Creation of a Mental Health Virtual Assistant for College Students

André Antunes
andre.r.antunes@tecnico.ulisboa.pt
Instituto Superior Técnico,
Universidade de Lisboa
Lisbon, Portugal

Pedro A. Santos (Supervisor)
pedro.santos@tecnico.ulisboa.pt
INESC-ID & Instituto Superior
Técnico, Universidade de Lisboa
Lisbon, Portugal

João Dias (Co-Supervisor)
joao.dias@gaips.inesc-id.pt
Faculty of Science and Technology,
University of Algarve, CCMAR &
INESC-ID
Faro, Portugal

ABSTRACT

Recent research has established the potential for Virtual Agents to act as non-judgemental interviewers, eliciting greater self-disclosure though trustworthiness and credibility, in a medical care environment.

We hypothesise these qualities might also be applied to mental health support applications, namely in higher education communities. It is a fact that current systems lack the capabilities to cope with the growing need for mental health support.

MHeVA was designed with the goal of creating a Mental Health virtual Assistant (MHeVA) to elicit self-disclosure regarding mental health issues and establish rapport with higher education students. In addition to this, MHeVA can be a tool to diagnose anxiety levels and help to mitigate the stigma related to seeking help about mental health issues.

In order to evaluate the agent's effect on students a user study was conducted where participants from different groups interacted with MHeVA. The study measured self-disclosure, rapport building, anxiety levels and stigma mitigation.

Our findings suggest that the agent received high levels of acceptance and engagement and was able to elicit self-disclosure in students, as hypothesised. The rapport building and stigma mitigation were met with average results, while the measurement of anxiety level showed potential and accuracy. While these are encouraging the MHeVA needs to be further tested.

KEYWORDS

Virtual Agents, Rapport, Disclosure, Mental Health

ACM Reference Format:
André Antunes, Pedro A. Santos (Supervisor), and João Dias (Co-Supervisor). 2022. Creation of a Mental Health Virtual Assistant for College Students. In . ACM, Lisbon, LIS, POR, 9 pages.

1 INTRODUCTION

Virtual Agents (VAs) have been a growing part of our society, becoming ever more capable of establish stronger emotional connections with humans [1, 2]. This ability has allowed these agents to establish trust and rapport and consequently be better at tasks that involve helping the human society. For instance, the use of such agents in pedagogic and professional experiences, has proven they can be a pedagogical asset as language trainers [3]. It as also shown their power to influence behavioural changes. The capacity to promote exercise in young adults [4] or even influence behavioural changes in a sensible environment, helping adults with schizophrenia to comply with their medication [5]. Besides day-to-day behaviour, VAs have proven to be capable of providing socio-emotional benefits. Using emotional and cognitive support, they are able to reduce the intensity of negative emotions in humans, such as anger and worry [6], offering an improvement in quality of life.

Acknowledging their capacities for rapport building and influence behavioural change, we turn towards the ability to take the role of interviewers and elicit self-disclosure from humans about their medical health related issues. In this domain, where patient's health is often at risk, the use of VAs to interview patients about their medical state was tested, proving that patients tended to be more honest in their answers, when interviewed by VAs, without external human involvement [7]. They have been accepted and trusted even when diagnosing Major Depressive Disorder (MDD) or tobacco and alcohol use disorders, with their credibility begging a factor to the self-disclosure that allowed for a successful diagnose [8].

Often they are equipped with modular architectures, taking advantage of advanced tools to perceive through audio and image, non-verbal behaviour and natural language understanding. Or dialogue managers, that allow for complex dialogue systems that mimic human conversational interactions. Using these tools, VAs are even able to promote self-disclosure about psychological distress [9].

These use cases, are only a small sample of the wide variety of studies, that have shown the positive impact VAs can have in our society and their capabilities to build rapport and improve interviewees willingness to self-disclose. So we focused our work in one promising aspect of the human life we believe they can improve: Mental health. More specifically, mental health in higher education students.

1.1 Mental Health in Higher Education

To test a VA in this field, we have understand the issue of mental health in higher education, where the mental health organizations have reported that colleges across the world are contending with rising rates of mental disorders [10]. It is aggravated by the fact that the demand for services on campus far exceeds the available resources, where it is impossible to provide quick and consistent access to professional psychologists, for every student, simply due to lack of capacity and resources [11, 12]. Additionally, there is yet

another problem: the stigma that is associated with somebody seeking psychological help, which affects the student community and makes them avoid seeking professional help, due to fear of judgement from other people [13]. Now if we apply all these problems to the current reality we live in, of a world coming back from a two year long pandemic, we see that besides worsening the previously enumerated problems, it also created new ones [14, 15]. The Federação Académica de Lisboa (FAL) made a study that revealed that up to 55% of Lisbon's college students have worsened their mental health condition and up to 83% didn't search for psychological help, from the available systems provided [14]. We now know that a lack of academic results, anxiety issues and even deterioration of mental health, become recurrent if the proper assistance is not provided. This is more than enough motivation for us to try to search for a solution, outside the conventional methods and, in our view, after studying their application, there are strong indicators that VAs can be used as a tool to mitigate mental health issues in the higher education community.

1.2 Objectives

The primary motivation for this work was to understand if Virtual Agents can be positive influences and helpful tools for supporting students suffering from mental health issues, namely anxiety. We mainly focused on their ability to promote self-disclosure, with emphasis on answering the following question:

- **Will Students Disclose with the Agent Regarding their Mental Health Issues?**

Self-disclosure is the sharing of any information about oneself to another entity [16]. Self-disclosure is beneficial in numerous settings, specially when that information is needed to improve the health of the one disclosing. In dealing with mental health problems, the psychologist needs to obtain information from the patient, in order to understand the problems that affect them. However, eliciting self-disclosure can be challenging due to factors such as fear of self-disclosure [17].

Furthermore, there are several factor that are related to the ability of eliciting self-disclosure and several ways the mental health support services can benefit from a VA. So in order to understand the depth of the support a VA can provide and what factors influence its performance, we will also tested: (1) Its acceptance and engagement, (2) the ability to build rapport with students, (3) its accuracy measuring anxiety levels and (4) its influence in mitigating the stigma related to seeking help about mental issues. These were the questions we found most relevant to answer, advised by a mental health support professional from Instituto Superior Técnico (IST)[11], that gave us valuable insight on the present situation of mental health in higher education, and that advised us on how to build this agent in accordance to mental health support practices and regulations.

2 MHEVA

Advised by the IST mental health support service, we designed the best way to deploy our VA, so it could best fit the environment and needs of the system already in place. The challenges that seemed more important to tackle were the reach of the support provided and early identification of mental health issues in students, in order to prevent their worsening. Thus, a mental health support assistant accessible to students and able to identifying mental health issues seemed the best format. Furthermore the agent should be able to advise students on how to deal with their mental health issues and forward them to mental health support professionals, if the issues were too severe. This would allow the agent to best support the system already in place and the work of the professionals on the field.

It was also important to narrow the agent's support to specific mental health issues. Anxiety was our choice, since it is one of the most common mental health problems that affects college students [11] and one less prompt to produce severe consequences.

2.1 Format of the Interaction

The scenario of our interaction between MHeVA and the student, will be a one on one interview, a widespread format of therapy used in mental health support, that allows for a better control of the interaction and better privacy. Both crucial factors to empower the interviewer and leave the interviewee more comfortable to self-disclose, since it isolates the conversation from any exterior distraction or judgement. This format follows on the footsteps of several studies already made, that had positive results [6–9], where a VA takes the role of the interviewer.

2.2 Behaviour

MHeVA will adopt a goal-based behaviour that will focus on two main goals:

- **Building Rapport**
- **Obtaining Disclosure**

Its decision making will revolve around the priority of these goals, acting in conformity with their correlation. In order to obtain willingness to self-disclose from the student, it first needs to earn their trust by building rapport. Consequently, the agent will create a Student State that will store its beliefs about the interaction and current relationship with the student. This concept is based on the Theory of Mind, which is the capacity of storing beliefs about the beliefs of another entity [18]. Through these beliefs, MHeVA can decide which action to execute. If it has enough rapport it will ask questions about anxiety issues, if not it will strengthen the relationship.

In order to build rapport MHeVA focuses on two types of exchange that have proven to be effective[19], courteous behaviour (honesty, civility, empathy) and information sharing behaviour (offering advice, sharing knowledge, asking questions). It tries to exchange names at the start of the interaction, so it can refer to the student by their name, something that has a positive effect on the user's perception of an Agent [20]. It is always courteous an professional, thanking every information disclosed and behaving with comprehension when the student is more reserved. And finally it focus on the information sharing behaviour to establish the perceived rapport. Every time MHeVA feels it needs to build rapport, it changes the subject to non-sensible topics about the student, so the exchange of information might improve their relationship and the student's perception of the agent.

In order to obtain willingness to self-disclose, besides building rapport, MHeVA ensures three requirements that strongly influence

this capacity in VAs: (1)Trust in terms of security and confidentiality, (2) Credibility, normally associated with accuracy and organizational credibility and (3) Ability to listen to the user and react to their utterances [21]. MHeVA clarifies that all the information shared during the interaction will be protected and that their identity will be kept safe and hidden. Secondly, achieves credibility by associating our study with the mental health support services from our college and explaining the purpose of this research. Thirdly obtains accuracy through the knowledge shared by our mental health support professional[11], that keeps every utterance, behaviour and information provided by MHeVA, in accordance with the correct practises and regulations of psycopedagogy support. Finally the ability to listen and react to the user utterances is obtained by follow-up interactions, that adapt to the student's answers, thus making sure they feel they are being listened too.

When the agent feels it has built enough rapport, it changes the questions asked to anxiety screened ones, in order to ascertain the student's levels of anxiety.

The interaction comes to a close when the agent has asked enough anxiety screening questions to render a verdict on the student's anxiety levels.

2.3 Dialog

The main modality of interaction is text-based dialog. We opted for a scripted dialog that presents MHeVA with choices to pursue its goals, and choices for the student to answer to the posed questions. The interaction is be turn based, where the agent asks questions and the student answers and none of them is be able to act before the other one finishes its utterance.

MHeVA interacts mainly by asking the student questions that fall in one of two categories: (1) Anxiety Screening Questions and (2) Rapport Building Questions. The first try to assess if the student has anxiety problems and if so, what is the seriousness of it. They cover a wide array of topics, related with the student's behaviours and physical or psychological consequences of anxiety. These questions were selected by a mental health support professional[11], based on the Generalised Anxiety Disorder Assessment (GAD), in accordance to their practises and all of them will help to measure anxiety levels. The latter have the objective of creating a trusting relationship between MHeVA and the student. They won't cover sensitive topics as the questions above, on the contrary, they will have day to day topics of conversation. They will allow the student to share information about himself, that isn't related with mental issues, strengthening the bond with the agent [19]. Additionally, in order to take full advantage of every shared information, this questions will be oriented towards obtaining some insight on the origins of anxiety, if the students have any. For example, academic, social and parental relationship topics are normally perceived as day to day topics, but are often related to problems of anxiety and depression. So innocently speaking about this issues, might help MHeVA achieve both rapport and knowledge about anxiety problems.

As for the student, for each question asked by MHeVA there are several available choices of answer. They compromise a representative range that although sometimes might not allow for precision, always allows for veracity. Furthermore, to every question there is a choice that allows the student not to answer to the given question (eg:"I don't feel comfortable answering this question"). This is mandatory, because in mental health interviews we should never force an answer from the interviewee. They should only disclose information of their own accord [11, 22].

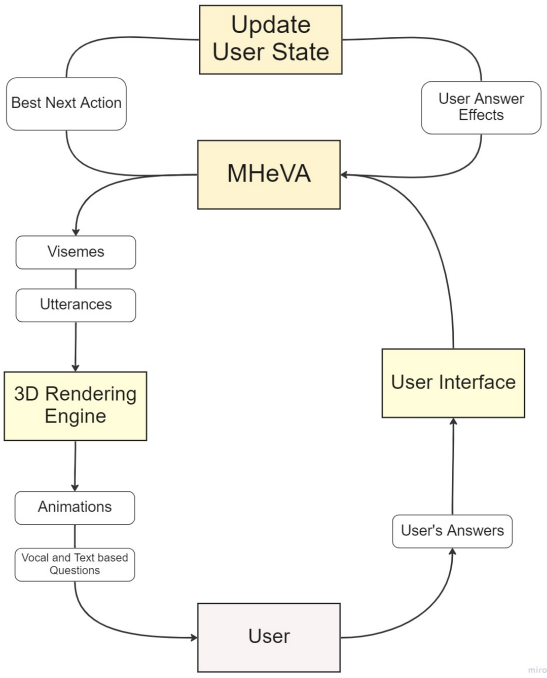

Figure 1: MHeVA's architecture

2.4 Sub-Tree Dialog Logic

Our intention with MHeVA's interaction was for it to be close to a mental health related conversation and not to be interpreted as a mental health questionnaire. Hence, the agent focuses on appropriate topics of conversation and avoids abrupt change of topic even to another socially sanctioned topic, to avoid endangering the healthiness of the interaction and the possible pre-established rapport[23].

MHeVA always complete the current topic, before moving to another one. So in order to facilitate this process we divided the dialog tree, into several sub-trees, that are identified by a main question, covering a specific topic. Additionally, the MHeVA's evaluation of its goals and the choice of what next question to pursue only happens at the end of each sub-tree. These sub-trees besides the main question have follow-up questions, reactive responses (eg: "Thank you for sharing this information") and informative feedback about mental health.

There are three types of sub-tree, excluding the introductory ones, that are connected by the main sub-tree hub, where the decision making is executed:

- **Rapport sub-tree**. A total of three sub-trees with day-to-day topics that focus on information sharing and rapport

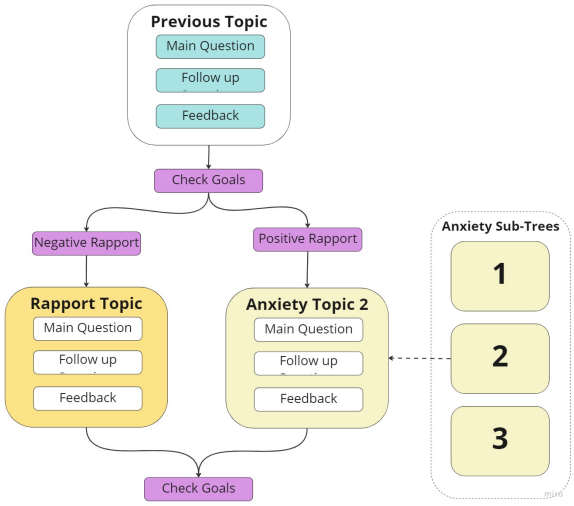

Figure 2: Sub-tree dialogue system

building. If the user shares information during these topics, MHeVA will increase the perceived rapport.

- **Anxiety sub-tree**.A total of six sub-trees focused on obtaining information about the user's anxiety levels, with the following issues: (1) Recent tiredness, (2) Feeling something bad and inexplicable is about to happen, (3) Recent feelings of nervousness, (4) Frequency of getting upset, (5) Anxiety attacks and (6) Sleeping difficulties. These issues were selected by a mental health support professional[11] and the user's answers increase or maintain the MHeVA's perceived anxiety levels.
- **Final Verdict sub-tree**. At the end of the interaction, MHeVA will give council to the student about their tribulations: Revealing if there are signs of anxiety, how serious they appear to be and how to best cope with them.

2.5 Appearance and Voice

Besides the spoken interaction, MHeVA is embodied with a human model and equipped with a computer-generated voice (CGV) and Lip-Sync animation.

This follows several studies' findings, that support that more anthropomorphic representations of VAs, improve their perceived competence and trustworthiness [24]. And that humanness of synthesis leads users to associate VAs with intelligence and personality [25, 26], enhancing one's perception that 'someone' is socially present and collaborating in the same space [27, 28].

2.6 Deployment

One of our objectives is to not just mitigate the stigma associated with seeking help but also to ease the access to mental health support, empowering and sharing the workload that the professionals on the field are saturated with. This means that MHeVA should be of easy access to any student. Thus its deployed in a computer build, since almost every student, if not every single one, has a personal one and is proficient in its use.

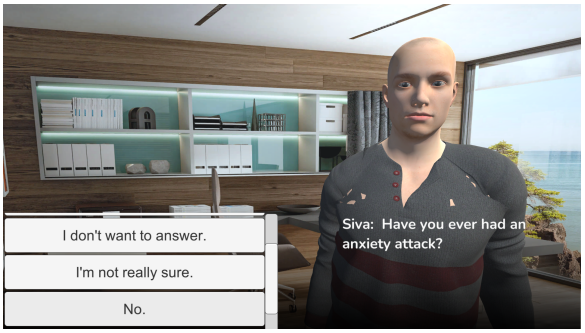

Figure 3: The final Unity 3D build of the MHeVA interaction.

3 RESEARCH METHOD

We realized two different testing phases. One in August and another in October. The first had a greater number of testers and was participated by college students, indiscriminately selected, so we could try to represent the general higher education community. And the second testing phase, participated by college students diagnosed with anxiety by mental health service professionals, to test if MHeVA was accurate in its anxiety evaluation.

To replicate the environment of the expected interaction, between a student with mental health issues and MHeVA, the testers interacted with our agent without interference from our team.

We chose a self-reported survey has the main tool for data gathering. Moreover, we complemented it with some data that we needed to obtain actively through scripting, namely general data about the interaction duration, MHeVA's verdict on anxiety levels and chosen answers, to name a few. This was obtained through a text file our agent outputted after the interaction (the identification of the testers was never registered and was kept anonymous).

The survey was constructed having in mind several questions that measure the success of our objectives. It was divided into two parts. The first one, filled before the interaction, focused on obtaining some basic information about the user, previous experiences interacting with Virtual Agents, mental health support and experiencing anxiety related issues. The second one, questioned the tester about their interaction with MHeVA, how they felt about the agent and the effects of the conversation.

3.1 Deployment

We deployed our agent through a web page, were the users had to download the build file and execute it. We left a four step instruction on the website for the users to know what to do and in what order. The survey itself was deployed via google forms and the link to it was attached in the web page.

We reached all the testers through personal messaging and discord servers, making a small announcement for students to help in our testing phase. Our target population was college students or recent graduates mainly from the University of Lisbon.

3.2 Participants

We managed a participation of 44 testers, 27 males and 17 females, from which 3 were recent graduates. The vast majority was from

IST, 30, while the rest came from different colleges, namely Faculdade de Belas Artes, Instituto de Educação, Faculdade de Medicina Dentária, Faculdade de Ciências e Tecnologias, Instituto Superior de Engenharia de Lisboa, Escola Superior de Educação de Coimbra and Faculdade de Ciências Sociais e Humanas.

4 RESULTS

We conducted a statistical analyses of our findings to render a comprehensive verdict. Quantitative variables were expressed with means (M), and standard deviations (SD), and qualitative variables were expressed using percentages (%). We calculated the correlation between variables using Pearson's Correlation (P), and Cramer's V Correlation (C).

4.1 Mental Health Support

The first data related to mental health helped us prove the insufficiency of the support provided in colleges and how students tend to avoid seeking help.

Even though the vast majority reported to have had problems related with anxiety or depression, almost half (43%) of that majority didn't seek help in dealing with this issues. This might be justified by the stigma that normally is related to the act of seeking mental health support, but the responses suggest very low levels of this preconception. We studied the association between both variables, through Cramer's correlation and it indicated a weak link (C = 0.199), between those who did look for support and those who had stigma against it.

4.2 Duration

The tested interactions varied in terms of duration. We managed to record it, since MHeVA kept track of it from the first utterance to the last and wrote it on the output text file.

The interactions lasted from 2 minutes and 16 seconds, to 10 minutes and 30 second (M = 4:24, SD = 1:25), providing reasonable variety.

We asked the users what they though about the duration of the interaction, using a 5-point Likert scale, from 1 (too short) to 5 (too long), revealing a general satisfactory opinion (M = 2.74, SD = 0.59). However when calculating the Pearson Correlation between the interaction time and the satisfaction rating, we found a weak correlation (P = 0.186).

4.3 Freedom of Answer and Clarity of Choice

Since we opted for a structured dialogue tree, mostly composed by pre-written choices, we had to test if this approach did not hinder the students' satisfaction and capacity of disclosure. We used a 5-point Likert scale, from 1 (Strongly Disagree) to 5 (Strongly Agree).

The prompt with the sentence "I had perfect freedom of answer", gathered moderate positive results (M = 3.37, SD = 1.02), revealing a certain satisfaction in the choices the students had at their disposal. This is not indicative that the testers wouldn't enjoy more freedom of speech, a natural processing or voice recognition systems. It only allow us to confirm that the freedom of choice did not generally impact negatively the testers' experience.

When having in mind the clarity of the available options and the easiness to chose one of them, the results were even more positive, for both clarity (M = 4.42, SD = 0.7) and easiness (M = 4.19, SD = 0.88).

4.4 Acceptance and Engagement

To measure acceptance and engagement of the students towards MHeVA, we utilized five prompts, rated in the base Likert scale already mentioned above.

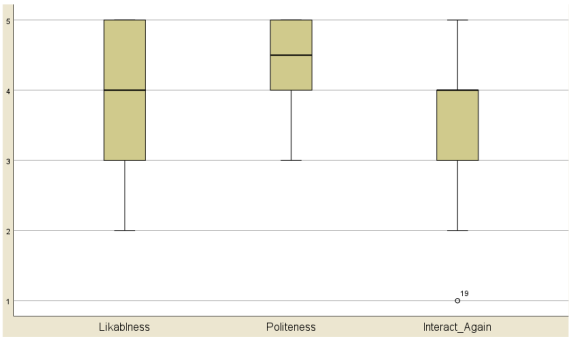

Figure 4: The acceptance and engagement ratings

We tested the likableness with the prompt "I liked to interact with SIVA", to understand if the interaction was enjoyable to the students and the politeness parameter with the prompt "I think SIVA was nice to me", since the politeness parameter is important for positive perception of a VA [20]. Both likableness (M = 3.95, SD = 0.86) and politeness (M = 4.39, SD = 0.69) had positive results, revealing the students perceived a level of politeness from SIVA and all in all enjoyed the interaction.

We also tested the students willingness to engage in a new interaction with MHeVA, a good measure of acceptance and engagement [8] (M = 3.70, SD = 0.95). The results can be seen in the figure (Fig.4).

Lastly to try and understand if the communicative abilities of MHeVA were well met and well understood, we checked if its messages were clear and natural. The ratings were considerably high for clarity (M = 4.57, SD = 0.55), revealing that the language and utterances used was clearly understood by the students. And although the second ratings were lower (M = 3.66, SD = 1.10) it still gave us grounds to conclude the utterances were perceived mostly as natural ones, as opposed to scripted and robotic.

To conclude the engagement measurements, we took advantage of the agent's output file, were we registered if during the interaction, the testers exchanged names with MHeVA (we did not store the actual names, only the positive or negative disclosure). From the 44 testers, 40 of them exchanged names (90%), allowing MHeVA to communicate on a first name basis.

4.5 Rapport

We measured the rapport in to three scales, following the definitions by Jonathan Gratch et al. [29], using a 5-point Likert scale.

First we measured the emotional rapport, using the items "I felt a connection with SIVA" and the related pair, "How was your relationship with SIVA during the first 5 questions?" and "How was your relationship with SIVA during the last 5 questions?", to understand the evolution of the relationship perceived by the students. The

results were a bit lackluster, with only two testers fully agreeing they had established a connection (M = 2.64, SD = 1.26). Moreover the perceived initial relationship ratings were reasonably lower too (M = 2.27, SD = 0.85). However there was a clear improvement on the perceived relationship by 0.87 points (M = 3.14, SD = 1.25).

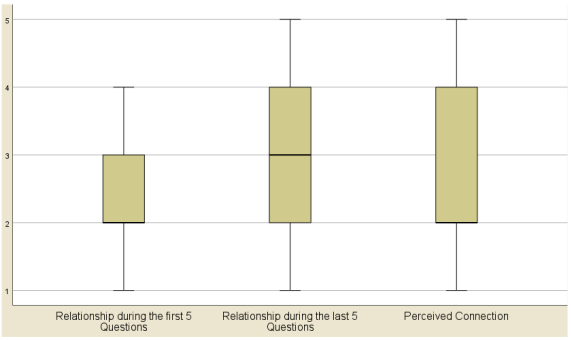

Figure 5: The participant's perceived rapport ratings between them and MHeVA

Further on, we measured the cognitive rapport, through the item "SIVA and I understood each other", from 1 (strongly disagree) to 5 (strongly agree). The results were slightly positive (M = 3.18, SD = 1.195), closely related to the perceived strength of the relationship in the last five questions.

4.6 Disclosure

We decided not just to measure the amount of disclosure obtained but how difficult it was for the students to share their issues. In order to obtain more accurate results, we gathered data not just from the survey, but from the output file, relating the interaction.

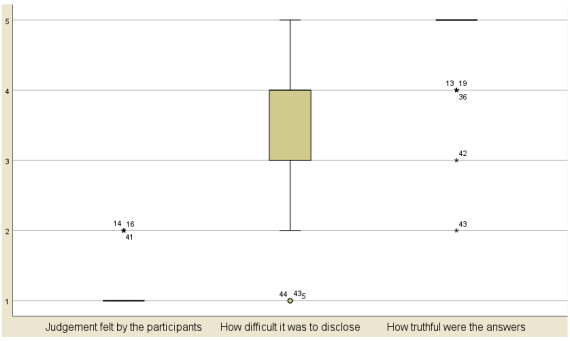

Figure 6: The participant's disclosure related ratings

It was important to understand possible judgemental factors regarding mental health disclosure, which often hinder the interviewers' job in this field. The majority of participants strongly disagreed with the item "I felt judged by SIVA" (M = 1.14, SD = 0.35), in a scale of 1 (strongly disagree) to 5 (strongly agree).

Participants were also asked to report their own feelings of disclosure and level of honesty behind their answers. The study obtained good values (M = 3.66, SD = 1.14) from the prompt "How difficult was it for you to open yourself about personal issues?", scaled from 1 (Hard) to 5 (Easy), with only 3 rating it 1. Truthfulness, a very important aspect to understand if the disclosed information was honest, obtained even more positive results (M = 4.70, SD = 0.63). However both these ratings are self-reported, so we can't be totally sure they are accurate.

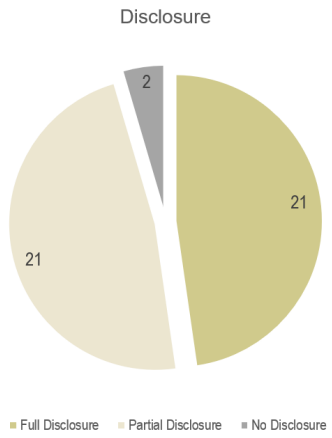

Figure 7: The participant's disclosure levels

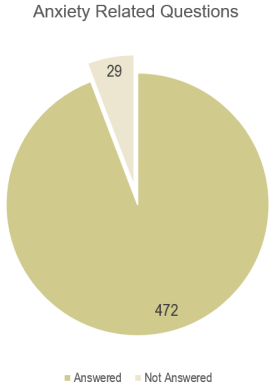

Figure 8: The participant's disclosure practical values

In order to properly confirm our findings without self-assessment questions alone, we have also looked into participants experience when interacting with MHeVA, through the registered the answers in the output file (except the name of the student, to assure anonymity). From 501 anxiety screened questions, including the main and the follow-up questions, 472 were answered, obtaining a very positive value of disclosure (94%). Additionally, from the 44 participants, 21 answered all anxiety related questions asked by MHeVA and only 2 out of 44, did not answer at least one anxiety screened question.

Table 1: Compared values between the self-reported anxiety and the MHeVA's evaluation

Anxiety Level	Didn't Know	No Anxiety	Had Anxiety
0-1	2	1	1
2-4	0	2	10
5-7	2	1	13
>7	1	0	11

Table 2: The calculated correlation between the self-reported anxiety and the MHeVA's evaluation

Nominal by Nominal	Value	Aprox. Significance
Phi	.749	.424
Cramer's V	.530	.424

4.7 Effects on Mental Health

To try and understand if our agent had the potential to be a positive influence, to the participants and to the mental health support system, we used two questions: (1) "Did your opinion about mental support improve after this interaction?" and (2) "Do you feel SIVA helped you with some issues you might have?".

The ratings achieved neutrality in stigma mitigation (M = 2.95, SD = 1.01) and neutral to low ratings in mental health understanding improvement (M = 2.59, SD = 1.04).

4.8 Ability to identify Anxiety

In order to understand if the final verdict given by SIVA and consequently its evaluation of anxiety levels were accurate, we had to cross examine some data obtained, by both SIVA and the students and posteriorly, test our Agent with students that were already diagnosed with anxiety problems, by professionals in the field of Mental Health Support.

In the first group of testers, from the 44 students, 4 were reported having no anxiety problems, 12 a not troublesome level, 11 having at least one anxiety crisis and 17 having experienced several anxiety crisis. This had to be compared with the self-reported question that had half (52%) of the students revealing they had anxiety problems. Of the 35 students who reported having had anxiety problems, 11 were evaluated with the maximum anxiety level (more than 7), 13 with the middle level (5 to 7), 10 with the low level (2 to 4) and 1 with no anxiety (less than 2) (Tab.1). It is important to note however, that the one who was reported with no anxiety did not answer anxiety screened questions, providing no measurements for MHeVA and consequently cannot be considered an evaluation of anxiety levels. We calculated the correlation between the self-reported anxiety and the anxiety level evaluation and found moderate levels as presented in the table (Tab.2).

At the time of writing this document, we have, so far tested our agent with two students that were diagnosed with anxiety. Participants were willing to disclose information, sharing both their names and answering 25 of the 26 anxiety screened questions they were asked by MHeVA. They scored 7 and 8 for anxiety levels, which places them on the most serious level, corresponding to have experienced several anxiety crisis. We also tested our agent with a student that was diagnosed as not having anxiety issues and for that student MHeVA did not detect any level of anxiety.

4.9 Discussion

We supported the several studies that claim an insufficiency of the support provided in colleges and how students tend to avoid seeking help. Nevertheless we didn't find correlation between stigma against seeking help for mental issues and the avoidance of this seeking. Being a self-reported form, there might be misjudgement on what stigma really is and how it affects one's actions, however, we prefer to look for answers in the question related to accessibility, were most of the participants thinks this kind of support is lacking in college.

Looking into the details of the interaction, freedom of answer and clarity of choice did not negatively impact the experience, even though we went for a more structured conversation. Also the duration wasn't a strong factor, since the satisfaction was positive and there was no correlation between both.

Also in the fields of acceptance and engagement, our agent ranked high in the self-reported survey. Namely in likeness, politeness and the prospect of a new interaction, a good measure of acceptance and engagement [8].

It was the capacity to build rapport that lacked proficiency. Although it managed to improve the perceived rapport during the conversation, the ratings were still a bit lower than anticipated. However, since the willingness to disclosure was so high, it might mean rapport and disclosure are not so well related as we initially though, at least in the field of mental health. Supporting the claim that factors such as Credibility, Security and Accuracy might be more influential to achieve this goal [21].

Our agent wasn't prepared to extend deep advises about mental health, so if someone already contacted a mental health support professional, they would feel a lack of depth from our agent's answers. This fact might explain part of agent's low score regarding this metric. However that does not mean it can't help with mental health issues if its properly equipped. Regarding improvement on the opinion about mental health support, the score reflects a slightly higher rating but not influential enough. This might be because the students did not associate MHeVA directly to the system already in place, since it is still a research tool. Perhaps if it shared more information about the mental health support channels in college, it would help improve this score.

Finally, looking at MHeVA's capacity to measure levels of anxiety, its results were promising. There was a certain correlation between the self-reported survey and the agent's evaluation. However this statistical study is merely theoretical and self-reported anxiety is a flawed data set, since the students are not qualified to diagnose anxiety. Additionally, even after MHeVA accuratly predicted the anxiety levels of three diagnosed students, we feel the number is simply to low to generalize the accuracy of our agent. These early findings are promising but there needs to be further testing.

5 CONCLUSIONS

In this thesis we looked at virtual agents as a possible answer to the ever growing problems in mental health services in higher

education. We tried to understand if college students would be willing to disclose mental health related issues with a VA deployed as a MHeVA and consequently provide grounds for a possible future deployment.

In order understand the current situation of the mental health support in higher education, we studied several works and studies on the subject and contacted a mental health support professional from IST[11]. Later to find out what architecture would best fit this agent and this interaction we undertook a deep research about already established VAs, incorporated in pedagogic and professional environments, were the focus was on establishing rapport with humans, influence behavioural change and obtain disclosure about healthcare related issues. To complement this research we further looked into psychopedagogy interviews, their practices and regulations, so we could frame the agent's behaviour and interactions so it could safely pose as a MHeVA, without endangering the students mental health.

Equipped with the necessary knowledge we constructed MHeVA's architecture and implemented it with the help of FAtiMA toolkit. A complex dialog tree with more than 400 utterances was created, divided into sub-trees that would group the dialog states of a given subject and allow for measured and natural changes during the conversation.

The agent's rationality and intelligence came from a combination of previous works on VA and current knowledge on psychopedagogy interviews. Where the agent changed between rapport building and eliciting self-disclosure, in accordance to its understanding of the Student's State, making use of the Theory of Mind concept.

Then to render a verdict on our agent and our findings, we conducted two testing phases, where college students interacted with MHeVA in a simulation of a one on one interview, and gathered data through a survey and registered values of the interaction.

The results obtained provided valuable information for the future insertion of VAs in the world of mental health support. We showed a VA, deployed as a Mental Health Virtual Assistant is capable of obtaining disclosure from college students about their mental health issues. This with a semi-linear dialog tree and providing low levels of emotional support. Our findings indicate that it can be accepted in that position by said students and has great chances of being able to detect anxiety levels, however we didn't collect sufficient data to ensure this last claim.

We also like to underline the importance of constructing a MHeVA under the guidance and counseling of a mental health support professional, that helped tailor the agent's behaviour and interactions, to best correspond with the requirements and practices necessary to deal with mental health issues. Consequently ensuring accuracy of the support provided and a greater degree of credibility and assurance to those who interact with a MHeVA.

6 FUTURE WORK

Although we consider our work to have successfully completed the objectives it was set out to, we also understand even better now, the amount of work that still needs to be done, before a VA can be finally deployed as a Mental Health Virtual Assistant and provide the so much need help, the mental health college services need.

There was from the early moments of implementation the possibility for our agent to deal not just with anxiety but also depression. We decided to focus on the first, because it is more common in college students and it does not result in such dangerous consequences as depression. So to deal with anxiety was a first good test before advancing into more serious matters. But now that our work is concluded and we already have some strong proven bases to deal with mental health issues, we can set our goals towards also dealing with depression. Identify it, evaluate its seriousness and provide adequate tools an insight for students to deal with it and mitigate its consequences.

Other component that we were also interested in developing was an emotional tracker and facial recognition, that would provide the agent with the ability to perceive emotion and eye gazing, through the use of a small camera. As we explained in this thesis' background, a lot of information about mental issues comes from non-verbal behaviour and it would be as interesting as it would be useful, if our agent could take advantage of those systems. It might provide even more information, not just about rapport building but identifying what issues might be originating the feelings of anxiety in a given student.

Lastly it would be really interesting if our agent would be able to process natural language and understand voice communication. For the same reason as the component above, voice cues give valuable insight to mental health interviewers and it would be a beneficial system, not just to identify the emotions the student is feeling but to also allow for better and more free flowing conversation.

These are but a few of possible improvements our agent could benefit.

In the end we are ever proud of what we managed to achieve with this work. This is a serious issue that endanger thousands of students in Portugal alone and that has been consistently overlooked. We believe VAs in the future can make a difference in this field and help the mental health services improve the lives of so many students. We dare to hope this work can pave the way to future studies and future progress in this area and that one day the now distant reality of a Mental Health Virtual Assistant might come to past.

ACKNOWLEDGMENTS

I would like to ackowledge the great help and dedication Manuel Guimarães and Carla Boura gave to this project. Carla Boura from the IST's Student Support service, that help MHeVA come to life in accordance to the current psychopedagogical practices. And Manuel Guimarães, a PhD student from GAIPS that gave me valuable feedback and guidance throughout this work.

REFERENCES

[1] Samuel Mascarenhas, Nick Degens, Ana Paiva, Rui Prada, Gert Jan Hofstede, Adrie Beulens, and Ruth Aylett. Modeling culture in intelligent virtual agents. *Autonomous Agents and Multi-Agent Systems*, 30(5):931–962, 2016.

[2] Ondřej Beran. An attitude towards an artificial soul? responses to the "nazi chatbot". *Philosophical Investigations*, 41(1):42–69, 2018.

[3] James C. Lester, Sharolyn A. Converse, Brian A. Stone, Susan E. Kahler, and S. Todd Barlow. Animated pedagogical agents and problem-solving effectiveness: A large-scale empirical evaluation. In *In Proceedings of eighth world conference on artificial intelligence in education*, pages 23–30. Press, 1997.

[4] Timothy W. Bickmore, Amanda Gruber, and Rosalind W. Picard. Establishing the computer-patient working alliance in automated health behavior change

interventions. *Patient education and counseling*, 59 1:21–30, 2005.
[5] Robert Brian Haynes, H. Mcdonald, and Amit X. Garg. Helping patients follow prescribed treatment: clinical applications. *JAMA*, 288 22:2880–3, 2002.
[6] Lisanne Pauw, Disa Sauter, Gerben van Kleef, Gale Lucas, Jonathan Gratch, and Agneta Fischer. The avatar will see you now: Support from a virtual human provides socio-emotional benefits. *Computers in Human Behavior*, 136:107368, 06 2022.
[7] Gale M. Lucas, Jonathan Gratch, Aisha King, and Louis-Philippe Morency. It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37:94–100, 2014.
[8] Pierre Philip, Lucile Dupuy, Marc Auriacombe, Fushia Serre, Etienne de Sevin, Alain Sauteraud, and Jean-Arthur Micoulaud-Franchi. Trust and acceptance of a virtual psychiatric interview between embodied conversational agents and outpatients. *NPJ Digital Medicine*, 3, 2020.
[9] Arno Hartholt, David Traum, Stacy Marsella, Ari Shapiro, Giota Stratou, Anton Leuski, Louis-Philippe Morency, and Jonathan Gratch. All together now, introducing the virtual human toolkit. pages 368–381, 08 2013.
[10] Randy P Auerbach, Philippe Mortier, Ronny Bruffaerts, Jordi Alonso, Corina Benjet, Pim Cuijpers, Koen Demyttenaere, David D Ebert, Jennifer Greif Green, Penelope Hasking, et al. Who world mental health surveys international college student project: prevalence and distribution of mental disorders. *Journal of abnormal psychology*, 127(7):623, 2018.
[11] Taguspark Campus Student Support Office Carla Eduarda de Sousa Boura Costa, IST. Personal interview conducted by andré antunes.
[12] Sofia Veiga and Helena Lopes. Serviços de apoio psicológico ao estudante de ensino superior: As experiências do centro de intervenção psicopedagógica e do gabinete de orientação e integração: confluências e especificidades. *Sensos-e*, 7(1):57–66, 2020.
[13] Patrick Corrigan. How stigma interferes with mental health care. *The American psychologist*, 59:614–25, 11 2004.
[14] Lusa. Covid-19. cerca de 55% dos estudantes do superior piorou o seu estado de saúde mental. *Diário de Notícias*, June 14, 2021.
[15] Marco André Pinto Ribeiro. *Impacto da COVID-19 no Planeamento do Estudo e na Saúde Mental nos Estudantes do Ensino Superior.* PhD thesis, Universidade de Coimbra, 2021.
[16] Paul C. Cozby. Self-disclosure: a literature review. *Psychological bulletin*, 79 2:73–91, 1973.
[17] Richard Balon. Self-disclosure in psychotherapy. *Annals of Clinical Psychiatry*, 19(3):205–207, 2007.
[18] Chris Frith and Uta Frith. Theory of mind. *Current biology*, 15(17):R644–R645, 2005.
[19] Dwayne D. Gremler and Kevin P. Gwinner. Rapport-building behaviors used by retail employees. *Journal of Retailing*, 84(3):308–324, 2008.
[20] T.M. Holtgraves, S.J. Ross, C.R. Weywadt, and T.L. Han. Perceiving artificial social agents. *Computers in Human Behavior*, 23(5):2163–2174, 2007.
[21] Christopher You, Rashi Ghosh, Andrew Maxim, Jacob Stuart, Eric Cooks, and Benjamin Lok. How does a virtual human earn your trust? guidelines to improve willingness to self-disclose to intelligent virtual agents. In *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*, pages 1–8, 2022.
[22] Isabel Leal. *A Entrevista Psicológica.* Fim de Século - Edições, Sociedade Unipessoal, LDA., Lisboa, 2008.
[23] Robert J. Shiller. Conversation, information, and herd behavior. *The American Economic Review*, 85(2):181–185, 1995.
[24] Li Gong. How social is social responses to computers? the function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, 24(4):1494–1509, 2008. Including the Special Issue: Integration of Human Factors in Networked Computing.
[25] Ewa Luger and Abigail Sellen. "like having a really bad pa": The gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 5286–5297, New York, NY, USA, 2016. Association for Computing Machinery.
[26] Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. "what can i help you with?": Infrequent users' experiences of intelligent personal assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, New York, NY, USA, 2017. Association for Computing Machinery.
[27] Carrie Heeter. Communication research on consumer vr. *Communication in the age of virtual reality*, 1:191–218, 1995.
[28] Chang S Nam, Joseph Shu, and Donghun Chung. The roles of sensory modalities in collaborative virtual environments (cves). *Computers in Human Behavior*, 24(4):1404–1417, 2008.
[29] Jonathan Gratch, Ning Wang, Jillian Gerten, Edward Fast, and Robin Duffy. Creating rapport with virtual agents. In *in 7th International Conference on Intelligent Virtual Agents. 2007*, pages 125–138. Springer, 2007.

Received 31 October 2022; revised 24 November 2022; accepted 24 November 2022