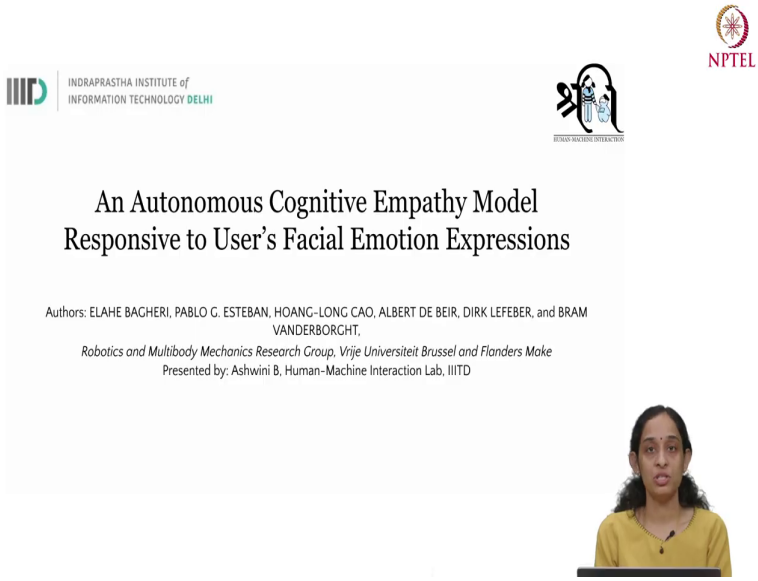
Affective Computing
Prof. Ashwini B
Department of Computer Science and Engineering
Indraprastha Institute of Information Technology, Delhi

Week - 10
Lecture - 34
Tutorial Research Paper Discussion

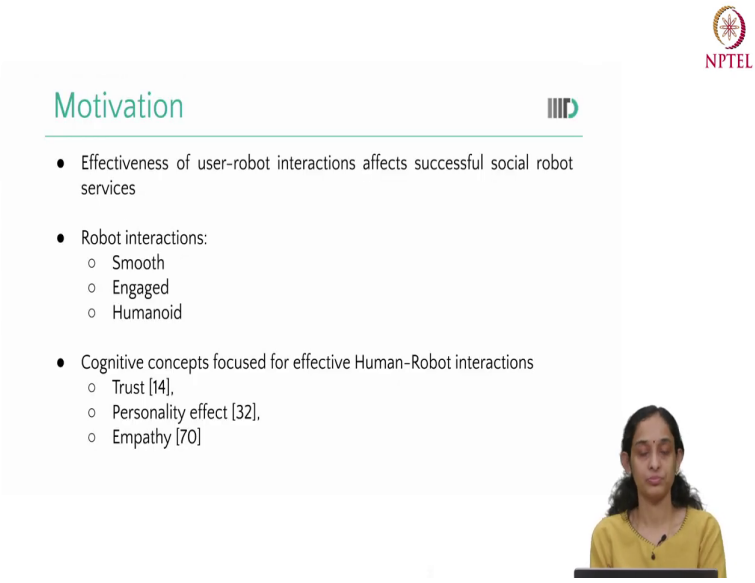
Good morning everyone. I am Ashwini. I am a PhD scholar under Dr. Jainendra Shukla at Human Machine Interaction Lab, IIIT Delhi.

(Refer Slide Time: 00:39)



Today I will be discussing the paper on An Autonomous Cognitive Empathy Model Responsive to User's Facial Motion Expressions by Bagheri et al.

(Refer Slide Time: 00:45)



Like in any interactions, whether it be human-human interaction or human-machine interaction, the effectiveness of the interaction depends upon how well the partners of interaction understand the intense and emotions of each other and respond appropriately. In a social robotics environment, the effectiveness of the interaction depends upon how the robot responds emotionally and empathetically towards the users or the target. To make this possible, the interaction has to be smooth, engaged, natural and human-like.

(Refer Slide Time: 01:25)

There are several studies that are worked on in this direction. Before going into the details of this paper, let us understand what is empathy, what are the factors that affect the empathetic responses, what are the different kinds of empathy and levels of empathy. Literature has defined empathy in different way.

For this study, the authors have adopted the definition in such a way that the empathy is defined as the responses of the empathizer towards the emotions of the target, which aligns with the definition given by Hoffman et al. There have been two kinds of empathy. One is the cognitive empathy and the second one is affective empathy.

In cognitive empathy, the empathizer perceives the emotions of the target in a rational or logical manner while in affective empathy, the empathizer considers or perceives the

emotions of the target more emotionally or a natural way. This could be considered as an intrinsic empathy.

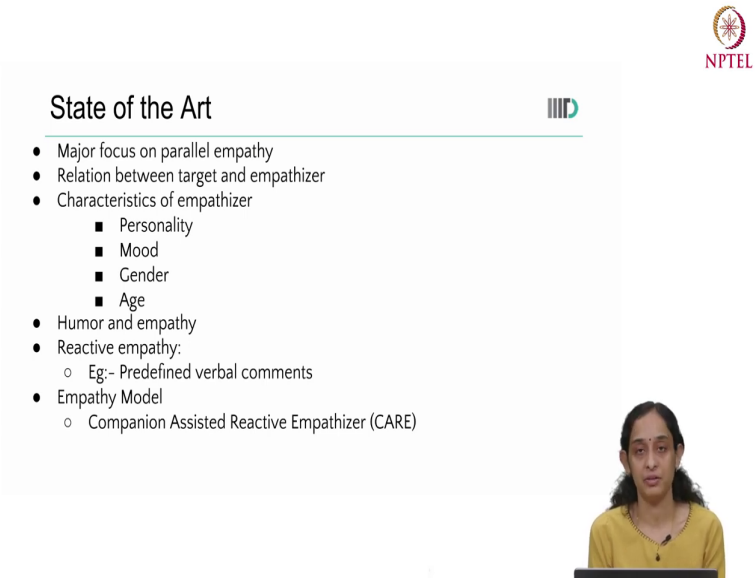
There are different factors that affect the emotional response, which also involves empathetic response. First one is the intrinsic features of the shared emotions. This is in some way representing the nature of the target's emotions. What is the emotion expressed by the target or the user, whether it is positive or negative, whether it is strong emotion or a certain emotion, what is the salience of the emotions expressed by the target.

Second one is the characteristics of the empathizer. This basically represents the personality of the empathizer, whether the empathizer is extrovert or introvert, the gender of the empathizer, the age of the empathizer, the mood of the empathizer, etcetera affects the emotional response.

Third one is the relationship between the empathizer and target. How well you understand the target depends upon what is the relationship that you have with the target or the user. You may not respond in the same way to a stranger as that of a friend. Third one is the situational context. It depends upon when and where and how you respond to the user's emotion.

The empathetic behaviors has been categorized into two levels, one is parallel empathy and reactive empathy. In parallel empathy, you mimic the emotions of the target or the user. That means if the user is sad, you respond in a way that aligns with the emotions of the user. In reactive empathy, you feel and empathize with the user in such a way that the user's distress is reduced. You uplift the positive emotions in the user while reducing the negative emotional energy in the user.

(Refer Slide Time: 04:48)



The existing literature has studied different kinds of empathy, parallel as well as reactive. Most of the works have focused on parallel empathy, where the emotions of the user is identified and the empathizer also aligns with the emotions of the user. Even though the reactive empathy has been studied, it has been limited to generating responses in the form of verbal comments.

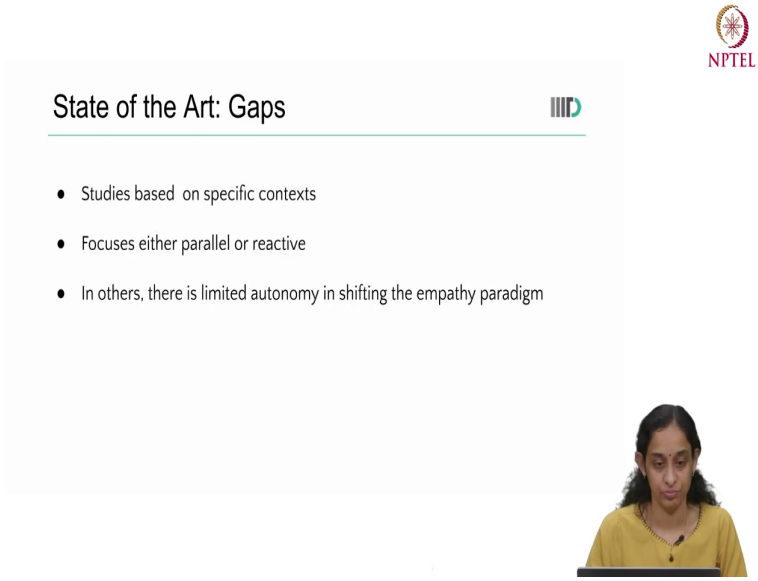
Machine learning techniques have been extensively used for developing models for empathizing, developing; machine learning models have been extensively tried for developing emotional empathetic model. One of those is companion assisted reactive empathizer.

In this, the empathizer is developed in a virtual environment with human trainers. The human trainers interact with each other during a virtual interaction platform. During these

interactions, they exhibit emotions and these emotions are understood and perceived by each one of these partners and they react appropriately to the emotions of the partner.

This has been used for training which also involves the physiological signals like heart rate, PPG etcetera from the interactive partners, and which is used to understand the emotions of the interacting partners, and used to generate a response in return to it. One of the drawbacks of this study is that we have to understand the different context, the different set of possibilities for these interactions to happen in order to predict the appropriate responses for each interaction in sessions.

(Refer Slide Time: 07:06)

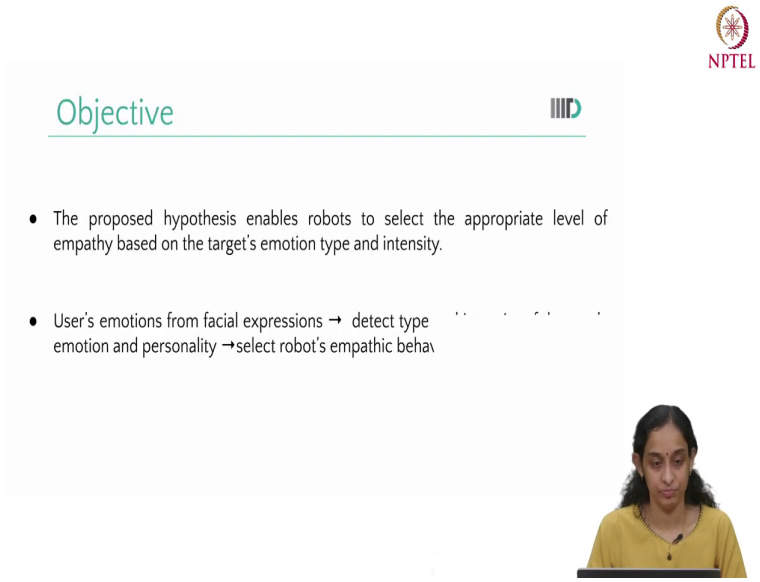


By looking into the literature, we could understand that there are many gaps like most of these studies are specific to a particular context. They have considered only a specific kind of

empathy model for example, either the parallel empathy or the reactive empathy or these studies lack autonomous decisions by the empathizer.

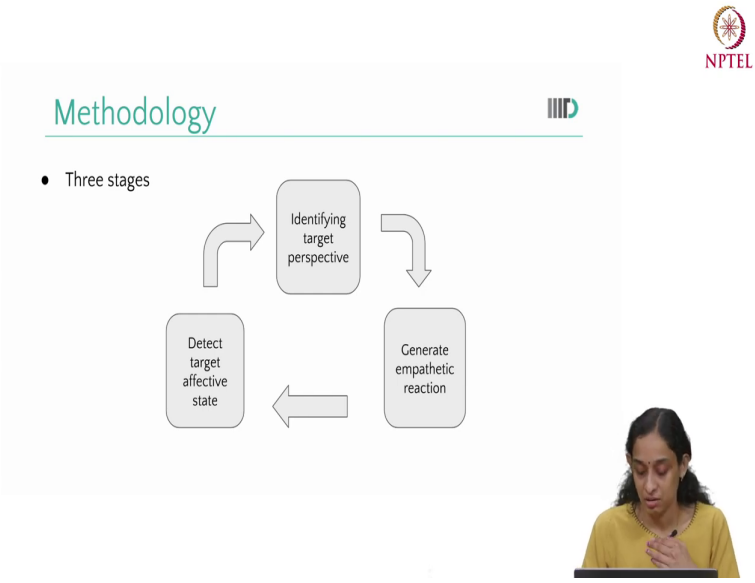
Whether to generate a reactive response or a parallel response has to be decided by the empathizer in real time scenarios. These studies could not make the empathizer empowered enough to make that decisions.

(Refer Slide Time: 07:52)



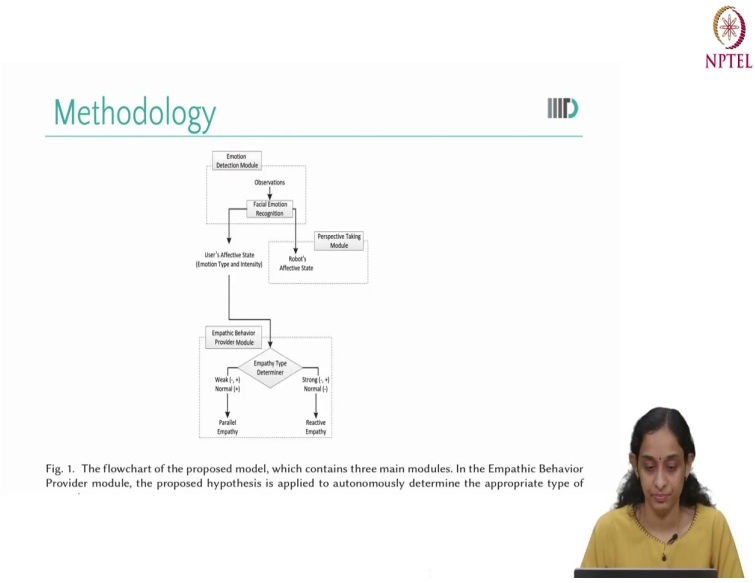
In this study, the authors have tried to generate a autonomous empathizer which understands or perceives the emotions of the target and generates appropriate reactive or parallel responses in response to the user's emotions. To understand or perceive the emotions of the target, here the authors have relied upon facial expressions. Facial expressions are one of the major components which aids the expression of emotions.

(Refer Slide Time: 08:33)



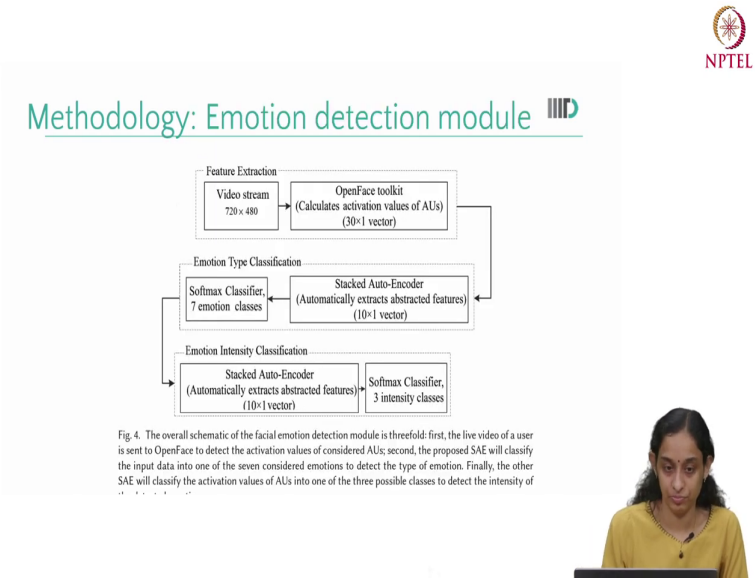
Coming onto the methodology, there are 3 stages in this process. First is to detect the target's affective state. Once that is detected, we have to understand what is the target's perspective of the emotion. Once that is understood by the empathizer, the empathizer has to generate an empathetic response to that emotion. And this process goes on until the interaction ends.

(Refer Slide Time: 09:11)



This work has 3 different modules. One is the emotion detection module, where the emotions felt by the target is perceived using facial expressions of the target. Now, the emotions of the target is perceived using this detected module. And finally, based on the detected emotions, the empathizer generates responses, empathetic responses to be specific, based on the emotions and the intensity of the emotions expressed by the user.

(Refer Slide Time: 09:51)



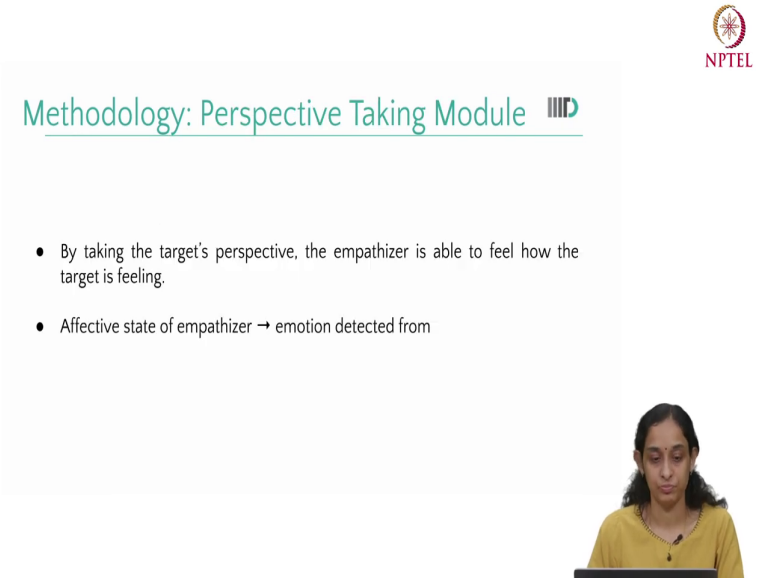
The first module or the first step in this process is to understand the emotions of the user. How will you understand the emotions? As I said, one of the major factors that helps us to express our emotion is our face or the facial expressions. To define the emotions using facial expressions, Ekman has defined an emotional model which has 6 basic emotions; happy, sad, fear, anger, surprise and disgust.

His emotional model is a categorical one and for the purpose of his study, the authors has used Ekman's emotional models. And Ekman has also defined that all these emotions could be expressed using different facial muscles which are called the facial action units. There are 55 facial action units which could express different emotions by activating or by relaxation or contraction of these facial muscles.

From the video stream using open phase toolkit, these facial action units are extracted which is sent to a stacked auto-encoder model and finally, classified into the different emotions using a Softmax Classifier. Again, once the emotions are identified, we have to understand the intensity of emotion to generate appropriate empathetic responses, which is again performed using a machine learning model consisting of a stacked auto-encoder and a Softmax Classifier.

I am not going into the details of the architecture; you can find it in the paper.

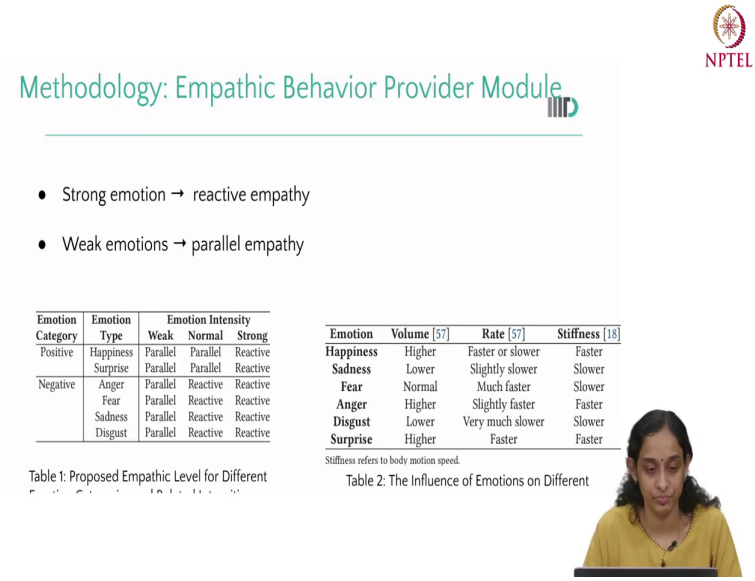
(Refer Slide Time: 11:49)



Once we detect the emotions using the facial action units, next is to understand the perceptions, emotional perceptions of the target. For representing the emotional perception of

the target, we assign the detected emotion to the target. And this is considered as the emotions expressed by the user with whom the empathizer is interacting.

(Refer Slide Time: 12:19)



The last step in this procedure is to generate the empathetic behavior by the empathizer. The authors have defined both reactive empathy as well as parallel empathy depending upon the intensity of the intensity of the emotions expressed by the user. For subtle emotions or positive emotions, parallel empathy is used and for strong emotions or negative emotions, reactive empathy is performed.

You can see the different emotional responses and the combinations used in these tables. For example, if the emotion is happiness and the intensity is high, then the emotional response is parallel. For example, if the emotion category is positive and the emotion type is happiness or

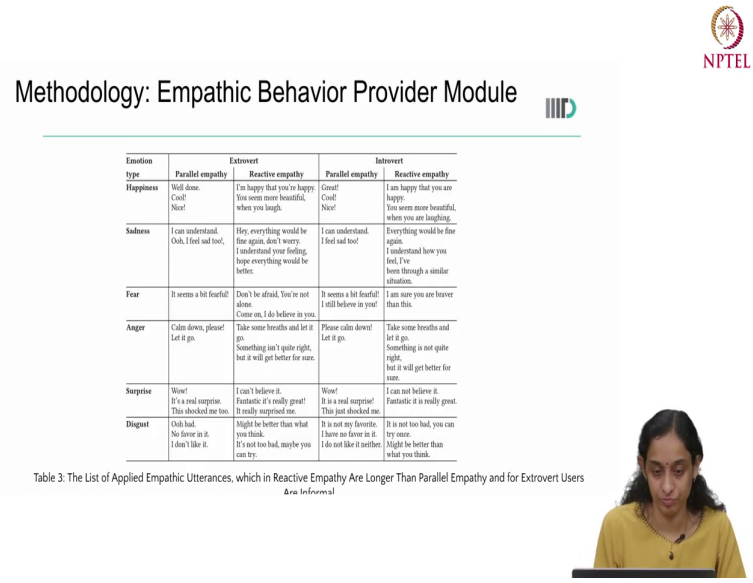
surprise, the based on the emotional intensity, whether it is weak or normal, the responses also vary.

Now, how should we generate the empathetic response? What are the changes that has to be made in the response behavior of the empathizer? You know, the emotions could be expressed or humans express their emotions in different ways. There will be some changes in their facial expressions.

The gestures that make also represent the emotional intense of the person, also the pitch and the tone of the voice they generate or the intonation of the pronunciation, intonation or the pronunciation or intonation of the emotional state. Considering that, in this study, the emotional empathetic response of the empathizer is defined using various parameters.

One is the stiffness of the body activities or stiffness of the joints of the interactive agent or the empathizer. Second is the pitch and intensity of the voice of the robot or the empathizer and also the eye color of the empathizer. These things are used to respond with parallel empathy. And if the empathizer is adopting a reactive empathy, then the eye color as well as verbal comments were used.

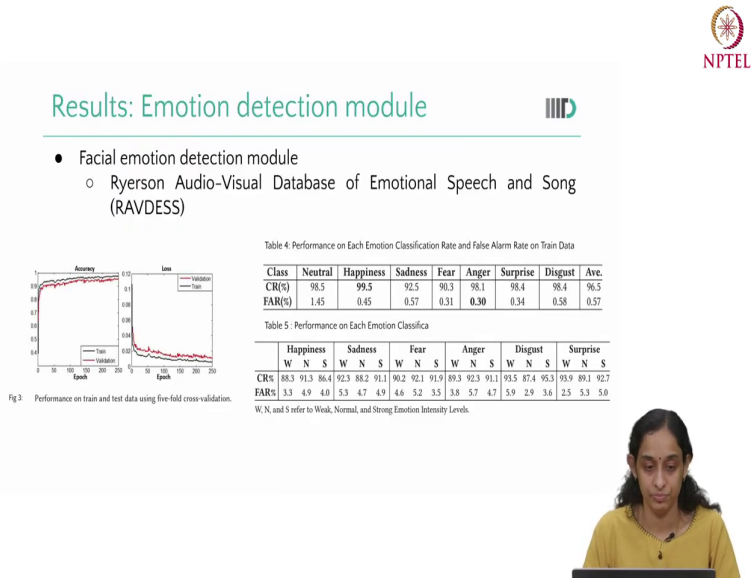
(Refer Slide Time: 15:10)



The empathetic behavior provider module: This, the empathetic behavior also defines the personality of the empathizer, whether the empathizer should create a introvert behavior or a extrovert behavior. This depends upon the similarity attraction principle used in psychology. Usually, people respond positively or people find it more interesting to interact with people who are similar to their traits.

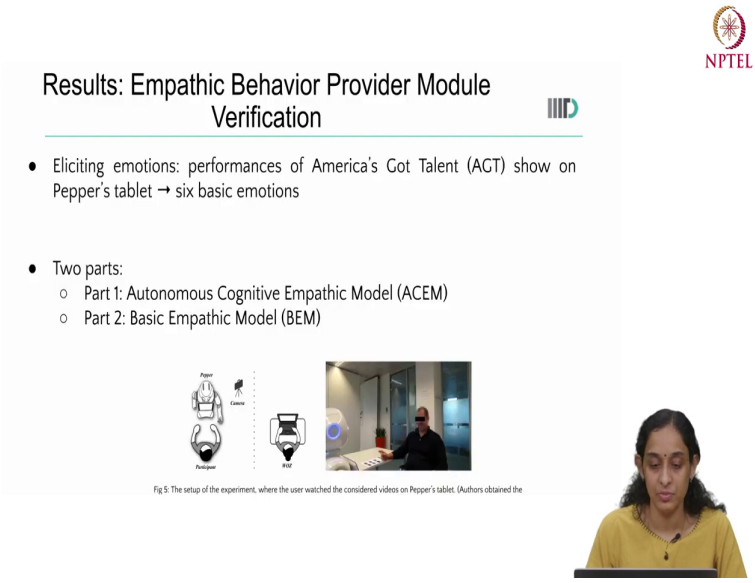
So, depending upon the personality of the target, the empathizer's behavior is also changed, and accordingly the speech, the eye color, the behavior, gestures, etcetera are also changed.

(Refer Slide Time: 16:07)



Coming on to the results. The facial emotion recognition module is trained on RAVDESS dataset, which is a popular dataset used for emotion detection.

(Refer Slide Time: 16:18)



After which a user study is conducted to decide whether the autonomous empathetic, autonomous cognitive empathetic model is better than the existing models or not. In this a user scenario or a interactive scenario is defined, in which the participants or the target is shown videos in different emotional categories aligning to the Ekman's basic emotion model.

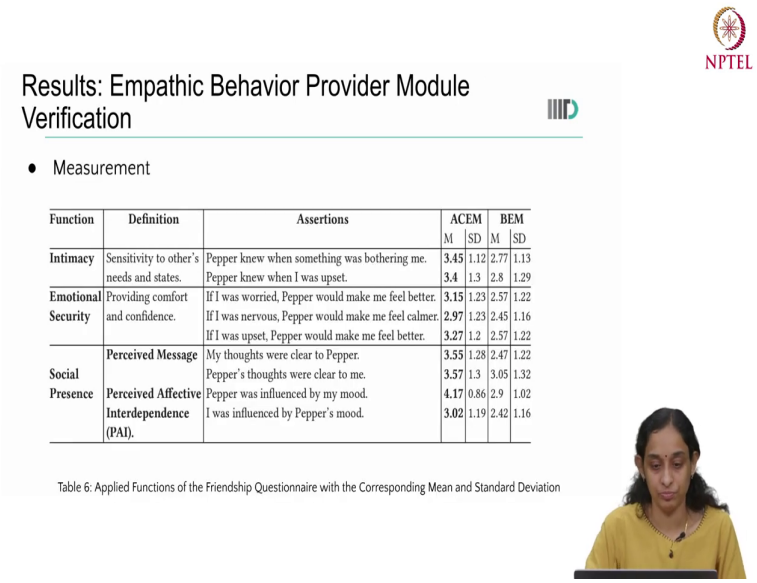
These videos were selected from America Got Talent Show, and during these interactions the different emotions are elicited in the users and a social robot, which is pepper robot in this case has been used to respond appropriately to the user's emotions. The study was conducted in two different parts.

One in which the empathizer or the social robot responds according to the autonomous cognitive empathetic model and in the second part the robot or the empathizer responds to a baseline model which is the basic empathetic model. The difference between the basic

empathetic model which is the baseline and the autonomous cognitive empathetic model is that, in basic empathetic model only the eye color and the verbal comment is produced by the empathizer.

While in ACEM or the autonomous cognitive empathetic model the empathizer changes the speech in donations, then the stiffness of the body, the eye color as well as generates the verbal comments depending upon the emotions and intensity of emotions of the user.

(Refer Slide Time: 18:13)

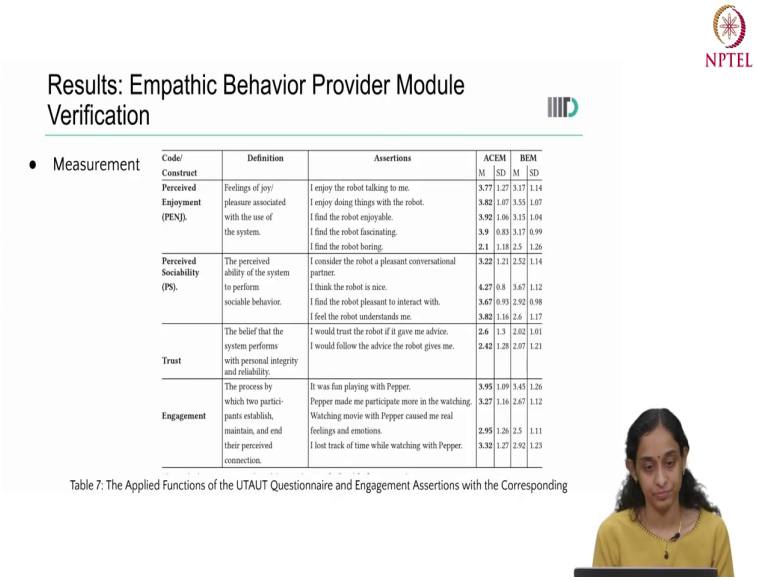


Function	Definition	Assertions	ACEM		BEM	
			M	SD	M	SD
Intimacy	Sensitivity to other's needs and states.	Pepper knew when something was bothering me.	**3.45**	1.12	2.77	1.13
		Pepper knew when I was upset.	**3.4**	1.3	2.8	1.29
Emotional Security	Providing comfort and confidence.	If I was worried, Pepper would make me feel better.	**3.15**	1.23	2.57	1.22
		If I was nervous, Pepper would make me feel calmer.	**2.97**	1.23	2.45	1.16
		If I was upset, Pepper would make me feel better.	**3.27**	1.2	2.57	1.22
Social Presence	**Perceived Message Interdependence (PAI).**	My thoughts were clear to Pepper.	**3.55**	1.28	2.47	1.22
		Pepper's thoughts were clear to me.	**3.57**	1.3	3.05	1.32
	Perceived Affective Interdependence (PAI).	Pepper was influenced by my mood.	**4.17**	0.86	2.9	1.02
		I was influenced by Pepper's mood.	**3.02**	1.19	2.42	1.16

And after this the responses of the empathizer is evaluated on different parameters. One is the intimacy. Intimacy shows how sensible or how sensitive is the empathizer towards the target or the user. Second one is emotional security. This means how well the participant or the user feels confident and comfortable in interacting with the empathizer or the robotic agent here.

Third is social presence. In social presence the users evaluated the empathizer based on its sociability; how well they relate this robot or the empathizer as a social entity.

(Refer Slide Time: 19:03)



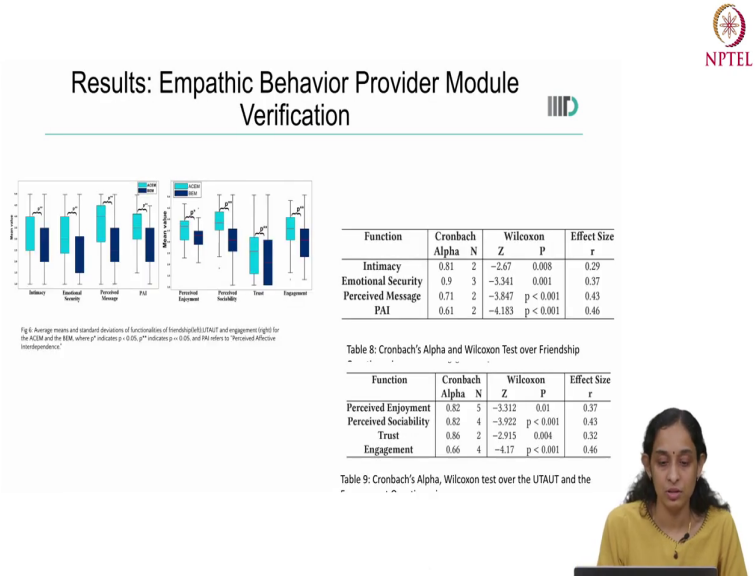
Code/ Construct	Definition	Assertions	ACEM M	ACEM SD	BEM M	BEM SD
Perceived Enjoyment (PENJ).	Feelings of joy/ pleasure associated with the use of the system.	I enjoy the robot talking to me.	**3.77**	1.27	3.17	1.14
		I enjoy doing things with the robot.	**3.82**	1.07	3.55	1.07
		I find the robot enjoyable.	**3.92**	1.06	3.15	1.04
		I find the robot fascinating.	**3.9**	0.83	3.17	0.99
		I find the robot boring.	**2.1**	1.18	2.5	1.26
Perceived Sociability (PS).	The perceived ability of the system to perform sociable behavior.	I consider the robot a pleasant conversational partner.	**3.22**	1.21	2.52	1.14
		I think the robot is nice.	**4.27**	0.8	3.67	1.12
		I find the robot pleasant to interact with.	**3.67**	0.93	2.92	0.98
		I feel the robot understands me.	**3.82**	1.16	2.6	1.17
Trust	The belief that the system performs with personal integrity and reliability.	I would trust the robot if it gave me advice.	**2.6**	1.3	2.02	1.01
		I would follow the advice the robot gives me.	**2.42**	1.28	2.07	1.21
Engagement	The process by which two participants establish, maintain, and end their perceived connection.	It was fun playing with Pepper.	**3.95**	1.09	3.45	1.26
		Pepper made me participate more in the watching.	**3.27**	1.16	2.67	1.12
		Watching movie with Pepper caused me real feelings and emotions.	**2.95**	1.26	2.5	1.11
		I lost track of time while watching with Pepper.	**3.32**	1.27	2.92	1.23

Table 7: The Applied Functions of the UTAUT Questionnaire and Engagement Assertions with the Corresponding

Next is perceived enjoyment which shows whether they whether the users enjoyed the interaction, what were their feeling about the interaction, whether it is positive or negative. And next is the perceived sociability. Again, this shows how this empathizer or the robotic agent could be used in a social interaction. Trust. Trust represents how well the empathizer could respond to the respond to the user reliably and what does the user understand about the integrity of the interactions.

Next is engagement. Engagement represents how well the interaction went, so that or how engaged the interaction were or how engaged the users were in the interaction so that the interaction can extend for a prolonged time.

(Refer Slide Time: 20:11)

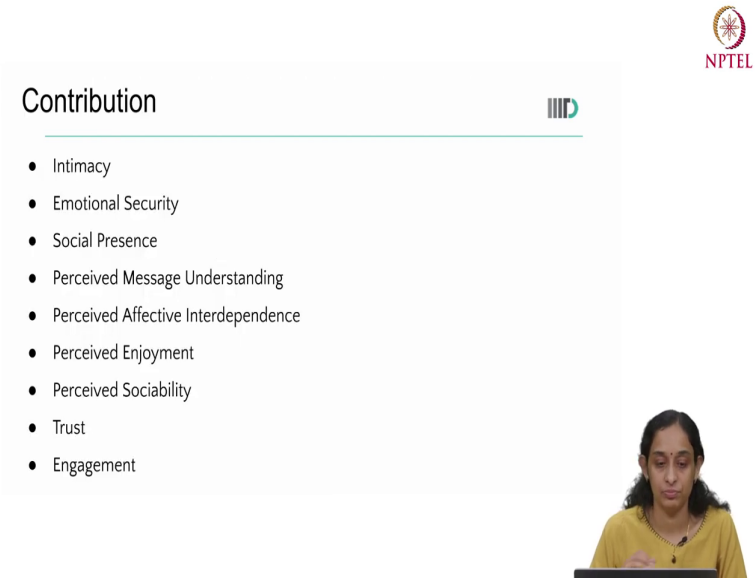


Function	Cronbach Alpha	N	Wilcoxon Z	P	Effect Size r
Intimacy	0.81	2	−2.67	0.008	0.29
Emotional Security	0.9	3	−3.341	0.001	0.37
Perceived Message	0.71	2	−3.847	p < 0.001	0.43
PAI	0.61	2	−4.183	p < 0.001	0.46

Function	Cronbach Alpha	N	Wilcoxon Z	P	Effect Size r
Perceived Enjoyment	0.82	5	−3.312	0.01	0.37
Perceived Sociability	0.82	4	−3.922	p < 0.001	0.43
Trust	0.86	2	−2.915	0.004	0.32
Engagement	0.66	4	−4.17	p < 0.001	0.46

So, the robotic or the empathizer target interactions were evaluated based on these parameters. And the results are shown in these tables. It is evident from the figures that the autonomous cognitive empathetic model performed far better than the basic empathetic model in most of these parameters.

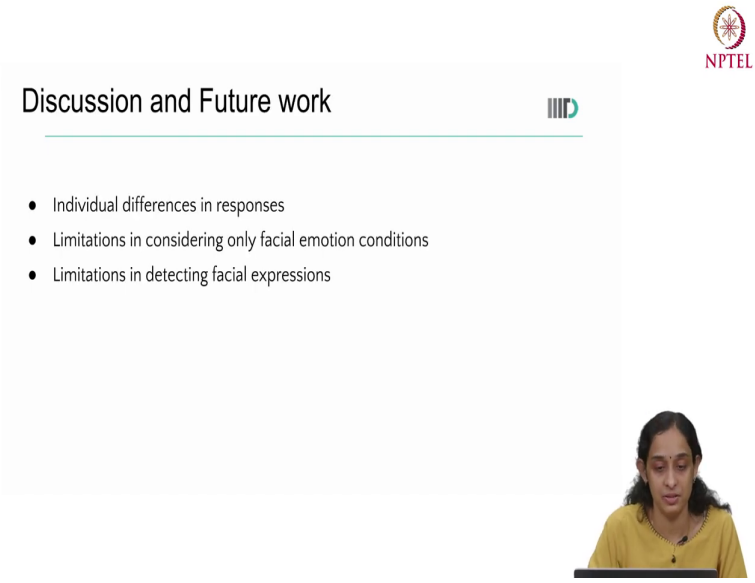
(Refer Slide Time: 20:42)



Coming on to the contribution of this paper. This model has been seen to provide more effective interactions in terms of intimacy, emotional security of the target. They found the interactive agent or the empathizer as more social and considered it as a social entity good for social interactions. They found that their emotions were understood better by the empathizer. And it also showed that their empathetic, their emotional responses were dependent upon how well the empathizer responded to their emotions.

And according to the empathizer's responses, their moods or their emotions also varied. And they enjoyed these enjoyed these interactions, and they were more engaged and they had trust in these interactions or in these interactive agents.

(Refer Slide Time: 21:58)

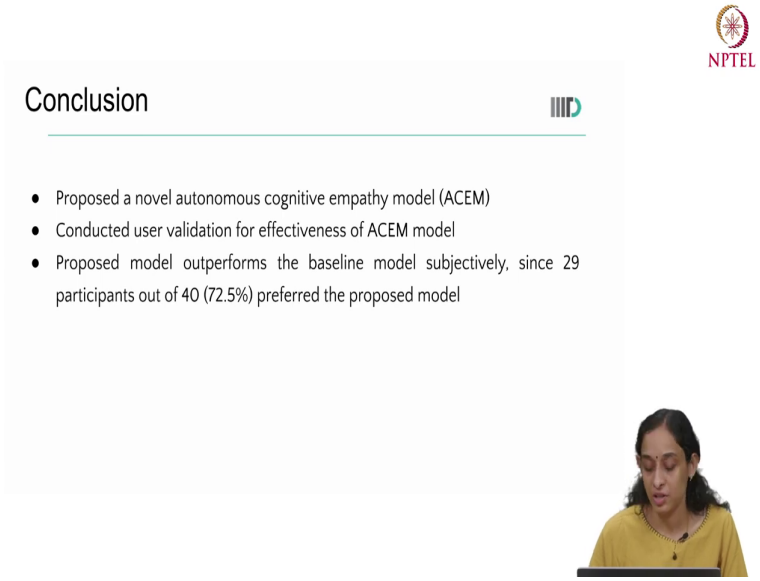


Coming on to some of the limitations of these studies, they having mixed responses on how well this, there have been there have been mixed emotions on the response behavior generated by the empathetic agent here. Some found that having more tactile responses like hugging or a touch on the shoulder might have made the interactions more better, must have made the interactions better.

And some thought that more expressions on the face of the empathizer or the Pepper robot might have been improved the interactions. Further, in this study only facial expressions were considered as an indication for emotions or intense of the user. Sometimes there may be other factors that can represents emotions better. Considering a holistic approach in the emotional perception will help in understanding the emotions of the target and react appropriately.

And this method, in general is restricted by the bottleneck of the performance of different facial expressions algorithms or facial expression detection algorithms.

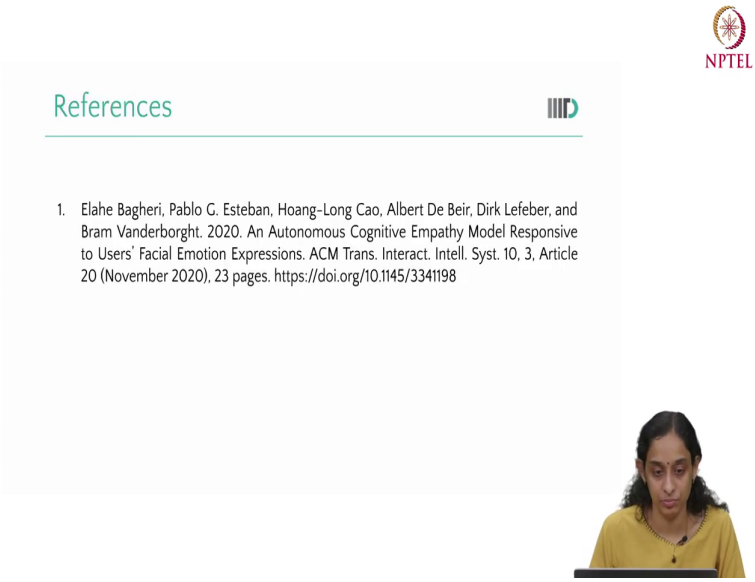
(Refer Slide Time: 23:28)



In short, this paper provides the in short, this paper studied a, in short, this paper explored an autonomous cognitive empathy model which could understand the emotional intense or moods of the user and generate a empathetic response system which is appropriate to the emotions of the user.

This study also conducted experiments to validate their claims, and the proposed method is found to be affective in making the interactions more engaged and affective using robots.

(Refer Slide Time: 24:23)



For more details, you can refer to Bagheri et al's paper.

Thank you.