An Affect-Aware Intelligent Tutoring System for EmoRegulators: A Restricted-Perception Wizard-of-Oz Approach

Pedro Rodrigues*
Instituto Superior Técnico,
Universidade de Lisboa

João Dias†
INESC-ID & Instituto
Superior Técnico,
Universidade de Lisboa

Ana Paiva‡
INESC-ID & Instituto
Superior Técnico,
Universidade de Lisboa

Abstract

This project addresses the problem of how to create an intelligent, adaptive system capable of being aware of, understanding and regulating players' emotions during an interaction session in EmoRegulators, a serious game that teaches players to be aware of their emotions and to regulate them. We conducted research into other affect-aware Intelligent Tutoring Systems to learn which emotion regulation strategies they use. Unlike many of these emotion-aware ITS which use a set of rules to activate the usage of emotion regulation strategies, we propose a model heavily based on restricted-perception Wizard-of-Oz studies, where a human expert plays the role of a tutor during a demonstration phase by controlling the facilitator (an instructional agent) in EmoRegulators, using an interface which was carefully designed to enforce that the expert has the same perceptions as the agent. We created a set of features that represent the game state and the affective state that could be used later in the training phase, using the data collected from users in the WoZ demonstrations. This data was pre-processed and fed to a selection of machine learning algorithms to study their performance in EmoRegulators specifically. Users reported an overall positive experience during the demonstrations and we achieved performances upwards of 88% in accuracy and 90% in F1 score for the kNN, decision tree and random forest using a Leave-One-Subject-Out approach. The very good generalization results for new subjects seem to show meaningful steps towards solving the posed problem.

Keywords: Emotion Regulation; Intelligent Tutoring Systems; Serious Games; EmoRegulators; Wizard-of-Oz Studies; Machine Learning

1 Introduction

Researchers have been attempting to use intelligent autonomous agents for educational purposes [8]. Their goal is to replicate a real teacher's success as best as possible in the form of tutoring systems or pedagogical agents, which intend to teach students a particular subject or method, or to perform specific tasks.

According to Malekzadeh et al., an "Intelligent Tutoring System (ITS) is a computer-based educational system that provides individualized instructions similar to like a human tutor. Typical ITSs determine how and what to teach a student based on the learner's pedagogical state to enhance learning." [21].

Additionally to considering the learner's pedagogical state it is equally important to address the learner's emotions. As humans we resort to emotions to guide our thinking process all the time. Ahn and Picard have recently found that emotions affect several cognitive and behavioral processes, such as decision-making [1]. Furthermore, the social cognitive theory tells us that learning is highly impacted by the affection-cognition correlation [17].

By now we all know how most times it is the methodology of a teacher that separates the great ones from the not-so-good ones. Not only that, but it is also their emotional intelligence and their ability to maintain good interpersonal relationships with their students, that often dictates the latter's overall success in learning [26].

To better understand the concept of emotional intelligence we can refer to Salovey and Mayer [33] who consider emotional intelligence as a subset of social intelligence that involves assessing one's own and others' emotions, as well as identifying them, and using that information to guide one's thinking and actions.

The way we manage emotions is also very important, especially in the context of relationships. Our ability to control which emotions we express, how intense they are and for how long they last can dictate how beneficial or harmful they can be in a given situation [16]. This is only a fraction of what is called emotion regulation, and Gross defines it as "the heterogeneous set of processes by which emotions themselves are regulated" (a more in-depth look at emotion regulation can be found in the following section). Managing our and others' emotions in such a way is also one of the four core skills that comprise emotional intelligence according to Mayer et al. [27].

Given the relevance of emotions to behavior, it should come as no surprise that researchers have already attempted to explore emotion-aware ITS in hopes to replicate the success of real teachers. D'Mello and Graesser's Affective AutoTutor [7], for example, is able to formulate empathetic replies to negative emotional states such as frustration and boredom, as well as attempting to help a student recover from a state of confusion when he/she doesn't understand something - a form of emotion regulation in itself. However, these affective capabilities didn't have different results compared to their regular AutoTutor. Despite this particular example, in his review of emotion regulation in ITS [21] Malekzadeh et al. concluded that emotion-aware ITSs have better results than their non-emotional counterparts, namely achieving more positive emotions [5, 36] and satisfaction and positive impressions [24] in the learners.

The above cited examples achieve emotion regulation in the user by monitoring facial-features, speech contours, body language, interaction logs, language and peripheral physiology, to name a few [4]. Focusing on the latter, The EmoRegulators [34] , a serious game for learning how to regulate emotions, attempts to teach humans to regulate theirs through a series of BEAR-based (see [30] for the original BEAR) sessions composing a game, by measuring physiological and biological data such as heart rate and electrodermal activity (EDA). Despite its ability to teach these emotion regulation mechanisms to the player, the EmoRegulators lacks the adaptability and the autonomy that ITSs should pack to thrive. The system isn't capable of acting upon stressful or frustrating emotions that the player might feel while performing the proposed sessions. This poses the following problem:

"How to create an intelligent, adaptive system in EmoRegulators, capable of helping to regulate players' affective states during an interaction session?"

Several affect-aware ITS at the time of writing use rule-based

*e-mail: pedromcordrigues@ist.utl.pt
†e-mail: joao.dias@gaips.inesc-id.pt
‡e-mail: ana.paiva@inesc-id.pt

systems to activate the usage of emotion regulation strategies after assessing a user's emotional state. Auto-tutor, for example, does this by mapping dynamic assessments of the students' cognitive and affective states with appropriate tutor actions [7].

On another note, teachers are also responsible for teaching classes that often have students with different learning curves and cognitive abilities; good teachers can guide both types of students at the same time with optimal efficacy and help them achieve good results, which is something ITSs have struggled with in the past due to expertise-reversal literature, which basically states that methods which promote learning for low-knowledge students can impair the knowledge of high-knowledge students [6]. Also noteworthy to achieve good results, and as corroborated by Moreno's conclusion, is the fact that the most successful key point of pedagogical agents - which ITSs are an example of - lies in the specific instructional method embedded in the agent [28].

This idea that meeting a student's specific learning needs in pedagogical environments is key, coupled with the aforementioned conclusions regarding the efficacy of emotion-aware ITS when compared to their non-emotional counterparts, brings us to the hypothesis we propose in this paper:

"The usage of emotion regulation strategies based on expert knowledge from human tutors will help emotion-aware intelligent tutoring systems better replicate their ability to understand and regulate users' emotions."

In the following chapters we delve into the concepts and state-of-the-art involved in ITS and serious games. The next section, Theory and Background, introduces some emotion regulation strategies that have been used by researchers in ITS. Then we take a look at the state-of-the-art when it comes to emotion regulation in ITS, emotion regulation in serious health games - including a detailed look at EmoRegulators - and work based on a well-known learning framework that is used to teach expert knowledge and methods to agents. In the fourth section, we describe our solution for the identified problem and then discuss results and draw conclusions regarding the conducted work.

2 Theory and Background

Several views on emotion regulation inspired decisions made in this work. Gross defines emotion regulation (in the sense of the regulation *of* emotions) as a diverse set of processes that regulate emotions, simply put. In 2001 [15] Gross proposed a process model of emotion regulation that divides emotion regulation strategies into five families according to the point in time when they have their primary impact on the emotion-generative process.

Other researchers have studied emotion regulation strategies from different perspectives. These concepts are introduced so that the reader can easily understand certain aspects mentioned in this article.

Attribution Theory While there are many attribution "theories", the common ideas are that people interpret behavior in terms of its causes and that these interpretations play an important role in determining reactions to the behavior [18]. In the particular case of AutoTutor, for example, it attempts to attribute the cause of a certain situation, such as the student feeling bored, to a particular factor, entity or event, such as the learning content being uninteresting, the tutor itself or the student.

Cognitive Disequilibrium Theory In his theory of cognitive development, Jean Piaget describes cognitive disequilibrium as a state of cognitive imbalance [29]. One experiences such a state of imbalance when acquiring information that requires us to develop a new schema (a building block of knowledge) or modify an existing one. For example, a child learning how to tie their shoes has to physically maneuver the shoe laces while thinking through the steps as they try to develop a new schema for shoe tying.

Cognitive Reappraisal Cognitive reappraisal is the process by which a person changes his/her interpretation of a certain emotional response by reinterpreting the meaning of the emotional stimulus [31]. This emotion regulation strategy involves (1) recognizing one's negative response and (2) reinterpreting the situation to suppress or overcome the severity of that response.

3 Related Work

3.1 Emotion Regulation in ITS

The problem of attempting to simulate human success in autonomous tutoring systems is something that goes a while back; several agents, models and architectures with strong psychological foundations have been developed to achieve this purpose. Thus, it is deemed important to understand how these systems regulate emotions, how they adapt and change over the course of a tutoring session, and how they decide whether it's necessary or beneficial to intervene.

3.1.1 AutoTutor and Affective AutoTutor

AutoTutor is an intelligent tutoring system that teaches concepts of Newtonian physics to students using adaptive dialog in natural language to resemble a human tutor as much as possible. Its affective version features two implementations, the Supportive and Shakeup AutoTutors, which detect and act upon students' affective and cognitive states. A set of affect-sensitive production rules map dynamic assessments of the students' cognitive and affective states with appropriate tutor actions. Of particular interest are the two main strategies used by these alternate versions to regulate students' negative affective states: the attribution theory and the cognitive disequilibrium theory, which address the feelings of boredom/frustration and confusion, respectively. In the supportive version, for example, the agent might say "Maybe this topic is getting old. I'll help you finish so we can try something new", in an attempt to indirectly address student's boredom while also focusing on a new topic, hence avoiding that the student becomes disengaged from the learning experience.

3.1.2 EER-Tutor

In 2008, given the affective gap in Intelligent Tutoring Systems, Zakharov et al. [37] developed an affect-aware pedagogical agent persona for an ITS that teaches database design skills, EER-Tutor [38], which uses feature tracking technology to learn the user's affective state in a dimensional approach. The authors found that the results of the experiments conducted advocate the presence of affective pedagogical agents, with the affect-aware agent demonstrating superiority over its non-affective counterpart and Strain and D'Mello [36] concluded in theirs that subjects in shallow and deep cognitive reappraisal conditions report higher arousal than those in a control scenario.

3.1.3 A web-based learning system

Strain and D'Mello [36] developed a web-based learning system that analyses the effect of cognitive reappraisal on students' self-reported emotions and performance results, considering both valence and arousal. They assigned participants randomly to three groups with different reappraisal conditions and found that the results favored the usage of cognitive reappraisal, as participants in groups of deep and shallow reappraisal condition reported higher arousal that those in the control group.

3.2 Emotion Regulation in Serious Games

The usage of virtual environments - such as video games and the internet - as more cost-effective treatment approaches in health and medical applications has become increasingly popular in recent years [16, 27]. Serious games are a particular example of those environments, with some performing emotion regulation strategies to help players manage their affective state.

3.2.1 EmoRegulators

EmoRegulators [34] is a serious game that helps adolescents learn to perform emotion self-regulation through a series of sessions based on a technology-enhanced version of the BEAR [30] protocol. In order to assess the user's emotional state, the game uses bio-sensors that measure heart rate, EDA, and muscle activation for the trapeze and bicep. The game is comprised of several sessions and exercises that teach physical regulation through an instructional agent, called the facilitator, which follows a scripted behavior. The game consists of two sessions:

Session 1 (Introduction): Presents the topic of coping resources identification via experimental, playful and fun exercises. It aims at increasing the awareness of self within the game.

Session 2 (Physical Regulation): Focuses on experimental exercises in order to practice relaxation coping skills and provide awareness of one's bodily sensations. During gameplay, users perform breathing, active/shaking and Progressive Muscle Relaxation (PMR) exercises and reflect about their safe places and their internal sensations.

During these sessions, the player earns points by completing the exercises and these are displayed below their personal box, the I-Box, which they learn about and customize in the introduction session. Players can also see their heart rate, which makes them aware of their physical state.

3.2.2 PlayMancer

PlayMancer [10] is an EU initiative to develop a video game prototype for treating specific mental disorders, namely eating disorders and impulse control disorders. Its goal is to increase emotional self-regulation skills in patients and to increase their self-control towards impulsive behaviors. Interestingly, it's the player's emotions that dictate the difficulty of the games. This information regarding the person's emotions is gathered through bio-sensors and facial gestures, and speech based emotion recognition is used to decide which skills and attitudes should be changed in order to achieve each task's therapy goals.

3.3 Wizard-of-Oz Studies

The usage of learning architectures in intelligent agents has become increasingly popular over the years. These agents are capable of performing very well in more complex situations, given they've been trained properly. Usually, though, these require a decent number of interactions, or demonstrations, to perform better.

3.3.1 Restricted-Perception Wizard-of-Oz Studies

Inspired by Learning from the Wizard (LfW) [19], Sequeira et al. [35] proposed a methodology to create social interaction strategies for Human-Robot Interaction (HRI) based on restricted-perception Wizard Of Oz (WoZ) studies. The interesting core idea is restricting the wizard's perceptions over the environment and the behaviors it controls according to the agent's inherent perceptual and acting limitations. This methodology is divided into three phases.

The first phase is data collection. The main purpose here is to prepare and perform restricted-perception WoZ studies by gathering useful knowledge about appropriate interaction strategies to be considered. The idea is to let humans that are experts on the given task to perform several interaction sessions with prospective end-users of the system. The collected data is used to build a set of task-related artificial intelligence modules, referred to as the Task AI. Once the Task AI has been implemented, appropriate interaction strategies can be discovered for the agent by performing the WoZ studies. In this next phase, the idea is to build an interaction strategy controller for an agent based on the previously collected data, corresponding to the Strategy Extraction phase, where they try to "infer" the decision process used during the restricted-perception WoZ studies.

One of the underlying problems in WoZ studies is the correspondence problem, where a direct mapping between sensors and actions of the agent and those of a human is not possible. The restricted-perception approach mitigates this problem, as previously explained, helping to significantly reduce the complexity of finding a correspondence between the task's state as observed by the wizard and the information available to the agent.

In this methodology, interaction strategies are learnt through a mapping function between the agent's state features and interaction behaviors given the wizard demonstrations in the restricted-perception WoZ studies. Similar to how the wizards have the responsibility of choosing which behavior to trigger and when to trigger it, so does their ML-based module. For this matter, ML algorithms are chosen to learn the mapping function, by using classification or clustering algorithms, for example. In the last phase of the methodology, evaluation studies are conducted to assess the performance of the agent being autonomously controlled while interacting with others during the given task. Based on the implementation of their methodology, the authors concluded their work was an improvement to the challenge of mapping between the expert's observations and the agent's perceptive capabilities - and it should definitely be considered when attempting to build successful interaction strategies applied to learning in ITS.

4 Solution

For us to solve the problem of *how to create an intelligent, adaptive system in EmoRegulators, capable of being aware of, understanding and regulating players' emotions during an interaction session*, we need to solve two distinct subproblems. First, we need to decide what may be some appropriate emotion regulation strategies applicable in the context of EmoRegulators based on literature we've just seen. And second, we need to solve the problem of how these strategies should be activated in EmoRegulators. It is hoped that by using restricted-perception Wizard-of-Oz studies, one can help emotion-aware intelligent tutoring systems to better replicate the behavior of expert human tutors. In the solution, the necessary steps to solve these problems are taken in order to make the in-game facilitator into an intelligent tutoring system rather than having a scripted behavior.

4.1 Emotion Regulation: Defining Appropriate Strategies for EmoRegulators

The task of deciding which emotion regulation strategies could be suitable in the context of EmoRegulators required that some research into the game components and exercises was conducted first. Early thoughts on how to solve this subproblem were that it could be interesting to experiment with both exercise or activity related strategies, as well as ones which could be used at any given moment in the game. But these strategies are supposed to complement the standard behavior of the facilitator, meaning each of them individually would likely occur few times during the full length of a game session. This means exercise-specific ones would see so little usage that it would be hard to assess their impact on gameplay.

As such, focus was shifted towards defining strategies that could be activated in any moment of the session while remaining relevant. The above literature suggests that the attribution theory and the cognitive reappraisal theory could help solve the problem of keeping players engaged during a game session of EmoRegulators; therefore we decided to experiment with these strategies.

In its classical version, the facilitator in EmoRegulators interacts with the player verbally, with a calm voice, following a script that explains each exercise. The least intrusive way to intervene when the sensors readings say the agent should do so is to have a set of empathetic or motivational interactions that it can choose from and say at the appropriate moment.

In the current version, strategies meant to apply the attribution theory will attempt to blame the feelings of frustration or boredom on external factors, such as the exercise or the agent itself, similar to what AutoTutor does. For example, if the sensors readings indicate the player might feel bored, the agent can say something like "This exercise is a bit long..." or "Am I talking too much?". Both examples, depicted in Figure 1, attribute the cause of boredom to other things, alleviating that pressure from the player, while making them aware of their emotional state. The first example can be used to keep the player engaged in a lengthy exercise, while the second one can recapture the player's attention during a tiresome explanation.

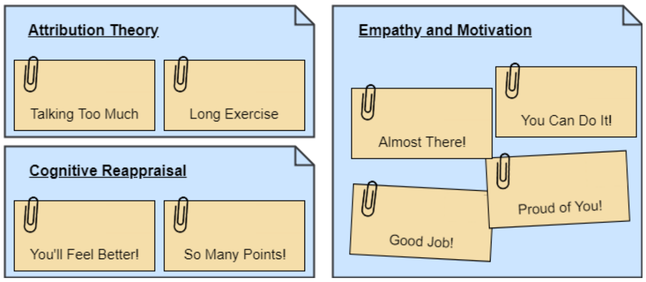

Figure 1: A graphical representation of the mapping between the emotion regulation strategies and their respective foundations.

In cognitive reappraisal, the idea is to say something that causes the player to acknowledge their negative emotional state and then re-evaluate the situation to overcome it. This is a very beneficial approach for two different reasons. The first is that it can potentially help the agent regulate the player's emotions, which is one of the goals for the agent. The second is that, as an added bonus, it makes the player aware of their emotions upfront which, remember, is important in a game like EmoRegulators that is meant to teach players to be aware of, and control, their own emotions.

In one of the strategies that follow this theory, the facilitator says "Even if you're not feeling very motivated, remember: you can earn a lot of points if you make this little effort!". This does two things: for one, it helps the user recognize their negative response to the exercise by suggesting he may be lacking motivation to continue; and, secondly, it emphasizes there is a reward for completing the exercise in an attempt to refocus the player's motivation to pursue additional points and to push through the session with a fresh mindset.

As an example of how this interaction could prove useful, consider an exercise where high physical intensity is required, such as the one where the player is required to dance for a few minutes. Since players can opt to simply ignore such an exercise by not moving or altering their physical activity, if the expert has information about the points available, and the sensors' readings say the HR is low, the expert can chose this interaction to try to convince the player to do the exercise as intended (by actually moving and dancing) increasing their HR and potentially improving the player's emotional state. This is a double-edged sword, however; if the sensors indicate more than just a low HR - boredom, for example -, it might be better to take a different approach and let the player ignore it, because they might not like the exercise they're about to do and maybe ignoring it will suppress or eliminate this feeling. This is entirely the expert's call, but making this strategy available can give the expert leverage for persuading the player to go through with the exercise - if that is the correct decision.

In addition to the attribution theory and cognitive reappraisal strategies, we decided to incorporate additional generic empathetic remarks that celebrate the player's performance in the game - "Good job!" and "Proud of you!" are examples of that. It is hoped that this can help maintain a positive arousal and attitude during the game, reducing the likelihood of feelings such as boredom and frustration taking place. The remaining interactions "Almost there!" and "You Can Do It!" are a motivational version of the aforementioned ones, implemented with the same purpose in mind; finally, "You'll Feel Better" also makes use of the cognitive reappraisal by highlighting that the purpose of EmoRegulators is to make players feel better, and that in doing these exercises they are working towards that goal.

The initial strategies were reviewed by the domain consultant who fulfilled the role of the expert in preliminary demonstrations (this is explained in more detail in the section regarding restricted-perception WoZ demonstrations) and by a psychologist. Their feedback was that some of these emotion regulation strategies were too long and that it could potentially make them ineffective. In those strategies where it was possible, they were reformulated to become shorter and straight to the point while maintaining the essence of their respective foundations (attribution theory and cognitive reappraisal, for example). As an example, one strategy was to say "I'm sorry... Am I talking too much? Maybe you've understood this exercise already. Should we start?"; indeed, this could be reduced to only the "Am I talking too much?" segment, which is the part that effectively applies attribution theory. This new iteration of shorter strategies were revised and approved by the domain experts and are the ones being used in this system at the time of writing.

4.2 Activation of the Emotion Regulation Strategies

Auto-tutor, a landmark in Intelligent Tutoring Systems, has the Support and Shake-up versions, which are affect-aware. In these systems, the activation of the emotion regulation strategies that the agent uses are decided according to a set of rules which guides this decision-making.

Looking at the literature, it was decided it could be interesting to combine the lessons taught by these systems in terms of what strategies can be used for emotion regulation, with the idea that the activation of these strategies should be entirely guided by human decision-making - specifically that of an expert - instead. This led us to explore Wizard-of-Oz studies but, more specifically, restricted-perception Wizard-of-Oz studies.

4.2.1 Methodology

As explained by the authors of the work [35] that explores this approach, by restricting the perceptions of the human acting as the expert in the demonstrations phase to match those of the agent, the learning curve is steeper (in the technical sense of the expression, meaning more is learned quicker). Our methodology was developed to use a restricted-perception WoZ technique to learn how to apply emotion regulation strategies in EmoRegulators.

In the case of EmoRegulators, the agent benefiting from the restricted-perception learning is the conversational instructor, also called facilitator, that is responsible for explaining the exercises and tasks in the game. The goal is to teach the agent to have expert-like decision making in a tutoring scenario, instead of following a script, so that it incorporates some level of interactivity and dynamism into its dialogue system. This is done by allowing the wizard to effectively control the agent, typically choosing from a set of possible strategies during an early demonstration phase.

We first needed to identify the information available both in the game and provided by the physiological sensors, and determine which features would or not be relevant to help the expert (and the system), to make decisions on when to perform the emotion regulation strategies, as well as if any additional ones besides these would be necessary. We decided in favor of adding a few others besides the ones given by sensors and game state, which are calculated from existing ones - but a better explanation of these features is provided in the following section.

A fundamental requirement and necessary assurance in conducting restricted-perception WoZ studies is to enforce that the human expert has access to the same information that the system will have to make decisions. For this reason, a separate WoZ interface was carefully designed, implemented and interconnected to the EmoRegulators game to ensure this requirement was respected and not violated. The reader can refer to section 4.2.3 to see how this interface was designed exactly. This interface was then used by the expert while conducting the restricted-perception WoZ demonstrations depicted in the leftmost parcel of Figure 2, the phase when data collection takes place.

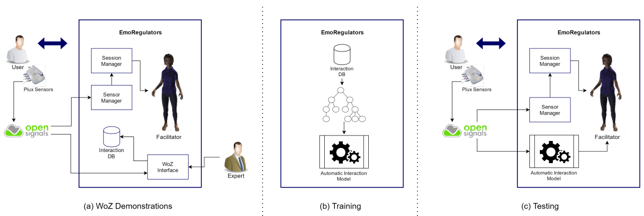

Figure 2: Our originally proposed model. In (a), a series of restricted-perception WoZ demonstrations are conducted with an expert. In (b), a learning algorithm is trained to create an automatic interaction model. Optionally, in (c), the created model can be tested with users for feedback.

After this demonstrations phase, the gathered data is pre-processed before being used to train an automatic interaction model using machine learning techniques. Finally, this model can optionally be put to test against new players without the presence of the wizard and using just the learned interaction model instead, which would correspond to step (c) of Figure 2. Steps (a) and (b) of this restricted-perception Wizard-of-Oz paradigm were implemented in this work with the previously discussed set of strategies meant to keep players engaged in EmoRegulators.

4.2.2 Game State and Affective State Representation

One very core aspect of this work was to define which features or information would be used to model the state of a session of EmoRegulators. These features would play a leading role both in the restricted-perception WoZ demonstrations and in the training phase. Furthermore, these features would need to be representative of the game state, as well as the affective state, in order to build a solid set of features for the successful application of emotion regulation strategies in EmoRegulators.

In regards to game state, the session, exercise, and points gained by the player were collected during play time. This information would help guide the expert's decision-making during demonstrations by providing context for the data collected from the game at every instant. This is complemented, of course, by features corresponding to the data collected from the bio-sensors. This information was also later complemented by collecting, for a given instant, time elapsed since the beginning of a game session and since the beginning of the exercise being done as well.

Finally, a few additional features calculated from these first ones went into modeling the affective state. One of these is a feature that represents arousal using some of the physiological data. Despite there not existing a direct mapping between emotions that suggest lack of engagement, such as boredom and frustration, and the data that can be collected from the bio-sensors in EmoRegulators, thankfully it has been shown that a significant correlation exists between psycho-physiological arousal (i.e., HR and EDA) and self-reported gameplay experience, which includes the feelings of frustration, boredom, and others [9]. Other claims have been made about the correlation of higher EDA readings and a higher level of frustration [9,23], and other studies report in favor of a correlation between higher HR and higher player arousal [22, 23]. Therefore, arousal, meant to help the expert understand the player's emotional state, was implemented as being represented on a 0 to 10 scale. This was done so that an arousal meter using this scale could be shown in the restricted-perception WoZ interface. As such, the following steps were taken to calculate this value:

i. Aggregate data from all past sessions of EmoRegulators (specifically, those conducted with players in past work on the game) for both HR and EDA;

ii. Calculate the mean and standard deviation for those variables;

iii. With every data update, in real-time, calculate the Z-Score [25] for both HR and EDA based on current measurements of these using the previously calculated statistics

iv. Adjust to a 0 to 1 value range and then scale this value to a 0 to 10 range.

However, due to certain problems in readings of EDA, we removed it from the calculation of arousal in the meter; the usage of only HR should still be plausible given that it has been proved that HR correlates to arousal [22, 23]. We later made some adjustments to this calculation so it would reflect common heart rate values of children of ages 14-18 instead of just using the data from the ones who participated in earlier tests with EmoRegulators. Based on a systematical review of normal heart rates in children [11] we adjusted the scale so the minimum value would reflect a heart rate of 60. We also used a known formula for maximum heart rate per age calculation [12] to arrive at the maximum value of 206 at the top of the scale, using age 14 in the formula as this is the age (for EmoRegulators' target audience) that maximizes that function. Finally, both calculations of arousal are weighed at 50% and rounded up to arrive at the final value for arousal, the one used as a feature in the work.

Besides the raw data that could be gathered from EmoRegulators, there were other features besides time and arousal which were considered to be fed to the algorithms in the training phase to come. Additionally it was decided that, at a given instant when data was collected, there should be a calculation of the average of the 10 most recent samples leading to that moment for heart rate, EDA, bicep activity and trapeze activity as features. So this resulted in four additional features called averageHR, averageEDA, averageM1 (for bicep activity) and averageM2 (for trapeze activity). These features could potentially come to play in the training phase for specific approaches which we present later in this document; since not all machine learning algorithms take into account feature dependency or connections between different features, this is one way to include variables that account for some sort of connection between them. Finally, the percent variance for HR between each two consecutive heart rate readings was calculated as a feature for each data observation as well.

4.2.3 Restricted-Perception WoZ Interface

In order to teach the system using a restricted-perception Wizard-of-Oz paradigm, some means of ensuring the perceptions of the expert were controlled and restricted had to be implemented. Two steps are taken to achieve this. The first one, is that the expert would be interacting with EmoRegulators by using a carefully designed interface. This interface could not provide any more information than the one the system would use during training or runtime classification (these phases are explained in detail further in this article). The interface, shown in Figure 3, focuses on the following:

i. Displaying information regarding the bio-sensors' readings (i.e., the basic player state features)

ii. Showing the current context of EmoRegulators (i.e. the game state features)

iii. Mapping the calculated arousal value in a 0 to 10 scale, called the arousal meter

iii. Displaying a set of buttons corresponding to the emotion regulation strategies the expert can choose from for the facilitator to say.

The second step was to conduct the restricted-perception WoZ demonstrations by ensuring the expert would be controlling the agent from a physically separate room than the one where the player would be interacting with the game. This was to guarantee that the expert would not make decisions by picking up on body language, facial expressions or other behavior while the game session took place.

A special detail that had to be taken into consideration was the number of bins being shown in the heart rate and EDA charts (in red and blue, respectively). Originally 30 values were shown at any instant in the interface, which had to be corrected to no more than 10 values, as to ensure this was consistent with the temporal window of 10 values (over 5 seconds) present in each observation that was being logged during game sessions.

Finally, it was concluded that if the audio played in the game wasn't reproduced in the interface as well, the human expert using the interface would not be able to effectively predict a user's emotional state. This change was accommodated as requested by the person playing the role of the expert in the earlier demonstrations in order to provide high-level context of what was happening in the game session at each instant. Note that, in any case, the expert was not watching the player or anything besides what is displayed on the interface; only listening to the dialogues.

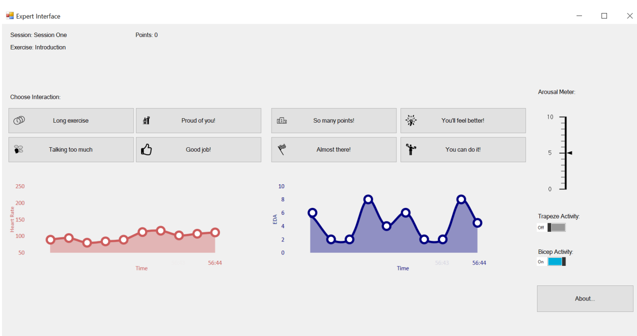

Figure 3: The expert interface used by the wizard to control the facilitator during the WoZ demonstrations.

When developing the interface, a number of decisions had to be made in regards to all the data traveling from the game to the interface and vice-versa. The most important one which has to be mentioned is the decision to register and send sensor data every 0.5 seconds. We feel this allows the expert to make an informed decision fast enough that it is still relevant when the user is presented the interaction resulting from that decision. There were also some changes that had to be made in respect to the audio management for the dialogue system of EmoRegulators, with the addition of the new speech utterances. The solution we found (which required actively notifying the players of it before the demonstrations started), was to allow for overlapping audio of two files, lowering the audio of any on-going instruction as the emotion regulation strategy audio was being played. A far from ideal solution, but the best compromise we came up with at the time given some limitations that the system had.

Design Decisions

The interface shown in Figure 3 can be divided into three sets of elements. The first one informs the expert about the current game state, by displaying textual information regarding the session, exercise and number of points obtained thus far. There are other things we could have included, such as the time occurred since the beginning of an exercise or session, but ultimately we decided to keep only the information regarding the game state since different players take different amounts of time to complete each task. Not only is there nothing to be derived from such data, but it could even be harmful in that the expert could feel obligated to take this variable into consideration, potentially clouding their judgment.

The second one is the set of possible interactions which one can choose from. These are distinctively split into two sub-groups of motivational and empathetic strategies just to ease distinction between the two for the interface user (the expert). We also consulted with a domain experts in the fields of Data Visualization and Human-Computer Interaction to validate an initial selection of the icons to accompany the name of the description in the buttons. Then we iterated further based on this feedback to create a new set of icons and we created a small questionnaire and asked researchers in this area to match the icons to the interaction descriptions. Using a Pearson Chi-Square test to measure statistical association between the icons and expressions, we obtained a value of $p = 0.006 < 0.05$ thus supporting the usage of this version of the icons, which are the ones currently present in the interface.

The third set is dedicated to sensor data. There are two charts, one showing heart rate data and the other one EDA. There are also two switches representing whether muscle activity is "on" or "off" (that is, if it is occurring or not) for the trapeze and bicep muscles. Additionally, we decided to include the arousal meter mentioned earlier, which attempts to measure arousal in a scale of 1 to 10 based on previous EmoRegulators sessions' data.

The way the interface is meant to be used is quite straightforward. Every half second, a request for updating the information is made from the interface to the game, and upon receiving a response, the new data is shown in the interface, updating it. The expert reads the information displayed in the interface and makes an assumption of what the player's emotional state might be. Then, in case the expert thinks they should intervene, they select the most suitable strategy and clicks the corresponding button. A request for that interaction is sent to the game via TCP and then it is mapped to its corresponding audio file to be played. The game sends a response as to confirm the requested audio will be played, and plays the audio. Upon receiving a confirmation response from EmoRegulators, the audio is reproduced on the interface side as well, adding to the expert's perception of the game state. In case no interaction is deemed necessary, data flows between the interface and the game through update requests every half second as per usual.

4.2.4 Data Preparation

As stated before, the strategy being activated in the game at every time interval is recorded: if there is no strategy being requested from the interface, an inaction is logged; otherwise, one of the strategies (whichever one was selected) is logged. What is bound to happen is that by the end of the full session, or even just one exercise, there will be a lot more samples recorded as inactions than any of the strategies, resulting in an imbalanced data set.

The problem with imbalanced data sets is that sometimes what one want to learn the most happens the least. One technique that can be used to address imbalanced classes is called oversampling. While it would have made sense to undersample the majority class (the inaction class), oversampling of the under-represented classes is advised when there is little data available to train with, which is why the latter was preferred in our solution. Knowing that realistically we will always have an excess of inactions compared to the possible strategies to choose from, we experimented with a 50% representation of inactions and 50% combined representation of the remaining classes. This means that each of those eight classes are still far less likely to be chosen over not performing an interaction, which we found was a reasonable compromise. We were also mindful of performing oversampling inside the cross validation loop (we discuss cross validation and our approaches to training in greater detail in the following section) as to not test a trained model on synthetic samples [2].

4.2.5 Training

To conduct the training phase of our solution, we decided to use Scikit Learn [3], an open-source machine learning library for Python. We progressively followed three approaches aimed at minimizing the likelihood of problems such as overfitting and which take advantage of the pre-processing steps taken in data preparation. These approaches were tested with five machine learning algorithms - k-Nearest Neighbors (kNN), a decision tree (DT), a random forest (RF), Gaussian Naive Bayes and a Neural Network - as to later choose the most appropriate and best performing one for EmoRegulators based on the data collected in the restricted-perception WoZ demonstrations.

k-Fold Cross Validation

As a staple to our training approach, we first wanted to test our selection of algorithms using a solid technique that could also deal with overfitting, a common problem in machine learning model training. As such, we decided to use k-fold cross validation using stratification, a method that consists of rearranging data to guarantee that each fold holds a class balance similar to the full data set. As for which value of k to pick, many scientific sources [13, 20, 32] typically recommend using a value of $k = 5$ or $k = 10$ folds as a good sweet spot. For this reason, we chose a value of $k = 10$ to perform 10-fold cross validation. For this assessment of our models' performance, we merged all the subject-oriented data into a single dataset and then performed stratified k-fold cross validation. This data set gives us overall performance information for different models on EmoRegulators, based on the features we chose to study.

Leave-One-Subject-Out (LOSO)

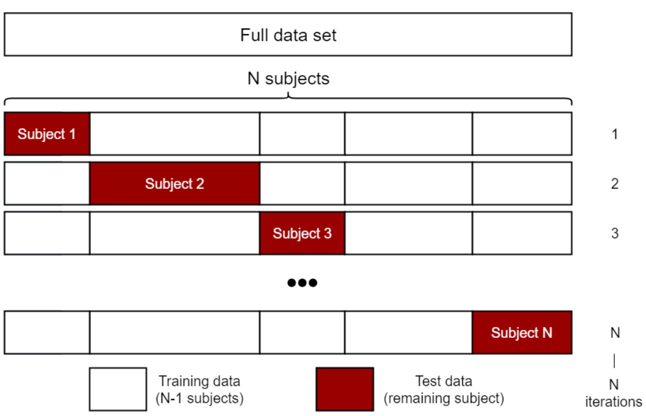

Figure 4: A subject-oriented approach to cross-validation. Note that unlike k-fold cross validation, the different folds may have different numbers of samples.

We also thought it would be interesting to evaluate how these models generalize to new subjects specifically. As such, for this approach we imported each subject's data logs separately and stored them as different data sets. We can refer to Figure 4 to understand how Leave-One-Subject-Out was conducted; each player's data set acts as a fold. Afterwards, the process seen in k-fold cross validation takes place: each of the k iterations consists of using the k^{th} player dataset as a test set, and using the remaining $k-1$ player folds as the training set. The process is repeated k times to allow a full rotation of test sets without any two test sets overlapping.

Leave-One-Exercise-Out (LOEO)

Just as it is interesting to look at how the model might generalize for new subjects, we additionally decided to see how models would be able to generalize for different exercises in EmoRegulators. This was a more experimental approach, but one we still wanted to study results for, that works just like the Leave-One-Subject-Out, but using the logs of each exercise merged from that of each player together as each fold.

4.2.6 Runtime Classification

Research was conducted with the purpose of finding tools that could facilitate Python integration with Unity3D, the game and content creation engine that EmoRegulators was built on. After researching, we ultimately decided to develop Python scripts to implement a simple socket-based communication with EmoRegulators. The way this works is similar to the way the restricted-perception WoZ interface interacts with the game for the demonstrations phase:

i. the sockets connect over TCP using the local address and a designated port;

ii. the Python socket script requests an update of EmoRegulators data every half second;

iii. the server socket on EmoRegulators side responds with the raw data (session, exercise, points and biosensor data) plus the time elapsed information in that instant;

iv. after 10 requests (that constitute the 5 second temporal window we describe in the data collection section) the Python script calculates the more complex features such as HR percent variance and averages from the 10 samples of raw data and then builds an observation with all the features;

v. the generated observation is passed to a predictive function of the machine algorithm chosen during model selection (this prediction is of course based on the previously trained data), returning the predicted classification of that observation;

vi. the Python script matches that class to the appropriate interaction and sends a request to EmoRegulators to activate the corresponding strategy in the game.

These 6 steps are continuously repeated to generate predictions autonomously until the game session ends. This automated procedure works for any of the proposed approaches - whether it's 10-fold cross validation, Leave-One-Subject-Out or Leave-One-Exercise-Out - with minimal computational cost, but works better for some models than others, as more complex machine learning algorithms can take longer to make a prediction.

4.3 Results

4.3.1 Restricted-Perception WoZ Demonstrations

Given how EmoRegulators is targeted at adolescents and how the demonstrations took place during school season, it was exceptionally hard to come up with a suitable schedule. The expert chosen to conduct these demonstrations was Professor Esther J. Schek, someone who is very familiar with EmoRegulators and the BEAR protocol, making her the most suitable candidate. A consent form allowing for the usage and processing of the adolescents' biometric data had to be signed by one of the their parents or their guardian. After finding a suitable schedule for both the expert and the subjects, we only managed to perform demonstrations with 5 of them. However, the collected data from these demonstrations was flawed and incomplete, which meant the researcher of this work had to play the role of the expert in a second round of demonstrations, as this happened at a point in time very close to the deadlines of this work and the researcher couldn't afford to be dependent on other parties' availability any more. The researcher carried out these demonstrations in the same way these were carried out by the first expert initially - following the same protocol (explained below) that was established for the initial WoZ demonstrations, acting as both parties involved.

The subjects were three boys aged 14, 15 and 18, and two girls who were both 17 years old, allowing for a good sample age-wise, despite smaller than preferable.

WoZ Protocol

A protocol was defined between the wizard and the researcher, so that there would be a standardized procedure across different demonstrations and so that communication between the two would be clear and effective in case any errors or problems arose during demonstrations. Before each demonstration, an explanation of EmoRegulators was provided. This consisted of an overview of the exercises, where some additional tips were provided regarding particularities of certain exercises. This was to ensure that there was minimal room for confusion so that interrupting a demonstration could be avoided. Afterwards, we would set up the sensors for HR and the two for muscle activation. Finally, the IP address of the computer EmoRegulators would run on was sent to the expert, so they could type it in an interface dialog to connect to the game. After a ready check from the wizard, the game session would begin, the expert would effectively connect and both would start monitoring the demonstration. While the expert monitored the session from the interface, the researcher's job was to keep an open communication channel with the former in case there was any problem on the expert's end, while at the same time checking every few minutes if the test subject needed anything. At the end of each demonstration, the researcher would warn the expert that the former would close the session, so the latter could do the same and await further updates. The researcher would then talk with the subject for about 5 minutes while removing the sensors, to get general feedback.

Participants' Feedback

In our initial planning of the work, the idea was to segment this into a work flow consisting of four phases, the last one being a final round of tests with players where they would be split into a control group, interacting with the original scripted version of EmoRegulators, and a condition group, interacting with this more interactive, automated version of the game. We wanted to accommodate this additional phase not because we needed it necessarily to increase model performance, but mainly for feedback on user enjoyment and emotional state in both versions. Unfortunately with the WoZ demonstrations happening so late in schedule, we weren't able to conduct this final comparison between those versions. What we did, though, was try to gather some informations in the demonstrations phase. While removing the sensors at the end of each one, the researcher asked the subjects for feedback on the tests in the form of a free discussion. The subjects reported an overall fun and enjoyable experience but generally expressed feeling bored in the beginning of the session. Two of the subjects reported that they had a tough time associating certain feelings to parts of their body, which is required in one of the exercises, even though the procedure is explained. Another critique that was made in this second round was how some interactions (sent from the interface to the game) would sometimes overlap with instructions, a consequence of the approach taken with the addition of the audio of the new strategies to EmoRegulators.

4.3.2 Model Selection

We evaluated our five ML algorithms in four metrics: accuracy, F1 score, precision and recall. We also present our choice for the model to perform runtime classification and discuss these results for 10-fold cross validation, and leave-one-subject-out and leave-one-exercise-out approaches.

Looking at the data from all approaches (please refer to the original document for the LOEO and k-fold cross validation approaches), kNN, decision tree and random forest all see consistently great performances. In all approaches, as expected, random forests have always at least marginally better accuracy and F1 scores than decision trees, which makes sense given that a random forest in sci-kit predicts the class with highest mean probability estimate across all trees that compose it. On the other hand, Gaussian NB and Neural Network had worse performances than the others; the first one is known to be a bad estimator, so the result isn't surprising. As for NN, we suspect this is because the algorithm tends to perform better with extensive and representative training sets [14]. Our results corroborate this as some training sets are more representative of the full set than others - at least in LOSO and LOEO where data isn't stratified.

Table 1: Performance estimation for the Leave-One-Subject-Out approach.

	kNN	DT	RF	Gaussian NB	NN
Accuracy	0.887	0.956	0.967	0.082	0.363
F1	0.909	0.945	0.950	0.134	0.490
Precision	0.933	0.935	0.935	0.932	0.930
Recall	0.887	0.956	0.967	0.082	0.363

Regardless of the choices made, one conclusion can be confidently drawn from these results: 3 out of the 5 models performed well for the Leave-One-Subject-Out approach, achieving scores of 88.7%, 95.6% and 96.7% accuracy for kNN, decision tree and random forest respectively. With such good performance being achieved on such a small sample size of demonstrations, we believe this was a very positive step towards solving our posed problem. Moreover, these results seem to suggest that taking a restricted-perception WoZ approach can lead to promising results in models' capabilities of replicating the success of tutors.

5 CONCLUSIONS

In our solution we discuss that the problem of how we can create an intelligent and adaptive system to help maintain engagement in EmoRegulators can be split into two subproblems, and we present a solution that uses emotion regulation strategies to attempt to improve the gameplay experience of EmoRegulators through a WoZ paradigm. We discuss our model, where we detailed how data was collected and prepared for training, as well as the thought process behind building the interface for the demonstration phase. These demonstrations, conducted with 5 adolescents between ages 14 and 18, allowed us to gather data that was then used for training with the five algorithms and a three-fold approach, allowing us to conduct a classifier comparison to decide which combination to use in runtime classification, made possible by using TCP socket communication to bridge EmoRegulators in Unity and the scripts for training and runtime scripts written in Python.

Finally, we discussed results for the culmination of the work conducted in this thesis. Through a small, free discussion in demonstrations, we learned that players reported an overall positive experience, but also expressed concerns regarding the dialogue system of the in-game assistant - an implementation decision that needs a pondered revision. We found that simpler algorithms such as kNN and decision trees were the best performers achieving results for accuracy and F1 score not lower than 88.7% regardless of whether data was trained using 10-fold cross validation, Leave-One-Subject-Out or Leave-One-Exercise-Out. With emphasis on achieving higher recall scores, as to reduce the number of interactions being mislabeled as inactions, we decided to either use Decision Tree or Random Forest, combined with the Leave-One-Subject-Out approach, given our goal to generalize well to new players, towards providing better individualized experiences.

These results in classification performance upwards of 88% with as few as 5 restricted-perception WoZ demonstrations conducted allow us to say with some confidence that we have taken meaningful steps towards solving our posed problem. The evidence in our results - for Leave-One-Subject-Out especially - using a restricted-perception WoZ paradigm also seem to corroborate our hypothesis that *the usage of emotion regulation strategies based on expert knowledge from human tutors will help emotion-aware intelligent tutoring systems better replicate their ability to understand and regulate users' emotions*.

References

[1] H. Ahn and R. W. Picard. Affective-cognitive learning and decision making: A motivational reward framework for affective agents. In *International Conference on Affective Computing and Intelligent Interaction*, pp. 866–873. Springer, 2005.

[2] Altini, Marco. Dealing with imbalanced data: Undersampling, oversampling and proper cross-validation. `https://www.marcoaltini.com/blog/dealing-with-imbalanced-data-undersampling-oversampling-and-proper-cross-validation`, 2011. [Online; accessed 3-May-2018].

[3] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pp. 108–122, 2013.

[4] R. A. Calvo and S. D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on affective computing*, 1(1):18–37, 2010.

[5] S. Chaffar, L. Derbali, and C. Frasson. Inducing positive emotional state in intelligent tutoring systems. In *AIED*, vol. 2009, pp. 716–718, 2009.

[6] S. Choi and R. E. Clark. Cognitive and affective benefits of an animated pedagogical agent for learning english as a second language. *Journal of educational computing research*, 34(4):441–466, 2006.

[7] S. D'mello and A. Graesser. Autotutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(4):23, 2012.

[8] S. Domagk and H. M. Niegemann. Pedagogical agents in multimedia learning environments: Do they facilitate or hinder learning? In *Proceedings of the 2005 conference on Towards Sustainable and Scalable Educational Innovations Informed by the Learning Sciences: Sharing Good Practices of Research, Experimentation and Innovation*, pp. 654–657. IOS Press, 2005.

[9] A. Drachen, L. E. Nacke, G. Yannakakis, and A. L. Pedersen. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*, pp. 49–54. ACM, 2010.

[10] F. Fernández-Aranda, S. Jiménez-Murcia, J. J. Santamaría, K. Gunnard, A. Soto, E. Kalapanidas, R. G. Bults, C. Davarakis, T. Ganchev, R. Granero, et al. Video games as a complementary therapy tool in mental disorders: Playmancer, a european multicentre study. *Journal of Mental Health*, 2012.

[11] S. Fleming, M. Thompson, R. Stevens, C. Heneghan, A. Plüddemann, I. Maconochie, L. Tarassenko, and D. Mant. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. *The Lancet*, 377(9770):1011–1018, 2011.

[12] S. M. Fox III and J. P. Naughton. Physical activity and the prevention of coronary heart disease. *Preventive medicine*, 1(1-2):92–120, 1972.

[13] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*, vol. 1. Springer series in statistics New York, 2001.

[14] M. W. Gardner and S. Dorling. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*, 32(14-15):2627–2636, 1998.

[15] J. J. Gross. Emotion regulation in adulthood: Timing is everything. *Current directions in psychological science*, 10(6):214–219, 2001.

[16] J. J. Gross. *Handbook of emotion regulation*. Guilford publications, 2013.

[17] A. Kapoor and R. W. Picard. Multimodal affect recognition in learning environments. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pp. 677–682. ACM, 2005.

[18] H. H. Kelley and J. L. Michela. Attribution theory and research. *Annual review of psychology*, 31(1):457–501, 1980.

[19] W. B. Knox, S. Spaulding, and C. Breazeal. Learning social interaction from the wizard: A proposal. In *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.

[20] R. Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, vol. 14, pp. 1137–1145. Montreal, Canada, 1995.

[21] M. Malekzadeh, M. B. Mustafa, and A. Lahsasna. A review of emotion regulation in intelligent tutoring systems. *Educational Technology & Society*, 18(4):435–445, 2015.

[22] R. L. Mandryk. Physiological measures for game evaluation. *Game usability: Advice from the experts for advancing the player experience*, pp. 207–235, 2008.

[23] R. L. Mandryk and M. S. Atkins. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International journal of human-computer studies*, 65(4):329–347, 2007.

[24] X. Mao and Z. Li. Implementing emotion-based user-aware e-learning. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pp. 3787–3792. ACM, 2009.

[25] M. Margulies, M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y.-J. Chen, Z. Chen, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057):376, 2005.

[26] R. P. Martín, P. F. Berrocal, and M. A. Brackett. La inteligencia emocional como una competencia básica en la formación inicial de los docentes: algunas evidencias. *Electronic journal of research in educational psychology*, 6(15):437–454, 2008.

[27] J. D. Mayer, D. R. Caruso, and P. Salovey. Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 27(4):267–298, 1999.

[28] R. Moreno. The role of software agents in multimedia learning environments: When do they help students reduce cognitive load. In *European Association for Research on Learning and Instruction Annual Conference, Padova, Italy*, 2003.

[29] J. E. Ormrod. *Educational Psychology: Pearson New International Edition: Developing Learners*. Pearson Higher Ed, 2013.

[30] R. Pat-Horenczyk, C. S. W. Shi, S. Schramm-Yavin, M. Bar-Halpern, and L. Tan. Building emotion and affect regulation (bear): Preliminary evidence from an open trial in children's residential group homes in singapore. In *Child & Youth Care Forum*, vol. 44, pp. 175–190. Springer, 2015.

[31] R. D. Ray, K. McRae, K. N. Ochsner, and J. J. Gross. Cognitive reappraisal of negative affect: converging evidence from emg and self-report. *Emotion*, 10(4):587, 2010.

[32] P. Refaeilzadeh, L. Tang, and H. Liu. Cross-validation. In *Encyclopedia of database systems*, pp. 532–538. Springer, 2009.

[33] P. Salovey and J. D. Mayer. Emotional intelligence. *Imagination, cognition and personality*, 9(3):185–211, 1990.

[34] E. J. Schek, F. Mantovani, O. Realdon, J. Dias, A. Paiva, S. Schramm-Yavin, and R. Pat-Horenczyk. Positive technologies for promoting emotion regulation abilities in adolescents. In *eHealth 360°*, pp. 169–174. Springer, 2017.

[35] P. Sequeira, T. Ribeiro, E. Di Tullio, S. Petisca, F. S. Melo, G. Castellano, A. Paiva, et al. Discovering social interaction strategies for robots from restricted-perception wizard-of-oz studies. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 197–204. IEEE, 2016.

[36] A. C. Strain and S. K. D'Mello. Emotion regulation during learning. In *International Conference on Artificial Intelligence in Education*, pp. 566–568. Springer, 2011.

[37] K. Zakharov, A. Mitrovic, and L. Johnston. Towards emotionally-intelligent pedagogical agents. In *International Conference on Intelligent Tutoring Systems*, pp. 19–28. Springer, 2008.

[38] K. Zakharov, A. Mitrovic, and S. Ohlsson. Feedback micro-engineering in eer-tutor. In *Proceedings of the 2005 conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology*, pp. 718–725. IOS Press, 2005.