Social Behavior and the Brain: An Introduction to Social Neuroscience
Dr Ark Verma
Department of Cognitive Sciences
Indian Institute of Technology Kanpur
Week - 02
Lecture – 10

Hello and welcome to the course Social Behavior and the Brain, an Introduction to Social Neuroscience. I am Dr. Ark Verma, an Associate Professor at the Department of Cognitive Science, IIT Kanpur. This is the last lecture of the second week where we were talking about retrieval of person knowledge from faces. And in this 10th lecture, we will talk about evaluation of faces through the process of overgeneralization. Let us move ahead and talk about it.

Now, as we saw in the previous lecture that results from competition modeling suggest that trade judgments are constructed from cues that have evolutionary significance. The primary dimension of valence basically derives from cues resembling expressions of anger or happiness. Now Freeland in 1994 argued that one of the functions of emotional expressions is to signal behavioral intentions. So we are not only judging whether a particular face is pleasant or unpleasant due to its facial expressions, but we are also able to figure out, we are able to reason the intentions that an individual might be harboring by having a look at his or her face.

And this, I'm sure you can appreciate, is a very important commodity in social interactions. If you can look at a person's face and you can basically decide whether the person should be approached or avoided at any given point in time, whether somebody is happy or sad at any given point in time, it may be absolutely useful in determining the course of your interaction going forward i gave this example you know in one of the previous lectures say for example you are waiting for your parent or let's say your partner at the end of the day to ask them for something maybe something more expensive maybe something difficult to get And they come into your house at the evening, they're slightly overworked, irritated, all of that. They've not said anything to you, but looking at their face when they enter the house, you can make out whether the person is in a happy space of mind or in a relatively irritated space of mind. And say for example, you can also get what their intentions are. Are they planning to just shut down, go rest? Or say for example, they are willing to engage in terms of maybe going for a movie together, for a restaurant.

All of those things we are able to compute from the faces of individuals. And this basically gives us a very interesting way of calibrating our social interactions with people. So, as I said, expressions of happiness signal to the perceiver that a person can be

approached at a given point in time. Expressions of anger signal that the person must be avoided or left alone at any point in time. Also, there is evidence that angry faces trigger automatic avoidance responses.

So we are very good at judging this you know and a lot of times you will see and it's a very interesting experiment if you can if you want to do it if you really don't want to talk to anybody at a point in your house you can just make a very somber expression at your face and you will see immediately people will sort of you know start avoiding you they'll start avoiding you to the extent that okay maybe this person needs to be left alone at this point in time and so on. So again it works both ways that ways. Now, consistent with the social cognition research suggesting that the valence evaluation of stimuli is directly linked to approach or avoidance behavior, this is already been demonstrated by Chen and Barg in their approach avoidance experiments. Now, how is this important? It is important from an evolutionary point of view because the cost of approaching an angry individual are much greater than the cost of avoiding a happy individual. See, if somebody is very happy or less happy and so on, even if you avoid them, it's not really a very costly proposition.

If somebody is very angry or frustrated and it is in a very bad state of mind, you would not want to approach them because you would not want to sort of invite their anger or give them an opportunity to vent their anger at you for that matter. And people can be violent verbally or physically when they are in an angry space of mind if you have not really said the right things in that moment. So these judgements can be used to explain the non-linearity of trustworthiness judgements as well. So remember we said that individuals are more sensitive towards the negative end of trustworthiness and happiness, anger continuum than they are at the other end where there are these all positive emotions are there. So, these judgments also were found to be sensitive to changes at the negative end as compared to the positive end of the transversal dimension as I just said and threat judgments for example in a similar way you know generated by this threat dimension that we were talking about at the end of the previous lecture.

Individuals were also found to be more sensitive to changes at the threatening end rather than the non-threatening end of the dimension. So again, you can see that how we calibrate our interaction, how we calibrate our decision or facial status, facial trait estimation from an individual is very dynamic and it's a very context sensitive situation. So the authors propose a very interesting hypothesis called the emotion overgeneralization hypothesis. What does this hypothesis say? It says that the similarity of facial features to emotional expressions is attributed to personality traits. And previous studies have shown that emotional expressions do affect trait judgments from faces.

So for example, if you are having a pleasant or a happy expression on your face, there is a higher chance that you will be judged as more trustworthy, more approachable and so on

as opposed to when you are having neutral expressions on your face. So that's what the research says. Smiling faces are perceived as more trustworthy than neutral faces and higher on affiliation, an attribute that is found to be similar to trustworthiness. People will show higher affiliation, people will deem you as more trustworthy, people will deem you as you know more approachable if you have a smiling face rather than if you are having a neutral expression on your face. Also judgments of anger and happiness from emotionally neutral faces are correlated with the judgments of trustworthiness.

So in that sense you can see again there is a continuum of trait judgments based on facial expressions and we will see this in more detail as we go ahead. So the authors sought to provide an extended replication of the modeling results that we saw in the previous lecture and therefore they collected emotion judgments of emotionally neutral faces that they had earlier collected. So they sort of brought together that data and then they started computing correlations between these judgments and emotional expressions and what did they find? they found that correlations for 55 out of 84 judgments so 14 traits then 16 emotion judgments like anger disgust fear sadness surprise and happiness were actually significant so in 54 out of 85 out of 84 times the facial expressions were used or were highly correlated with facial trait judgments Okay, just in a bit more detail, the judgments of happiness were more positively correlated with all positive trait judgments.

So if you're perceived as happy, if you're having a happy expression on your face, people are more likely to estimate you having all those positive traits, you know, punctuality, responsibility, approachability, trustworthiness, honesty, all of those kind of things. okay and negatively correlated with all the negative judgments if you're happy uh you're bearing a happy expression on your face you are less likely to be interpreted as being have as having negative uh personality attributes like untrustworthiness violent aggressiveness and those kind of things so see it it sort of all falls together that where are we making these judgments from we are making these judgments based on the facial emotion expressions okay also the valence component derived from the trade judgments were also strongly correlated with the judgments of happiness and anger moderately correlated with the judgments of disgust and weakly correlated with the judgment of sadness so you can see while these correlations are there and they are very important in predicting the trade judgments the correlations are strongest for positive relatively weaker for disgust envy those kind of things and weakly correlated with sadness so it sort of weakens a little bit when you go towards the negative or so called negative emotions Here is the figure here you can see that you know this is happy so happy is most correlated with trustworthy and then but if you go to let's say for example you know mean weird and these kind of things the it sort of comes down and you know so this is this is a interesting thing you can see this is all significant at 0.05 level so these are color maps you should pay some attention to this. Moving on, faces that were evaluated positively were perceived as happier and more surprised and were perceived as less

angry, less disgusted, less fearful and less sad. So bearing or let's say if you are having positive emotional expressions, you are more likely to be getting positive trait judgments. That is as straightforward as it is. Now the dominance component, which is the second component in their computational model, was found to be correlated with judgments of anger, surprise, sadness and fear.

Dominant faces were perceived as more angry faces, less sad, less fearful and less surprised. than submissive faces. While these findings are consistent with the emotion overgeneralization hypothesis, they are also consistent with the hypothesis that these correlations can be accounted for by common semantic properties of emotion in trade judgments rather than by perceptual similarity. Now this is a very interesting point. What people are saying is that It is good that these emotional expressions and the trait judgments are coming to be highly correlated but what is the cause for this high correlation? On one hand the proposal from their side is that it is because the features of the face are sort of correlating when you are looking at say for example a very trustworthy face versus a face bearing positive expressions.

The features and the way they are arranged, these dynamic features as well as invariant arranged in a particular way in a particular kind of correlated manner and that is why the correlations are coming as such. The other thing at a slightly more deeper level if you see is that there are also these are also common semantic properties in a conceptual network in a semantic network you will see that feeling happy or feeling or being trustworthy and all of these lie very close to each other in the semantic conceptual space and it could probably be the case that in a sense these are properties along the same continuum these are different facets of the same underlying property of being positive and that is why these correlations are coming out to be very high this is this is an interesting proposition let us see how it you know goes so what they say is that expectations about the relation between emotional stimuli such as smiling as an expression of happiness and personality traits such as sociability may actually lead to strong associations between emotional expressions and trait judgments which is something that we are just discussing which is something that is turning out to be the case actually Now this idea is actually very consistent with research on implicit personality theory which people had and they said that you know people have these assumptions about the relationships between various traits. So you know just like basic emotions are sometimes portrayed on a continuum, different traits can also be portrayed on a continuum where positive traits are clustering on one end whereas negative traits are clustering on the other. And this is basically why they also resorted again to a different kind of a principal component analysis. Let us look at what they found.
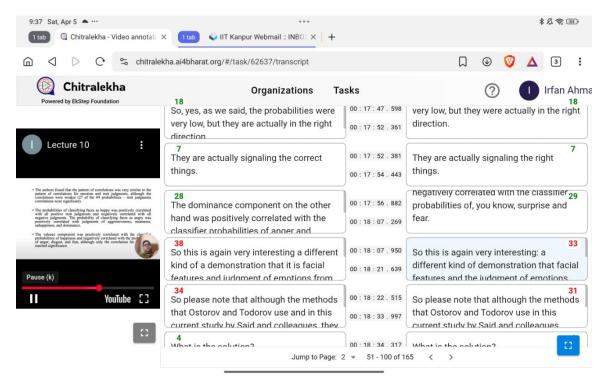
So the dimensional structure of emotional judgments was actually found to be very similar to the dimensional structure of the trait judgments in these two different PCAs. and the first PC derived from the PC of emotion judgments was found to be highly

correlated with the you know with the dominance component whereas the second PC of the emotion judgment was more strongly correlated with the you know valence actually second is more with valence the first is more with the first is more with valence second is more with dominance there's a bit of a typo here Now, to rule out the possibility that the relations between trait and emotion judgments can be accounted for entirely by semantic similarities, the authors actually used an emotion classifier to categorize the emotionally neutral phases. Now see what they are trying to do here is they are trying to have a computational model do this. Computational models obviously don't understand the semantics in the way or represent the semantics in the way humans do. So they are basically trying to feed this classifier with facial features and these faces and see if the classifier can perform these classifications successfully on the basis of facial features rather than on the basis of semantic similarity which obviously it does not share that with us.

So, the authors use the Bayesian network classifier to detect the presence of subtle presence of features resembling emotions in faces and this classifier you know accepted a feature vector containing the displacement between automatically chosen landmarks in the same landmarks of a prototypical neutral phase as input and it gave out a set of probabilities corresponding to each basic emotion. So, we will basically take some kind of a you know feature vector in the phases as an input and they compared it with the prototypical neutral phase you know benched it around that point and it got the classifier gave out set of probabilities that this is the probability that this face has this basic emotion or that basic emotion and so on and this is being done on basis of features. As the authors applied this classifier to neutral faces, the output probabilities were actually very low. So it was not doing this job at a very high probability but these probabilities were able to predict trade judgment from faces. So it seems that trade judgments can actually be performed on the basis of facial features rather than the semantic similarities which obviously this Bayesian classifier would not have access to.

Also, the authors found that the pattern of correlations was very similar to the pattern of correlations for emotion and trade judgments. Although the correlations were relatively weaker, 27 out of 84 probabilities, the trade judgment correlations were rather significant. The probability of classifying faces as happy just to sort of give you more details the probability of classifying faces as happy was found to be positively correlated with all positive trait judgements and negatively correlated with all negative trait judgements. Also, the possibility of classifying faces as angry was positively correlated with judgements of aggressiveness, meanness, unhappiness and dominance. So you can again see that the classifier is just using facial features, it is classifying for facial emotions, basic emotions it was giving an output for, this much probability this is anger, this much probability this is happiness and these outputs are still correlating very well with the trait judgments.

So that is what they wanted to actually demonstrate. Now, the valence component of their computational model was found to be positively correlated with the classifier's probabilities of happiness and negatively correlated with the probabilities of anger, disgust and fear, although only one of these correlations actually reached significance. So, yes, as we said, the probabilities were very low, but they are actually in the right direction. They are actually signaling the correct things. The dominance component on the other hand was positively correlated with the classifier probabilities of anger and negatively correlated with the classifier probabilities you know surprise and fear.

So this is again very interesting a different kind of a demonstration that it is facial features and judgment of emotions from facial features which are correlated and contributing towards you know trait judgments of the human face. So please note that although the methods that Ostorov and Todorov use and in this current study by Said and colleagues, they differ in different ways but they actually converge on very similar solutions. What is the solution? We judge basic emotions from facial features and this judgment of facial emotions actually correlates very highly with trait judgments about these faces. So just to elaborate a bit more, faces with positive valence were more likely to be classified as happy and less likely to be classified as angry and disgusted than faces with negative valence. So you can see emotion judgment goes along with trait judgment.

Also highly dominant faces were more likely to be classified as angry and less likely to be classified as fearful than highly submissive faces. So you can see again there are these basic underlying factors in how we are processing faces and how our processing of these

faces of different aspects of faces are highly correlated with each other. Now to the extent that facial features signaling positive valence or trustworthiness are similar to expressions of anger and happiness, it should be possible to demonstrate that the facial features affect the perception of emotional expressions. So you know the basic facial feature should be able to affect how we perceive faces as happy, sad, disgusted, envious and so on. So they wanted to investigate this.

What did they do? They based on prior trustworthiness judgments, the author selected trustworthy and untrustworthy faces and they created a dynamic stimulus in which the faces expressed either happiness or anger. So trustworthy faces, untrustworthy faces could display either happiness or anger. Now although the authors added the same amount of emotional intensity to faces, trustworthy faces expressing happiness were perceived as, you know, higher than untrustworthy faces also untrustworthy faces expressing anger were perceived as expressing the same emotion this is the figure you can see on this continuum you have the same phase you have congruent morph so these are by the way morphoses generated from this initial basic phase and you can see that they already lie on a similar continuum Now the authors also did something interesting. They manipulated the changes in trustworthiness during animation when these were morphed and started to move. For example, incongruent animations, it happened that an untrustworthy or a trustworthy face gradually was morphed into untrustworthy into trustworthy or trustworthy into untrustworthy morphing was done.

To the extent that these trade judgments are actually an overgeneralization of cues resembling expressions, changes that are in the direction of the expressed emotion should amplify the intensity of the perceived emotion. So if you are going from trustworthy to untrustworthy, judgments of anger should magnify. If you are going from untrustworthy to trustworthy, judgment of happiness should amplify. And in the incongruent direction, this should sort of minimize a little bit.

So that's what it is said. So in contrast changes in the opposite direction should dampen the perceived intensity of the emotion. Now, this is exactly what the authors found. For instance, when an untrustworthy face was changed or morphed into an untrustworthy face, the same angry expression was perceived as angrier. Now, you are going from trustworthy to untrustworthy. So, the same angry expression is now being more intensely interpreted.

It was perceived as more angry. On the other hand, when an untrustworthy face is changing to an untrustworthy face, there was no change in the identity of the face. So, they are not changing identity, just the judgment of intensity of that expression sort of is magnified. Similarly, when an untrustworthy face changes into a trustworthy face, the same angry expression is now being interpreted as less angry. So, you can see when you are going in an opposite direction, it is dampening the perception of anger in the

individuals. based on these morphed faces now this is this is all good this is now that they are sort of established that judgment of expressions and judgment of traits actually go very closely together there is also something that the authors sort of wanted to you know investigate in some detail So, to investigate the similarities in the neural codes of perceived trustworthiness and expressions of happiness or anger, the authors used a behavioral adaptation paradigm.
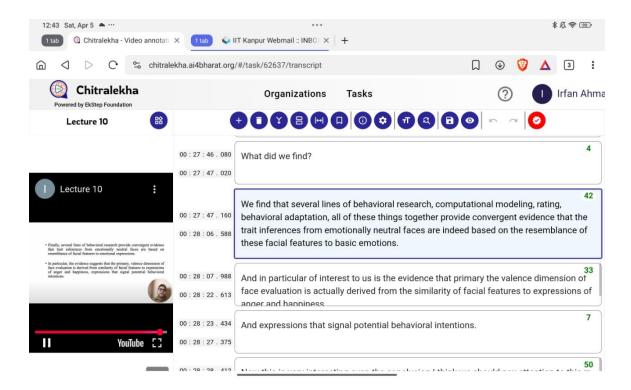
What is the behavioral adaptation paradigm? So, I am sure some of you might know what sensory adaptation is. If you are receiving the same sensory stimulus for an extended period of time, your sensory receivers will get adapted or saturated and then there will be a minor temporary aversion from gaining more from there and a bit of an inclination to go towards the opposite end. So this adaptation paradigm has actually been utilized earlier to investigate other dimensions of the neural representation of faces including viewpoint invariance, gender, attractiveness and expression. And the central point as I said of this paradigm is that extended exposure to a stimulus dimension results in fatigue of the neural population that represents that particular stimulus. So subsequent exposure to stimulus along the same dimension should result in a perceptual shift away from the adapting stimulus.

So if you are showing happy, happy, happy, happy, happy again, then the judgment of happy should become lower. and the judgment of the other face it will sort of increase. So in Webster and colleagues they showed that androgynous faces which could be interpreted as a male or a female were seen as distinctly male after extended exposure to female faces and distinctly female after extended exposure to male faces so again you're doing the same thing again and again the system also gets bored it sort of shifts the direction to the other side Now, if trustworthiness works in the same way, if trustworthiness is indeed an overgeneralization of perceiving positive features, you know, anti to anger and in convergence with happy facial expressions, then something interesting should happen here. The author should be able to influence this evaluation. evaluation of trustworthiness by first adapting the neural populations that support the perception of these expressions.

You show anger, anger, anger again or you show happy, happy, happy again. If you have saturated the system, it will take a decision on the opposite end of the continuum. So what did they do? In the pre-adaptation stage of the experiment, participants were randomly assigned to one of three adaptive conditions. First is passive viewing of angry expressions, passive viewing of fearful expressions which is the control condition and passive viewing of happy expressions for 66 seconds. After this adaptation phase, so there you are saturating the system by showing happy, happy, happy, happy, or fearful, fearful, fearful, fearful, or angry, angry, angry, angry.

After this adaptation phase has ended, the participants were asked to rate the trustworthiness of these phases again. The test phases were reduced slightly because this is a sensory stimuli. They have been reduced to size to 80% of the adaptation phases to just prevent interference from the low level adaptation effects. But more interest are these results that I am going to talk to you about now. As expected, adaptation to angry faces resulted in higher trustworthiness ratings whereas adaptation to happy faces resulted in lower trustworthiness ratings.

So if the system is getting saturated, it tries to take a decision on the other end. If you're showing happy, happy, happy again or let's say angry, angry, angry again, then the system tends to look at the more positive side and start sort of building oh this is a more highly trustworthy phase if you're showing happy happy happy again it says oh it does not seem very happy anymore because you know all pleasant things or all good things the system will get saturated will get bored of that and will try and seek and escape out of there that's basically you know a bit of both you know sensory biology and you know some practically practical observation Now, as expected, adaptation to angry faces resulted in higher trustworthiness rating, whereas adaptation to happy faces, yeah, this is something I just said, so let's move on. What do we see in the last two lectures? What did we find? We find that several lines of behavioral research, computational modeling, rating, behavioral adaptation, all of these things together provide convergent evidence that the trait inferences from emotionally neutral faces are indeed based on the resemblance of these facial features to basic emotions. And in particular of interest to us is the evidence that primary the valence dimension of face evaluation is actually derived from the similarity of facial features to expressions of anger and happiness. And expressions that signal potential behavioral intentions.

Chitralekha
Powered by EkStep Foundation

Organizations   Tasks   Irfan Ahma

Lecture 10

| 00 : 27 : 46 . 080 | What did we find? | 4 |
| 00 : 27 : 47 . 020 | | |

Lecture 10

- Finally, several lines of behavioral research provide convergent evidence that trait inferences from emotionally neutral faces are based on resemblance of facial features to emotional expressions.
- In particular, the evidence suggests that the primary, valence dimension of face evaluation is derived from similarity of facial features to expressions of anger and happiness, expressions that signal potential behavioral intentions.

▶ YouTube ⛶

| 00 : 27 : 47 . 160 | We find that several lines of behavioral research, computational modeling, rating, behavioral adaptation, all of these things together provide convergent evidence that the trait inferences from emotionally neutral faces are indeed based on the resemblance of these facial features to basic emotions. | 42 |
| 00 : 28 : 06 . 588 | | |

| 00 : 28 : 07 . 988 | And in particular of interest to us is the evidence that primary the valence dimension of face evaluation is actually derived from the similarity of facial features to expressions of anger and happiness. | 33 |
| 00 : 28 : 22 . 613 | | |

| 00 : 28 : 23 . 434 | And expressions that signal potential behavioral intentions. | 7 |
| 00 : 28 : 27 . 375 | | |

| 00 : 28 : 28 . 412 | | 50 |

Now this is very interesting even the conclusion I think we should pay attention to this in the sense what we are saying is that the facial expressions yes they affect our trade judgments but they are also you know contributing towards the judgment of behavioral intentions as well. So the face stimulus is actually a very important social stimulus. We have seen so far the different dimensions on which we evaluate faces. We have seen the different mechanisms through which this might be done. But the take home message is to be able to understand and adapt to it and understand how these very basic decisions are actually correlated with each other.

So, thank you for attending this lecture. I will see you in the next week in the first lecture with the role of amygdala in judging facial expressions. Thank you so much.