Script

Introduction:

Hi, my name is Tom Robinson, I’m here to talk to you about my project titled Human interaction with voice-assistant agents. For this project I have created an application for Raspberry Pi which can be placed in a home and record conversations that contain conflicts. My project comes under the HCI umbrella and has been supervised by Petr Slovak.

Contents:

As you can see from the slides I will be covering; the problem my project is addressing, what I aimed to achieve and how, why the project is important, the work that has already been done in this field and the difficulties of creating an effective system like this. I will then move into the achievements I reached in the creation of the project, following that I will discuss the results of testing the system and then the contributions I have made to the research area. After discussing these areas, I will demonstrate the various features of the system and finish with conclusions.

Problem Definition:

So, the problem my project is trying to solve is that of creating a way for therapists to get high quality conversations that contain specific techniques that have been outlined as either poor and good parenting. This is important as the parent child relationship has been identified as a key factor in the development of mental health in young people, so capturing the pressure points of this relationship should give a professional deeper insights into the relationship between them. Involved in this problem is that of outlining when these conversations occur so that there can be less human processing of data

Aims and Objectives:

After defining the problem, I set out the aim to detect when conflicts were occurring in the home, record these conversations and then send this data to a therapist where this can then be discussed at length with the family to see where they can improve. When I set out the objectives you can see on the screen here I had not defined how to detect conflicts, as whatever kind of system I implemented these were goals this system should hit. However, shortly after laying out these objectives I decided on detecting shouting as the indicator of a conflict as I felt this was the best way to it hit these goals

Motivations:

The motivation for this project comes due to current techniques being limited in how certain the conversations heard are natural. Many experiments take place under human examination, which can cause the participants to act unnaturally, which will in turn provide false or limited data for inspection. Having conducted an extensive literature review, I failed to come across a paper that has linked a parent encoding manual to an emerging technology, like SSP.

Background:

As I just mentioned, I completed an extensive literature review without finding a paper linking DPICS to an emerging technology. I was also struggling to find many projects that were focused on the *prevention* of mental health issues occurring, even though there is so much research into the fact that mental health issues are still growing and prevention is the best way to combat this

Difficulties:

The main difficulty with a project like this is the correct identification of a conflict, shouting is a clear sign of a conflict but is not exclusively the definition of a conflict, there may be some conflicts in which there is no shouting and there are some techniques that are identified in the DPICS manual that wouldn’t be heard in the conversations surrounding someone yelling. Another difficulty with a system like this is getting people to participate in testing as personal conversations are something people are unlikely to be forward in wanting to be recorded

Achievements:

Within this project I have achieved the creation of a device that will be able to be placed into the home, relatively inconspicuously, and will be able to detect when conflicts are occurring in the house as long as there is some shouting contained within the conversation. When the device has detected shouting it will start recording the ensuing conversation, but it will also attach a certain amount of the conversation that was occurring before the shouting was detected, this will allow the clinician to have a better understanding on what the home dynamic was like before the shouting occurred. The addition of this data gives us a whole stream of conversation as the recording will continue as long as there is any kind of conversation heard up to a certain point where there has been silence for 5 seconds. This is an indication the conversation has either ended or the participants have moved away from the device, although the device will have detected the trigger which will still be useful to discuss.

I decided to detect conflicts as it is often the case that when a conflict occurs people’s emotions are heightened and are likely to show some of the techniques outlined in the DPICS manual. Conflict resolution has also been highlighted as an area that is important to address in the development of children and how they handle social interactions as they grow up. This can either be from a direct conversation with the child OR even any kind of interaction that is in earshot of the child. Detecting conflicts and recording the ensuing conversation will allow the therapist to see if the conflict was successfully resolved, again aiding in the prevention of mental health issues in young children

To partially address the issue of privacy in this project I decided it was best to have the system upload the files to a shared online space. This would allow the participants access to their own data and allow them to remove any data they felt misrepresented the family. During my research one project I found allowed the participants access to delete the files kept on them but found that only one out of three families did so and they only did this on one occasion, so this feature should not greatly affect the quality of information gathered. I selected Dropbox as the medium to upload the files to as when viewing the files on there it gives you a breakdown of the recording in decibel level heard at each second. This in itself is a useful tool as it allows participants and clinicians to see where the shouting instances occurred, on top of this tool there is also a possibility of attaching comments to specific seconds of the recording. This feature gives an opportunity for the therapist to leave specific pointers and advice to the patients in between sessions.

Exporting the software to Raspberry Pi was a relatively easy process due to using python to create it, there is still some changes that had to be made, although these were fairly minor. As mentioned before the use of Raspberry Pi allows the device to be rather inconspicuous but also using a popular named brand to capture the conversations may raise confidence in the system and possibly make participants more willing to accept the device in the home.

Results:

As you can see from the slide, the results for effectiveness of the system, in regard to how the system can be used in a therapy session are yet to be tested due to not implementing the system with a family currently in therapy. I decided this kind of testing wasn’t completely necessary due the lengthy process of getting approval from the college and then finding a family willing to participate in my project on top of this. This process would have been an unnecessary hurdle due to my project being focused on the building of the software and the detection of conflicts rather than the actual content of the conversation.

Contributions:

The main contribution I’ve made to the area of HCI are the comparisons I have drawn between SSP and DPICS. Social Signal Processing is an area that is making great leaps in the detection of human interactions and various conversation dynamics, including mood and identifying who is leading the conversation. SSP also aims to detect underlying intentions of interactions on top of the more obvious indicators. These comparisons are something that may bring many more technological advances to the detection of mental health issues in young people. This will give the therapist an even better understanding of the home dynamic and could also be used in the automatic detection of the parental techniques in DPICS, which would further remove the need for humans to be involved in the process of identifying the appropriate interactions.

The actual device has contributed to this area of research as the use of audio to detect conflicts and parental techniques was one that seemed to be most lacking from what I found during my review of current systems. Although this system will not record *every* conflict there is a high chance that this system can be used as an example of how audio detection can help in the prevention of mental health issues developing due to its ability to successfully identify when shouting occurs

Demonstration:

I will now demonstrate how the system works and the various features I have included

Conclusions:

To conclude, I judged how successful my system was on matching the detection technique selected to the aims and objectives I set out earlier. Shouting is something that I think will more often than not be a marker of a conversation that contains information useful for therapists. It is a unique point in conversation that will not overload those processing conversations stored on the system. Having the system on Pi gives the device the best chance to be placed in range of good conversation and be accepted in the home.

The comparisons between SSP and DPICS give the chance for future systems to be built that will increase the amount of software available to contribute to the pre-treatment of mental health issues in young people.