

# Extended Abstract Virtual Tutor - Virtual Tutoring Agent using Empathy and Rapport Techniques

André Guilherme Carrasqueira Lima

Instituto Superior Técnico

**Abstract.** Going to University is a big change for someone and the student normally enters University feeling motivated. However, when the difficulties start to arrive and the reality that the effort required to succeed is much bigger than on the previous steps of their academic lives starts dawning on the students that just entered University they tend to lose motivation and cannot find the right methods of dealing with this loss of motivation, some of them fail courses they would on other circumstances find easy. Some universities have decided to implement a tutoring program in which the users are accompanied by a real tutor of the faculty. We intend to add to this program by delivering a tutor that is always available for the student and will help them emotionally. We studied what factors will help the student's emotional state and we arrived at rapport, empathy and relationship. We did an exploratory work on which one was more important by showing different behaviors for a virtual tutor agent in an app environment with the hopes of finding the most important factor among these 3. However we came to the conclusion that these three factors are equally important.

**Keywords:** Empathy, Rapport, Relationships, Virtual Tutors, Education, Emotional Intelligence, Technology, Virtual Agents, Virtual Assistants, Future Education.

## 1 Introduction

It is part of a student's life to find themselves with problems of adaptation to their new Academic life when they reach University.

To counteract this issue, some universities started having more office hours with the course professors so that students feel motivated to ask questions and clear all misunderstandings and questions that came up during that week. Some universities even started a tutoring program in which one professor is in charge of a group of students. This professor is normally not from any

course these students are attending but they are tasked with keeping the students' morale up and help them be more motivated and prepared, often talking about study methods and take preventive and corrective measures in case any of the above mentioned problems happen.

As we come into a more digital age where everybody is connected to each other and has a smartphone everywhere for apps at their daily convenience, we thought it was a good idea to contribute to the tutoring program by developing a helper for the tutor, a sorts of "Virtual Tutor" that would always be there for the student for the more daily tasks and coaching, leaving the more urgent tasks to the real tutor.

Our Virtual Tutor is inserted in the "*Tutoria Virtual*" program. A program from "*Fundação para a Ciência e Tecnologia*", a Portuguese public agency for science and technology. This program studies the pedagogical impact of Encapsulated Virtual Agents in online learning environments as Virtual Tutors available for each and every student for online support[30]. The Virtual Tutor was requested and will be used by "*Universidade Aberta*", a Portuguese e-learning course university, they want the Virtual Tutor as a tool for their students to use as a complement for their studies.

With the development of a prototype of the concept and the expected behavior we wanted the validity of our concept and idea as well as establishing the foundations of the empathic agent and the surrounding tool. We also wanted to find out whether or not this idea would have meaningful impact in our potential user's lives. We also investigate whether the relationship and rapport building aspects as well as empathy had an impact on the agent's reception by its potential users in order to find out if these factors were relevant to keep the users interacting more with the Virtual Tutor.

As the brainstorming for this idea first started we decided to gather the most relevant aspects of emotional agents and use them on the context of virtual tutoring. We got to a conclusion that the three most relevant aspects of emotional agents that we thought would be best used in this context were: **Empathy**, **Rapport** and the development of a **Relationship**. Aside from the

*Virtual Tutor* app we decided to break apart these three factors into smaller atomic behaviors that could help an emotional agent and explore whether or not there was a difference in the importance of these behaviors and a relative importance between these three factors in the context of a tutoring agent.

On the following pages we present the theoretical knowledge and relevant similar works that allowed us to reach the grounds for the Virtual Tutor as well as the prototype's architecture and all the different prototype steps and phases and their respective testings.

## 2 State of the art and related work

These related works and theoretical backgrounds are divided in four different topics that are the main topics that characterize our solution and testing areas: Virtual Tutoring Agents, Relationships, Rapport and Empathy.

### 2.1 Virtual Tutoring Agents

[25] proposed a paper in which they would develop a framework for a reliable and intelligent software that should have served a kind of tutoring function within a virtual team, their research's objective was to attribute an agent real human characteristics and behavior which in turn would enhance the agent's believability as a tutor and as a conversational agent alike. For that effect the agent would have the duties of a human tutor within a virtual team, these duties are:

1. Interrogator - Asking questions to the members of the group so they can answer them. These questions should help the members of the group in reaching a common goal.
2. Reviewer - Analyzing the answers provided by the team's members in whether these answers are correct or not. If any disagreement between the team's members occurs, the agent advises them to communicate in order to reach an agreement.
3. Monitor - Keeps a record of the answers from all members and of their communication with each other.
4. Instructor - giving individual instructions and helping the members who can not keep up with the progress of their group-mates.

[25] also argues that to corroborate the aforementioned changes in the educational system, new types of educational mechanisms and services need to be developed and supplied. These

services need to fulfill the following requirements: customization, adaption and support to the users when they deal with new technologies. [5] argues that agents have emerged to provide better solutions to these requirements as compared with already existing technology and they can also influence different aspects in educational systems: they supply new educational paradigms, support theories and can be very helpful for both teachers and learners. [25]'s work is non-hierarchical environment in the sense that everyone can teach and everyone learns. According to [21] in the beginning the main trend was to build an agent to replace a teacher in all the ways a teacher teaches. Now the environments are more exploratory and that gives the students more responsibility for their learning process. [35] believes that the roles of an agent of this type must be to monitor, control and catalyze social knowledge building among the community of learning.

In [25]'s proposal every agent has its mental model which contains: emotions, personality, attitudes, objectives and knowledge. The proposed framework has plans, questions, topics and actions.

A plan is a course of actions in order to achieve one of the agent's objectives.

An action can either be primitive or complex. Complex actions are called decisions.

Each question has a set of possible answers and to each possible answer a reply the agent will give if the answer is provided to it. When a question is asked, the agent will look for patterns of words and compare it with topics in its own base of knowledge so he can give the proper answer.

Each Topic can have zero or more subtopics. The learning material for a semester is a set of topics per lecture and the number of lectures. Topics may have precedence (which means a topic b may only be available after teaching a topic a). Each topic has an allocated time for it in lectures, associated questions, goals, results and consequences.

In our proposed virtual tutor we will be inclined also to provide a form of coaching and stabilization of the emotional status to our students.

*Coaching* is defined on [12] and [32] as a form of development between a coach and a client or coachee. The coach helps the client achieve professional or personal growth through training, advice and guidance. Most of the times the coach does not train the client in the specific subject they are accompanying, instead they take care of the client's state of mind and guide them to

harness their potential and making the most of it.

The virtual tutor will use a set of questions and pre-programmed answers to keep up with the student's academic progress. It will analyze and monitor their answers with the option to warn the human tutor responsible for the student if needed, and will also instruct the student in academic and non-academic levels (i.e. it can tell students to study and teach them but can also encourage them). We agree that a virtual tutor should not replace a human tutor or teacher, the virtual tutor will be a coach and a guide to the student while also shedding some lights on academic matters.

## 2.2 Relationships

In order to make a virtual companion for a human, we have to find out the factors behind the forming of a companionship bond between humans so we can simulate it in the agent.

According to [9] a relationship is generally defined as "a persistent construct, incrementally built and maintained over a series of interactions that can potentially span a lifetime". [7] refers that these interactions can be placed in two groups: generalized patterns for an interaction based on the stereotypes of society (doing something that society believes you are supposed to do for a relationship) or interactions specific to that two-people relationship (some friends have specific interactions that they do not have with anyone else). Because of this last type of relationship it is difficult to say when a relationship starts or ends or even what defines the relationship between those two people. For instance, two friends may talk rudely to each other and insult themselves while they talk to each other and laugh it off, for some this can be appalling but for the integrants in that relationship is a perfectly normal behavior.

Since we are trying to prove that a motivated and happy user will interact with our virtual tutor more often, it is relevant to look at the way relationships between humans are formed because according to [6] relationships are a relevant aspect for "human happiness, physical and mental health", in [7] we also see that some of the benefits of a relationship for a human being are that relationships can provide emotional support, emotional and instrumental support and other forms of nurture.

Levinger [24] proposes that human-human relationships go through five steps in that relationship's lifetime: 1- acquaintance, 2- buildup, 3- continuation, 4- deterioration, 5- termination. Since we are going to be focused on the creation

and continuation of a good relationship with a human being, we will need to focus on the first three stages that Levinger proposed.

Starting with acquaintance, the process of meeting someone for the first time, it is common sense that the first impressions will dictate our willingness to get along with the person we have just met. That is why most people already know that first impressions are the most important aspect of starting to build a relationship with someone. If the first impressions are good, we are more willing to get to know and even agree with that person, however if our first impressions of that person are not that good, we will be less willing to have any kind of relationship with that person.

As for the second and third stages proposed by Levinger, they have to do with the maintenance of a relationship. We maintain a relationship by ensuring we feel satisfied by its existence.

In [8], Bickmore bases himself on the human-human relationship research to point out some strategies that agents can use in order to maintain a relationship with a human. These strategies include performing new activities with the human, meta-relational communication (i.e., talk about the relationship), empathy, reciprocal self-disclosure, the use of humor, talking about past and future events, continuity behaviors such as greetings, and emphasize commonalities and de-emphasize differences, to increase solidarity and rapport. On that same article, Bickmore states that empathy plays a major role really an important factor for close relationships.

To study that exact factor, [13] shows us a framework for the creation of a realistic social game-agent companion called *ERiSA*, a framework that grants the agent the ability to establish simple social relations with a human throughout the course of their interactions together. This framework is built with five different components: the Sensing Component, the Interpreter Component, the Behavior Component, the Agent Components and the Game Components. The Sensing Component provides modules for the sensory inputs received from the user. In *ERiSA*'s case the sensory inputs are face recognition, facial expression recognition and speech recognition. This Face Recognition module tells the users apart making the agent recognize each user by their faces. The Facial Expression Recognition is based on FACS and six basic emotions [2].

The Interpreter Component receives data about the player: their facial expressions, utterances and identity. Their utterances are interpreted by a Verbal Interpreter Module while

their facial expressions are interpreted by a Non-Verbal Interpreter Module.

The Behavior Component defines the agent’s behavior based on the information of both aforementioned interpreters and the agent’s Interaction Rules. It is then sent to a SAIBA compliant embodied system [22] by the Action Selector Module. After that the Action Manager updates all necessary modules.

The Agent Component is divided in memories of internal rules, states and personality and a Social Relationship Module that manages the relationship between the two players at the time whether they are human or agents. Finally, the Game Components are everything necessary to have knowledge of the game such as rules and game states.

[13]’s agent has a personality based on the OCEAN personality model [33]. Although there were some limitations to this experience, they found out that an agent using this particular framework was able to establish a relationship with the users.

They developed a formula to implement a social relationship based on emotion, likes and personality traits. However we want the people to feel attached to the virtual tutor therefore we believe it is better not to use such a model because it could lead to aspects where the virtual tutor becomes less empathic due to being mistreated by the user. Using these studies on relationships

## 2.3 Rapport

The feeling of harmony, sync and flow that humans feel when they are engaged in a good conversation with someone is what we know as rapport [17]. Also according to [17], rapport mechanisms are what keeps the interaction interesting for both parts.

In [34]’s study about Rapport, they establish that in order to create this feeling of Rapport three factors must be considered: positivity, mutual attentiveness and coordination. If positivity is a strange word you should consider it as the capacity of an agent to be friendly towards the user. Only when these three aspects are working together can we build rapport. Tickle-Dengen and Rosenthal[36] mention that some of these rapport mechanisms include smiles, head nods, mutual attentiveness and coordination ( for example mimicking posture or using the same gestures). They also mention the three aspects of rapport as study [34]. They describe positivity as a feeling of approval and friendliness, these feelings can be achieved through humor or self-disclosure. They describe mutual attentiveness as the feeling we get when the other is solely focused on what we are saying. And coordination is described as the feeling of sync and predictability (such feelings can be induced by mimicking posture or the respect of social stereotypes). Studies [15], [14], [37], [16] and [11] each show a different area of social interaction that rapport facilitates: negotiations, management, psychotherapy, teaching and care-giving.

When starting any kind of relationship with someone, positivity is a massive decisive factor for building rapport and carries a huge importance because both interactors are behaving the way society expects them to: greet and be polite. However, as time and the relationship progress, positivity will lose its importance because both interactors expect less friendliness from each other than on the first meeting[36][39]. [39] also states that as time moves on and the relationship becomes stronger, both interactors may lose what we see as supposed behavior (or society imposed stereotypes) and form their own way of interacting. We can see it clearly if we think about those friends that treat each other with sarcasm and/or insults.

Eye-contact and gaze is one of the ways to build this feeling of rapport.[26] made this experience in order to see the actual effects of gaze and eye-contact on people by a mutual gaze agent that synchronizes gaze behavior with pre-recorded voice and gestures. They concluded that participants would recall a story better if the agent looked at them more often. But also

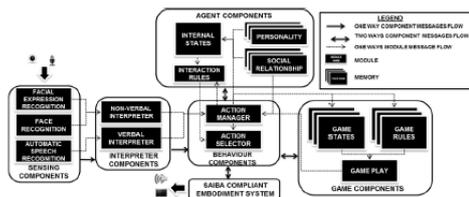


Fig. 1: [13]’s ERiSA’s framework overview

we will try to make the agent create a relationship with the user in order to make the user want more interactions with the agent. With this we hope that the agent will improve the users emotional and affective stages by making so that the agent can become the user’s friend and they can rely on it for their good and bad moments. To increase the relationship, the Virtual Tutor will keep a memory of former interactions and call the user by their name. It will also make some self-divulging comments to increase fondness and we believe that as time and interactions happen it will become friends with the user.

women felt better when the agent gazed less at them.

According to [39], in a context of tutoring, if the tutor is a friend of the student, the student will prefer if the tutor is a bit rude and more informal. However if the student and the tutor have no prior friendship whatsoever the student will prefer a more traditional, conservative approach. We plan on making the tutor behave more familiar with the student as the number of interactions grows, in order to emulate this information. Meaning that the more familiar the tutor becomes to the student, the more familiar and informal its speech will be. We plan to take this method into our Virtual Tutor so that as the number of interactions grow, the Virtual Tutor takes a more informal even "rude" approach with the user.

[18] created a Robot with a Rapport module that followed the three main aspects of building rapport. The agent should adapt to the surrounding society, behaving the way society expects it to. Their rapport module has the following components:

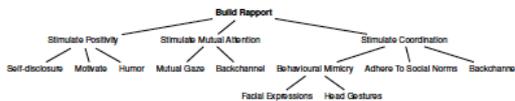


Fig. 2: [18]’s Goal Tree of the Rapport Model

In order to enhance mutual attentiveness their agent would produce listener behavior and make eye contact with the listener. Their *Mutual-Gaze Effector* is based on the works of [4]. It swaps establishing eye-contact with the participant and the game during established periods of time depending on whether a task is being performed or not and on whether the user is an introvert or an extrovert. They also use a *Backchannel Effector* based on the works of [27] to analyze pitch variations during interaction to produce listener behavior.

To enhance positivity the agent will make context-related vocalizations to motivate or praise the interaction partner. These utterances can also be of the self-disclosure type as suggested by [20] that self-disclosure has an important role in closing relationships between two strangers. And to enhance the final aspect of rapport: coordination, the model is composed by three different parts: behavioral mimicry, basic conformance to social norms and the already explained backchannel. The model mimics head gestures and facial expressions, these animations are proportional to the perceived emotional in-

tensity. [26]’s agent adheres to these following social norms:

- Introduces itself when it meets someone for the first time.
- Greets the interactor before starting the interaction.
- Avoids invading someone’s personal space.
- Says "please" when making a request.
- Says "thank you" when a person does a task asked by it.

This work showed no statistical proof that the applied rapport strategies increased likability. However they postulate that if they could have increased the sample size, they might have obtained a clear confirmation of said improvement on likability. However when they analyzed the video feed of the interactions with the users they noticed more frequent positive reactions from the participants in the rapport condition.

With the addition of rapport to our virtual tutor we plan on making it more human-like, in order to gain the user’s trust and also more engaging to talk to. We believe that if the user is more trusting of the virtual tutor and is able to speak freely, the user will be more motivated to continue interacting with the tutor and, if this project is successful, the user will also have an improvement in his or her school-life. We plan on having nods and smiles on our virtual tutor’s avatar in order to improve rapport.

## 2.4 Empathy

One of the most important components of the virtual tutor will be an empathy behavioral component, we give a major importance to empathy because [3] mentions that empathy is argued to facilitate the development and creation of social relationships. Since we want the user to establish a good social relationship with our virtual tutor, we believe empathy is key on achieving this.

According to [38], empathy increases fondness, similarity and affiliation. That is why it facilitates the development and creation of social relationships. Empathy also plays a crucial role in cooperative and prosocial behaviors such as helping and caring [19].

[19] gives us a broad definition of empathy as "an affective response more appropriate to someone else’s situation than to one’s own" and [31] adds that not only affective processes are influenced, but also cognitive and prosocial behaviors. An example of the former is for example how we take actions to reduce the other’s worries or distress. Work [10] showed that empathic agents are often perceived as more likable, caring and trustworthy than non-empathic agents.

Study [28] also shows that empathic agents can foster empathic feelings on their users. In [29] Paiva et al. shows us that empathic virtual agents can stimulate empathy in children helping them deal with bullying situations.

On the same work as we have the broad definition of empathy ([19]), Hoffman also states that empathy results from five different kinds of behavior: motor mimicry, classical conditioning, direct association of cues from the person we are empathizing with, mediated association and perspective-taking. Work [19] suggests that from these five, perspective taking is the most empathy arousing one.

[23] conducted an experiment in which they made a robot with an empathic artificial intelligence to study long term interactions with empathic robots. They evaluated this agent through a series of chess games and chess puzzles with children. This robot's A.I had a chess heuristic, which means it could look at the game state and find out who had the most advantage and what the best move normally was. By doing this taking into account the user's point of view (putting itself in the user's state). Because of this and some visual recognition, the agent was able to infer the user's affective state and act accordingly. The point of the robot was not to be a competitor but to be a friend to the children by cheering them on and giving them support on their plays or indecisive times. Some children said the iCat (the robot in question) was teaching them to play and at the end of the experiment some children even considered the iCat a friend. There were many factors that made the iCat achieve these results, but one of the major factors that distinguished the iCat was the empathy component. With and without the empathy component the results were different and the empathic iCat was perceived as being more friendly and more trustworthy. Most of the children that interacted with the iCat wanted another interaction. These results are part of why we think having an empathic virtual tutor will keep the users engaged throughout their interaction cycles. A state recognition and consequent perspective taking is one of the most crucial aspects of our work. With it we can infer from the user's grades, upcoming projects and past interactions whether the user's emotional state is positive or negative. As mentioned already, an empathic agent is normally perceived to be more likable by the users, we plan to use this likability to improve the frequency that the student uses the proposed virtual agent tool and as a result perhaps improving the user's results in their studies.



Fig. 3: I. Leite's iCat interaction cycle overview

### 3 The Final Prototype

Our Virtual Tutor app was built on Unity3d using scripts on the C# programming language and using Unity's user interface options. We chose Unity3d because it had compatibility with 2D programming as well as a component to develop the UI - user interface component of an app or game, it is also an engine that can port the prototype to mostly any platform and it works well with the necessary Android plugins, for example to implement the idea of the integration with the Google calendar for the study plans.

#### 3.1 Architecture

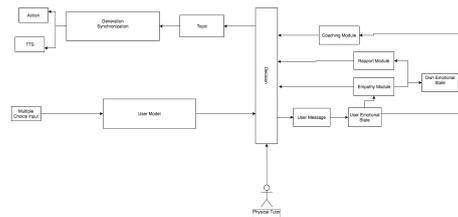


Fig. 4: Each agent's architecture

Figure 4 describes each agent's interaction loop with the user. This loop starts with an input given by the user as a choice of one of the possible choices colored green. This action gets saved as a user message and will be pondered upon in order to take a decision. This message will, if needed, alter the user's perceived emotional state. That user's perceived emotional state will be transported to the *empathy module* that will combine the user's emotional state with the agent's own personality and if needed it will cause an alteration on the agent's emotional state. The user's emotional state will also

be taken into account while doing coaching activities, in this case defining the optimal required time for the user to study for that next week, as we assume that the student's results are a manifestation of their own emotional state. As it is the *Coaching Module* will also contribute to the final decision based on the user's emotional state. The user's expression of it's own emotional state is decided by the *Rapport Module* and one of the facial expressions is chosen in order to express it, as well as other factors to build rapport, on a future time nods and smiles to acknowledge the conversation that is happening. As such, the user's message passes through the three modules and these three modules each take a weight on the decision of a topic of answer. A topic, as previously explained, is a set of lines, that each have their content and the agent that speaks them, the possible inputs for the user's response and all of the sentences timings. This topic will then be expressed through TTS and if the expression chosen by the *Rapport Module* is different than the one the agent is already making, the agent's expression will change with a small animation.

### 3.2 Dialogue

One of the key factors of the implementation on structuring dialogue was made by creating an object we called *Topic*, each *Topic* would be the highest object to dialogue in the hierarchy and would contain a set of *Lines* and a set of *Inputs*. This *Topic* served as a divider of the interaction in subjects. For example, when the Virtual Tutors greet the user, they are in the topic related to greetings, and in that topic there is information on the *Lines* they have to speak relevant to the greet the user and there is also information regarding the input choices the user has. Every time the user choses one of the input options the TOPIC is changed. Each *Line* object has information about which string of text needs to be "said", which tutor needs to "say it" as well as the start and end timing of that line, like a movie script or a movie subtitle program. Because we need to know which tutor has to "say" which sentence we have an *Agent* object that knows to identify its own position to know which agent it is and its own emotion. Lines also have their own timing, we implemented this timing feature for two main reasons: timeouts and animations. Timeouts were implemented as a transition to a *Topic* in which the users would not respond to the tutors after a fixed period of time after the last *Line* of that *Topic*. This was made as personality trait and a proactive attitude the agents can take. Concerning animations, we decided to give slight animations to the appearance

of the speech bubbles and the changes on the facial expressions in order for this change to occur more naturally to increase the believability of our agents.

### 3.3 Checkpoint



Fig. 5: Example of Checkpoints in the UI

We also created the *Checkpoint* object to encapsulate both the formative evaluations and the achievements deemed relevant by the course's responsible professor. We developed an abstract class called *Checkpoint* and these "checkpoints" would have an associated grade to it or they would just be items to check on a checklist. The evaluations correspond to pretty much all evaluation moments, be it exams, tests, mini-tests, essays or even projects, i.e. everything that can have a score. On the other hand, the important events correspond to events that can have a state of either true or false, as items on a check list.

### 3.4 Sprites and Expressions

On the final prototype we decided to make our character sprites, the images that would make the Tutors' representation for the user, from an expressions database by Cohn-Kanade[1] obtained through a student's license in order to have more natural and standardized expressions for our tutors. Each tutor had an array of

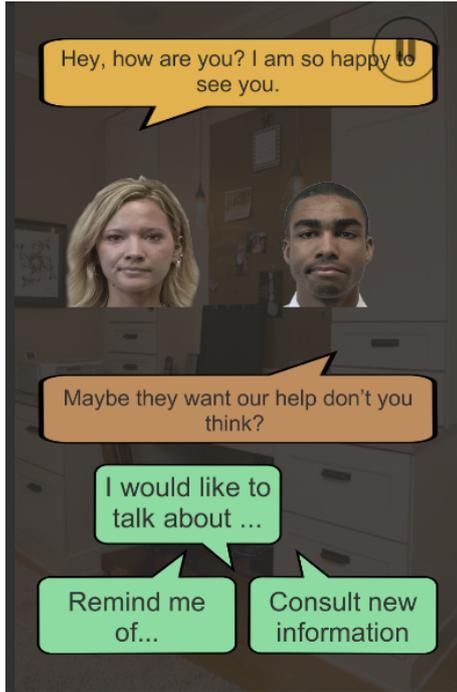


Fig. 6: The final Avatars

Sprites for the expressions, each sprite would then be corresponded to one *enum type* object that would help us control the tutor's expressed emotion. For instance: if the tutor was supposed to change their expression to show they were feeling sad we would make the Agent's expression enum variable become "SAD" which its numeric value would then correspond to the array index of the sprite correspondent to the sad expression. We found this decision key for our application and for the exploration work that we wanted to achieve because if the agents' expressions we chose for each case were not standard expressions of an emotion and chosen solely on our interpretation of the image they could lead to wrong interpretations from the part of the users and also if the expressions were not believable enough the users might not take the tutors as serious as we want the users to take them.

### 3.5 User Interface

We implemented some objects we called *Hook*, one for each Unity3d UI prefab that represented each different screen, and their respective *Control* objects in order to synchronize the UI and make the event functions that happened on the click of the mouse or when some values were altered.

### 3.6 Integration

We wanted the users to be able to interact with the Virtual Tutors in a way that was comfort-

able for their daily lives, their smartphones. This objective was reached using android plug-ins on the Unity3d C# scripts as alternate functionalities for the functions while they are operating on a device hosting the Android operating system. We implemented this measure in order to have the integration we mentioned in previous chapters so we could augment the positive experience felt by the users, simplify that same experience with familiar concepts and take advantage of programming for the Android operating system.

### 3.7 Sound

We decided to use a TTS module with pre-recorded sentences using an offline version of the Ivona TTS software we acquired so that the agents could say what they wanted to say in emotionless words instead of emotional gibberish devoid of any content. We believe the TTS to be important while interacting in a quiet environment because it provides hearing help to the user. As such, the user does not need to be as focused on reading the agents' lines and can pay attention to other details in the interaction such as the agents' expressions.

## 4 Experimental Protocol

We used the final prototype of the *Virtual Tutor* app as the grounds for the experiment to examine if there was any significant statistical difference between six behaviors we gathered each pair based on one of the three pillars our agent's emotion intelligence was based on: the theory of relationship, rapport and empathy. These 6 components were presented to the users as a short video and the users were able to grade them according to the importance they thought each component had, and also they were required to provide small description of maximum three adjectives to each observed behavior. For that we had to make some alterations to our agents' functionalities in order to isolate each of the behavioral components we wanted to explore and we just emphasized that specific behavioral component. We made these alterations mostly on the dialogue script and we normalized the scenario. This experiment took place online from July 31 to August 20.

The scenarios would all have the same common points: the user would login and interact with the tutors, after pleasantries and greetings, the tutors would warn the user that their Algebra's test grades had already been published, the user would then see it and notice that they had

gotten a 7.0 out of 20.0 for the most recent Algebra test, the video will then show both agents' reactions to the result depending on the behavior that we are trying to emphasize, different behaviors will have of course different reactions although the context is always the same.

For the relationships we focused on the phases of acquaintance and maintenance because we wanted to see how the users would react as to being treated more or less familiarly by a tutor that would impose this familiarity in their manners of speech, as such we wanted to know which kind of familiarity degree would ensure a more pleasant feeling or reaction from the user's part or if there was any difference in the way the relationship was emulated. As such we chose these two degrees: acquaintances and good friends.

The two clips that followed we chose the two most important rapport building factors in our context and we hope to know which of these is more important for the context of Virtual Tutoring: mutual attention through back-channeling or positivity from self-disclosure. One of the clips was focused on positivity through the strategy of self-disclosure and the other mixed coordination and mutual attention by a back-channeling through dialogue. Although the research done in the state of the art section shows us that positivity, mutual attention and coordination are fundamental in order to build rapport and have a harmonious interaction flow we want to discover which of these two factors is more important in this determined context, that means which one people notice more on the context of education.

The last theme we wanted to test was empathy. Based on the state of the art we gathered that empathy is achieved especially on two levels[23][3]: posture mimicry and perspective taking. Each of the clips focused on one specific aspect of empathy generation and will show these factors' importance on generating empathy with the user.

All of these videos were recorded and submitted to six different Google Form questionnaire that were only different in the order the videos were shown as to not make unconscious prejudicial thoughts while watching each video making sure that after a big number of participants the results we would gather were not dependent on the order the videos were shown and the answers we would gather would be balanced. We used a web page that would redirect each user to one of the respective questionnaires with the use of Roman Square strategy to avoid any kind of bias.

Each questionnaire was divided in two parts: demographic questions and the video analysis. On the demographic questions we asked the standard questions of age, gender, the partic-

ipant's familiarity with apps, the participant's familiarity with virtual assistants, the participant's familiarity with the calendar app on their smartphone and the participant's familiarity with NPC's (Non-playable characters in games). On the video analysis part the participant would watch each video at a time and after each video the participant would answer two questions: one a scale from 1 to 9 the importance of the observed behavior and three adjectives to characterize the observed behavior.

## 5 Results

### 5.1 Demographics

Our participants' age varied from 15 to 42 years of age. With an average value of 23.60 and a mode of 24 years of age. This means the most of our participants were around the intended target audience for the *Virtual Tutor* app. We had 18 female participants and 12 male participants which accounts to 60% female participation and a 40% male participation. 18 of our participants were University students and 5 were students elsewhere. That means that roughly 83% of the participants were students. Aside from asking each participant's gender, age and whether or not they were a student in University we asked four more questions about each participant's demographic background that we deemed relevant to their point of view regarding their experience with the Virtual Tutor's videos. These four questions were:

- How often do you use mobile apps?
- How often do you use the calendar app on your smartphone?
- Have you ever had experience with virtual assistants like Apple's "Siri", Microsoft's "Cortana" or Amazon's "Alexa"?
- How often do you play games that require you to interact with Non-Player Characters (virtual characters that you don't control)?

These control questions were used to later check whether or not the user's experience and familiarity with each factor had any relevance into their reactions to the virtual tutors.

With the answers to these questions we have gathered that 22 of the 30 people that answered the questionnaire use mobile apps every day which is around 73.3 %, with only one person of the 30 answering that they never use any kind of mobile apps.

We also gathered that the majority of the participants occasionally use their own calendar apps in their smartphones, 70% to be exact,

whereas 10% never uses it and 20 % uses it every day. This means 90% of the sample is familiar with the calendar app in their smartphones which we can extrapolate that most people are.

Regarding the familiarity with Virtual Assistants, the results were that none of the participants in the sample interacts with any of the Virtual Assistants given as examples every day and that only 10 of the 30 participants actually interact with them. That is 66.7 % of the people in the sample that never interact with a virtual assistant.

Finally, for the last question we can see that to 17 of the 30 participants, which corresponds to 56.7% it does not pose a relevant difference whether or not they have to interact with NPC's in a game. However, the percentage of people that prefer to interact with NPC's in a game was really close to the percentage of people that prefer not to interact with NPC's in a game.

## 5.2 Importance

		Statistics					
		Acquaintance_Importance	Maintenance_Importance	SelfDisc_Importance	BackChannel_Importance	Mimicry_Importance	Perspective_Importance
N	Valid	30	30	30	30	30	30
	Missing	1	1	1	1	1	1
	Median	6,50	7,00	5,00	6,50	6,00	7,00
	Mode	7 <sup>a</sup>	7	5 <sup>a</sup>	7 <sup>a</sup>	3 <sup>a</sup>	8

a. Multiple modes exist. The smallest value is shown

Frequency Table

		Acquaintance_Importance			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not important at all	1	3,2	3,3	3,3
	2	2	6,5	6,7	10,0
	3	1	3,2	3,3	13,3
	4	1	3,2	3,3	16,7
	5	5	16,1	16,7	33,3
	6	5	16,1	16,7	50,0
	7	6	19,4	20,0	70,0
	8	6	19,4	20,0	90,0
	Fundamental	3	9,7	10,0	100,0
	Total	30	96,8	100,0	
Missing	System	1	3,2		
	Total	31	100,0		

Fig. 7: Statistic information on the Importance variable in the six videos

Each video's importance was perhaps our most relevant indicator because our objective was to ascertain if any of the selected behaviors based on the three pillars of our research: empathy, rapport and relationship, was in fact more important than any of the others and if there was a relative order to them. From the participants' evaluation of the importance of each video we gathered the frequency table in figure 7. This figure showed us that the behavior that shows the highest mode and median values among the values of the importance of the shown behaviors in the questionnaires was the empathy strategy of perspective taking. It had a mode value of 8 out of 10 and a median value of 7 out of 10. Although the behaviors relative to the acquaintance strategy of the relationship and the rapport strategy

of back channeling had the same mode, they had slightly lower medians, meaning that these behaviors were, in frequency, voted as slightly less important by the public.. The behavior with the weakest mode was the one regarding the empathy strategy of emotional mimicry and the one with the lowest median was the behavior regarding the rapport strategy of self-disclosure.

## 5.3 Shapiro-Wilk test to check normality

We started our statistical tests by doing a Shapiro-Wilk test in order to realize whether or not the data of each video's importance had a normal distribution. Since the data did not have a normal distribution in at least one of those cases we would be forced to make non-parametric tests to ascertain whether or not there were relevant variants in the data for each of these cases.

## 5.4 Friedman General test

### Friedman Test

	Ranks	Mean Rank
Acquaintance_Importance		3,58
Maintenance_Importance		3,23
SelfDisc_Importance		3,20
BackChannel_Importance		3,72
Mimicry_Importance		3,28
Perspective_Importance		3,98

### Test Statistics<sup>a</sup>

N	30
Chi-Square	5,126
df	5
Asymp. Sig.	,401

a. Friedman Test

Fig. 8: Friedman test results on the variance of all behaviors shown in the videos

As such, the only data that had a normal distribution was the result from the video that emphasized the self disclosure strategy of building rapport by increasing positivity, we had to do non-parametric tests in order to find out

whether any of the behaviors shown in the videos had a statistic relevant greater importance than any of the others. In order to know for sure, we performed a Friedman test with all of the answers referent to all of the six behaviors' importance for the participants.

As we can see from figure 8, the value of p on the Friedman test was higher than 0.05, meaning that the values for the behavior's importances shown in the videos were considered equally important.

### 5.5 Wilcoxon signed-rank test in pairs

	Wilcoxon signed rank tests					
	Acquaintance	Maintenance	Self-Disclosure	Back-Channelling	Ministry	Perspective Taking
Acquaintance	X	0.913	0.2	0.662	0.409	0.706
Maintenance	X	X	0.597	0.452	0.859	0.316
Self-Disclosure	X	X	X	0.055	0.782	0.136
Back-Channelling	X	X	X	X	0.180	0.818
Ministry	X	X	X	X	X	0.238
Perspective Taking	X	X	X	X	X	X

Fig. 9: Matrix of significance of all the pairs of Wilcoxon signed rank tests performed

Even if in general all of the behaviors' importances were considered equally important, we decided to compare each video to check between every two behaviors if one was considered more important than the other. For that we did a series of Wilcoxon signed-rank tests with every possible combination of two videos. As we can see from figure 9 all the value of p on the Wilcoxon signed rank tests were higher than 0.05, meaning there are no relevant statistical differences between the importance of any pair of behaviors in the whole behavior set. Meaning there were no behavior was statistically more relevant than any other behavior when comparing the two individually.

### 5.6 Friedman test with a split the dataset

With all of the Wilcoxon signed-rank tests performed we could finally conclude that generally speaking with all our samples there was no statistical difference between any importance of any of the shown behaviors. Since this took into account all the data, we decided we could deepen the search for differences in the shown behavior's importances but now splitting the sample by the demographic groups trying to find patterns. We split the dataset several times by each category and found that even only accounting for people with the same demographic backgrounds there was no emerging pattern of any dominant behavior's importance.

## 5.7 Discussion

As it stands, there is no clear relative importance on any of the utilized behaviors for any of the three pillars of our research, making them equally important. However, we can not say that this data alone can clear the insecurity of the videos being explanatory by themselves for several factors: all of the tutors behaviors received a considerable amount of negative adjectives that some happened because the participants expected more from a tutor than just what it did, or they thought they were being dishonest or intrusive, in our opinion one of the factors is the ever present difference in how two people experience the same behavior that will lead to different and sometimes polar opposite reactions .Some behaviors that were not supposed to show emotion were deemed as emotional or sentient because the people felt the tutors were caring for them as they were just doing their job of tutoring, so the participants added some sort of caring persona to a unit who is responsible for another. Finally, some participants in our study did not see much difference between the behaviors shown in each of the video as some participants decided to leave similar adjectives or synonyms as answers to each video they were shown. This fact is also proven by some acquaintances directly pointing out to us after filling the questionnaires phrases like "I don't know if I did it right. They all seemed the same to me". As a result we feel that some videos were either too short to make the participants notice the differences in them. Another of our hypothesis is that the subject of the questionnaire was too vague, when we ask what the person thought of a behavior used by the agent and how important they think that behavior is without even revealing what kind of behavior we are studying it can sometimes lead to people seeing things that are not there, and we are afraid this factor may have also taken an influence with our sample. Perhaps there were too many characteristics to be tested and we decided to divide into six behaviors, perhaps it would have been better to study the difference between the importances of empathy, rapport and relationship building as just three separate behaviors in a Virtual Tutor.

## 6 Conclusions and Future Work

In this document we covered our gathering of requirements and the further implementation of a prototype of a virtual tutor with two complementary personalities that use strategies of rapport, empathy and theory of relationships to better interact with a human student in order

to perform tutoring and coaching, we presented early stages of the prototype to get concept validation and received positive results. After having a full interaction cycle developed, we had some usability tests done in order to discover interface and experience issues as well as concept validation with students, who are a different type of customer than the integrands of the committee to whom we presented the concept. The final version of the prototype that is described in full in this document was used as the basis of a study to identify which of the three factors that stood as pillars of our study was the most important in the educational context of our app. We come back to our hypothesis defined in the beginning of this document of examining whether or not one of the six characteristics based on rapport, empathy and the relationship between the user and the agent was statistically more important than the others, or if there was any specific order of importance among them. With the Friedman Tests and the Wilcoxon signed rank tests we reached the conclusion that there is no statistical difference among any of the six behaviors presented in the respective videos from the questionnaires filled by the users. However there may have been many inconsistencies in the videos that impeded us on getting accurate results for this experience.

As future work we would suggest making an in depth study on the preference of rapport as a whole, empathy as a whole and the relationship as a whole to find out if in general one of these is more important than the others. These factors should be on a more closed environment than "please explain what you thought of the tutor's behavior" because the question may be misleading. As far as the prototype goes, we would be happy if it was continued as a full app to support students throughout their university years and perhaps even younger students with some slight changes to it.

## References

1. October 2017.
2. T.R. Almaev and M.F. Valstar. Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition. *Affective Computing and Intelligent Interaction*, 2013.
3. C. Anderson and D. Keltner. The role of empathy in the formation and maintenance of social bonds. *Behavioural and Brain Sciences*, 25(1):21–22, 2002.
4. S. Andrist, B. Mutlu, and A. Tapus. Look like me: Matching robot personality via gaze to increase motivation. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015.
5. L. Aroyo and P. Kommers. Preface - intelligent agents for educational computer-aided systems. *Journal of Inteactive Learning Research*, 10, 1999.
6. E. Berscheid. Interpersonal relationships. *Annual Review of Psychology*, 45(1):79–129, 1994.
7. E. Berscheid and H.T. Reis. Attraction and close relationships. *The Handbook of social psychology*, pages 193–281, 1992.
8. T.W. Bickmore. *Relational Agents: Effecting change through human-computer relationships*. PhD thesis, Massachusetts Institute of Technology, 2003.
9. T.W. Bickmore and R.W. Picard. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2), 2005.
10. S. Brave, C. Nass, and K. Hiutchinson. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *Internation Journal of Human-Computer Studies*, 2005.
11. M. Burns. Rapport and relationships: The basis of child care. *Journal of Child Care*, 2:47–57, 1984.
12. P. Chakravarthy. The difference between coaching and mentoring. *Forbes*, 2011.
13. A. Chowanda, P. Blancheld, M. Flintham, and M. Valstar. Erisa: Building emotionally realistic social game-agent companions. *Intelligent Virtual Agents*, 2014.
14. J.W. Cogger. Are you a skilled interviewer? *Personnel Journal*, 1982.
15. A.L. Drolet and M.W. Morris. Rapport in conflict resolution: accounting for how face-to-face contact fosters mutual cooperation in mixed-motive conflicts. *Experimental Social Psychology*, 36:26–50, 2000.
16. D. Fuchs. Examiner familiarity effects on test performance: implications for training and practice. *Early Childhood Education*, 1987.
17. J. Gracht, N. Wang, J. Greten, E. Fast, and R. Duffy. Creating rapport with virtual agents. *Intelligent Virtual Agents*, 2007.
18. B.A.P. Henriques. *Establishing Harmonious Relationships Between Robots and Humans*. PhD thesis, Instituto Superior Técnico, 2016.
19. M.L Hoffman. Empathy and moral development: Implications for caring and justice. *Cambridge Univ Press*, 2001.
20. S.H. Kang and J. Gratch. People like virtual counselors that highly-disclose about themselves. *Studies in health techonology and informatics*, 2011.
21. J. Kay. Learner control. user modelling and user adapted interaction. *Kluwer Academic Publishers*, 2001.
22. S. Kopp, B. Krenn, S.C. Marsella, and et al. Towards a common framework for multimodal generation: The behavior markup language. *Int'l Conf. Intelligent Virtual Agents*, pages 21–23, 2006.
23. I. Leite. *Long-term Interactions with Empathic Social Robots*. PhD thesis, Instituto Superior Técnico, 2013.

24. G. Levinger. Development and change. *Close Relationships*, pages 315–359, 1983.
25. B.F. Marin, A. Hunger, S. Werner, S. Meila, and C. Schuetz. Roles of an intelligent tutor agent in a virtual society. *Proceedings of the 2005 Symposium on Applications and the interest*, 2005.
26. B. Mutlu, J. Forlizzi, and J. Hodgins. A storytelling robot: Modeling and evaluation of human-like gaze behavior. *Proceedings of the 2006 6th IEEE-RAS International Conference on Humanoid Robots*, 2006.
27. R. Niewiadomski, E. Bevacqua, and C. Pelachaud. Greta: An interactive expressive eca system. *Aamas*, 2009.
28. A. Paiva, J. Dias, D. Sobral, R. Aylett, P. Sobrepepe, S. Woods, C. Zoll, and L. Hall. Caring for agents and agents that care: Building empathic relations with synthetic agents. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, 1:194–201, 2004.
29. A. Paiva, J. Dias, D. Sobral, R. Aylett, S. Woods, L. Hall, and C. Zoll. Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, 2005.
30. FCT Fundação para a Ciência e Tecnologia. Projectos de investigação.
31. S.D. Preston and F.B.M. de Waal. Empathy: Its ultimate and proximate bases. *Behavioural and Brain Sciences*, 25(1):1–20, 2002.
32. Jane Renton. *Coaching and Mentoring: What they are and How to Make the Most of Them*. Bloomberg Press, 2009.
33. G. Saucier and L.R. Goldberg. The language of personality: Lexical perspectives on the vector model. *The ve-factor Model of personality: Theoretical perspectives*, pages 21–50, 1996.
34. H. Spencer-Oatey. (im)politeness, face and perceptions of rapport: Unpacking their bases and interrelationships. *Politeness Research*, 2005.
35. S.M.F.D. Syed-Mustapha. Agent mediated for intelligent conversational channel for social knowledge building in educational environment. *Proceedings of the fifth international conference on Information Technology*, pages 533–538, 2004.
36. Tickle-Dengen and R. Rosenthal. The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, 1(4):285–293, 1990.
37. P. Tsui and G.L. Schultz. Failure of rapport: Why psychotherapeutic engagement fails in the treatment of asian clients. *American Journal of Orthopsychiatry*, pages 90–104, 1985.
38. F. De Vignemont and T. Singer. The empathic brain: how, when and why? *Trends in Cognitive Sciences*, 10(10):435–441, 2006.
39. R. Zhao, A. Papangelis, and J. Cassell. Towards a dyadic computational model of rapport management for human-virtual agent interaction. *Intelligent Virtual Agents*, pages 514–527, 2014.