UNIVERSITÉ DE GENÈVE

FACULTY OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

MASTER'S THESIS

Submitted by Abdallah Itani

Under the supervision of Professor Laurent Moccozet

December 2020

Empathic UI: Building Interfaces That Support Users' Emotions

---

Acknowledgements

First, before all, I would like to thank my parents, whom I would not be here without. I thank you for your unconditional love and support throughout my journey and the choices I made.

Also, I thank the University of Geneva for allowing me to have the opportunity to study and complete my master's degree abroad, and everyone who has helped me either directly or indirectly during my study course and with any of my academic accomplishments.

Finally, I would like to express my deepest gratitude to my supervisor, Professor Laurent Moccozet, for his guidance, encouragement, and useful critiques of this work and for providing me with different ideas and perspectives to complete this thesis, especially during the difficult times the paper was written in.

Abstract

Building a user-friendly UI is a hard task and it requires understanding the needs and emotions of the user beforehand. One of the main goals of measuring affect in UI is to better design future products for users, by capturing how they use the current products in their daily lives, and empathic UI design is a step forward towards building interfaces that feel with the user and know what he is going through. A lot of interfaces that we use today lack this element of empathy and instead adapt sympathy or nothing at all, which will often keep users frustrated and unaware of what is going on, or how to further proceed in the current task at hand. The presented work emphasizes the idea of empathy in UI by exploring interfaces that respond empathetically to users by adding visual feedback, to tailor the interface for the user by encouraging and training him when faced with difficult parts of the interface and congratulating him when these difficult tasks are completed.

Contents

1. Introduction

Building user interfaces that suit the needs of all users and copes with all their emotions is not an easy task, it requires a deep understanding of the feelings of the users when using the interface and performing specific tasks, and of what they are struggling with. User interfaces that provide no feedback whatsoever and leave the user alone to figure out how to navigate the interface and perform difficult tasks without any visual/auditory indicators can make it even harder for the user to use the interface and can induce negative emotions and feedback. According to Don Norman, negative affect can make it harder to do easy tasks, while positive affect can make it easier and encouraging to do difficult tasks, and affect is the experience of feeling, emotion, or even mood. Affect is the core of emotions, mood, and attitude towards things, and it has a major impact on how users perform tasks and solve problems they face. Therefore, measuring affect is not uncommon for interface designers as it is an indicator of how the users are feeling when using an interface or performing a difficult task and understanding the emotions of users will eventually lead to a better design choice and will further improve UI design and the interactions between user and machine. Products should follow good human-centered design since negative emotions will make people less able to cope with the difficult tasks and less flexible in their approach to problem-solving which will push them away from the problem, while positive emotions make people more tolerant of minor difficulties and more flexible in finding solutions and being creative [1]. Anne Aula and Veikko Surakka [2] also mention that positive emotional feedback had beneficial effects on human behavior and physiology in the interaction between user and interface. All of this means that UI design should move forward towards designing interfaces that can not only prevent inducing negative emotions, but also induce positive feedback, and the feeling that the interface feels and understands what the user is going through and re-assures him that the task in hand is indeed difficult. Interfaces designed for humans should aim at integrating human emotions.

This is where the designers can implement Empathy in their design. Empathy is the ability to understand, mirror, and share another person's expression and needs, in interfaces, it translates to the designer being able to understand the user's frustrations, hopes, fears, goals, limitations, reasoning, and more. It allows understanding the user to create solutions that will improve the quality of life for them when using the UI by removing unnecessary suffering, or by providing feedback to encourage/congratulate the user. Practicing Empathy in UI and providing affective feedback includes first evaluating the level of affect users have, then using

this evaluation to implement empathic feedback and responses in the interface that can cope with the user's emotions. Song and Yamada [3] discuss the use of color, sound, and vibrations to evaluate how humans interact affectively towards a robot, others like Anna DeWitt and Roberto Bresin [4] proposed the use of different sound design strategies applied in applications that implement affective interaction which give the users auditory feedbacks when specific functions are completed, which would enhance the users' potential to extract meanings from the interfaces and understand what's going on, like sound in SMS messages, or photography applications. Anne Aula t & Veikko Surakka [2] also evaluate how auditory sound can affect the users when solving difficult math problems, psychologically, and physiologically. Sokolova and Antonio [5] discuss how light and different color combinations affect user emotions. But most of these propositions have their limitations. These include usually high cost, technical difficulties, the lack of natural interactions between user and interface, the limitations of the problems being given to the participants in the tests, and the inconsistency between different participants of different age groups, particularly how children react to visual feedback (light and color) differently to adults.

It is undoubtedly required, to significantly improve the interaction between man and interface, to apply affective communication in the implementation and design processes, to create "Empathic UI", that integrates human affect, or human emotions in interfaces, and makes it easier for humans as users to utilize these interfaces and perform tasks without negativity induced emotions like frustration and stress.

This paper will explore a new approach in affective feedback, by building an Empathic UI that adds response elements, either visual or auditory, that will empathically respond to a user. The idea is not to adapt the interface to the emotions of the user, but rather to provide emotional support by encouraging/congratulating the user. The interface will be materialized by building a Maze game with Unity, with multiple levels of increasing difficulty, and the player will be asked to finish all the mazes. Focusing on the affective feedback from interfaces and how it impacts the users' abilities to complete tasks, we wish to explore the following questions:

1. Will providing empathic feedback to the users help them complete the mazes?
2. Should empathic responses be present in every game/interface?

1.1 Thesis Structure

The paper consists of five parts, the first one includes the introduction to the work, and to the idea of affect, affective communication, human emotions in interfaces, which is vital to understanding our proposition and why it matters.

The second part includes the literature review of papers related to the topic of affective UI and empathic interfaces and feedback, in which we will discuss the methods of measuring and evaluating affect from users, and methods of implementing affective feedback to cope with the emotions of the users.

The third part of the paper will include a discussion on Empathy, its importance in UI design, and how we intend to integrate it into our game.

The fourth part will discuss the implementation of the Maze game built with Unity, the different chosen methods of measuring the emotions of the users and providing visual feedback to the users, and the setup of the survey we used to evaluate the users after completing the game. The section will conclude with the results and discussion of the results.

The final part of the paper provides a conclusion and discusses possible future works that can be taken into consideration to further improve the research we have done, the methods of providing feedback, and the methods of evaluation.

2. Literature Review

2.1 Introduction

In this part of the paper, we discuss several proposed methods to measure human affect. These methods can be self-assessment methods in which the participant in the activity currently performing a task will provide affective feedback indicating how he feels about the task at hand or can be other emotion measurement techniques that use physiological signs and body language to deduce the current emotional state of the user. Self-assessment techniques like pictorial tests and questionnaires consist of providing interactions in which the user explicitly expresses his emotional state.

Then we will discuss the different methods for providing affective feedback back to the users, which means sending back feedback to the user, either visual or auditory, with the aim of evaluating how different feedback can modify his emotional state, negatively or positively. These methods are crucial for the notion of Empathic UI since the emotional state of the user will define his ability to complete the task at hand.

The emotion measurement methods vary from one another, and only recently did some researchers start to question whether the classical self-reporting method is good enough to evaluate affective interactions between users and interfaces. The main argument is that self-report methods that are conducted after the experiments are unsatisfactory in estimating the moment-by-moment or real-time experience of a user during his task, and therefore unreliable in assessing their emotions during that moment [6], and some questionnaires might be unreliable or impractical since they put a lot of cognitive load and overall burden on the users and take a lot of time to instruct and complete [7]. Besides, and most importantly for us is that emotion measurement techniques that use physiological signs and other monitoring devices, and affective feedback methods that are focused on interactions using facial expressions and language, even though reliable, are not fully accessible and practical for all: They generally suffer from high cost and technical difficulties and can be a hassle or an intrusion for both the experiment conductors and the participants. In this paper, we aim to use emotion measurement methods that are reliable, low-cost, and widely accessible.

We will conclude this part with a summary of the discussed methods, and the limitations that these techniques present which might make them impractical or inaccessible in some cases.

2.2 Methods for Measuring Affect

Measuring human affect plays an important part in the relation between humans and computers, more specifically to our topic, between user and interface. Human emotions play a key role in how users interact with technology. As Don Norman [1] previously mentioned, negative affect can make it harder to do easy tasks, while positive affect can make it easier and encouraging to do difficult tasks, which means that emotions influence how we perform as users while handling a certain piece of technology/interface. One example of this is adaptive games that cope with the user's emotions to enhance his gameplay [8].

The goal of measuring affect is to eventually evaluate it and do something useful with it. In our case, we aim to encourage/congratulate the player in our Maze game to make the player feel better at a difficult level and help him get over it. Assessing these emotions is becoming more important than ever since it allows us to develop better interfaces, and better understand how the users interact with the interfaces.

The human affect can be measured in multiple different ways, it can be sand self-assessment technique using pictorials and questionnaires, or it can be done by measuring the physiological sign like heart rate and skin conductance, measuring the behavioral changes of the users with their computer and body language, or measuring cognition such as assessing judgment and reasoning. In this section, we discuss methods used by several authors to measure human affect.

2.2.1 The AffectButton

The human affect can be measured in two different ways: Implicit Affect Measurement or Affective Self-Report. Implicit Affect Measurement automatically infers affective information from the user by measuring and evaluating the behavior of the user, meaning the physiological signals. While Affective Self-Report is when the user is asked to provide his judgment on how he feels, his mood, or preferences, which can be done in several ways verbal or picture-based methods, and even questionnaires. Joost Broekensn and Willem-Paul Brinkman [8] [9] discuss a new digital interactive self-report method for measuring the human affect behavior, which is far ahead of the typical methods used is self-report techniques, which are sometimes impractical since the user is often required to be instructed on how to fill up a questionnaire for example.

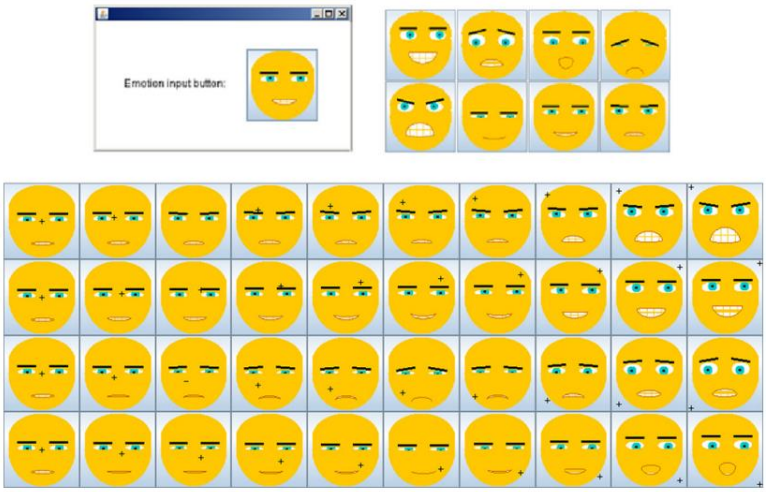


Figure 1: The AffectButton in a simple window (left), its extreme affective states (right), and four example trajectories from neutral to extreme PAD states with corresponding mouse pointer location (down) [8]

The AffectButton, as seen in Figure 1 (left), is a static UI component visible on a page, represented by a round button-shaped face, or "Emoji", that changes dynamically based on the coordinates of the user's mouse pointer on it. It measures affect in terms of 3 dimensions: *Pleasure*, *Arousal*, and *Dominance*, which have values ranging between -1 and 1, and the user should select the face that is most appropriate to his current state of emotions to express his affect levels, the selected face by the user will set the values of the 3 dimensions mentioned above represented by a triple i.e. (1, 0.5, -1). The authors of the paper picked the 3 dimensions as a basis for the measuring system behind the AffectButton since the entire number of emotions cannot be represented clearly in the affect space. The theory they have used, called ***PAD***, is the idea that **P** is related to positivity versus negativity, the **A** represents activation, and the **D** whether the environment is imposing any influence on the user, or the opposite. An example of how this theory works is the representation of "anger" by negatively-arousing-dominant emotion, or "fear" by a negatively-arousing-submissive emotion. P and D will be controlled by the mouse's pointer coordinates.

Since only 2 dimensions can be represented on the 2D face, the authors decided to drop the Arousal dimension and only keep Pleasure and Dominance and leave Arousal to be derived from the start of the inner border (-1.0) to the start of the outer border (1.0), as seen in Figure 2 (right). In the outer border, A is always 1.0, while P and D are mapped to their nearest point on the inner border. It is worth saying that the outer border's only purpose is to provide space

for the user to move so that the user will not move outside the AffectButton when trying to express extreme affect, to make it easier for them to select them without exiting outside the box, and extreme emotions are ones that are associated with high Arousal affect such as rage and terror.

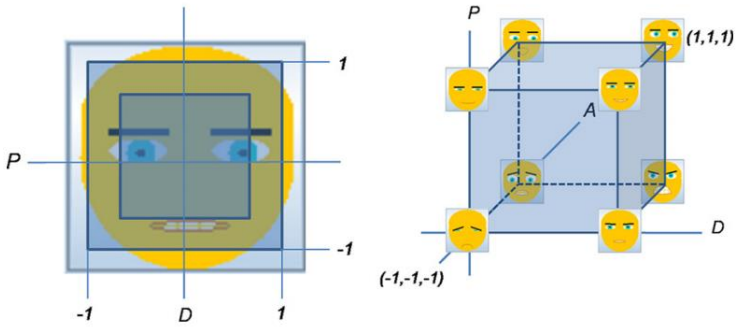


Figure 2: The 2D coordinate mapping to PAD space (left), and the location of the extreme prototypes in PAD space that are used for interpolation (right) [8]

In Figure 2 (left) we see that the D dimension mapped to the Y-axis, and the P dimension mapped to the X-axis, whereas the mouse pointer moves along the Y-axis, the face becomes more open and looks the user in the eye, creating a more dominant posture, if the pointer was moving upwards while moving down makes the face look down creating a submissive posture, where each facial expression is defined in terms of eyebrows, eyes and mouth configuration of the displayed face.

The authors conducted several experiments to test the AffectButton's Reliability which measures consistency, Validity which measures the thing intended to be measured, and Usability. These experiments include rating preferences, words, music, and avatars, and comparisons with the SAM Self-Report method.

They have found that the AffectButton offers good Reliability in terms of measures on PAD factors as indicated by rater-total correlations[1], good Validity indicated by the overall correlations between AffectButton measurements and SAM measurements as well as validated stimuli, a large diverse sample size, and the consistent predictive power of AffectButton ratings for related constructs such as preferences and music. Finally, Usability reports showed that the AffectButton is no different than other rating preferences, requires less effort, and is easier to

[1] Rater-total correlation: The extent to which a particular subject agrees with how all other subject in that study rated the stimuli.

use than other methods that complicate the procedure and require instructions like questionnaires.

The main conclusion that can be drawn from the analysis is that the AffectButton is a reliable, valid, and usable means for gathering affective self-report, contrary to other methods of Affective Self-Report which are often lacking in terms of Usability and require a lot of instructions before being used. But the AffectButton itself has its limitations, such as not being able to conclude the correlations between the different dimensions, and not being able to measure the different categories of emotions.

2.2.2 Self-Assessment Manikin (SAM)

Teah-Marie Bynion and Matthew T. Feldner [10] discuss a pictured oriented questionnaire called the Self-Assessment Manikin (SAM). It has been developed to measure emotional responses from the users to an object or event, using the same 3 dimensions discussed in the AffectButton [8] which are Arousal, Pleasure, arousal, and Dominance.

The SAM contains 5 images for each dimension of the three dimensions (Figure 5), that are to be rated by the users on a 9- or 21-point scale. The users are required to place an “X” next to the figure that most represents how they are currently feeling. Valence is represented as faces ranging from smiling to frowning (positive-neutral-negative), Arousal is pictured as faces ranging from eyes wide open to closed eyes (high arousal-low arousal), and Dominance as small figures to large figures (controlled-in control). Arousal also uses additional imagery to represent high-arousal vs low-arousal as images of explosive-like bursts to small pinpricks (holes). (Figure 4)

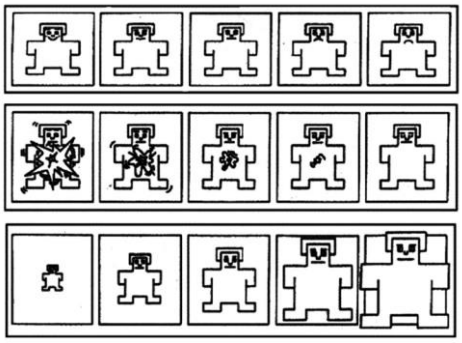


Figure 3: The Self-Assessment Manikin (SAM) showing, from top to bottom, the pictorial scale for valence, arousal, and dominance [8]

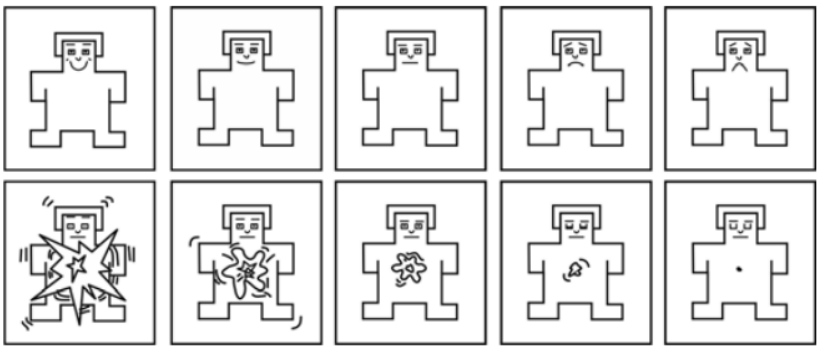


Figure 4: The Self-Assessment Manikin (SAM) showing Pleasure (top) and Arousal scale (bottom) [7]

One of the advantages of SAM is that it is picture-based, so it can be used with any users from any nation and with any demographic since it is basically language-free. Moreover, it is brief, and due to that feature, it can be used to capture emotional responses from the participants using a wide array of simulation methods. SAM is an easy to control, a non-verbal method for quickly assessing the PAD dimensions associated with a user's emotional reaction to some sort of event, and can be employed with a variety of demographics, including non-English speaking subjects, children, etc.

The limitations of affect measuring questionnaires, including cognitive load, overall burden, and time-to-complete can be overcome using methods that rely on visuals, and this is where SAM is very good at. Although the SAM is a very efficient picture-based affect measurement tool, it has its limitations: SAM requires considerable explanation and instructions before the participants can effectively report their feelings for each of the 3 dimensions separately, meaning that this limitation overshadows the brevity and the language-free feature of SAM. It is also intrusive, like verbal self-reports, which will force the users to stop what they are doing to select the picture to express their affective status.

2.2.3 Pick-A-Mood

The authors Desmet, Altenburg, and Romero [7] discuss a character-based pictorial scale that is used by the user for self-assessing their emotional state or mood. Being able to measure mood is extremely important, but measuring mood is different from measuring emotion since inherently mood and emotion are two unique phenomena, but the influence of

both mood and emotions on the user is undeniably similar. Having information about the mood of the user allows designers to optimize UIs and the interactions between user and interface, since mood influences which products people choose to interact with, how they want to interact, and what kinds of information they process and receive during these interactions. Mood measuring will allow future interfaces to be designed with the intent of making the users happier. There have been many mood-inspired product designs that use a method of real-time mood measurement, which is done either by sensors that measure physiological signals like in the case of wearable devices (wristbands or rings) (Figure 5) that will change color according to the mood measurement, which is done by measuring the temperature of the person wearing that ring.

Figure 5: A wearable ring that changes color according to the person's mood[2]

The authors of this paper further discuss the advantages and disadvantages of the different types of measurements, including verbal, behavioral, physiological, and pictorial methods. And they eventually choose the pictorial approach for a given number of reasons such as speed and accuracy, the ability to use these pictures among different cultures, and so on. But they notice not one single pictorial scale available measures distinct mood types but only emotions. So, the authors develop Pick-A-Mood, which is a cartoon-based pictorial self-report scale that rapidly and intuitively enables the measurement of eight distinct mood types. It consists of three sets

[2] Mood ring functionality information here

of cartoon characters (male, female, and robot characters respectively) (Figure 6) where each set includes eight expressions that display distinct moods (Figure 7).

Figure 6: The 3 sets of cartoon characters [7]

Figure 7: The 8 mood expressions (Left: neutral; top left to right: calm, relaxed, cheerful, excited; bottom left to right: sad, bored, tense, irritated) [7]

An experiment was conducted where the participants were asked to generate a label for each of the visual expressions, then select a label from a set of predefined labels (excited, cheerful, relaxed, calm, bored, sad, irritated, tense, and neutral) and assign it to the characters, finally rate the expressions on the basic affect dimensions of valence and arousal using a 5-point scale.

They concluded that the participants were able to select the correct label for the various character expressions and that the eight characters represent a wide range of moods that can be used to measure a user's mood state accurately. But some overlap did occur within the basic mood categories. However, the neutral expression was ambiguous. It was used by the authors to serve as a baseline mood, but some participants labeled the neutral expression as "tense" or related it with "astonishment". And this is because the participants did not know that they could choose neutral as a mood label. So, the neutral label must be used with precaution to avoid confusion, or some sort of instruction might be given to the participants beforehand.

2.2.4 Keyboard and Mouse Behavior

The authors Khan, Brinkman, and others [11] discussed estimating the mood of computer users by detecting and analyzing their behavior on mouse and keyboard. Computers that recognize mood or emotions are considered to be affective and more humane, and with the aim of enhancing the interactions between users and computers, more methods are needed to reliably detect the affect of the users and respond accordingly which may comfort or encourage them. Given the different methods of measuring mood/emotions of the users, and their different drawbacks such as cost or obstructing users, the suggested method of using keyboard and mouse analysis to estimate mood is an alternative way that stays clear of these drawbacks.

The authors conducted 2 experiments to test their method. During normal PC usage, in the first one, the mouse and keyboard usage logs were recorded and stored over several days, and the participants were also required to self-report their mood using the SAM [10] method, to record their emotions on the PAD dimension scale. The self-reports were also recorded and stored. In the 2nd experiment, music was used to manipulate the mood variation of the users, and SAM was replaced by a galvanic skin response measurement (GSR).

The results of both experiments show that measuring mood based on keyboard and mouse interactions is possible for each user. The method suggested by the author is not obstructive and does not require the users to wear special measuring equipment or be interrupted, and it is also widely accessible and practical and requires no additional costs. However, creating a generic mood prediction method is not possible, and is difficult to achieve since each user behaves and interacts differently. And there are also ethical issues that are still in question. Users generally dislike the idea of being tracked and knowing that their computer is logging their keyboard and mouse inputs might be intrusive for the privacy of some users.

2.3 Methods for Providing Affective Feedback

After measuring affective emotions from users, comes the evaluation process of these responses, which will allow us to develop a better way to design interfaces that cope with the emotions of the users not efficiently and most importantly empathically. The evaluation process will help us figure out what kind of emotions certain tasks can provoke, and whether these tasks and their emotional stimulation are helping the user through encouragement and positive attitude or are simply making it harder for the user to complete even the simplest tasks. This

attitude will eventually lead to designing better products in the future that can understand what the user is going through and can support him emotionally with the task in hand by outputting some sort of feedback that can be auditory or visual.

So, in the section of the paper, we will discuss several already explored methods to provide affective feedback back to the users, to see how these kinds of outputs can influence the participants' emotions and their capability to complete the task at hand. Providing affective feedback to the users will help us understand which factors induce positive/negative emotions in the users so that we can develop better products in the future that do not contain any negatively inducing tasks and focus purely on making the users feel good about themselves, and comfortable when facing difficult situations that require critical thinking and creativity.

2.3.1 Sound Design for Affective Interaction

Anna DeWitt and Roberto [4] Bresin discuss a new method for sound design that can be applied in interfaces taking advantage of affective interactions between user and application. Sound design is one of the many affective feedback types that can be used to talk, feel with, and influence the user, it can be positive, enriching, and can quickly capture our attention, but also can be negative, disturbing, and makes us turn away. However, it has mostly been applied to the visual part of the feedback, and not exclusively depending on audio, meaning that sound is usually accompanied by some sort of visual feedback, and sound might become somehow less important in this case, affecting the user less.

Sound carries a lot of information with it, and this calls for new ways to integrate it into affective interfaces as a new way of giving back feedback to users. The authors expand on their proposition by showing how sound works as an affective feedback method and working towards narrowing the gap between real-world sound and sound that we experience in interfaces and applications to further enhance how much information sound can bring to the users who will extract their meaning. Sonification[3] addresses the proposition of using sound as an affective feedback method to cope with the emotions of users, similarly to how music is an answer to some of the emotions we feel.

[3] Sonification: The use of non-speech audio to deliver information or perceptualize data.

2.3.1.1 The Marble and the Container

The first demonstration of sonification is with SMS messages, with the SMS message being the information carried by sound. The authors carry out this demo using an Affective Diary (AD) that focuses on emotions, affects, and physiological signs of the user, with the purpose of getting them to focus on their emotional and bodily status. The SMS messages are visualized by concentric circles and appear on a timeline that shows the time of receipt, and the text of the SMS appears when pointing to one of the circles. The sound that is accompanied by the SMS will be the sound of a marble impact into a container, and the difference in the impact sounds will be configured based on the impact force, weight, material, by the user. And the affective aspect of this example would be that the different sounds of impact will be representing the importance of the SMS message received and the sender, also, the number of SMS messages received can be determined by the user if he shook his device. Following that shaking action, the device would output a unique sound of marbles shaking and clicking together into the virtual container, giving the user an idea of how many marbles or messages he has received.

2.3.1.2 Digital Camera and Photos

Another example that is prevalent today is the camera shutter sound effects. When a user takes a picture using a non-mechanical camera, a shutter sound effect is played as the user takes a picture to let him know that the picture was indeed taken and avoid any feelings of confusion or frustration, the user should not be guessing whether the photo was taken. Also, when scrolling through an image gallery, each photo can play its own unique audio file that was recorded during the time the photo was taken to help the user remember the situation as if he were taking the photo by himself, or for enhancing the sense of presence in the environment where the photo was taken. We notice that these audio feedback features are already present in today's newest smartphones especially the contextual audio from images.

The authors of the paper further dive into more examples of sonification and affective feedback to show the importance of sound in affective communication and the interactions between users and interfaces. There is no denying that sound design is crucial in the design of interfaces and modern applications and can affect the users on an emotional level depending on the different sounds they perceive. But we should be careful in sound design since it can unintentionally

confuse the users, for example not knowing why different sounds are being played, when the sounds are playing, or where the sounds are coming from, so this requires instructing the user on what the different sounds will mean, and when these sounds will be playing. Also, the user should always have the option to customize the different sounds and potentially turn them off completely since sounds playing all the time might get annoying, and this will result in negatively affecting the user and causing unnecessary stress.

2.3.2 Color and Light in Affective Computing

Marina Sokolova and Antonio Fernández-Caballero [5] discuss in their paper how light and colors in interfaces and applications might be an affective factor that can influence the users on an emotional level, and why they should be considered significant and fundamental in the interactions between the two ends.

The authors support the idea that human emotions must be integrated into the interactions between man and machine since these emotions are what affect the users' thinking and problem-solving abilities, and performance when doing difficult tasks.

Various studies were carried out to prove the importance of color in influencing humans' emotions, for example, one experiment consisted of showing drawings made by psychologically healthy and troubled people, in chromatic (colored) and achromatic (no color) versions, and each participant in this experiment had to guess which drawing was made by the psychologically healthy, and which ones by the unhealthy. The conclusion was that the participants made more correct guesses for the chromatic versions of the drawings than the achromatic ones, and that healthy people use more colors for their drawings than the unhealthy ones which use limited colors [12]. Also, the colors surrounding our environments induce emotional changes in people, like different room paintings that are proven to be related to affecting emotions [13]

Light is also another factor that can influence a user. Exposure to bright light and luminance distribution affects our mood, and it can be used to improve emotional states. For instance, several studies have concluded the clinical relevance of light exposure in natural environments for affecting mood and sleep [14]. However, unlike color that is always present in our environments, thus constantly influencing us like in the cases of painted rooms, light is an active factor, which could be labeled as a more proactive and energetic factor.

The combination of both light and color is beneficial for the aims of affective computing. Exposing users to different lighting and colored environments can affect their emotions, mood, and problem-solving skills, which can either enhance the performance of the users or make it worse. But diverse demographics and cultures might react to color-light combinations differently, so the combinations of these two must be used with caution.

2.3.3 Empathic Feedback Interfaces

In this section, we discuss several papers that adopted empathic elements to provide appropriate empathic feedback to the participants, and we show the methods used to measure the emotions of the users, and how they use these measurements to link the emotions of the users with the empathic feedback provided by the elements.
And this is to evaluate how having empathic interfaces can affect users emotionally and psychologically, hence, affecting their performance in the task they are doing.

2.3.3.1 Empathic Interface Agent

The authors Predinger, Becker, and Ishizuka [6] show the effect of having animated character-based UI elements (called agents) on users' emotions if equipped with affective and empathic responses. They will be measuring the physiological signs of the users, like skin conductance, to estimate the participants' emotional experiences during the task they are doing. The physiological approach has been chosen since the physiological signs have always provided consistent and rich information regarding emotions, for example, Ekman [15] investigated the effects of several emotions on physiological signs like heart rate and skin temperature and found that heart rate increases with the emotions of fear and anger. Also, measuring physiological signs allows for real-time estimation of the users' emotions, which are uncontrollable by them so they cannot be easily faked, and the measurement of these signs does not impede the primary task at hand. Following the measurements of the physiological signs, the life-like agent will provide empathic responses to the users that respond to their emotions accordingly.

The users will be participating in a card game of "Skip-Bo[4]" against a virtual opponent, in which the goal is to get rid of the eight cards on the pay-off piles to the right side of the table by playing them to a shared white center stacks. The hand and stock cards must be used strategically to achieve this goal and win the game. **Figure 8** shows the complete setup of the experiment.

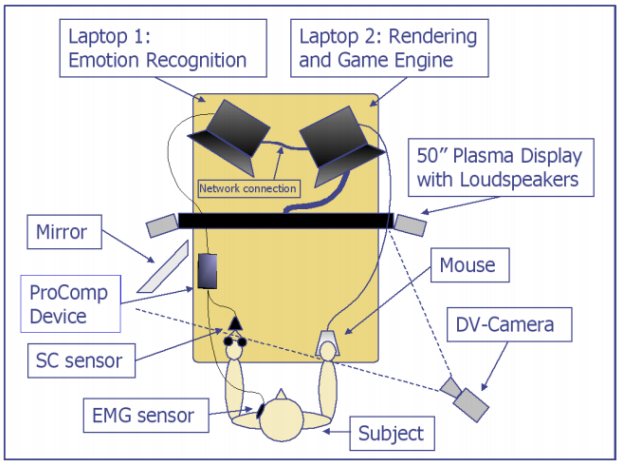


Figure 8: The setup used by the authors [6]

The life-like agent used will not be outputting and speech as it is seen not necessary by the authors, but it will generate affective sounds such as grunts and hums. Feedback was provided by the agents whenever the user was selecting or moving cards, it also gave visual feedback by looking at the cards selected by itself or the user for a short period of time and then looking straight ahead again in the direction of the user, and it also performed other movements in other scenarios. The visual feedback provided by the agent is represented by the facial expressions seen in **Figure 9**, which are constructed using the PAD dimensions (Pleasure, Arousal, Dominance).

[4] Skip-Bo: Full instructions to the card game can be found here.

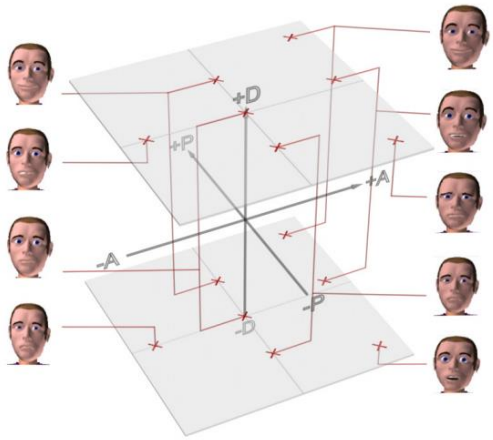


Figure 9: The ten emotion categories in PAD-space (Pleasure, Arousal, Dominance) and the seven attributed facial expressions used for the agent [6]

After considering both positive and negative empathic responses from the agent, the authors concluded with their experiments that displaying positive affect is significantly more arousing or stressful than displaying negative effect, according to skin conductance measurements. Also, the valence of users' emotional responses is consistent with the valence of the emotion expressed by the UI virtual agent. However, although providing promising results, the methods used by the authors to measure the emotions of the users are not very practical and widely accessible for everyone since they are of high cost. They can also hinder the users and might be intrusive.

2.3.3.2 The Empathic Companion

Helmut Prendinger and Mitsuru Ishizuka [16] discuss how interfaces that integrate affective elements can have positive effects on the users of that interface in the specific shown application. They describe the notion of employing lifelike characters in the interface as interaction partners of the user and serve a specific goal which is to educate/train the user, by displaying social cues including speech, communicative gestures, and emotional expressions.

The element adopted by the authors is called the "Empathic Companion", which is a character-based UI element that accompanies users by measuring their physiological signs, for reasons already mentioned reasons [6] [15], and deducing their current emotion in real-time, and copes with their emotions accordingly by immediately providing empathic feedback in form of

synthesized speech by the lifelike character, unlike traditional empathic interfaces that use text. And this is because users are more likely to interact with virtual characters rather than just text. The companion is called Empathic since it shows concern about the users' affect and emotions, which is the definition of empathy.

Figure 10: Job interview scenario showing the interviewer (left) and the Empathic Companion character (right) providing empathic feedback to the user after asking an unpleasant question [16]

All of this is done in a virtual job interview setting (Figure 10). The companion will have a main goal of preparing job applicants for virtual interviews, making them less stressed and better performing. The interviewer will most likely ask unpleasant questions to the applicant and the companion's job is to be empathic with the applicant and feel with him the unpleasantness of that question. The authors concluded from their work that overall, the users receiving empathic feedback are less stressed when being asked an interview question, although they could not show any results of statistical significance.

2.4 Analysis and Synthesis

2.4.1 Summary

The different methods of measuring the human affect vary greatly, some use self-assessment tests, others user pictorial scales, and some use physiological signs. We elaborated on three different techniques that are used to measure human emotions and affect, these are the

AffectButton, SAM, and Pick-A-Mood. Each method offers to measure a user's emotional state (mood for the case of Pick-A-Mood) with accuracy and reliability. The AffectButton is a reliable, valid, and usable method for gathering affective self-report. SAM is picture-based, so it can be used with any person from any nation and with any demographic since it is basically language-free, also it is brief and easy to complete and control. Finally, Pick-A-Mood is the first pictorial scale that offers mood measuring, and it has proven that it represents a wide range of mood expressions that users can identify easily.

Similarly, for the methods of providing affective and empathic feedback, we discussed Sound Design, Color and Light, and Empathic interface elements, each providing different ways to interact with the users empathically and provide affective support according to their emotional status. Sound in interface design has shown its importance in affective communication and the interactions between users and interfaces and can significantly affect the users on an emotional level depending on the different sounds they perceive. Also, the combination of both light and color is beneficial for affective computing, since exposing users to these different combinations can affect their emotions, mood, and problem-solving skills. Lastly, empathic agents deployed in interfaces to support users emotionally greatly influence the users, whether these agents are showing positive or negative expressions, it reflects on the user.

2.4.1 Gaps

The explored methods for measuring affect and providing affective feedback are not perfect, they present their own sets of limitations.

In the methods of measuring emotions and mood, for the case of the AffectButton, it was not able to conclude any correlations between the different dimensions (Pleasure, Arousal, and Dominance), and it was incapable of measuring the full spectrum and categories of emotions. SAM although an improvement over verbal questionnaires that put a cognitive load and a burden on the users in terms of time to complete, still requires considerable explanation and instructions before the participants can effectively report their feelings for each of the PAD dimensions. And it remains intrusive to the users and forces them to stop the task at hand to select the picture best suitable for their emotional status. Pick-A-Mood also reported some overlap within the basic mood categories, meaning that some users were unable to correctly place labels over the mood expressions represented, especially for the neutral mood expression

which confused some of the participants, meaning that Pick-A-Mood requires instructing the participants to some extent, regarding which expressions to label or not.

And for the methods of providing affective feedback, sound design should be proceeded with caution and attention, since overdoing sound design can unintentionally confuse or irritate the users, which might do the opposite of what the initial intent was. Similarly, for the case of light and color, their influence is not universal, different demographics and cultures might react to color-light combinations differently, so the combinations of these two must be used with caution too. Finally, empathic agents present in interfaces although providing promising results, are not very usable and widely accessible for everyone since they are costly and complicated, and the methods used to measure emotions before outputting feedback are mildly intrusive and obstructing to the users. They also require further research and studies since no statistical significance was shown by any of the authors that discussed these methods.

2.5 Conclusion

The different methods we explored for both measuring and providing affective feedback both offer a lot of advantages and conveniences but are also not perfect, they all have their own set of limitations that cannot be ignored. But they do offer a great step towards normalizing affective communication and improving the relationship and the interactions between users and interfaces, and they provide us with great insight on how to implement our own empathic interface, how to measure the emotions of our players accurately, and eventually relate these two aspects together to further reinforce the bond between empathic responses and human emotions, which will allow us to build better interfaces.

However, we must pay attention to the set of limitations discussed and choose wisely a precise method for measuring the emotional (or mood) state of a player, and a method for providing empathic feedback. We will further discuss what we have chosen in section 4 of the paper.

3. Overview of Empathy

3.1 What is Empathy?

Empathy is the ability to understand, mirror, then share another person's expressions and needs. In user interfaces and experience, empathy translates to the interface being able to understand the user's emotions, frustrations, hopes, fears, goals, limitations, reasoning, etc. which will allow creating better interfaces that will improve the quality of life for the users by coping with their emotions and mood and removing unnecessary suffering. Some examples of empathy in UI can be a designer using a website made for the blind, blindfolded, and figuring out what needs to be changed, or a teaching app that congratulates and encourages users on difficult levels which is indirectly telling the user that the application knows that this level is indeed challenging.

There exists a fundamental difference between **Empathy** and **Sympathy** as shown in Figure 11. Sympathy is the acknowledgment of the suffering of others, which translates to acknowledging that the user is having a difficult time, and it is more widely practiced than empathy. It simply does not put the designer of the interface in the shoes of the user, and in no way, does it have the same impact as Empathy when helping the user get through the task he is struggling with, or offering any emotional support. Sympathy ignores what the user is going through most of the time [17].

Figure 11: Spectrum showing the increased engagement for the different types of feedback [18].

3.2 Empathy in Interface Design

To practice empathy, we must first capture the user's emotions/mood when they are performing a certain task. This is done using the methods discussed to measure the human affect in **section 2.2** of the paper. We can either measure the affect of the users during or after the task has been completed and deciding which method to go with depends on a lot of variables such as cost and accessibility. Ultimately, the final goal is to measure user affect to influence the user emotionally.

Integrating empathy into interfaces would be a part of affective computing, which is an interdisciplinary research field that makes efforts to integrate issues dealing with emotions and the interaction with computers, where the idea that emotional states can affect human rational thinking and problem solving is at the core [5]. According to Picard, affective computing "relates to, arises from, or deliberately influences emotions" [19]. Emotions unconsciously regulate and control many aspects of our lives and they are usually triggered by something, usually a stimulus [20]. In interfaces and applications, most of these triggers are difficulties in using the UI or performing challenging tasks, winning the game, etc.

The emotional changes a user can experience when handling a difficult or challenging task can have a major influence on his ability to further complete it. This influence can be either positive or negative and affects the performance accordingly. If it is negative the user will be discouraged to continue with the task and will make it harder to complete, on the other hand, if it is positive, the user will be more encouraged, patient, and creative with the problem he is facing. And the positive influence is what empathy is all about, and what we aim to do in our paper, which is to encourage the users in challenging tasks and congratulate them when they complete them which will motivate them more into completing even harder challenges. This will help reduce the negative affect the users develop when faced with challenging tasks.

There have been many scientific findings that show that humans treat computers in the same way that they treat human colleagues, by being polite or feeling angry or irritated. This means that a computer that provides feedback to the user in a specific tone, that corresponds to this user's emotional state, is more likely to be positively accepted than a technical automated response and affecting the user. The principal goal of affective computing is trying to humanize computers by making them cope emotionally with the users interacting with them [21].

Empathic interfaces have been shown to improve the interaction between user and interface, by leaving the users less frustrated or stressed in the case of a difficult task. Prendinger discussed that the role of empathic agents deployed in the interfaces is not limited to that of a teacher instructing a student, but rather to support the user in achieving a goal, such as receiving emotional support in coping with a challenging problem. The general idea is that if a user's emotion could be recognized by the interface, then the interactions would become more natural, enjoyable, productive, and motivational. The interface will be able to cheer up a frustrated user, encourage him, and eventually congratulate him, and hence react in ways that are more appropriate than simply ignoring the user's affective state as is the case with most current interfaces that adopt sympathy over empathy [6]. Also, in the investigation of Klein and others [22] which is most closely related to our paper on empathic interfaces, they describe the design and evaluation of an interface that implements strategies aimed at reducing negative affect, one of which is empathy. They showed that the deployed empathic agent used in a simulated network game scenario helped undo some of the users' negative feelings after deliberately frustrating them by simulated network delays inserted into the course of the game. This means that empathy shown by the agent helped reduce their frustrations and stress levels significantly.

To conclude, empathic interfaces could cause significant improvements in the relationship between user and interface, contributing to enhance the usability of interfaces, and design better products in the future. Empathy can be integrated into any system or application and can improve the interactions with the user and reduce the amount of negative affect such as frustration or stress when dealing with difficult interfaces.

4. Methods and Implementation

4.1 Introduction

We aim in this paper to demonstrate an Empathic UI in action, that can actively respond to users empathically, and to collect information on whether empathy can influence users into solving the problem they are facing. We also explored the different methods of providing empathic feedback to the users whether it was sound, text, graphics, and methods of measuring the mood/emotions of the user during their gaming session or after the game. To do so, we have created a Maze navigation game with the Unity Engine.

4.2 The Maze

4.2.1 Game Elements

The Maze is a first-person game where the goal of the player is to navigate through the maze and reach the end of the maze. It includes 2 different maze levels, the first level is an easy level maze, while the second level maze is of intermediate difficulty. The levels were built in 3D by referencing 2D maze maps found online using Unity's Pro-Builder tool. The game we built is not lacking on anything as far as casual games go, a Main Menu, Settings, Controls, Pause, Levels screens, and a Win scene are all included into the game so it can feel like a fully built one as much as possible. The different screens of the game are shown in Figure 12.

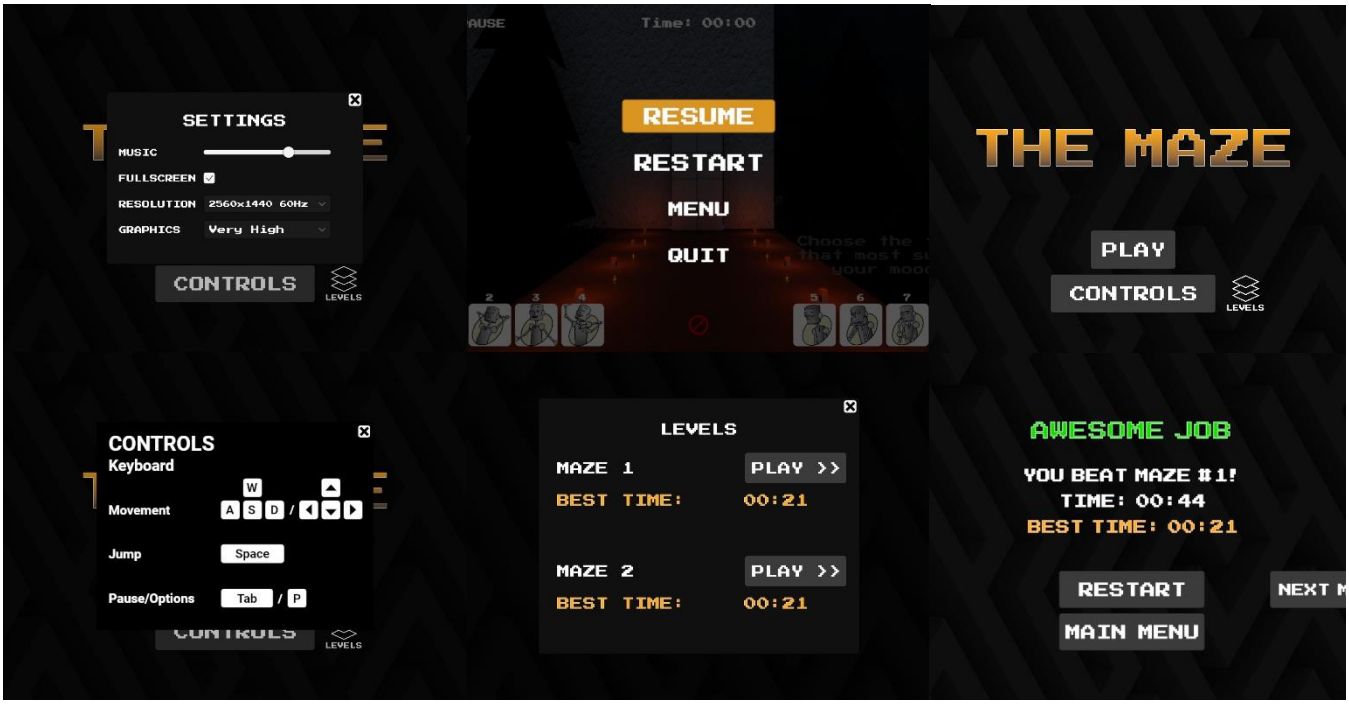


Figure 12: The Settings Menu, Pause Menu, Main Menu (from top left to top right), and the Controls Screen, Level Selector, Win Screen (from bottom left to bottom right).

When the player loads up the first level, he is presented with a set of on-screen instructions that will be discussed in section **4.2.2** of the paper**.** To make sure the player reads the instructions and does not simply start the maze without reading, a barricade was placed on the entrance of the maze that opens when the on-screen instructions in level 1 are finished. The timer starts as soon as the player steps into the maze, which is triggered by the blue floor tile. The player will

know that he has reached the end of the maze by seeing a very noticeable glowing green wall, when the player traverses the green wall, the timer is stopped and recorded, and a current time and best time are kept track of (Figure 12, 13). From there the player can decide whether to load level 2, which will not have on-screen instructions or go back to the main menu.

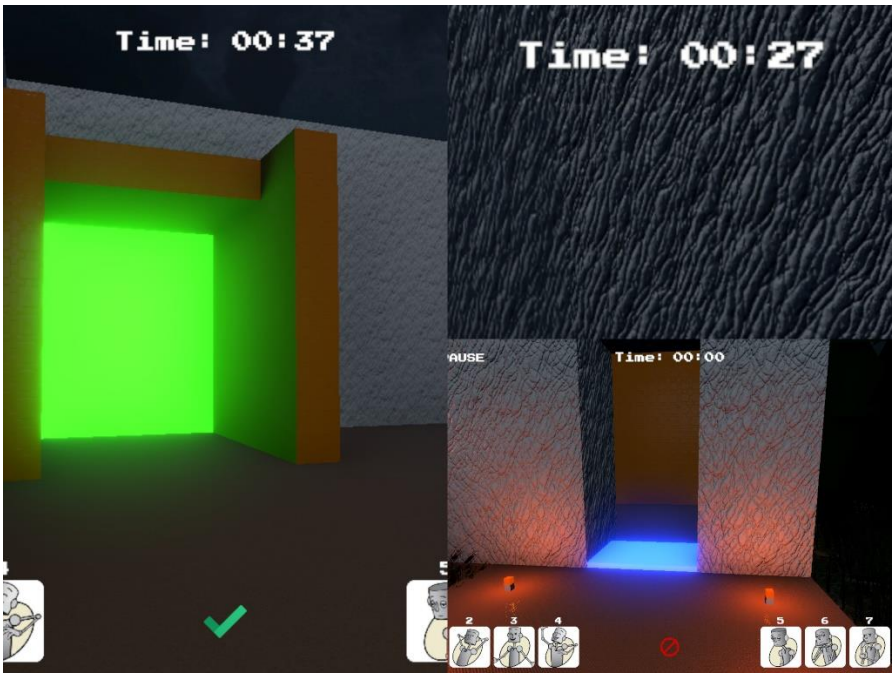


Figure 13: The green end level wall (left), the blue floor tile that triggers the start timer (bottom-right), and the on-screen timer (top-right)

4.2.2 Choice of Empathy Activation Methods

To integrate empathy into our game, we looked at all the different options of measuring the mood/emotions of the player whether it was based on physiological signals, or a self-report method, and the different ways of providing empathic feedback to the player.

To display empathic feedback, we decided to implement text and picture-based feedback into the game, which will be the empathic responses the game will output in the correct scenarios. And regarding the method of measuring the mood or emotions of the player, we decided to integrate Pick-A-Mood [7], as discussed previously in section **2.2.3** of the paper, directly into the game as a feature, so the player would feel like it was part of the game, instead of having an evaluation after the gaming session has concluded.

The decisions of our choices were motivated by the results of the research we have made previously in the paper. Given the limitations of affect questionnaires, including cognitive load, overall burden, and time-to-complete, they can be overcome using pictorial methods like Pick-A-Mood, that enable respondents to report their affective state quickly, intuitively, and accurately, and when cross-culturally validated, they can be used reliably across cultures, because they will not lead to translation complication Compared to verbal self-report scales, pictorial methods are well suited to situations in which participants have limited time or motivation to express their affective state [7], which is suitable in our game scenario. The main advantage of Pick-A-Mood is that it short (only 8 pictures), can be used to measure subtle and distinct mood types, is reliable and easy to administer, because it does not require dedicated instruments or training, and the users only require a very short instruction phase (maze level 1) on how to use this method in our game. Unlike SAM that requires considerable explanation before participants can effectively report their feelings for each factor separately and is too big to be able to implement it into the game interface with ease.

We have also decided to abandon the choice of using physiological methods to measure the players' mood/emotions. Although these methods allow for continuous measurement in real-time, are not intrusive to the user's experience, because they do not require the respondents to interrupt what they are doing and are not biased by cognitive or social appeal constraints, they are costly and therefore not accessible to everyone. Also, interpreting the data from measuring using physiological techniques is not straightforward, and requires specialized training, and the data analysis requires a lot of time and effort due to different factors like mental and physical activity, apart from moods and emotions that influence the measured patterns of psychosocial signals [7].

Time, Dead-Ends, and Mouse Tracking

We implemented different types of empathy activation triggers into the game, some are automatic, and some are manually triggered by the player. Time, Dead-Ends, and Mouse Tracking are the automatic empathy activators integrated into the game.

First, regarding empathy activation with time, when a player takes too long to solve a level, empathy-related to time will be activated and the following texts *"Looks like you're taking too long to solve this maze ...*", *"... it's okay. Mazes take time and are difficult to solve.*", and "*Just take a deep breath and calm down. You can do this*." will be displayed in succession on screen

for a total duration of 10 seconds. The text chosen to be displayed was written in a way to comfort the player that mazes are indeed difficult and might take time to solve and encourage him to further continue playing and solve the maze. The time threshold is currently set at 1 minute and 30 seconds, and empathy will be activated when a player exceeds that threshold.

Second, when the player reaches a Dead-End area inside the maze, for the first time, both the texts "*Oh no! You hit a dead end*.", "*It's going to be fine! It's not unusual to get stuck in a maze*", "*Relax, backtrack, and memorize your mistake. You can do this!*", and a graphical image (Figure 14, left) will appear on-screen, and the player reaches more dead ends further into the game, only the graphical image will be shown to the player without the text. This choice was made to test which method (text or pictorial) of providing feedback to the player was better in empathy and letting the player know that the game knows what he is going through.
An important thing we wanted to test is that although the Dead-End picture (Figure 14, left) can be seen as more informative than empathic, since the first time it is displayed it has empathic text coupled with it, we wanted to see if the players will later associate the empathy they have received from the text, with the image alone. The text chosen in the Dead-End scenario will first let the player realize that he is not the only one that gets stuck in mazes, re-assures him of the difficulty, and encourages him to continue in the end.

Finally, regarding Mouse Tracking, the game will detect when the player makes extreme mouse movements such as rotating clockwise or anticlockwise around himself and will display a graphical image (Figure 14, right), letting the player know that the game knows he/she is lost or confused.

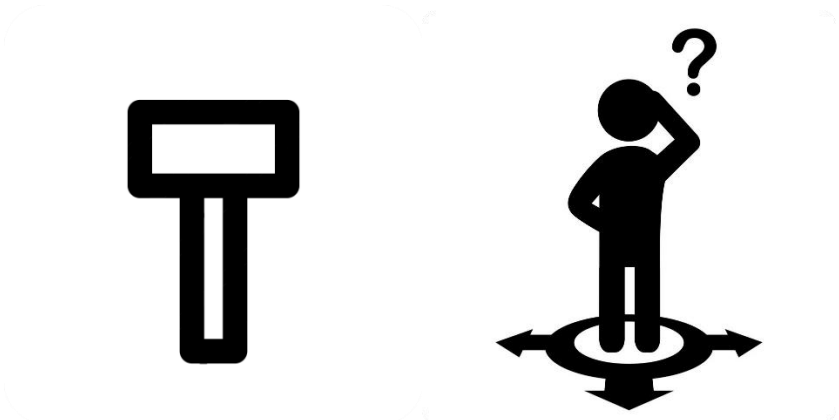


Figure 14: The picture displayed on-screen when a player enters a Dead-End zone (left), and the one when the player is lost/confused according to his mouse movement.

We implemented the Mouse Tracking empathy activation in a way to avoid any accidental empathy activation triggered by users' mouse movement. Empathy from mouse tracking will only activate if a player rotates his mouse more than 720° degrees or less than -720°, if the player exceeds any of those 2 thresholds, empathy from mouse tracking will be activated, then the rotation tracker of the player will be reset to 0° degrees to avoid any further activations that are not meant to happen, and make sure that it can be activated again later. We are convinced with the threshold we have chosen since no one playing casually will rotate around themselves and then some unless he is lost/confused. The threshold can be increased if we notice reports of accidental empathy activations in situations where no empathy is required.

Pick-A-Mood Buttons

When the player loads level 1 of the maze, he is presented with a set of on-screen instructions that will guide him through the process of selecting the mood buttons, like how and when to use them, and most importantly, why to use them. The player is told to use these buttons with his keyboard number keys (1-8) whenever he feels like letting the game know what his current mood is. So, instead of using an emotions/mood self-report method after the game experiment, we have decided to implement Pick-A-Mood directly into the game mechanics. Pick-A-Mood is integrated into the game so that the player can actively self-report his mood while navigating through the maze. The 8 different mood categories are shown as buttons on a level screen, and the player can select a button that corresponds to his current mood using the keyboard number keys (1-8). When a player selects a mood, text-based empathic feedback will be shown on screen, and it will be related to the mood the player has selected, and the button he selected will be greyed out. The feedback showed between the positive moods, which are the moods from 1-4, and the negative moods from 5-8 are slightly different, as showing empathy for positive moods can be analyzed by the player as more congratulating while showing empathy for the negative moods will make the player feel okay with struggling with the levels by displaying text that acknowledges the difficulties in mazes and also encourages the players to keep going. For example, when a player chooses the Calm mood using key #1, a text is displayed reassuring the player of his selection of the mood and congratulating him on being calm, and when he chooses the Irritated mood using key #8, corresponding empathic text will also be shown, empathizing with the player by

acknowledging that mazes are hard, and encouraging him to continue playing. (Figure 15)

Figure 15: The empathy displayed when Calm mood (1) is chosen by the player (left) and when Irritated mood (8) is chosen (right).

The player can only select a different mood every 10 seconds, and his ability to select a mood is shown by the different symbols (Figure 16) between the mood buttons. Also, when any of the other text-based feedbacks are being displayed (Time, Dead-Ends), the ability to choose a mood is locked, to make sure the player reads the current showing text-based feedback on the screen before choosing a new mood. Each mood button can only be selected once, and it will be greyed out when selected. The choice of implementing this constraint is that the maze levels will not take a lot of time to complete, and the player cannot possibly switch between the same moods that very often or go back to his previous moods in a short period of time.

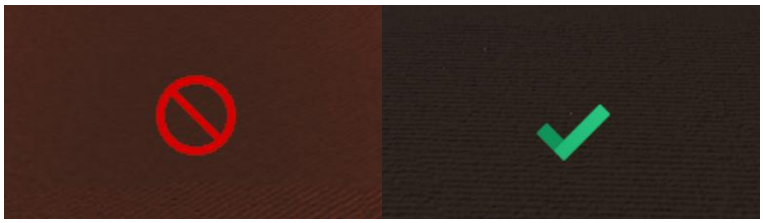


Figure 16: The symbol letting the player know he cannot choose a mood (left), and the one that he can (right).

We also implemented these different types of empathy activation in a way to prevent them from overlapping on top of each other, but not for all of them. For example, the mood buttons will be locked when the Dead-End Zone empathy is being displayed, but the unique picture-

based Mouse Tracking empathy activation will activate regardless of what other empathy methods are active, since it is purely pictorial. And if the Mood-button/Dead-End empathy is active, and the player exceeds the time threshold, the Time empathy will wait until the previous activated empathy method is over.

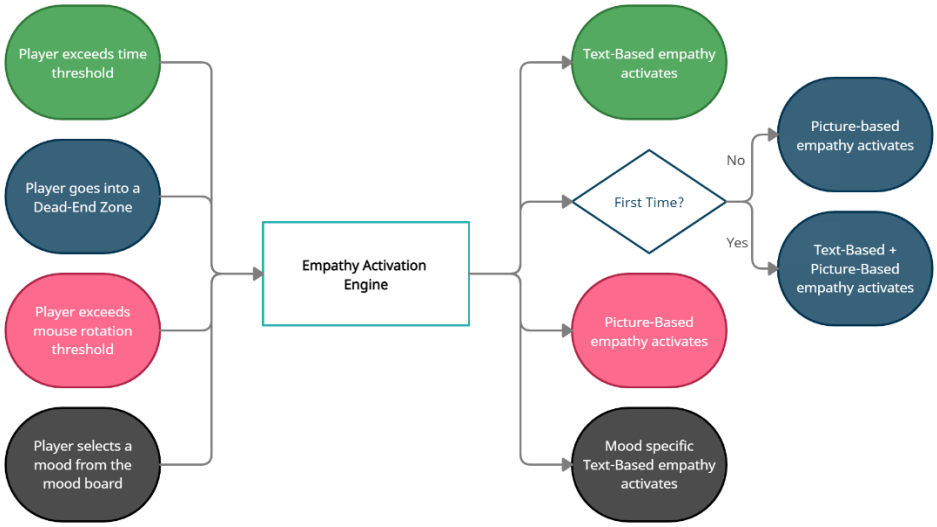


Figure 17: Diagram summarizing the 4 types of empathy activation methods implemented into The Maze game.

4.3 Experiment and Results

4.3.1 User Survey Setup and Participants

To evaluate the efficiency of the empathy activation methods we developed, we designed a survey that will question the participants on the empathic aspects of the game and some casual questions regarding their game performance.

Anyone who has basic knowledge playing games on a computer can participate in our experiment by playing the game and filling out the survey. In total, 18 participants have completed our survey, aged between 19 and 27, mostly male, and have no abilities or skills needed for completing the survey. The only technical requirement of filling out the survey is completing the 2 levels of the game regardless of the difficulty. The survey is divided into 5 sections that include multiple choice, linear scales from 1 – 5 (1 is strongly disagree, 5 is strongly agree), free writing questions, and does not require instructing to the participants or a set of guidelines.

The **first section** will ask the players the casual questions, they are asked to rate the difficulty of both levels, if they were able to complete both mazes, and if the instructions given during level 1 were clear.

The **second section** will ask the players about the graphical images (Figure 14) they have seen during gameplay, that is if they have seen them in the first place, what they think these images meant, and if these images helped them finish the maze by letting them know what was going on. They are also asked about the text that is associated with Figure 14 (left) the first time it showed up, and if it was better than the image in terms of communication and feeling with the player. This question serves as a comparison between the 2 types of feedbacks that are present inside the game, text-based and image-based, and whether they can associate the empathy received from the text the first time, with the Dead-End image if displayed further into the maze.

Section 3 of the survey inquires the players about the mood buttons, whether they have decided to use them at all when playing, and if the responses given by the mood buttons have supported them emotionally and helped them get through the maze. Also, in this section, we assess the level of obstruction/distraction of these mood buttons, and the accuracy of the moods present, meaning whether they depict all the possible moods a player can have. The number of times the mood buttons have been used by the players here does not matter, the only thing that is important for us to assess whether the mood buttons method is an effective and non-obstructive one is whether the players have willingly decided to use the buttons or not.

Then in **section 4**, we talk about time. The player is asked if he has taken more time than usual to finish one of the levels and to assess the feedback he received if he did take a long time, in terms of emotional support and encouragement.

Finally, **section 5** wraps everything up, it asked the participants which feedback type they have preferred in terms of empathy, obstruction, and preference. They are also asked if they felt like the game displaying empathy and communicating with the player, if the congratulating message was good enough after completing a maze, and whether empathy should be incorporated in most games/applications.

The results from this survey should give us enough insight on the empathic elements integrated into the game and their efficiency in supporting players emotionally and encouraging them to complete the level, which type of empathy feedback is more appropriate for the game and

answer our research questions regarding the usage of empathy and the push it can give to players to complete a maze level.

4.3.2 Results

A total of 18 participants played the game and completed our survey. In **section 1** of the survey, 14 out of the 18 participants (77.77%) voted that level 1 was not difficult to solve, on the 2nd level, all the participants (100%) voted that it was difficult to solve. 17 out of the 18 participants (94.4%) solved both levels in the game, and 16 of them agreed that the instructions provided by the game on the first level were easy to understand.

In s**ection 2** of the survey concerning the images, 16 participants (88.9%) saw the picture (Figure 14, right) displayed to them in the game, and when asked about the meaning, most of the participants (55.5%) were correct about their answers by saying that this image means they are lost, confused, and do not know where they are going, that they can ask the game for help, and moving in circles. But also, some interesting answers were mentioned: being at a crossroads, visiting the same path more than once, and that they should change the path to win the level. 9 out of 16 participants (56.3%) voted that the image being displayed was helpful in figuring out what happened.
Then, 17 participants (94.4%) saw the picture (Figure 14, left) displayed when playing the game, and almost everyone corrected interpreted the image by stating it was a dead-end and that they should ask the game for help, with 1 rogue answer of interpreting it as a hint. 11 participants (68.8%) thought that this image was helpful while playing, and all the participants (100%) voted that the text that displayed with the image on the first time reassured them of what has happened and 13 of them (81.3%) thought that it was better than the image in terms of communication and feeling with the player.

In **section 3** of the survey regarding the mood buttons, all the players (100%) have used the mood buttons during the gameplay, with 13 (72.2%) agreeing that the text-based feedback shown by the game after using the mood keys encouraged them or helped them emotionally in some way to finish the maze. Also, 16 participants (88.8%) voted that the mood buttons were not obstructing them and that they felt like a part of/feature of the game. And, regarding the accuracy of Pick-A-Mood, 14 participants (77.77%) voted that the 8 moods accurately depict all the moods that they can possibly have, and nothing felt missing.

Almost all the participants (94.4%) took over 1 minute to solve one of the two levels in the maze, as asked in **section 4** of the survey, and these participants noticed feedback provided by the game. 15 participants (83.8%) voted that the text-based feedback provided by the game when they took too long to solve a level encouraged them to solve the maze or reassured them that mazes are difficult.

Lastly, in **section 5** of the survey which discusses empathy and the players' experiences, 7 participants (38.9%) preferred text-based feedback, 2 participants (11.1%) the picture-based feedback, and the remaining 9 (50%) voted for both. However, 9 participants (50%) voted for the text-based feedback being more obstructing to them, 3 participants (16.7%) voted for the picture-based being more obstructing, 5 participants (27.8%) voted for none of them being obstructing, and 1 participant (5.6%) voted for both being equally obstructing. Also, 15 participants (83.3%) agreed that the feedback provided by the game, either the picture or text-based feedback, felt like the game was communicating with them, with 13 participants (72.2%) saying that this type of communication encouraged them to solve the maze levels.
When explained what empathy was in the survey **(Appendix A, Section 5, 7)**, 11 participants (61,1%) believed that the communication between them and the game was empathic, and 15 participants (83.3%) voted for the text-based feedback being more empathic than the picture-based feedback. The last question in empathy regarding whether empathy should be deployed in all applications/games we use, 14 participants (77.7%) voted for yes, and the other 4 (23.3%) for maybe, and that it depends on the game/application since in some games it might be unfavorable, but in the case of The Maze game, it is suitable to adapt empathy. Also, one of the participants wrote that showing text coupled with the picture for the first time, then showing the picture alone was good since later the player will know what the picture means and its empathic role.

Finally, for the questions regarding congratulating a player, 15 participants (83.3%) agreed that the game appropriately congratulated them after they finished a level. We also included a non-required free writing question in which the participants can tell us how we can improve the congratulating part of the game. All 18 participants added their answers, some of the answers mentioned were flashy screens, a rewards system, showing the path they took, sound effects and animations, and most noticeably a suggestion to show how X amount of people have given up, but this specific player decided to finish it.

4.3.3 Discussion

The results collected for the 18 participants helped us retrieve a lot of valuable information related to our work that will be used to answer our research questions and improve our research in the future. The overall results of The Maze game experiment were positive and helped gain valuable insight into the different types of empathic feedback deployed into the game, the mood measurement method, Pick-A-Mood. The results of the **1st section** of the survey prove that the instructions provided to the player in the first level of the maze were easy to understand, even though they were designed with brevity in mind: each instruction display on the screen successively for a total time of 15 seconds. Also, the players validated the increasing difficulty of the maze levels, as designed.

The **2nd section** of the survey gave us a lot of insight into the images we used in the game to display empathy to the players. The first picture (Figure 14, left) was correctly identified by almost all the participants, and this is due to the scenario in the game in which this picture will show up, and due to its familiarity with street signs. The text that is coupled with this image when first displayed, allows us to compare the 2 types of empathy feedback directly. The results indicate that the text-based feedback was better than displaying an image alone in terms of communication and feeling with the player. The 2nd image (Figure 14, right) was also correctly identified by the majority of the players.

The results also indicate that both images used were helpful in reassuring the players about their situation in-game and that our choice of images was rather good. Also, most of the players saw these images display on their screen, and this is due to the increased difficulty in the 2nd level, and our requirement to solve both maze levels before completing the survey, which guarantees that the players will activate and experience all 4 empathy feedback activation methods integrated into the game.

In **section 3** of the survey, the results showed that integrating the mood buttons into the game rather than having it as a post-game self-report method was successful. All the players used the mood buttons even though it was not mandatory, they were given a choice to let the game know how they feel, which means that the players felt like the mood buttons were part of/ a feature of the game. The results of this section also indicate that the mood buttons offered empathic feedback to the players that were helpful to them in terms of encouragement and emotional support, were not obstructing, and were accurate in depicting all the moods they can have. Pick-

A-Mood has proven to be well suitable for our game due to its brevity, ease of use, minimal instructions, and can be thought of as part/feature of the game.

Section 4 of the survey allowed us to investigate empty activation via time. The 2nd maze level was designed with difficulty in mind, and this is to allow all our empathy methods to be activated, and the results have validated our design choices. The feedback displayed by time empathy activation has shown to be helpful to the players, encouraging, and reassuring.

The final section of the survey, **section 5**, allowed us to analyze empathy in our game and the feedback types. Overall, the text-based feedback has proven to be more empathic, but also more obstructing than the picture-based feedback which is less empathic and less obstructing. However, the overall feedback by the game has shown to be empathic by the players, communicating with them, and helped them get through the maze by encouraging them and displaying emotional support. The results have shown that empathy in fact plays an important role in emotionally supporting the players by encouraging them in the game depending on their current mood and performance, and that empathy should be more widely available in more applications/games that we use daily. Both the text-based and picture-based feedback were accepted by the players, but the level of obstruction should be taken into consideration. They have also shown that our way to congratulate the player after finishing a level was good and appropriate but can be improved, hence why all the participants added comments on how to improve it in a non-required free text question.

These results allowed us to answer the research questions we have proposed in the paper. We have found that empathy played a role in supporting the players emotionally while playing, since most of the players acted positively towards the text and picture-based feedbacks, used the mood buttons willingly and felt that the game was encouraging them to solve the mazes. Also, we can say that empathy should in fact be present in most of the applications/games/etc. we use daily given the overall positive feedback on empathy and its presence in the game. The results also helped outline which mood measuring method, and which type of empathic feedback is more suitable to be implemented in these types of games.

Off-Survey Discussion

Other interesting answers from the participants also showed us how displaying empathy can be annoying or insulting by some players. 2 participants have reported separately that they

felt like the game was mocking them, especially when they took too long to solve the maze. They felt like the text-based feedback was making fun of them for taking too long, which made them feel uncomfortable or bad at the game, and that other players were better. Which is the opposite of what our intention was with the empathic feedback. One of these players reported that text-based feedback when playing a game was too distracting and that they did not care about their emotions, and their only goal was to finish the maze in the fastest time. Another player reported that showing the timer in the game was adding unnecessary pressure to complete the level.

5. Conclusion

The research we have made in our paper gave us a lot of insight into the importance of empathy and the improved interactions it enables with the users. We have built a complete maze navigation game and implemented text-based and picture-based empathy, and integrated Pick-A-Mood directly into it as an almost perfectly non-obstructive and non-intrusive game feature like most of the results indicate, differently to what Pick-A-Mood should be, which is a post-experiment self-report method. The game we made has 4 different methods for activating empathy, all of which will properly support the players emotionally by encouraging them to continue playing and acknowledging their moods. We were able to answer our research questions since we have found that empathy played a role in supporting the players emotionally while playing and noted that empathy should in fact be present in most of the applications/games/etc. we use daily. This research also helped outline which mood measuring method, and which type of empathic feedback is more suitable to be implemented in games.

Empathic interfaces are the future, and they are necessary if our goal is to improve the interaction between users and interfaces and design better products that are easier to use, less frustrating, and more supporting.

5.1 Limitations and Future Work

Several limitations of our research are worth mentioning. The first important limitation is that self-report methods like Pick-A-Mood, cannot be used for continuous measurement because the participants must interrupt their activities to record their responses. Moreover, the

response from the players can potentially be biased because they rely on the respondents' ability and willingness to report their feelings. Second, similarly to the verbal self-report, pictorial self-reports are also intrusive, as the players must stop what they are doing to select visuals to express their affective state.

We intend to further improve the work we have achieved here, by expanding the game into more levels and add more features that have been requested by the players such as better ways to congratulate the player when completing levels. Second, we want to further improve the method of self-reporting in game and make the interactions between the mood selections and the game more natural, since some of the players did find Pick-A-Mood to be obstructive and not a part of the game. Also, we want to make sure that the empathic feedback provided by the game will not be seen by some players as mocking or demeaning, and this can be done by improving the sentences displayed to the players and reformulating them and curating the sentences to each player according to the different moods he has selected before by adding different levels of intensity to each sentence. In addition, we want to expand on one of the features embedded into the game in section 4.2.2 (Dead-End Empathy) which tests if the players can associate the empathy from text with the pictures that are coupled with it when the pictures are displayed without the empathic text. We also intend to add sound into the game. Since sound plays an important role in the gaming experience, we will investigate the possibility of using sound to provide empathy for the players in different scenarios and compare the sound-based feedback to text and picture-based ones and potentially explore different combinations. Finally, we could potentially log all the players mood selection during their gameplay by also taking time and their position in the maze into account, and then use this data to integrate a machine learning engine into the game that can estimate the mood of the players without any input from them. This would be a way to overcome the lack of continuity from Self-Report mood measuring techniques and make the game more seamless. Logging the different feedback activation methods coupled with time and the location of the players can also be used to better analyze the different types of activation methods we have used and identify hotspots in the maze levels where the methods activate the most, and analyze how the players' moods evolve during gameplay.

References

[1] Norman, D. A., “Emotion and design: Attractive things work better”. *Interactions Magazine, ix* (4), 36-42. (2002).

[2] Aula A., Surakka V. “Auditory Emotional Feedback Facilitates Human-Computer Interaction”. In: Faulkner X., Finlay J., Détienne F. (eds) People and Computers XVI - Memorable Yet Invisible. Springer, London. (2002).

[3] Sichao Song and Seiji Yamada. 2017. “Expressing Emotions through Color, Sound, and Vibration with an Appearance-Constrained Social Robot”. In Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17). Association for Computing Machinery, New York, NY, USA, 2–11.

[4] DeWitt, Anna & Bresin, Roberto. “Sound Design for Affective Interaction”. (2007). 523-533.

[5] Sokolova, M.V.; Fernández-Caballero, A. “A Review on the Role of Color and Light in Affective Computing”. *Appl. Sci.* (2015), *5*, 275-293.

[6] Prendinger, H. et al. “A Study in Users' Physiological Response to an Empathic Interface Agent.” *Int. J. Humanoid Robotics* 3 (2006): 371-391.

[7] Desmet, Pieter & Vastenburg, Martijn & Romero, Natalia. (2016). Mood measurement with Pick-A-Mood: review of current methods and design of a pictorial self-report scale. J of Design Research. 14. 241-279.

[8] J. Broekens and W. Brinkman, “AffectButton: A method for reliable and valid affective self-report”, International Journal of Human-Computer Studies, Volume 71, Issue 6, (2013), Pages 641-667, ISSN 1071-5819.

[9] J. Broekens and W. Brinkman, “AffectButton: Towards a standard for dynamic affective user feedback”, 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, Amsterdam, (2009), pp. 1-8.

[10] Bynion, Teah-Marie & Feldner, Matthew. “Self-Assessment Manikin”, (2017).

[11] Khan, Iftikhar & Brinkman, Willem-Paul & Hierons, Robert. (2013). Towards estimating computer users’ mood from interaction behaviour with keyboard and mouse. Frontiers of Computer Science. 7

[12] Jue, J.; Kwon, S.M. “Does colour say something about emotions? Laypersons’ assessments of colour drawings”. Arts Psychother. (2013), 40, 115–119.

[13] Harleman, M. “Colour emotion in full-scale rooms”. In Proceedings of the AIC 2004 Color and Paints, Interim Meeting of the International Color Association, Porto Alegre, Brazil, 3–5 November 2004; pp. 223–226.

[14] Dumont, M.; Beaulieu, C. “Light exposure in the natural environment: Relevance to mood and sleep disorders”. Sleep Med. (2007), 8, 557–565.

[15] P. Ekman, R. W. Levenson and W. V. Friesen, “Autonomic nervous system activity distinguishes among emotions”, Science 221 (1983) 1208–1210.

[16] Prendinger, Helmut & Ishizuka, Mitsuru. “The Empathic Companion: A Character-Based Interface That Addresses Users' Affective States”. Applied Artificial Intelligence. 19 (2005): 267-285.

[17] Gibbons, S. “Sympathy vs. Empathy in UX”. (2019).

[18] Alli, E. “UX design and empathy: are we doing it right?”, UX Collective, 2019.

[19] Picard, R. “Affective Computing”; The MIT Press: Cambridge, MA, USA, 1997.

[20] Strongman, K. “The Psychology of Emotion: From Everyday Life to Theory”, 5th ed.; John Wiley & Sons: Hoboken, NJ, USA, (2003).

[21] Reeves, B.; Nass, C. “The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places”; CSLI Publications: Stanford, CA, USA, (1996).

[22] Klein, J., Y. Moon, and R. Picard. “This computer responds to user frustration: Theory, design, and results”. Interacting with Computers. 14:119–140. (2002).

[23] Unity Real-Time Development Platform, unity.com.

Appendix A

Survey Questionnaire[5]

The Maze - An Empathic Game

Postgame survey! The following questions are to be answered after you have completed the first 2 levels of The Maze. These questions are related to the feedback you have received from the game during your maze escape and your overall experience.

* Required

Section 1 of 5 – General Questions

1- Was level 1 difficult to solve? *

Strongly Disagree 1 2 3 4 5 Strongly Agree

2- Was level 2 difficult to solve? *

Strongly Disagree 1 2 3 4 5 Strongly Agree

3- Did you solve both levels? *

Yes No

4- "The instructions provided by the game in the first level were easy to understand." *

Strongly Disagree 1 2 3 4 5 Strongly Agree

Section 2 of 5 – Picture-based Feedback

Picture 1

[5] Full survey results: The Maze - An Empathic Game (Responses) - Google Sheets

1- Did you see the above picture (Picture 1) show up while you were navigating the maze? *

Yes No

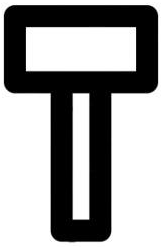
2- If yes, what do you think it meant when the game showed it to you?

[Short Answer Text]

3- Did it help you in any way in figuring out what was going on?

Yes No

Picture 2



4- Did you see the above picture (Picture 2) show up while you were navigating the maze? *

Yes No

5- If yes, what do you think it meant when the game showed it to you?

[Short Answer Text]

6- Did it help you in any way in figuring out what was going on?

Yes No

7- The first time it showed up, text was also displayed with it, did this text help reassure you in some way what has happened?

Yes No

8- Was it better than displaying the picture in terms of communication and feeling with the player?

Yes No [Other]

Section 3 of 5 – Mood Buttons

1- Did you use any of the mood buttons with your number keys (1-8) while playing? *

Yes No

2- "The text-based feedback shown by the game after using the mood keys encouraged me or helped me emotionally in some way to finish the maze."

Strongly Disagree 1 2 3 4 5 Strongly Agree

3- "Using the mood buttons was not obstructing me while solving the game, and felt like a part of/feature of the game."

Strongly Disagree 1 2 3 4 5 Strongly Agree

4- "The mood buttons accurately depict all of the moods that I can possibly have, and nothing felt missing."

Strongly Disagree 1 2 3 4 5 Strongly Agree

Section 4 of 5 – Time

1- Did you take over 1 minute to solve some levels of the maze? *

Yes No

2- If yes, did you notice any feedback provided by the game?

Yes No

3- "The text-based feedback provided by the game when I took too long to solve a level encouraged me to solve the maze, or reassured me that mazes are difficult"

Strongly Disagree 1 2 3 4 5 Strongly Agree

Section 5 of 5 – Empathy and Experience

1- Which method of feedback did you prefer: text-based, picture-based, or both? *

Text-Based Picture-Based Both None

2- Which type of feedback was more obstructing to you? *

Text-Based Picture-Based Both were Equally Obstructing None of the Feedback was Obstructing

3- Given the feedback provided by the game, either the picture or text-based feedback, did you feel like the game was communicating with you? *

Yes No Sometimes

4- Did this type of communication encourage you to solve the levels? *

Yes No [Other]

5- Did the game appropriately congratulate you after you finished a level? *

Yes No [Other]

6- How would you improve congratulating a player? *

[Short Answer Text]

7- "Empathy is the ability to understand and share the feelings of another". Did you feel like the communication between you and the game was empathic? *

Yes No Maybe [Other]

8- Which type of feedback displayed more empathy from the game? *

Text-Based Picture-Based

9- Do you think more games/applications should have this aspect of empathy, which can improve communication between users and interfaces? *

Yes No Maybe [Other]