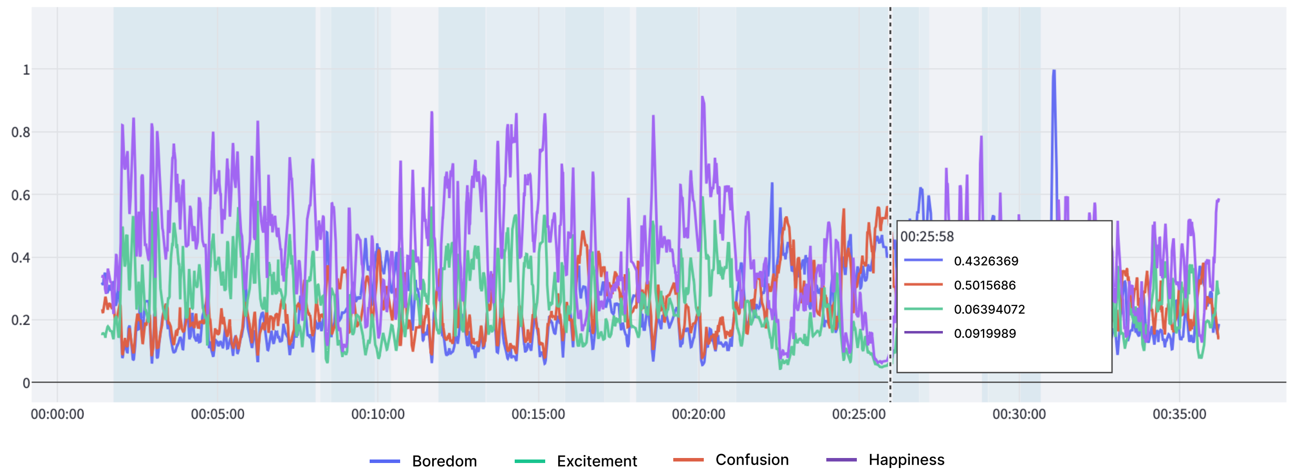


Figure 7: AI Facial Expression Intensity Analysis, 0.0 to 1.0 scale (Frustration, Confusion, Boredom, Happiness).

Confusion, which ranged from 0.09 to 0.55, with specific peaks occurring at Steps 6-7, 13, 15, and 22-23. Notably, the highest recorded Boredom level (0.99) occurred when a respondent was engaged in reading the description on the interface page.

However, it's worth noting that the AI-based approach to emotion detection has its limitations and does not always accurately convey human emotions. We observed that it has difficulty distinguishing smiles arising from surprise, discomfort, nervousness, and genuine joy from interacting with the app. For example, at Step 13 of the CJM, where the respondent finishes utilizing the main feature of the app, the AI analysis indicated a high level of happiness (0.83) for a respondent. Yet, the observer notes a drop in excitement and a low level of happiness at this point. This discrepancy highlights challenges in relying solely on AI for emotional analysis.

5 USER STUDY. THE APP MARKETPLACE

5.1 Participants

We recruit 15 participants (7 females, 8 males; average age 37.4) through an online research platform with diverse worldwide participants. The pre-screening process is designed to evenly distribute participants based on the existing segmentation of user profiles. This segmentation is determined by Software Awareness (Light, Moderate, and Power users), which considers the number and types of desktop applications installed (excluding pre-installed applications) and the amount of time spent on the computer per week (ranging from 10 to 30+ hours). We recruit 5 Light users, 5 Moderate users, and 5 Power users. Additionally, they had no prior experience using or exploring this App Marketplace. All studies are again conducted remotely via video conference calls, requiring participants to have an active front camera and screen sharing, so we can record and analyze their facial expressions and and actions. The platform compensates the participants upon completing the 60-minute study sessions.

5.2 Customer Journey Map

The defined Customer Journey contains 20 steps and begins with an overview of the main page on the App Marketplace website (Steps 1-5), then progresses to signing up (Steps 6-8), activating a free trial (Steps 9, 10), downloading (Step 11), and installing the desktop product App Marketplace (Steps 12, 13), completing an overview (Steps 14-16), and finally installing and experiencing a Vendor's app (Steps 17-20). Throughout the entire customer journey, users have one primary task to complete: finding a solution to enhance their screenshotting experience on their computers. Consequently, they need to locate the appropriate Vendor's app, install it, and take a screenshot. The whole journey can be divided into 3 main stages: the first is the website experience (Steps 1-11), the second is the Desktop App Marketplace (Steps 12-16), and the third is the Vendor's app (17-20). The Emotional Spectrum questionnaires are administered after completing the Sign-up process (Step 10) and upon finding the proper Vendor's app within the Marketplace (Step 16). At the end of the journey (Step 20), an assessment of overall user delight is conducted.

5.3 Incorporating the Observer Data

Average emotional change. We observe that the average magnitude of emotional fluctuations throughout the customer journey is insignificant, ranging from 2.4 to 3.33 (Average = 2.86); that is, there are generally no extremely negative or positive moments. However, certain steps do show more noticeable fluctuations. It starts from Step 3, which is finding the app that will help to complete the task on the website. The average rate of Excitement there is 3.33 (+0.47 from the average 2.86). The second emotional uplift happens at Steps 14 and 15, where the participants finish the installation process and start the onboarding to the App Marketplace, rated 3.07 (+0.21). The third uplift occurs at Step 16, 3.13 (+0.27) when choosing an application to perform a task; there is a slight emotional uplift, characterized by joy and intrigue, stemming from

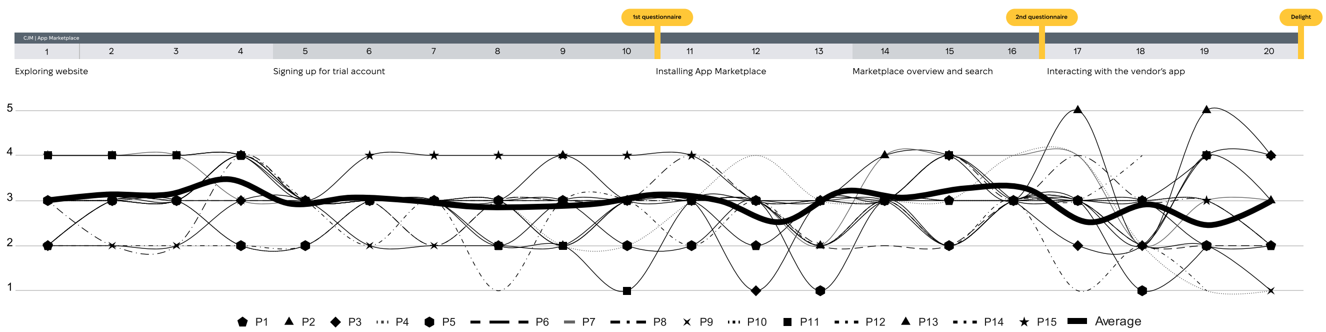


Figure 8: Customer Journey Map for the user study of the App Marketplace use, observer's assessment of emotional dynamic changes from strong negative to strong positive intensity for each participant, and average intensity dynamics.

the anticipation that the user had found what they were looking for. We observe an emotional drop almost immediately when users experience the number of required steps for app installation (Step 17), granting permissions for screen recording and full disk access (Step 18), rated 2.4 (-0.46), and eventually, the difficulty some participants experience in finding the app's interface at Step 19, rated 2.71 (-0.15), and the necessary functionality within the application at Step 20, rated 2.33 (-0.53).

Emotional changes per user. We also combine the observers' assessments (from 1 to 5) with the participants' direct speech and identify the reasons for such changes in their reactions. If a specific behavior or attitude is observed in even one participant out of a sample of 15 people, it could likely be observed within the broader population.

5.3.1 Emotional drops generators. We observe 7 steps that trigger the most frustration among 5 users (given the lowest rate of 1): sharing the name (Step 8); sharing credentials (Step 10); downloading the App Marketplace application (Step 11); granting permissions (Step 13); installing the Vendor's app (Step 17); interacting with the Vendor's app (Steps 18-19). The most common reason for these is a mismatch between the expected actions needed during the installation process of the App Marketplace and the Vendor's app. Participants expected to be able to install apps without providing any personal details. Additionally, technical bugs were encountered in Step 8 and Step 10, where input fields were disabled.

Positive emotions drivers. The 2 most exciting steps (rated 5) for participants were when they found the right Vendor app in the marketplace (Step 16) and saw the full functionality of the Vendors' apps (Step 19), realizing these could perform more complex and useful actions.

We observe that the same step, such as Step 19 "Launch vendor's app," could act as both an emotional drop generator and a positive driver for different participants. Users who are more knowledgeable about software (Power users) and with prior experience using utility apps were more likely to experience positive emotional uplifts (given rates 4 and 5). Meanwhile, those less familiar with such apps (Light users) tended to be more frustrated and negative (given rates 1 and 2).

5.4 Incorporating Self-Reported Data

First Questionnaire. Generally, participants' self-assessments of their emotions begin with low levels of negative emotions: Frustration = 1.53, Confusion = 1.73, Boredom = 1.33) and medium levels of positive ones: Excitement = 4.2, Happiness = 4.2, Contentment = 3.93. We observe that Contentment is the emotion that participants felt less intensely than Excitement and Happiness, by 0.27 points (see Table 3). Respondents were more likely to give negative emotions the lowest rating of 1 while showing reluctance to assign a higher positive rating of 7 to positive emotions. However, we recommend analyzing the data as percentages of given ratings, ranging from 1 to 7, for each emotion, rather than the absolute numbers.

Table 3: First Questionnaire results. 1 to 7 scale, distribution of the number of respondents per each answer, % of total respondents

Emotion	1	2	3	4	5	6	7
Frustration	80%	7%	0%	7%	7%	0%	0%
Excitement	7%	7%	13%	27%	33%	7%	7%
Confusion	60%	13%	20%	7%	0%	0%	0%
Happiness	7%	7%	13%	27%	27%	20%	0%
Boredom	87%	0%	7%	7%	0%	0%	0%
Contentment	20%	0%	7%	20%	47%	7%	0%

Only one participant, P3 (Moderate user), rated their positive feeling (Excitement) with the highest score of 7 in the first questionnaire. The lowest score of 1 for positive emotions, was given by only one participant, P9 (Medium user), for both Excitement and Happiness. Contentment received the lowest score of 1 from three participants: P2 (Power), P11 (Moderate), P14 (Light). When examining the extreme ratings for negative emotions, we do not find instances of the highest possible rating of 7. In the first questionnaire the highest given rate is given to feeling Frustration, and it's 5 (only 1 participant, P12. Moderate). For Confusion and Boredom only one participant, P4 (Power), gave the highest rate of 4.

Table 4: Second Questionnaire results. 1 to 7 scale, distribution of the number of respondents per each answer, % of total respondents. Arrows show the change compared to the First Questionnaire (increased, decreased)

Emotion	1	2	3	4	5	6	7
Frustration	↓ 67%	7%	7%	0%	7%	13%	0%
Excitement	13%	7%	13%	↓ 20%	13%	↑ 20%	13%
Confusion	↓ 53%	13%	13%	7%	0%	13%	0%
Happiness	7%	7%	7%	27%	7%	↑ 33%	13%
Boredom	87%	7%	7%	0%	0%	0%	0%
Contentment	20%	13%	13%	↑ 27%	0%	↑ 27%	0%

Second Questionnaire. After the second measurement, users were generally more conscious of their judgments. The level of negative emotions, Frustration, and Confusion, slightly increased: Frustration = 2.13 (+0.6 points), Confusion = 2.27 (+0.54 points), and only Boredom marginally decreased = 1.2 (-0.13 points). The level of Excitement remained nearly unchanged at 4.27 (+0.07 points), while Happiness saw an increase to 4.6 (+0.4 points). Only Contentment experienced a decrease, moving to 3.53 (-0.4 points) (Table 4; Figure 9): No observed differences were found in the self-reported data among Power, Moderate, and Light users.

Delight. Almost half of the participants (7 out of 15) expressed delight with the App Marketplace, selecting ratings of 6 and 7 (Figure 10). As a result, the average rating amounted to 5.2. However we consider evaluating data in dynamics and in comparison rather than reporting absolute average values.

5.5 AI Facial Analysis Results

Our observations indicate that Confusion is the predominant emotion encountered throughout the journey, generally ranging from 0.44 to 0.55, with its most intense occurrences (at Steps 4, 19, and 20) reaching up to 0.66. The second most intense emotion identified is Boredom, with a maximum of 0.64 observed at Steps 4, 10, and 12. Happiness was noted for having nearly 4 significant spikes, varying from 0.58 to 0.67 at Steps 1, 4, 10, 12, and 20, yet it remained low, between 0.04 and 0.13, during other times, but remains low

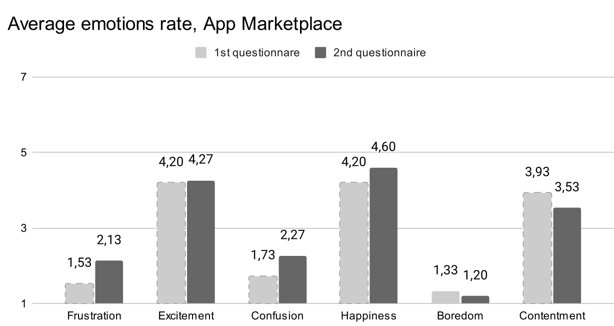


Figure 9: Average emotions rate, data from First and Second Questionnaires compared, 1 to 7 scale.

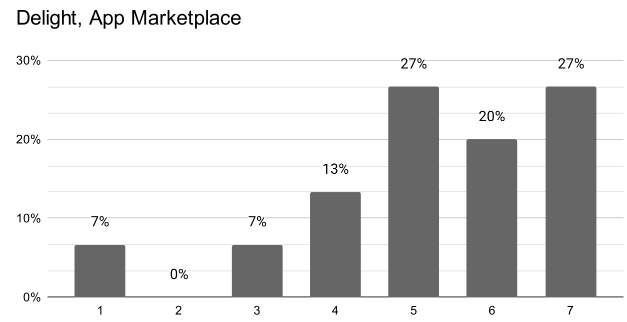


Figure 10: Delight rate. 1 to 7 scale, % of total respondents.

0.04-0,13 at all other times. Excitement was observed to be the least intense emotion, with its intensity ranging from 0.03 to 0.56 and displaying 5 notable spikes at the same steps as Happiness, which ranged from 0.42 to 0.57 (at Steps 1, 4, 10, 12, and 20). (Figure 11)

Conducting a reality check for data accuracy through human analysis is essential. For example, AI-based technique recognized the highest levels of positive emotions, such as Excitement (0.57) and Happiness (0.67), during a moment when the user was, in fact, experiencing significant frustration, manifested through a forced smile at Step 10. This misinterpretation arose when the participant faced a bug that blocked them from inputting their payment information and activating the trial at Step 10.

6 DISCUSSION

The objective of our multi-method study was to identify interaction points within the primary user flow of software products use that could elicit negative emotions, with the aim of mitigating these effects. Additionally, the study sought to identify positive triggers within these emotional aspects to facilitate scaling. Below, we discuss our findings.

6.1 Approach Comparison

Retrospectively, we assessed each approach on four criteria (value, complexity, accuracy, and potential) using a scale of low, average, and high, as detailed in the Table 5:

- *Value* refers to the quality of data obtained through the methodology, which could be transformed into actionable insights.
- *Ease* relates to the consumption of time and resources.
- *Accuracy* denotes our evaluation of the data's veracity.
- *Potential* pertains to the scalability and adaptability of the approach, indicating future focus areas.

For value, we conclude that observational methods and Customer Journey Maps that include direct user feedback provide the most actionable insights for improvement. Self-reported data, while less actionable, offers a broader understanding of emotional change dynamics and can be compared with subsequent measurements. AI-based FER insights failed to yield novel information.

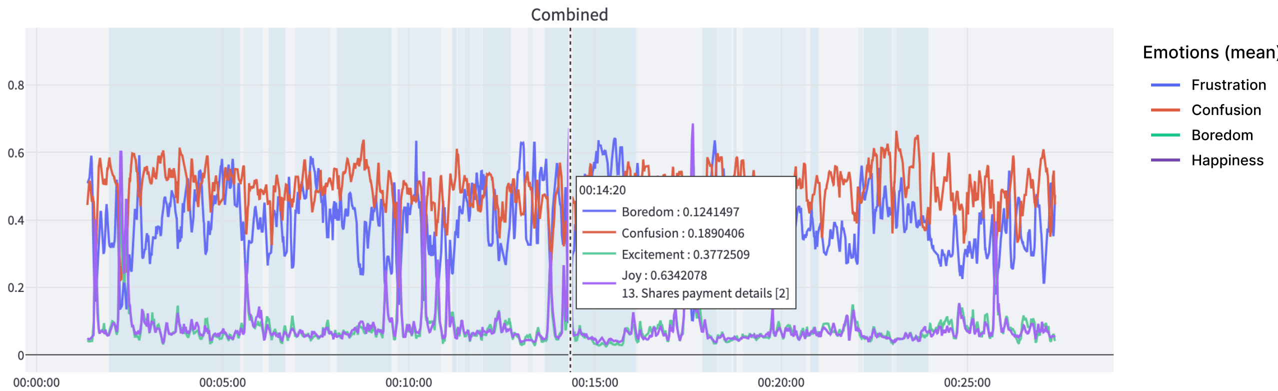


Figure 11: AI Facial Expression Intensity Analysis, 0.0 to 1.0 scale (Frustration, Confusion, Boredom, Happiness).

Table 5: The results of internal assessment of used types of approaches, conducted by the study team

Method	Value	Ease	Accuracy	Potential
CJM, including direct speech	High	Average	High	High
Observation marks	High	High	Average	Average
Self-Reported Data, Emotions	Average	High	Average	Average
Self-Reported Data, Delight	Average	High	Low	Low
AI facial analysis	Low	Low	Low	High

The AI approach turned out to be the most complex one, requiring significant resources. We find self-reporting and observational methods to be the least complex, demanding minimal researcher effort. Given the biases considered (rationalization, recall, social acceptability), we deem direct feedback in CJMs as the most accurate, whereas self-reported Delight data displays inconsistencies with other methods. AI-based FER technique we employed was also rated lower in accuracy, as we observed that despite AI's proficiency in facial expression recognition, the lack of contextual understanding leads to false positives. AI-based FER technique, however, received the highest mark in potential. We anticipate that with technological advancements, AI could become more efficient and provide immediate post-session results. Contextual understanding, crucial for empathetic engagement and broader emotional spectrum analysis, remains a human domain. CJM also received a positive potential rating, suggesting future optimization to reduce complexity of its application. The scalability of the Delight [1, 7, 19, 32] measurement received a *low* mark, as this metric exists in the industry, but there is no consensus on how to model delight. There is limited potential for scaling this method in a specific study, but we recognize its potential for quantitative measurement in a live product.

Overall, employing multiple approaches yielded a more comprehensive overview than using them imdependently.

6.2 Study Outcome

Through questionnaires and observations, we gained insights into customers' emotional interactions with the software, delineating a detailed emotional journey. We identified emotional shifts within this journey, highlighting the most and least engaging moments. Emotional lows pinpointed areas requiring immediate improvement to enhance user experience, which we have incorporated into the UX Backlogs for future software updates. Positive emotional responses were documented as best practices, guiding future product modifications.

7 CONCLUSIONS AND FUTURE WORK

The subjective nature and the complexity of emotional responses make them difficult to quantify and measure. Our study showed that orchestrating several emotion assessment methods can enrich collected data. We find the most important part of this study is its direct results application to improve user experience of the Utility App and the App Marketplace. Part of the study gave us actionable insight into what can be improved. Another part gave data that can show the emotion dynamics change over time. Our approach refers to continuous testing, collecting data, mitigating negative emotional factors, and enhancing positive ones, that leads to higher user engagement, higher user satisfaction, loyalty, retention, and brand perception.

Our future work will focus on continuously extracting actionable enhancements and evaluating their impact on emotional feedback, with measurements planned annually or after major updates. In future evaluations, we will incorporate previous findings and explore advancements in AI to improve the efficiency of our analysis. Another possibility is considering sentiment analysis over user's comments and interview transcripts, as a complementary emotion assessment technique.

Our initial cycle covered two software types, demonstrating the applicability of our methodology. In the next phase, we plan to extend our research to additional products, e.g., to software from the

Security category, with the aim of further refining the applicability of our approach.

ACKNOWLEDGMENTS

We thank Nataliia Stulova for providing assistance in preparing this manuscript. *We thank the Armed Forces of Ukraine for providing security to complete this work.*

REFERENCES

[1] 2022. Exploring Alternative Antecedents of Customer Delight | Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior. (March 2022). https://jcsdcb.com/index.php/JCSDCB/article/view/102

[2] Juliana Alvarez, Pierre-Majorique Léger, Marc Fredette, Shang-Lin Chen, Benjamin Maunier, and Sylvain Senecal. 2020. An Enriched Customer Journey Map: How to Construct and Visualize a Global Portrait of Both Lived and Perceived Users' Experiences? *Designs* 4, 3 (Sept. 2020), 29. https://doi.org/10.3390/designs4030029

[3] Bärbel Bissinger, Anja Beer, Christian Märtin, and Michael Fellmann. 2023. Emotion Recognition via Facial Expressions to Improve Virtual Communication in Videoconferences. In *Human-Computer Interaction*, Masaaki Kurosu and Ayako Hashizume (Eds.). Springer Nature Switzerland, Cham, 151–163. https://doi.org/10.1007/978-3-031-35599-8_10

[4] Callum Bradley, Luis Oliveira, Stewart Birrell, and Rebecca Cain. 2021. A new perspective on personas and customer journey maps: Proposing systemic UX. *International Journal of Human-Computer Studies* 148 (April 2021), 102583. https://doi.org/10.1016/j.ijhcs.2021.102583

[5] M. M. Bradley and P. J. Lang. 1994. Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1 (March 1994), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

[6] Israel C Christie and Bruce H Friedman. 2004. Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. *International Journal of Psychophysiology* 51, 2 (Jan. 2004), 143–153. https://doi.org/10.1016/j.ijpsycho.2003.08.002

[7] Jane Coine. 2007. *Spoil 'em rotten! : five-star customer delight in action.* New York : iUniverse. http://archive.org/details/spoilemrottenfiv0000coin

[8] Chirag Dalvi, Manish Rathod, Shruti Patil, Shilpa Gite, and Ketan Kotecha. 2021. A Survey of AI-Based Facial Emotion Recognition: Features, ML & DL Techniques, Age-Wise Datasets and Future Directions. *IEEE Access* 9 (2021), 165806–165840. https://doi.org/10.1109/ACCESS.2021.3131733

[9] Charles Darwin. 1872. *The expression of the emotions in man and animals.*

[10] Pieter Desmet. 2018. Measuring Emotion: Development and Application of an Instrument to Measure Emotional Responses to Products. In *Funology 2: From Usability to Enjoyment*, Mark Blythe and Andrew Monk (Eds.). Springer International Publishing, Cham, 391–404. https://doi.org/10.1007/978-3-319-68213-6_25

[11] Paul Ekman. 1992. Facial Expressions of Emotion: New Findings, New Questions. *Psychological Science* 3, 1 (Jan. 1992), 34–38. https://doi.org/10.1111/j.1467-9280.1992.tb00253.x

[12] Paul Ekman, Robert W Levenson, and Wallace V. Friesen. 1983. Autonomic Nervous System Activity Distinguishes Among Emotions. *Science* 221, 4616 (Sept. 1983), 1208–1210. https://doi.org/10.1126/science.6612338

[13] Lin Feng and Wei Wei. 2019. An Empirical Study on User Experience Evaluation and Identification of Critical UX Issues. *Sustainability* 11, 8 (Jan. 2019), 2432. https://doi.org/10.3390/su11082432

[14] Nico H. Frijda. 1986. *The emotions.* Editions de la Maison des Sciences de l'Homme, Paris, France.

[15] M. Rosario González-Rodríguez, M. Carmen Díaz-Fernández, and Carmen Pacheco Gómez. 2020. Facial-expression recognition: An emergent approach to the measurement of tourist satisfaction through emotions. *Telematics and Informatics* 51 (Aug. 2020), 101404. https://doi.org/10.1016/j.tele.2020.101404

[16] Udayan Grewal and Vishal Gupta. 2024. AI-Based Facial Emotion Recognition. In *Data Science and Applications*, Satyasai Jagannath Nanda, Rajendra Prasad Yadav, Amir H. Gandomi, and Mukesh Saraswat (Eds.). Springer Nature, Singapore, 205–216. https://doi.org/10.1007/978-981-99-7820-5_17

[17] Richard L. Hazlett and Joey Benedek. 2007. Measuring emotional valence to understand the user's experience of software. *International Journal of Human-Computer Studies* 65, 4 (April 2007), 306–314. https://doi.org/10.1016/j.ijhcs.2006.11.005

[18] Dra. Asunción Hernández-Fernández, Elísabet Mora, and María Isabel Vizcaíno Hernández. 2019. When a new technological product launching fails: A multi-method approach of facial recognition and E-WOM sentiment analysis. *Physiology & Behavior* 200 (March 2019), 130–138. https://doi.org/10.1016/j.physbeh.2018.04.023

[19] PhD Jeff Sauro. 2019. What Is Customer Delight? https://measuringu.com/customer-delight/. Accessed: 2024-03-29.

[20] Brodie Kershaw, Stuart Favilla, and David Sly. 2022. Biosensor measures of human emotion. *DRS Biennial Conference Series* (June 2022). https://dl.designresearchsociety.org/drs-conference-papers/drs2022/researchpapers/111

[21] Seunghyun Kim, Byeong Seon An, and Eui Chul Lee. 2023. Comparative Analysis of AI-Based Facial Identification and Expression Recognition Using Upper and Lower Facial Regions. *Applied Sciences* 13, 10 (Jan. 2023), 6070. https://doi.org/10.3390/app13106070

[22] T. Kunin. 1955. Faces Scale [Database record]. https://doi.org/10.1037/t05610-000. Accessed: 2024-03-28.

[23] Agnieszka Landowska and Jakub Miler. 2016. Limitations of emotion recognition in software user experience evaluation context. In *2016 Federated Conference on Computer Science and Information Systems (FedCSIS).* 1631–1640. https://ieeexplore.ieee.org/abstract/document/7733472

[24] Richard D Lane, Geoffrey L Ahern, Gary E Schwartz, and Alfred W Kaszniak. 1997. Is Alexithymia the Emotional Equivalent of Blindsight? *Biological Psychiatry* 42, 9 (Nov. 1997), 834–844. https://doi.org/10.1016/S0006-3223(97)00050-4

[25] Peter J. Lang, Margaret M. Bradley, and Bruce N. Cuthbert. 1997. Motivated Attention: Affect, Activation, and Action. In *Attention and Orienting.* Psychology Press.

[26] Ganix Lasa, Daniel Justel, Itsaso Gonzalez, Ion Iriarte, and Ester Val. 2017. Next generation of tools for industry to evaluate the user emotional perception: the biometric-based multimethod tools. *The Design Journal* 20, sup1 (July 2017), S2771–S2777. https://doi.org/10.1080/14606925.2017.1352788

[27] Gaël Laurans and Pieter M.A. Desmet. 2017. Developing 14 animated characters for non-verbal self-report of categorical emotions. *J. of Design Research* 15, 3/4 (2017), 214. https://doi.org/10.1504/JDR.2017.089903

[28] William L. Libby, Beatrice C. Lacey, and John I. Lacey. 1973. Pupillary and Cardiac Activity During Visual Attention. *Psychophysiology* 10, 3 (May 1973), 270–294. https://doi.org/10.1111/j.1469-8986.1973.tb00526.x

[29] Yuwen Lu, Yuewen Yang, Qinyi Zhao, Chengzhi Zhang, and Toby Jia-Jun Li. 2024. AI Assistance for UX: A Literature Review Through Human-Centered AI. https://doi.org/10.48550/arXiv.2402.06089 arXiv:2402.06089 [cs].

[30] Catherine Maréchal, Dariusz Mikołajewski, Krzysztof Tyburek, Piotr Prokopowicz, Lamine Bougueroua, Corinne Ancourt, and Katarzyna Wegrzyn-Wolska. 2019. Survey on AI-Based Multimodal Methods for Emotion Detection. Springer, pp 307. https://minesparis-psl.hal.science/hal-02135811

[31] Albert Mehrabian and James A. Russell. 1974. *An approach to environmental psychology.* The MIT Press, Cambridge, MA, US.

[32] Richard L. Oliver. 1980. A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research* 17, 4 (Nov. 1980), 460–469. https://doi.org/10.1177/002224378001700405

[33] Catherine Ordun, Edward Raff, and Sanjay Purushotham. 2020. The Use of AI for Thermal Emotion Recognition: A Review of Problems and Limitations in Standard Design and Data. https://doi.org/10.48550/arXiv.2009.10589 arXiv:2009.10589 [cs, eess].

[34] Ekman Paul and Wallace V Friesen. 1978. Facial action coding system: a technique for the measurement of facial movement. *Consulting Psychologists* 4 (1978).

[35] Robert Plutchik. 1980. A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION. In *Theories of Emotion.* Elsevier, 3–33. https://doi.org/10.1016/B978-0-12-558701-3.50007-7

[36] Amanda D. L. Roberts, Simon Wessely, Trudie Chalder, Andrew Papadopoulos, and Anthony J. Cleare. 2004. Salivary cortisol response to awakening in chronic fatigue syndrome. *The British Journal of Psychiatry* 184, 2 (Feb. 2004), 136–141. https://doi.org/10.1192/bjp.184.2.136

[37] Michael D. Robinson and Gerald L. Clore. 2002. Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology* 83, 1 (2002), 198–215. https://doi.org/10.1037/0022-3514.83.1.198

[38] James A. Russell. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39, 6 (1980), 1161–1178. https://doi.org/10.1037/h0077714

[39] James A. Russell. 1994. Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin* 115, 1 (1994), 102–141. https://doi.org/10.1037/0033-2909.115.1.102

[40] James A. Russell, Jo-Anne Bachorowski, and José-Miguel Fernández-Dols. 2003. Facial and Vocal Expressions of Emotion. *Annual Review of Psychology* 54, 1 (Feb. 2003), 329–349. https://doi.org/10.1146/annurev.psych.54.101601.145102

[41] James A. Russell, Anna Weiss, and Gerald A. Mendelsohn. 1989. Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology* 57, 3 (1989), 493–502. https://doi.org/10.1037/0022-3514.57.3.493

[42] Fahad Ahmed Satti, Jamil Hussain, Hafiz Syed Muhammad Bilal, Wajahat Ali Khan, Asad Masood Khattak, Ju Eun Yeon, and Sungyoung Lee. 2019. Holistic User eXperience in Mobile Augmented Reality Using User eXperience Measurement Index. In *2019 Conference on Next Generation Computing Applications (NextComp).* 1–6. https://doi.org/10.1109/NEXTCOMP.2019.8883528

[43] Lin Shu, Yang Yu, Wenzhuo Chen, Haoqiang Hua, Qin Li, Jianxiu Jin, and Xiangmin Xu. 2020. Wearable Emotion Recognition Using Heart Rate Data from a

Smart Bracelet. *Sensors* 20, 3 (Jan. 2020), 718. https://doi.org/10.3390/s20030718

[44] Gerhard Stemmler, Marcus Heldmann, Cornelia A. Pauls, and Thomas Scherer. 2001. Constraints for emotion specificity in fear and anger: The context counts. *Psychophysiology* 38, 2 (March 2001), 275–291. https://doi.org/10.1111/1469-8986.3820275

[45] Yanika Tueanrat, Savvas Papagiannidis, and Eleftherios Alamanos. 2021. Going on a journey: A review of the customer journey literature. *Journal of Business Research* 125 (March 2021), 336–353. https://doi.org/10.1016/j.jbusres.2020.12.028

[46] Ricardo Cordeiro Galvão Sant'Ana van Erven and Edna Dias Canedo. 2023. Measurement of User's Satisfaction of Digital Products through Emotion Recognition. In *Proceedings of the XXII Brazilian Symposium on Software Quality (SBQS '23)*. Association for Computing Machinery, New York, NY, USA, 62–71. https://doi.org/10.1145/3629479.3629488

[47] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54, 6 (1988), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

[48] Shaun West, Oliver Stoll, Marika Østerlund, Petra Müller Csernetzky, Felix Keiderling, and Christian Kowalkowski. 2020. Adjusting customer journey mapping for application in industrial product-service systems. *International Journal of Business Environment* 11, 3 (2020), 275. https://doi.org/10.1504/IJBE.2020.110911

[49] S. J. Westerman, E. J. Sutherland, L. Robinson, H. Powell, and G. Tuck. 2007. A Multi-method Approach to the Assessment of Web Page Designs. In *Affective Computing and Intelligent Interaction*, Ana C. R. Paiva, Rui Prada, and Rosalind W. Picard (Eds.). Springer, Berlin, Heidelberg, 302–313. https://doi.org/10.1007/978-3-540-74889-2_27