





ARTICLE



<https://doi.org/10.1057/s41599-024-02869-x>

OPEN

A theory of emotion based on a universal model

Yong Huang ¹ 

The complexity of emotions has thus far limited our understanding of them. To obtain a clear understanding of the nature of emotion, this paper proposes a novel emotion theory and establishes a universal model of the conscious world in the human brain, the substanguage and interaction model (SIM). Based on an analysis of the possibilities of the interaction process in the SIM, two basic emotions that are indecomposable factors within all emotions—hope and fear—are identified. A questionnaire survey reveals that this basic emotion exhibits high acceptability. Based on emotion theory, this paper reasonably explains the phenomena of facial attraction and infantile facial preference and discusses the psychological reasons for phonocentrism, the phenomenon of preferring the spoken word over the written word. In addition, this paper explores the possibility of artificial intelligence possessing self-emotions. Emotions are relevant to many areas of human knowledge, as well as to everyone's daily lives, and the simple, clear way to understand emotions provided in this paper may be instructive for everyone.

¹School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China. email: huangy@buaa.edu.cn

Introduction

Everyone has emotions, and people can fall into the confusion over their emotions. It is crucial to have a certain level of understanding of human emotions. The discussion of human emotions humans has been ongoing for thousands of years. For example, there is much discussion about emotion in the Buddhist classic *Shurangama Sutra*.

Psychologists and philosophers have continually discussed emotion. James (1884) suggested that bodily changes directly follow the perception of an exciting fact and that our feeling of the same changes as they occur is the emotion. Cannon (1927) noted that the lower part of the brain, which neuroscientists call the thalamus, controls emotional experience. Arnold (1960) emphasized that the evaluation of the external environment is the direct cause of emotion. Lazarus (1966) viewed emotion as a comprehensive response, arguing that no emotion can be determined by a single component. Young (1973) defined emotion as a violent disturbance of an emotional state or process that arises from a psychological condition and shows physical processes of smooth muscles, glands, and general behaviour. Izard and Malatesta (1987) defined emotion as a special combination of neural processes that guide specific expressions and corresponding specific feelings. Sroufe (1996) suggested that emotions are subjective responses to important events characterized by physiological, experiential, and external behavioural changes. Ekman et al. (1967, 2003) revealed the correlation between facial expressions and emotions. Damasio (2011) noted that emotions are complex programs of actions triggered by the presence of certain stimuli, whether external to the body or from within the body, when such stimuli activate certain neural systems. Feelings of emotion are perceptions of emotional action programs. These various discussions can help people understand emotions from different perspectives.

Although scholars have discussed and researched emotions at length and many words have been used to describe different emotions, people's understanding of emotions is still quite limited. This is because emotions are very complex phenomena. Scholars tend to describe emotions from the outside rather than directly stating what emotions are.

Psychologists have tried to identify the fundamental elements of emotion to gain a deep understanding of emotions. Based on a mathematical analysis of people's ratings of a large number of emotional terms, Plutchik (1980) suggested that basic human emotions include anger, disgust, fear, sadness, anticipation, joy, surprise, and trust. Ekman and colleagues (1987) reported seven basic emotions: sadness, fear, anger, disgust, contempt, happiness, and surprise. Since their study, several psychologists, including Ekman et al. (1999, 2011), Levenson (2011), and Panksepp et al. (2007, 2011), have continued to adjust the composition of basic emotions. For example, Levenson (2011) added interest, relief, and love to the list of basic emotions. Some components of these basic emotions proposed by scholars are the same or similar, but some components differ. The question is thus where the boundaries of these basic emotions are. For example, based on experience, the distinction between interest, relief, and love seems clear, but are there not common elements in these basic emotions? What is the evidence for and against these emotions having common elements? We need to answer this question in terms of the nature of the emotion, not in terms of the phenomenon. Different reasons may cause similar phenomena, and similar reasons may produce different phenomena. For example, people cry, most often because they are sad but sometimes because they are happy.

Neuroscientists have explored the physiological mechanism underlying the generation of emotions (Lindquist et al. 2012). They have carefully measured how emotions emerge from the

interactions that take place within the nervous system. Although current research is far from sufficient for a complete understanding of the topic, scientists have made some useful findings. For example, neuroscientists have found that the amygdala plays an important role in emotions (LeDoux, 1996; Whalen, 1998). Based on this research, scholars view some emotions, such as fear, anger, sadness, interest and joy, as basic emotions (Meng, 2005). This result is similar to the findings obtained by traditional psychologists. Such research has taken a tremendous step forwards in advancing people's understanding of emotions, but it is also inevitably intertwined with people's empirical understanding of emotions.

With regard to the concept of basic emotions, two questions need to be considered in depth. 1. What is an emotion? 2. What is the meaning of "basic" in the context of basic emotions (Lyons, 1999; Ortony and Turner, 1990)? Emotions are some of the most common phenomena in the consciousness of humans and have not yet been clearly defined. To define emotion, a suitable tool must be employed, and this is the universal model discussed in the next section.

Universal model

The world is full of substance and language, and we use the concept of "substanguage" to summarize the two. Each substanguage has a relationship with some other substanguage, which we call an interaction; interactions include verbs, prepositions, and conjunctions. People cannot point to anything and say it is interaction, so the term interaction actually refers to nihilism. Substanguage and interaction are both metaphysical concepts (Huang, 2018), and they are the most abstract of words. These terms serve as names for the two basic categories of the world; one category includes all things, while the other category is nihilism.

According to these two concepts, we know that any simple or complex system is composed of several substanguages and their interactions, as shown in Fig. 1. Since these two are the highest-level concepts, this substanguage and interaction model (SIM) is a universal model. For a physical system, substanguage as shown in the figure refers to various substances, while interaction refers to force. SIM is not a specific model for a specific object; it does not involve specific content. However, this model is extremely abstract and is applicable to a variety of concrete goals. The model simply reflects the fact that the essence of knowledge is distinction, and this distinction is essentially the distinction

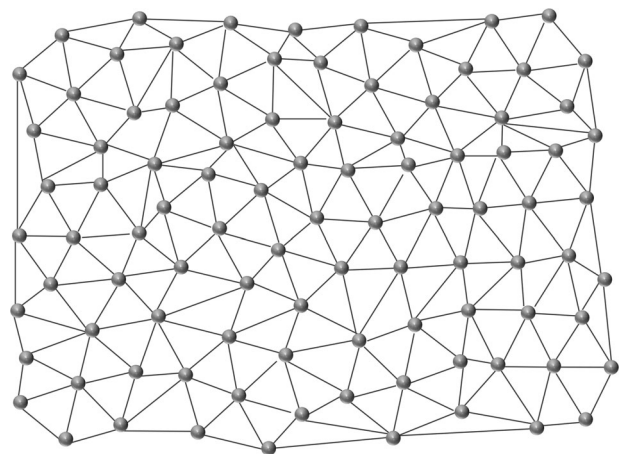


Fig. 1 The substanguage and interaction model. The dots in the figure represent the substanguage, and the lines between the dots represent the interactions.

between “thing” and “nothing”. In this paper, this distinction is also that between substanguage and interaction.

Because substanguage and interaction are very abstract words and refer to two high-level concepts, they can be interpreted only in an interconnected manner; that is, substanguage is a substanguage in interaction, and interaction is an interaction between substanguages. For example, in the sentence “the flower is red”, “flower” and “red” are substanguages, and “is” indicates some kind of interaction. Every substanguage has its own meaning, but it has no independent meaning. When we talk or think about a substanguage, we always associate it with other substanguages. In the sentence “the flower is red”, meaning is generated due to the interaction between “flower” and “red”, indicating that “interaction creates meaning” (ICM). Similarly, the new interaction between the two substanguages also entails new meanings. Meaning is also a substanguage; e.g., in the above example, “the flower is red” is also a substanguage.

In the above discussion, we have explained this universal model, and we can use it to build a new theory of emotion.

Emotion theory

SIM can also be viewed as a model of the conscious world in the human brain, which we call the network of substanguage (NS). To use terminology with which people are typically familiar, the substanguage in the NS is consciousness in the human brain, whereas interaction is thinking. People’s thinking entails interactions among different substances. The NS is always engaged in internal interactions, and it also interacts with the external world through human tissues and organs. The fluctuation that occurs in the interaction process is called emotion, and this is the definition of emotion.

Four basic interaction phenomena are included in the NS. The first phenomenon is called a “metaphor”. This term refers to the two original substanguages in the NS, which are distant from each other and have no direct interaction; however, they now have an interaction, as shown in Fig. 2. The two substanguages in the NS obtain new meanings because of new interactions, which is

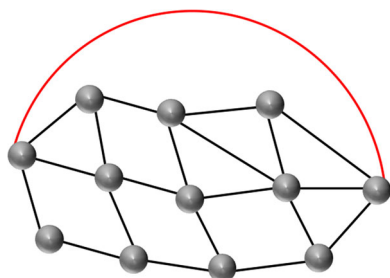


Fig. 2 Schematic diagram of a metaphorical interaction. The red line in the picture marks a metaphorical interaction.

equivalent to producing a new substanguage. For example, in the sentence “a girl is like flower”, “girl” and “flower” are irrelevant substanguages, but now, after this metaphor is read, they interact in the NS of human beings, which is a metaphor. If the metaphor of girl as flower is familiar, it is no longer a metaphor, and the red line in Fig. 2 becomes a black line.

The second phenomenon is called “ripples”, which indicates that the NS receives a new substanguage from the outside. After the new substanguage is embedded in the NS, it generates new meaning with its surrounding substanguage because of the ICM, and the meaning of its surrounding substanguage also undergoes some changes, which may be transmitted to a substanguage that is farther away.

The third phenomenon is called a “vortex”. This term refers to the disappearance of an object in the NS. This substanguage does not merely disappear; the substanguage around it also loses part of its meaning due to the absence of an object of interaction. This phenomenon is similar to pulling the plug at the bottom of the bathtub such that the water flows away from the hole to form a vortex. This vortex is also the location of meaning loss. To resist such a loss of meaning, people often construct a new substanguage at the vortex called “hope”, as shown in Fig. 3.

The fourth phenomenon is called “invisibility”. This notion refers to an interaction between the NS and an external strange substanguage that does not leave any trace in the NS, which is equivalent to no interaction and does not cause changes in the NS.

The interactions within the NS are usually complex, and the abovementioned situations occur simultaneously. The NS continues to appear in the form of new substanguages in interactions, and some substanguages disappear. Since our discussion does not involve specific content, the new substanguage generated through metaphorical interaction, the new substanguage entering from the outside and the hope generated in opposition to the vortex are essentially the same. All of these are new substanguages appearing in the NS.

Although the interaction process in the NS is very complex, it can exhibit only one of the following two of patterns: 1. Substanguage A (SA) interacts with another substanguage B (SB), and SB exists in the NS, or 2. SA interacts with SB, but SB does not exist; thus, the object of interaction of SA is “none”. This point is illustrated in Fig. 4.

In the second case, when SA interacts with “none”, the emotion that it produces is similar to that experienced when one misses a step when going down stairs or sees the abyss at the edge of a cliff. We use the word “fear” to mark the fluctuations that occur when substanguages interact with “none”.

In the first case, we can distinguish between two more specific cases. Case 1a: The object of interaction exists around SA because people have become accustomed to this kind of interaction relationship, and no emotion emerges at this time. Case 1b: The

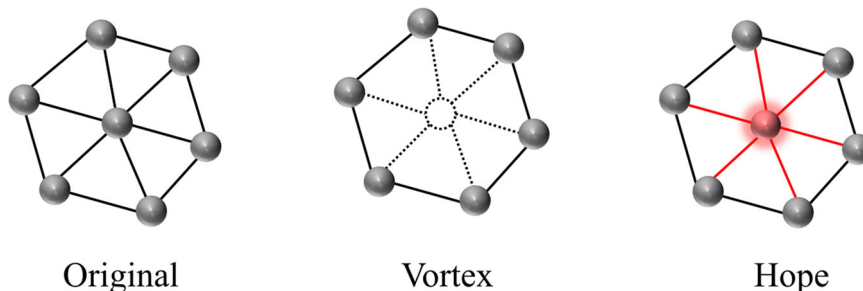


Fig. 3 A sketch of the vortex and hope. When a substanguage is lost in NC, it appears as a vortex, where people construct “hope” to resist the loss of meaning.

object of interaction is the new substanguage. We use the word “hope” to mark the fluctuations generated when the old and new substanguages interact. In case 1b, if SA interacts frequently with SB, SB gradually becomes the substanguage around SA, and this interaction no longer generates emotion. For example, people experience beauty when they go to a new scenic spot, but if they live there for a long time and the scenery simply becomes their surroundings, they no longer have special feelings about it.

Based on the above analysis, we identify two basic emotions—hope and fear—that are indecomposable elements of all emotions. “Hope” and “fear” correspond to two simple, clear, and easily distinguishable scenarios, which is why we call them basic emotions.

Hope refers to a person’s psychological motivation; people always try to interact with things that can elicit hope within them but avoid interacting with things that can cause fear within them. Hope and fear are always intertwined; even if someone likes watching horror movies, he or she has this preference because he or she can overcome the fear factor and be attracted by the hope factor. Our discussion also shows that as long as the brain is working, it is possible to generate emotions. In many cases, people do not feel that they have emotions, which indicates that the emotions that people perceive usually must reach a certain intensity.

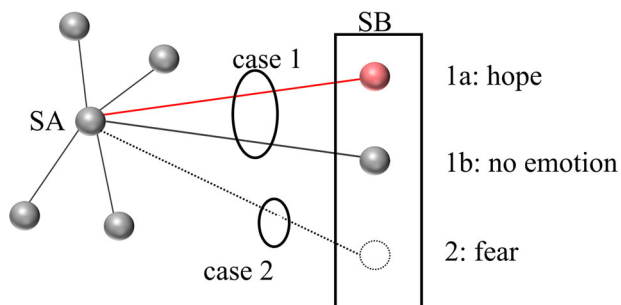


Fig. 4 Schematic diagram of basic emotions. There are only three kinds of interactions that can occur in NC, and two fluctuations lead to two basic emotions: hope (case 1a) and fear (case 2). The other interaction does not produce fluctuations and therefore does not produce emotions (case 1b).

Through the discussion in this section, we have established a new theory of emotion. This method of understanding emotions offers a tool for understanding the various phenomena associated with emotions. In the next three sections we will discuss three seemingly distinct questions, all of which can be answered clearly by our theory.

Facial attraction phenomena

It should be noted that in the discussion of the previous section, the two basic emotions of “hope” and “fear” are defined according to two different fluctuations in the NS, and they are not exactly the same as the natural emotions of “hope” and “fear”. Here, “natural emotion” refers to the superposition of a variety of basic emotions at every moment in people’s brains. However, when naming these two kinds of fluctuations, this paper also fully considers the relationship between them and humans’ natural emotions of “hope” and “fear”. The words people use to describe natural emotions are often not very clearly defined, but people usually understand what they are.

We distributed two anonymous survey questionnaires via social media. Since the questionnaire had only one question, which was not complicated, we conducted this survey only on a small scale, targeting students from our department. The survey results are shown in Fig. 5. The results show that after being given an appropriate explanation, the participants exhibited a high level of acceptance of the basic emotions proposed in this article. It also shows that the method of naming these two basic emotions in this paper is reasonable.

After these two basic emotions are identified, they are used as tools for conducting a binary analysis, which can help us understand more psychological phenomena more easily. For example, what is hate? Obviously, hate contains elements of both hope and fear; people experience fear of something and hope to try to make the frightening thing disappear. A major type of problem that people encounter in this context is choice, and the choice between hope and fear represents the basic choice that people must make. The most basic choice for an individual is to choose as much hope as possible and avoid as much fear as possible.

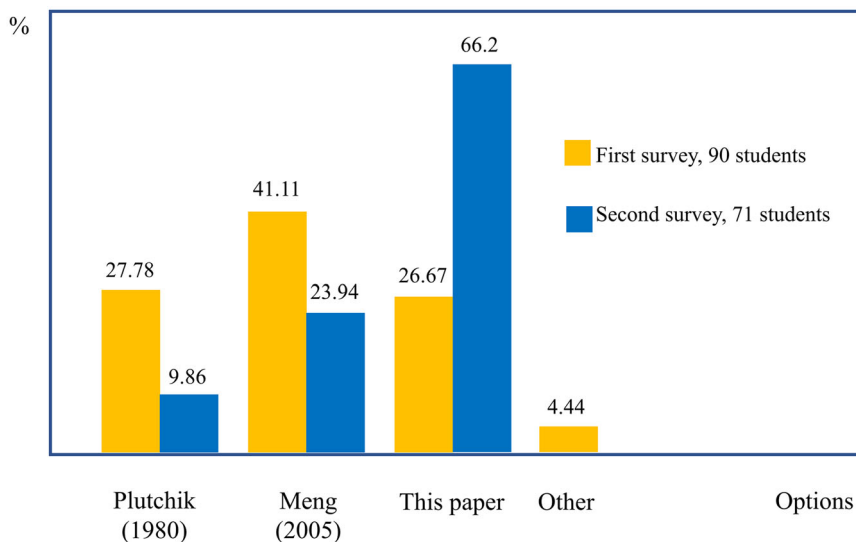


Fig. 5 The choices of different students in the two questionnaire surveys. We provided three sets of options for basic emotions, with the first set being drawn from Plutchik (1980), the second set being drawn from Meng (2005), and the third set being the results presented in this paper. If the respondents disagreed with all options, they could choose “other”. In the first survey, we merely provided these options without informing the respondents of their sources. In the second survey, we briefly explained the background of these options.

We use this theory of hope and fear to analyse a common phenomenon: as some people are considered beautiful, what does it mean to be beautiful? Although many factors affect people's judgement regarding whether a face is beautiful, psychological experiments have shown that people's judgements regarding whether a stranger's appearance is beautiful are surprisingly consistent and that people of different races have the same understanding of the women of their own race as beautiful (Cunningham et al. 1995; Miller, 2012). Rhodes and Little et al. found that facial features have similar effects on attractiveness judgements across different cultures, and their research also revealed a "universal" preference for face shape (Little et al. 2011; Rhodes, 2006). Han et al. (2018) reported that preferences regarding skin colour are not universal. Psychological experiments have also shown that when computer imaging technology is used to produce a new face that is an image composed of features from different local people's faces, the resulting "average" face is considered more attractive than almost all the individual faces of which it is composed (Miller, 2012; Rhodes, 2006; Rhodes et al. 2002; Rubenstein et al. 2002; Valentine et al. 2004).

Why are average faces attractive? When a person looks at a stranger's face, many new elements enter his or her NS and become new substanguages. There are various reasons why people think that some of the new elements are more hopeful and attractive than those of an average face. Therefore, the fact that the average face is more attractive reflects only statistical results. Psychologists' finding that the average face is attractive is undoubtedly important, but the question of why the average face is more attractive has not been answered in sufficient depth.

The face is likely the most common image that people have seen, and many face-related memories are stored in people's brains. It is obvious that people's memories of faces are not comprised of complete two-dimensional images but rather memories of a network of faces (NF). A NF includes various symbols (substanguage) that reflect the characteristics of faces and the relationships among symbols (interactions). Adults can view these symbols as referring to the shape of the eyes, nose, and lips as well as the face shape, skin colour, etc. The interactions and relationships among symbols refer to their arrangement in space, such as the distance between two eyes. "Interaction" cannot be perceived by people. Instead, what people can perceive is a substanguage; "eye-interaction-eye" as a whole is a substanguage and a symbol. For average faces, because these symbols are at an average level, they are surrounded by the most adjacent symbols.

When a person sees a strange face, he or she must recognize and understand it, a process that is based on his or her existing facial memories. Undeniably, all kinds of symbols in human memory are originally located in the NS, but according to the characteristics of human thinking, when people think about something or awaken something in their memory due to external stimulation, these thoughts or memories become part of their current consciousness, and other things in the NS are included in the depth of the memory. The substanguages in the NS are drawn from the depths of memory to the current working state of external interaction. These substanguages also take on new characteristics to suit the current state—by passing from their vague state in the depths of memory to their current clear state, they cause people to experience a feeling of hope, which is similar to the abovementioned "metaphor" phenomenon. People with average faces can awaken the most memories and elicit a stronger sense of hope. This phenomenon explains why people find average faces attractive.

The explanation in this paper is reasonable; it does not introduce additional assumptions, and it is more concise than

existing explanations. Some scholars have suggested that the preference for average faces has a specific biological basis (Langlois and Roggman, 1990; Thornhill and Gangestad, 1993), which is a possible explanation for this phenomenon; however, additional evidence is needed. Halberstadt and Rhodes noted that in humans' evolutionary past, a preference for previously seen stimuli was exhibited rather than a preference for average faces per se, which would be only a byproduct of this more general preference for the familiar (Halberstadt and Rhodes, 2000). This view is similar to that in this paper.

Every strange face that people see, including so-called beautiful or mediocre faces, has some "new" characteristics, and it also generates a sense of hope in people's brains. A beautiful face is associated with a high degree of arousal, and people thus feel a sense of brightness. At this time, people have a strong sense of hope; that is, they feel beauty.

Infantile facial preference phenomenon. Like adults, infants also show certain facial preferences. Research on infants has shown that infants prefer to look at people's faces more than other visual stimuli (Morton and Johnson, 1991; Simion and Di Giorgio, 2015; Valenza et al. 1996). This phenomenon is exhibited by babies as young as a few hours or even minutes (Slatera and Quinn, 2001). Babies spend more time staring at face images than at other images. They also spend more time staring at "beautiful" faces than at "unattractive" faces.

Geneticists have proposed various theories to explain the facial preferences exhibited by infants (Banks and Ginsburg, 1985; Cassia et al. 2004; Cassia et al. 2008; de Schonen and Mathivet, 1990; Johnson et al. 2015; Kleiner and Banks, 1987; Morton and Johnson, 1991; Nakano and Nakatani, 2014; Nelson, 2001; Turati, 2004). These explanations can be divided into two categories. One category claims that infants are affected by evolution and heredity and that there are special facial recognition mechanisms in the brain at birth that guide infants to pay more attention to faces during early infancy. The other category posits that there is no innate "face recognition mechanism" in infantile brains and that infants merely have a broad recognition tendency. Thus, among the various stimuli associated with early infantile experiences, because infants have the most opportunities to come into contact with faces or since the face simply conforms to the stimulation rules of infantile preference, infants exhibit a continuous preference for faces (Chen and Zhu, 2006; Johnson, 2001).

Simion and Di Giorgio, (2015) provided a number of examples that seem to support the first type of view. The authors noted that "2-day-old newborns, despite their lack of experience, orientation preferentially toward face or face-like configurations are more than to other, equally complex, non-face stimuli" (Simion and Di Giorgio, 2015). However, these examples do not show that newborn children have no visual experience; rather, at most, they indicate that newborns have no visual experiences that are similar to those of adults. For example, Sugita (2008) found that newborn monkeys, who lacked any visual experience with faces, manifest a preference for faces over objects. We noticed that in Sugita's experiment, the person who fed the monkey merely wore a mask. If the contour of the head is also viewed as part of the facial features, it cannot be said that the monkey entirely lacked facial visual experience.

Although we do not rule out the possibility of the first explanation, we believe that more rigorous evidence on this topic is still needed. In the following, we explain the infantile face preference phenomenon according to the method proposed in this article.

A newborn baby has no visual memory, and his or her NS is blank; however, the mechanism of "thinking" is already available.

After the baby is born, his or her eyes begin to receive external information. He or she does not know anything, and only some colour and graphic information is input into his or her brain. If these symbols are isolated or if their relationship is not fixed, even if they can leave a mark on the NS, they are easily forgotten. As observed in the experiences of adults, it is difficult for people to find or remember anything when they look at the chaotic images contained in the colour blindness test chart. If a number is included in the disorderly image, even if it is intermittent, it is easy to find and remember.

If the baby sees something similar, such as the outline of a face (such as the masked person in Sugita's experiment), that image is composed of many symbols in the NS, and the relationships among these symbols are also similar. Because babies often see the outline of the face, these symbols and the relationships among them have been repeated many times, and produced a deeper memory in the baby's NS. The next time an infant sees the outline of a strange face, more memories are evoked and a greater sense of hope is produced; thus, the infant will stare at it. The situation is similar if the baby sees not only the outline of the face but also various complete faces.

The physiological conditions of newborns and adults are quite different. The newborn's brain is not fully developed, and his or her visual acuity and contrast sensitivity are also very poor. The faces that newborns see are not clear. Therefore, do newborns and adults like to watch beautiful faces for the same reason?

The NF that is seen by an infant is different from that seen by an adult. The infant is not familiar with noses or mouths. What he or she sees is merely colourful blocks and certain simple shapes (Johnson, 2011; Quinn et al. 2001). The baby cannot build a whole image of these things in his or her brain. Of course, the symbols that enter the NS of adults are actually similar to those associated with infants, but the processing mechanism of the adult brain has been further developed, enabling the adult to quickly organize those colourful blocks, shapes and other symbols into the eyes and nose and even a complete face. Adults can organize the various symbols that constitute the nose into noses and the various symbols that constitute the face into faces because they have seen many similar things, and their brains have developed a processing mechanism for whole facial images.

We have no reason to think that these situations are different. Generally, the images that newborn babies see most frequently are human faces. Although the number of images they see is far less than the number of images adults have seen, the average effect of these images is similar to that of the average face in the eyes of adults, that is, a beautiful face. Thus, beauty in the eyes of infants and adults is essentially due to the sense of hope that emerges after more sublanguages are awakened in their memory.

Psychological reasons for phonocentrism

Although some phenomena appear to be cultural phenomena, they are essentially psychological phenomena. For example, Derrida once criticized the phenomenon of phonocentrism (Derrida, 1967), which emphasizes the priority of spoken words over written words. In his view, many past thinkers, such as Plato, Aristotle, Rousseau, Hegel, and Saussure, have praised spoken words while denigrating written words. The reason for this attitude is that spoken words are the original form, the symbol of inner emotions, whereas written words are derived, accidental, special and external. In this way, past thinkers seem to have made choices based on a cultural concept.

It is undeniable that spoken words are the most primal factor, but does this notion of "primal" have a special meaning? If we consider the difficulty of inventing written words, we also have

reasons to attach greater importance to written words. This phenomenon also occurs, for example, in China, where pictographic characters are used. The ancient myth of "Cangjie creating characters" demonstrates Chinese people's reverence for written words.

Did people develop a preference for spoken words based on this cultural concept, or did they form this cultural concept because they preferred spoken words? Even if people develop a preference for spoken words because of certain cultural concepts, these cultural concepts should also conform to people's psychological habits; otherwise, these concepts cannot be sustained.

People's preference for sound may be a widely acknowledged fact, but we must determine the reasons for such a preference. We analyse the interactions between people and sound or written symbols and find that the dependence of the relationships between people and these two symbols on time and space differ.

Sound symbols are presented in strict accordance with the progression of time. When people hear a word, they must usually hear the sound of the entire word to know it. When they hear a sentence, they must also hear each word in order for the sentence to be complete in terms of meaning. Each individual sound that people hear cannot constitute a complete meaning, and most of the time, an individual sound can even be meaningless. Therefore, the process of listening to spoken words always relies on time.

The process by which people read written words is different from that by which they listen to sounds. Although the process of reading written words is also based on time, the temporality of the process is not pure. For example, when people see the word "time", they do not see the word in sequence as t-i-m-e; they see the complete word at a glance. People can even see several words at a glance.

Another example is the following interesting passage: "Aoccdnrig to a rscheearch at Cmabrigde Uinervtisy, it deosnt mttar in waht oredr the ltteers in a wrod are, the olny iprmoent tihng is taht the frist and lsat ltteer be at the rghit plcae. The rset can be a toatl mses, and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe." (<https://www.mrc-cbu.cam.ac.uk/people/matt.davis/Cmabrigde/>). Many of the letters in the words in this sentence are out of order, but people may not notice the problem if they do not read the passage carefully. This demonstration also shows that reading written words reflects the temporality of reading in the correct order as well as the spatiality of reading multiple characters at the same time with the human eye.

Because at any given moment, the only thing that enters the auditory system is an incomplete and separate speech symbol, people always exhibit a psychological mode of "waiting-realization" when they listen. People are always waiting for the arrival of the next note and supplement the meaning of the previous notes after it arrives. People have expectations for the future while waiting for the next note, and as the meaning of the sound is presented over time, their inner hopes are also realized. Therefore, purely based on the characteristics of the symbols themselves, it is easier for sounds to generate hope than for written words to do so, which is a psychological explanation for why people prefer spoken words.

This discussion is just one example of an analysis of cultural phenomena using the emotion theory in this paper. We hypothesize that hope precedes knowledge or culture, which means that human knowledge is gradually produced and selected in hope. The reason for this assumption is also very simple: if individuals did not have knowledge in the first place, they must have gradually produced knowledge based on some human instinct. So far as we can imagine, this instinct is a choice between hope and fear. This is such a big question that readers may be able to give a more in-depth answer to this question than we can.

Can artificial intelligence with self-emotion be achieved?

Emotions are not just about people; the machines of the future may have emotions. Artificial intelligence (AI) is a hot topic in the field of science and technology, and it is expected that it will greatly change human life in the future. Different scholars have provided different definitions of AI. Russell and Norvig (2011) summarized the definitions of AI into four types in their book: technology that thinks humanly, thinks rationally, acts humanly or acts rationally. Although there are different definitions, it can also be argued that the first definition is the most important because the last three definitions are based on thinking humanly. If AI cannot think humanly, it cannot truly think rationally, act humanly or act rationally. Thinking humanly also implies that intelligent machines have emotions similar to those of humans because emotions are a fundamental fact of being human. As Minsky once noted, "The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without emotions" (Minsky, 1985).

Some scholars have focused on affective computing (Noroozi et al. 2021; Picard, 1997). Through intelligent algorithms, AI can recognize human emotions to a certain extent, but the essence of what AI obtains through recognition is not emotion but language. If we expand the meaning of language and claim that all kinds of symbols constitute language, then the facial expressions, body expressions and vocal expressions (Dael et al. 2012; Du et al. 2014; Friedhoff et al. 1962) that a computer detects all count as language. Whether the relationship between the expressions recognized by computers and human emotions can be correctly matched depends on people's knowledge of the relationship between these two factors. Even if this relationship is completely correct, the computer cannot generate its own emotions. At most, it can be considered to have a strong ability to observe expressions. Can AI generate its own emotions as people can? How can it do so? Answering these questions requires a clear understanding of the nature of emotions.

Based on the discussion of the nature of emotions in this paper, we propose that AI with self-emotions may be achievable. If we are familiar with the method of emotion generation, we can design the corresponding functions of AI in an appropriate way, and AI can generate emotions. AI with self-emotion is not a computer that contains a prebuilt model of emotional response. It should be like a baby and thus start with a clean slate. AI should have many storage cells to simulate the NS; these memory cells should initially be empty, but they should also be able to simulate the memory characteristics of neurons. The interactions among these storage cells should reflect the two basic interaction states of hope and fear.

The emotions generated by AI receive some weight in its cognition and decision-making, which then affect its cognition and decision-making. The weight of AI's emotion in cognition and decision-making is also the result of its own evolution. For example, when a person has more knowledge, knowledge factors account for a large proportion of the person's cognition and decision-making, while emotional factors account for a small proportion thereof. AI should also reflect this feature.

Although AI that can generate self-emotion can be realized in principle, we are still far from developing the technology needed to realize this goal, and we will encounter various unexpected difficulties in this task. If AI can simulate a baby's attraction to beautiful faces and produce the corresponding expression of staring at those faces, in our view, this achievement can be considered a great success in the early stages of this process.

Conclusion and prospects

This paper presents a novel emotion theory based on a universal model. We offer a clear definition of emotion and deduce the

basic emotions of hope and fear, which are completely different from the interpretations of previous psychologists. The function of this article is similar to that of using sound waves to describe sound, as the basic emotions identified by traditional psychologists are foundational to human experience, much the way the seven-tone scale is foundational to expressing basic tones in music.

Based on the emotion theory thus developed, we have reasonably explained the phenomena of facial attraction and infantile facial preference and discussed the psychological reasons for phonocentrism. In a paper, we cannot use this tool of binary analysis to analyse more psychological phenomena, but the clear and simple approach to understanding human emotions that it provides shows that it is a promising tool. This theory of emotion is based on a solid foundation, not on the experience of only a particular population, and is therefore applicable to the analysis of all emotion-related phenomena.

The universal model is also applicable to the nervous system in the brain; each neuron in the brain connects with thousands of other neurons (Stern, 2022), and brain connections determine the brain's functional organization (de Schotten and Forkel, 2022). We believe that this paper can also inspire neuroscientists to understand the biological basis of emotion more accurately.

Data availability

Data sharing is not applicable to this research, as no data were generated or analysed.

Received: 28 June 2023; Accepted: 21 February 2024;

Published online: 04 March 2024

References

- Arnold MB (1960) *Emotion and Personality*. Columbia University Press, New York
- Banks MS, Ginsburg AP (1985) Infant visual preferences: A review and new theoretical treatment. *Adv Child Dev Behav* 19:207–246
- Cannon WB (1927) The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology* 39(1/4):106–124
- Cassia VM, Turati C, Simion F (2004) Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychol Sci* 15(6):379–383
- Cassia VM, Valenza E, Simion F et al. (2008) Congruency as a nonspecific perceptual property contributing to newborns' face preference. *Child Dev* 79(4):807–820
- Chen G, Zhu B (2006) Review on the theories of infant face preference. *Adv Psychol Sci* 14(4):625–630 (in Chinese)
- Cunningham MR, Roberts AR, Barbee AP et al. (1995) "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *J Pers Soc Psychol* 68:261–279
- Dael N, Mortillaro M, Scherer KR (2012) Emotion expression in body action and posture. *Emotion* 12(5):1085–1101
- Damasio A (2011) Neural basis of emotions. *Scholarpedia* 6(3):1804
- Derrida J (1967) *De La Grammatologie*. Les Editions de Minuit, Paris. English edition: Derrida J (1997) *Of Grammatology* (trans: Spivak GC). Johns Hopkins University Press, Baltimore
- de Schonen S, Mathivet E (1990) Hemispheric asymmetry in a face discrimination task in infants. *Child Dev* 61(4):1192–1205
- de Schotten MT, Forkel SJ (2022) The emergent properties of the connected brain. *Science* 378(6619):505–510
- Du SC, Tao Y, Martinez AM (2014) Compound facial expressions of emotion. *Proc Natl Acad Sci USA* 111(15):E1454–E1462
- Ekman P (1999) Basic emotions. In: Dalgleish T, Power MJ (eds) *Handbook of Cognition and Emotion*. John Wiley & Sons, New York, p 45–60
- Ekman P (2003) *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. Henry Holt and Company, New York
- Ekman P, Cordaro D (2011) What is meant by calling emotions basic. *Emot Rev* 3(4):364–370
- Ekman P, Friesen WV, O'Sullivan M et al. (1987) Universal and cultural differences in the judgments of facial expressions of emotion. *J Pers Soc Psychol* 53:712–717

- Ekman P, Sorenson ER, Friesen W (1967) Pan-cultural elements in facial displays of emotion. *Science* 164:86–88
- Friedhoff AJ, Alpert M, Kurtzberg RL (1962) An effect of emotion on voice. *Nature* 193(4813):357–359
- Halberstadt J, Rhodes G (2000) The attractiveness of nonface averages: Implications for an evolutionary explanation of the attractiveness of average faces. *Psychol Sci* 11(4):285–289
- Han CY, Wang HY, Hahn AC et al. (2018) Cultural differences in preferences for facial coloration. *Evol Hum Behav* 39(2):154–159
- Huang Y (2018) *Substangue and Interaction*. Science Press, Beijing (in Chinese)
- Izard CE, Malatesta CZ (1987) Perspectives on emotional development. In: Osofsky JD (eds) *Handbook of Infant Development*, 2nd edn. Wiley, New York, p 494–554
- James W (1884) What is an emotion? *Mind* IX(34):188–205
- Johnson MH (2001) Functional brain development in humans. *Nat Rev Neurosci* 2(7):475–483
- Johnson MH, Senju A, Tomalski P (2015) The two-process theory of face processing: modifications based on two decades of data from infants and adults. *Neurosci Biobehav Rev* 50(SI):169–179
- Johnson SP (2011) Development of visual perception. *Wiley Interdiscip Rev- Cogn Sci* 2(5):515–528
- Kleiner KA, Banks MS (1987) Stimulus energy does not account for 2-month-olds' face preferences. *J Exp Psychol -Hum Percept Perform* 13(4):594–600
- Langlois JH, Roggman LA (1990) Attractive faces are only average. *Psychol Sci* 1(2):115–121
- Lazarus RS (1966) *Psychological Stress and the Coping Process*. McGraw-Hill, New York
- LeDoux JE (1996) *The emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon & Schuster, New York
- Levenson RW (2011) Basic emotion questions. *Emot Rev* 3(4):379–386
- Lindquist KA, Wager TD, Kober H et al. (2012) The brain basis of emotion: A meta-analytic review. *Behav Brain Sci* 35(3):121–143
- Little AC, Jones BC, DeBruine LM (2011) Facial attractiveness: Evolutionary based research. *Philos Trans R Soc B- Biol Sci* 366(1571):1638–1659
- Lyons W (1999) The philosophy of cognition and emotion. In: Dalglish T, Power MJ (eds) *Handbook of Cognition and Emotion*. John Wiley & Sons, New York, p 21–44
- Meng ZL (ed) (2005) *Emotional Psychology*. Peking University Press, Beijing, p 40 (in Chinese)
- Miller RS (2012) *Intimate Relationships*, 6th edn. McGraw-Hill Education, New York, p 77–78
- Minsky M (1985) *The Society of Mind*. Simon & Schuster, New York
- Morton J, Johnson MH (1991) CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychol Rev* 98(2):164–181
- Nakano T, Nakatani K (2014) Cortical networks for face perception in two-month-old infants. *Proc R Soc B- Bio Sci* 281(1793):20141468
- Nelson CA (2001) The development and neural bases of face recognition. *Infant Child Dev* 10(1-2):3–18
- Noroozi F, Corneanu CA, Kaminska D et al. (2021) Survey on emotional body gesture recognition. *IEEE Trans Affect Comput* 12(2):505–523
- Ortony A, Turner TJ (1990) What's basic about basic emotions? *Psychol Rev* 97(3):315–331
- Panksepp J (2007) Criteria for basic emotions: Is DISGUST a primary "emotion"? *Cogn Emot* 21(8):1819–1828
- Panksepp J, Watt D (2011) What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emot Rev* 3(4):387–396
- Picard RW (1997) *Affective Computing*. MIT Press, London
- Plutchik R (1980) A language for the emotions. *Psychology Today* 13(9):68–78
- Quinn PC, Eimas PD, Tarr MJ (2001) Perceptual categorization of cat and dog silhouettes by 3-to 4-month-old infants. *J Exp Child Psychol* 79(1):78–94
- Rhodes G (2006) The evolutionary psychology of facial beauty. *Annu Rev Psychol* 57:199–226
- Rhodes G, Harwood K, Yoshikawa S et al. (2002) The attractiveness of average faces: Cross-cultural evidence and possible biological basis. In: Rhodes G, Zebrowitz LA (eds) *Facial Attractiveness: Evolutionary, Cognitive and Social Perspectives*. Ablex, Westport, p 35–58
- Rubenstein AJ, Langlois JH, Roggman LA (2002) What makes a face attractive and why: The role of averageness in defining facial beauty. In: Rhodes G, Zebrowitz LA (eds) *Facial Attractiveness: Evolutionary, Cognitive and Social Perspectives*. Ablex, Westport, p 1–33
- Russell SJ, Norvig P (2011) *Artificial Intelligence: A Modern Approach*, 3rd edition. Pearson Education, Harlow
- Simion F, Di Giorgio E (2015) Face perception and processing in early infancy: inborn predispositions and developmental changes. *Front Psychol* 6:969
- Slatara A, Quinn PC (2001) Face recognition in the newborn infant. *Infant Child Dev* 10(1-2):21–24
- Sroufe A (1996) *Emotional Development: The Organization of Emotional Life in the Early Years*. Cambridge University Press, New York
- Stern P (2022) No neuron is an island. *Science* 378(6619):486–487
- Sugita Y (2008) Face perception in monkeys reared with no exposure to faces. *Proc Natl Acad Sci USA* 105(1):394–398
- Thornhill R, Gangestad SW (1993) Human facial beauty: Averageness, symmetry, and parasite resistance. *Hum Nat- Interdiscip Biosoc Perspect* 4(3):237–269
- Turati C (2004) Why faces are not special to newborns: An alternative account of the face preference. *Curr Dir Psychol Sci* 13(1):5–8
- Valentine T, Darling S, Donnelly M (2004) Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces. *Psychon Bull Rev* 11(3):482–487
- Valenza E, Simion F, Cassia VM et al. (1996) Face preference at birth. *J Exp Psychol -Hum Percept Perform* 22(4):892–903
- Whalen PJ (1998) Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Curr Dir Psychol Sci* 7:177–188
- Young P (1973) Feeling and emotion. In: Wolman BB (eds) *Handbook of General Psychology*. Prentice-Hall, Englewood Cliffs, p 649–771

Author contributions

YH was responsible for the conception, analysis, and writing of this paper.

Competing interests

The author declares no competing interests.

Ethical approval

The Ethical approval was requested by Yong Huang at Oct 16, 2023. Approval was obtained from the Bioethics and Medical Ethics Committee of Beihang University (No. BM20230257) at Oct 20, 2023.

Informed consent

All anonymous participants were aware of the purpose of the survey prior to the survey and volunteered to answer the questionnaire. The questionnaire was dated May 2, 2023.

Additional information

Correspondence and requests for materials should be addressed to Yong Huang.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024