Experimental Study on Abstract Expression of Human-Robot Emotional Communication

Jianmin Wang [1,2], **Yuxi Wang** [1], **Yujia Liu** [1], **Tianyang Yue** [1], **Chengji Wang** [1], **Weiguang Yang** [1], **Preben Hansen** [3] **and Fang You** [1,2,*]

1   Car Interaction Design Lab, College of Arts and Media, Tongji University, Shanghai 201804, China; wangjianmin@tongji.edu.cn (J.W.); wangyuxi@tongji.edu.cn (Y.W.); liuyujia@tongji.edu.cn (Y.L.); michaelyue0812@tongji.edu.cn (T.Y.); laoji@tongji.edu.cn (C.W.); yangweiguang@tongji.edu.cn (W.Y.)
2   Shenzhen Research Institute, Sun Yat-sen University, Shenzhen 518057, China
3   Department of Computer and Systems Sciences, Stockholm University, 114 19 Stockholm, Sweden; preben@dsv.su.se
*   Correspondence: youfang@tongji.edu.cn; Tel.: +86-21-69584745

Abstract: With the continuous development of intelligent product interaction technology, the facial expression design of virtual images on the interactive interface of intelligent products has become an important research topic. Based on the current research on facial expression design of existing intelligent products, we symmetrically mapped the PAD (pleasure–arousal–dominance) emotion value to the image design, explored the characteristics of abstract expressions and the principles of expression design, and evaluated them experimentally. In this study, the experiment of PAD scores was conducted on the emotion expression design of abstract expressions, and the data results were analyzed to iterate the expression design. The experimental results show that PAD values can effectively guide designers in expression design. Meanwhile, the efficiency and recognition accuracy of human communication with abstract expression design can be improved by facial auxiliary elements and eyebrows.

Keywords: facial expressions; emotion model; emotion expression design

1. Introduction

Facial expressions play an important role in understanding people's intentions and emotions toward each other in the process of communication. This paper studies general expressions for the Chinese population, using them as the main participants of the design and experiment. With the development and exploration of artificial intelligence technology, virtual images are increasingly used in different contexts such as private homes, healthcare, cockpits, and various service industries [1]. The demand for virtual images has been further increased from simple facial animation to more intelligent emotion expression; these emotion expressions will bring changes in human emotional state and further influence human behavior and emotional state [2]. Therefore, emotion expression features have become the focus of research [3], but few researchers have studied emotion recognition and emotion expression in robots [4]. In addition, designers need to follow a set of scientific and systematic methods to evaluating visual image designs across environments and scenarios. Most of the current designs have limitations as they are designed based on the subjective feelings of the designer and cannot be evaluated by quantitative methods.

In this research, based on the present facial expression design of virtual images in intelligent products, we studied how to reasonably extract the abstract expression features of virtual images and correspond the PAD emotion values to virtual images for better interaction and communication with people. For example, the expression design of service robots can improve the effect of human–robot communication and enhance user experience (UX). In this study, the virtual image was designed to better enable humans to

communicate with robots and allow humans to understand abstract emotion expression more widely. In addition, a quantitative evaluation method of virtual image based on the PAD three-dimensional emotion model was proposed, which can help designers better design and iterate the emotion expression of virtual image facial expressions and improve the effectiveness of human–robot communication. In view of this research, the study performed the following work:

(1) In the first part of this paper, we introduce the application scenarios of existing virtual images in detail and overview the development of facial expressions and three-dimensional (PAD) emotion model.

(2) In the second part of this paper, we summarize, discuss, and design the facial emotion features of virtual images.

(3) In the third part of this paper, we use the PAD emotion model to conduct experimental evaluation and analysis to verify whether the facial expression design of avatars conforms to human cognition.

2. Literature Review

Since the appearance of virtual assistants (VAs), the cooperation between human beings and VAs has gradually increased. The market scale of VAs in public services and medical care has been expanding [5,6]. The VA based on emotion design affects human life and behavior. The emotion expression function of VAs also plays an important role in many application scenarios in real life. In order to effectively improve UX and human–robot interaction (HRI), VAs can use non-verbal behaviors, especially facial expressions, to express emotions and play an important role in improving human–computer interaction (HCI) [7].

2.1. The Development of Emotion Expression Design in Virtual Images

In the research and exploration of the emotion expression design of virtual images, the facial design of virtual images serves as an important factor that influences user behavior and response [8]. Han [9] studied three robot facial design styles—iconic, cartoon, and realistic—and showed that iconic or cartoon-like faces are preferred to realistic robot faces. Fernandes [10] used the eyes and mouth to design the facial expressions of 2D cartoon faces to achieve the desired emotions. In Tawaki's research [11], it was proposed that when communicating with people, the robot's face conveys the most important and richest information. The author simplified facial expressions into face elements that are two eyes and a mouth, which are geometrically transformed to form facial expressions. Albrecht I et al. proposed [12] that the facial expressions of VAs can enhance the expressiveness of voice reminders. For example, in the aspect of intelligent vehicle interaction [13], the vehicle VA can not only reduce the driver's distraction but also enhance the deep understanding of the information conveyed [14]. However, the current research direction of VAs is more on how to help users complete certain tasks but less on how to facilitate emotional communication between humans and robots. Therefore, it is a great challenge in design to make virtual images show the same emotional expression as human beings and improve communication with them [15].

2.2. Facial Emotion Evaluation Model

Facial expression and its relationship with emotion have been widely studied. The existing emotion description models mainly include the following types: Dimensional Emotion Model, Discrete Emotion Model, and Classification Model [16]. For instance, in 1971, Ekman and Friesen [17] used the Discrete Emotion Model to express different emotions and put forward six basic emotion types: anger, disgust, fear, sadness, happiness, and surprise. These six emotions can be combined, and various compound emotions such as depression, distress, and anxiety can be derived.

In contrast to the Discrete Emotion Model, the bipolar pleasure–arousal scale of Mehrabian and Russell [18] is widely used. Russell [19] made the Ring Model of emotion

classification in 1989 and argued that people's emotions can be divided into two dimensions: happiness and intensity. Plutchik [20] designed the Emotional Wheel Model. He suggested that people's emotions have three dimensions: strength, similarity, and polarity. This model emphasizes that when emotions show different strengths, they will influence each other to produce similar or opposite emotions. Wundt [21] first proposed the Three-Dimensional Theory of emotion, whose dimensions are pleasant–unpleasant, excited–calm, and tension–relaxation. Furthermore, Osgood [22] suggested three dimensions to describe the emotional experience, which are evaluation, potency, and activity. The changes among each dimension are continuous. In 1974, Mehrabian and Russell [23] put forward the PAD Three-Dimensional Emotional Model with the highest recognition so far in which P (pleasure, indicating the positive and negative state of individual's emotion), A (arousal, indicating the individual's neurophysiological activation level), and D (dominance, indicating the individual's control over the situation and others) are the three dimensions.

The PAD emotion model has been used in various research of emotion calculation. Becker [24] used PAD space to quantify the emotional calculation of virtual humans and used coordinates or vector Q (P, A, D) to represent the emotion state. As shown in Figure 1, the origin of coordinates indicates that the emotion is in a calm state. Therefore, we can quickly find the basic emotional points and some indistinguishable emotional points in the three-dimensional PAD emotional space such as sadness −P−A−D and shame −P+A−D, which are divided into different emotional spaces by different arousal degrees. In this paper, three-dimensional emotional space with the best comprehensive performance was selected to define emotions.
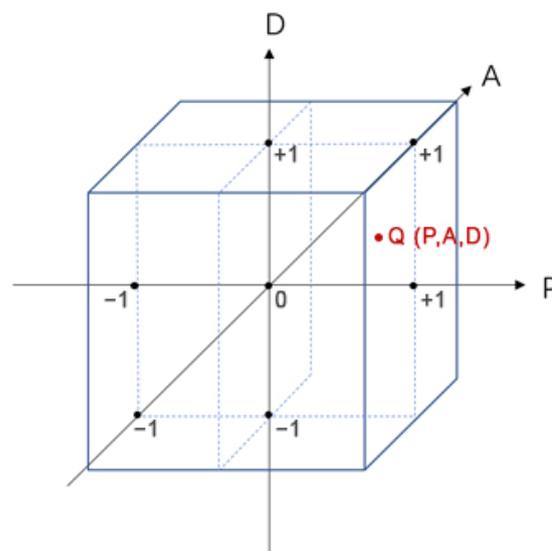

Figure 1. PAD three-dimensional emotion space.

The models above only study emotions from a qualitative perspective. In other words, based on the levels of pleasantness and arousal, it is not possible to conclude whether a participant felt either fascinated or proud, and thus this article uses the PAD three-dimensional emotion space model to study emotion values, integrally and comprehensively define the position of emotions in the space, and conduct quantitative analysis.

3. Emotion Expression Design of Virtual Image

In the process of interaction between people and VAs, the facial expression is an effective way to express emotions. Mehrabian [25] provided a formula: emotion expression =7% language communication +38% pronunciation and intonation expression +55% facial expression. It shows that most people convey emotional information through facial

expressions in nonverbal communication, and facial emotion expression is an important communication channel in the process of face-to-face interaction.

DiSalvo et al. [26] conducted a study on the influence of facial features and size of virtual images on human-like perception. They suggested that the facial expression design of virtual images should balance three factors which are human-ness, robot-ness, and product-ness. The "human-like nature" allows users to interact with VAs more intuitively; the "robot-ness" is the expectation of a virtual image's cognitive ability; and the "product-ness" allows people to regard the virtual image as a tool or equipment. Manning [27] introduced the triangle design space for cartoon face, which means that the more iconic the face of the avatar such as a simplified emoji is, the more it can enlarge the meaning of facial emotion and make people focus on information instead of media. At the same time, the more symbolic the avatar's face is, the more people it can represent, the stronger its credibility is, and the more it can avoid the influence of the "Uncanny Valley" theory [28]. If a robot looks too much like a human, it can be scary and even offensive. Designing animals and robots with human characteristics while maintaining their non-human appearance is much safer than using more anthropomorphic animals or entities. For designers, the safest combination seems to have an obvious non-human appearance but can express emotions like humans [29].

The eyes are the organs that can directly reflect people's emotions in facial expressions. Tawaki also explored ways to simplify faces, showing that two eyes and one mouth are important components of facial expression and can effectively convey emotion [11]. The gaze direction of the eyes can convey the visual focus of the VA and show its behavioral motivation. The emotions of virtual images are divided into two types: positive emotions and negative emotions. When the concave face of the eyes is upward, it means that the virtual image of which is in a positive emotion such as happiness, gratitude, and pride. When the eyes of negative emotions are concave upward, it indicates that the virtual image is in a negative emotion such as shame, sadness, and anger. When the eyes are slightly narrowed, it indicates that the avatar is in a mild and relaxed state. Mouth changes are one of the criteria for judging emotion expression. When the corners of the mouth turn up, most of the emotions are positive emotions such as joy, hope, gratitude, and admiration. A larger upward radian represents more pleasure. Otherwise, when the mouth turns down, it shows negative emotions such as sadness, shame, hostility, and other emotions.

House et al. [30] found that eyebrows made great contributions to the importance of perceiving users' emotions. Ekman [31] found that eyebrow movement is related to communication, and the type of facial emotion expression action depends on the dynamic of environmental eyebrows to a great extent. It can be seen that eyebrows cannot be ignored in the design of facial emotion expression.

Emoticons are auxiliary elements in visual symbols. On the one hand, they can express the emotions of virtual images more accurately through hand gestures, items, and so on. On the other hand, they can convey emotions that cannot be accurately transmitted to users visually, thus satisfying the functions and creating a relaxed and pleasant atmosphere. In the usage of emojis, users of different ages, countries, and cultural backgrounds will have the same understanding of the meaning of the emoji, thus achieving the unity of cognition. Robots can use a combination of emoticons and facial expressions to achieve a higher degree of anthropomorphism [32].

We adopted and extended the PAD emotional dimension method to construct the emotions expressed by facial expressions. Emotions are divided into categories according to the eight spatial principles of PAD [33]. Young's research [32] involved simplified and exaggerated facial expressions and hand gestures that could provide powerful expressive mechanisms for robots. In this study, we refer to the mapping relationship between expression parameters and facial expression features in the animation Tom and Jerry [34] and Disney classic animated characters, and it mainly includes four parts: eye feature, mouth feature, eyebrow feature, and emoticon feature. The iconic design elements that best meet the avatar expression are extracted. Emojis' visual emotional symbols are referenced

in the design of emoticons including facial expressions, hand gestures, and functional symbols. As shown in Table 1, facial emotion features are analyzed and summarized.

Table 1. Summary table of emotion expression features of facial design expressions.

Emotional Space Dimension	Emotion	Eyebrow	Eye	Mouth	Emoticons
Exuberant	Love	Eyebrow level	Eyes wide open, Notch down	Mouth up	Love
	Joy	Eyebrow level/Raised eyebrow level	Slightly squinted, Notch down	The corners of the mouth face up slightly	Flush
	Happiness	Eyebrow level/ Raise eyebrows	Half-squinted eyes, Notch down	Grin open	Blushing/Hands raised
	Pride	Raised eyebrow level	Eyes wide open, Notch down	The corners of the mouth substantially upward	Stars/V gestures/medals
	Gratification	Eyebrow level, Slightly larger spacing	Slightly squinted, Notch down	Corners of the mouth up	Blushing
Dependent	Hope	The two eyebrows face down at the ends, like '/\\'[1]	Slightly squinted, Notch down	Corner of mouth slightly upward	Stars/hands closed
	Liking	The two eyebrows face down at the ends, like '/\\'	Slightly squinted, Notch down	The corners of the mouth face up slightly	Blushing/Love
	Gratitude	Eyebrow level, Slightly larger spacing	Slightly squinted, Notch down	The corners of the mouth face up slightly	Blushing/Love/ Hands closed
	Admiration	Eyebrow level/Raised eyebrow level	Slightly squinted, Notch down	The corners of the mouth face up slightly	Great gestures/ fist gestures
Relaxed	Relief	The two eyebrows face down at the ends, like '/\\'	Slightly squinted, Notch up	The corners of the mouth face up slightly	Blushing
	Satisfaction	Eyebrow level, Raised eyebrow level	Slightly squinted, Notch down	Corners of the mouth up	Blushing/give the OK sign
Docile	Mildness	The two eyebrows face down at the ends, like '/\\'	Slightly squinted, Notch down	The corners of the mouth face up slightly	Blushing
Bored	Resentment	Raise both ends of the eyebrows, Raise and lower eyebrows	Eyes squinted, Notch up	The corners of the mouth are tilted down	Hold up head with hand
	Pity	Eyebrows down, like '/\\'	Slightly squinted, Notch up	The corners of the mouth face down slightly	Tears, Spread your hands
	Distress	Raise the ends of the eyebrows, Spacing becomes smaller	Half-squinted eyes, Notch up	The corners of the mouth face down sharply	Tears, Wipe away tears
	Fear	Eyebrows down, like '/\\'	Wide-eyed/slanted, Notch up	The corners of the mouth face down sharply	Hold up head with hand
	Shame	Eyebrows down, like '/\\', Spacing becomes smaller	Half-squinted eyes, Notch up	The corners of the mouth face down slightly	Blushing
	Remorse	Eyebrows down, like '/\\', Spacing becomes smaller	Half-squinted eyes, Notch up	The corners of the mouth face down slightly	Hold your forehead
Disdainful	Disappointment	Eyebrows down, like '/\\', Spacing becomes smaller	Half-squinted eyes, Notch up	The corners of the mouth face down sharply	Sigh
	Fear	The eyebrows are raised up, like '\\/'	Wide-eyed/slanted, Notch up	Mouth wide open	Hold your face with both hands
Anxious	Reproach	One eyebrow raised at both ends, Eyebrows Level	Half-squinted eyes, Notch up	Mouth wide open	Letters and symbols
Hostile	Disliking	Eyebrows slanted and raised	Wide-eyed, slanted, Notch up	The corners of the mouth face down sharply	Letters and symbols
	Anger	Eyebrows slanted and raised	Wide-eyed, Notch up	The corners of the mouth face down sharply	Angry symbol
	Hate	Eyebrows slanted and raised	Wide-eyed, Notch up	The corners of the mouth face down sharply	Flames of fury

[1] A Chinese character figuratively describes the eyebrow.

4. Evaluation of Facial Expression Design of Virtual Image

To verify the rationality of the design and evaluate whether the expression design of the virtual image accords with the user's cognition, we need to measure the emotional changes of the users when they see the design. There are many measurement methods such as the PAD semantic difference scale in the self-report measurement method, SAM self-emotion measurement method, and PrEmo [35] method, etc. In this study, a non-verbal self-assessment manikin (SAM) was used to evaluate three dimensions: pleasure, arousal, and dominance. The SAM model has five human figures in each dimension. There are scoring points between each human figure. As shown in Figure 2, the scoring range of each dimension is from 1 to 9, with 1 representing the lowest and 9 representing the highest.
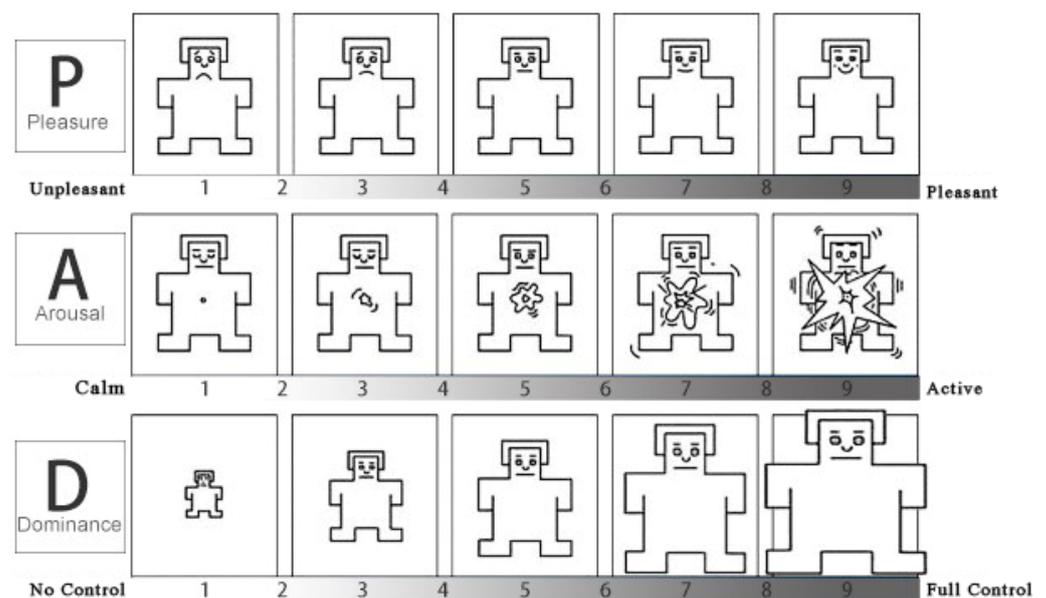

Figure 2. Self-Assessment Manikin (SAM) Scale [36]. The numbers in the figure represent the scoring range of each dimension, with 1 representing the lowest and 9 representing the highest.

Bradley and Lang [36] demonstrated that when rating the dimensions of a group of pictures, although the conclusions of using the semantic difference scale and SAM self-emotion evaluation scale were almost identical, the participants passed the cartoon picture information in the SAM measurement. Therefore, we can assume that the score of the SAM is the same as the value of PAD. Instead of text information, describing the emotion type with non-text characteristics can eliminate systemic errors caused by language semantic differences.

Through SAM evaluation, the average scores of three dimensions of emotion words and emotion expressions were obtained. Psychological experiments show that the intimacy degree between the emotional state of the participants and basic emotions [37] is the emotional tendency of the minimum distance called PAD [38] whose values of basic emotions are obtained through two rounds of experiments which can effectively evaluate the relationship between the user's emotional tendency and the degree of inclination. At any coordinate position Q (P, A, D) in PAD's three-dimensional space, the corresponding emotional state is Q. The PAD's emotional tendency is the distance between the 3D emotional space and the coordinated position between the emotional state of the participants and the basic emotions. The smaller the distance value is, the higher the degree of the participant's emotional tendency will be. The formula for calculating the coordinate distance of emotional space is as follows (z is a positive integer):

$$Li = \sqrt{(P - pi)^2 + (A - ai)^2 + (D - di)^2} \ (i \in z) \tag{1}$$

Li is the shortest distance between emotional state and basic emotions in three-dimensional emotional space. *P*, *A*, and *D* are the values of emotional state Q in three-dimensional emotional space. *pi*, *ai* and *di* are the coordinate values of basic emotional types Qi. According to Formula (1), the distance values between the emotional state of the participates and basic emotions are calculated which and denoted as L1, L2, L3 ... For example, Q = Pride (0.68, 0.61, 0.58), Qi = Pride (0.72, 0.57, 0.55) and their minimum distance value *Li* is calculated as 0.06 by the formula.

Through the data collection of the PAD value of the emotion word, the position of each emotion word in the PAD space is obtained, the position of the PAD value of the emotion word is used as the basic emotion, and the minimum distance between the PAD value of the two experimental expression designs is calculated. If the minimum distance of the expression design corresponding to the emotion word is less than or equal to other expression designs, it shows that the expression design conforms to the human's emotional cognition. Otherwise, it shows that the expressions do not conform to human emotion perception and need to be adjusted and iterated.

5. Method

There were two experiments in this study. The first experiment was divided into the emotion measurement task of emotion words and the emotion measurement task of expression. The second experiment only included the emotion measurement task of expression. Based on the results of the first experiment, we conducted a design iteration and proposed a design solution. The second experiment served as supplementary verification to design guidelines and specifications. The experiments used the Ortony, Clore, and Collins (OCC) model [39] to extract rules from emotion cognition to generate emotion words. Then the numerical calculation of the PAD dimensional space was performed for the emotion word and expression design, and the numerical relationship between emotion words and expression design was determined and verified through a quantitative study. PAD emotion space mapping was used as a bridge to match human perceptions of expression design with emotion words to ensure the accuracy of abstract expression design.

5.1. Participant

The initial task of the first experiment recruited 153 student volunteers, including 54 male participants and 99 female participants, whose ages ranged from 20 to 26, with an average age of 23.71. In the second task, 33 college students volunteered to participate in this experiment, including 9 male users and 24 female users, with an average age of 23.1.

The second experiment recruited 50 college students, including 26 male participants and 24 female participants, with an average age of 24.02. Participants in both experiments were required to be free of color blindness and color weakness.

5.2. Design of Experiment

To reduce external interference, the test was carried out indoors. People with cultural and geographical differences interpret facial expressions differently [40]. The research and experiment were conducted in China; thus we defined the Chinese people's understanding of emotions by scoring the emotion words with PAD [11]. In the emotion measurement task of emotion words in the first experiment, we translated the emotion words of PAD into Chinese-English bilingual languages. Participants scored the emotion words to obtain the accurate value of the emotion words in the emotional space, which was used as the benchmark for the mapping of the emotions in the affective space.

In the emotion measurement task of facial expressions in the first experiment, we designed facial expressions and used regular reasoning about events, objects, and attitudes to obtain the emotion word set in the OCC model of emotions. For commercial confidentiality reasons, the design used in this article was slightly adjusted in the thickness of the brush strokes. Facial expression designs 1-20, as shown in Figure 3, represent different emotions: happiness, hate, satisfaction, gratitude, reproach, distress, pride, fear, mildness,

pity, boredom, shame, disappointment, hope, resentment, love, gloating, anger, relief, and admiration. Before the start of the experiment, the participants were randomly divided into groups, with about 16 people in each group. After the experiment started, the expressions were presented to the participants in random order, and the participants scored the expressions on the SAM scale.
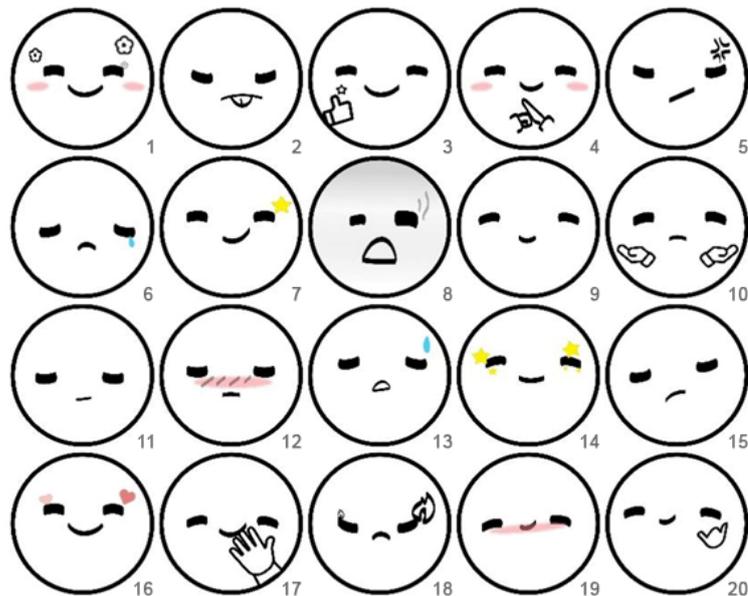

Figure 3. First edition expression design. (The numbers in the figure mark the expressions, as described in paragraph 2 of 5.2).

We modified and iterated the first-edition expression design based on the results of the first experiment. The expression iteration design is shown in Figure 4 below; the numbers in Figure 4 represent the same emotions as in Figure 3. The expression and emotion measurement task design of the second experiment is the same as the first one.
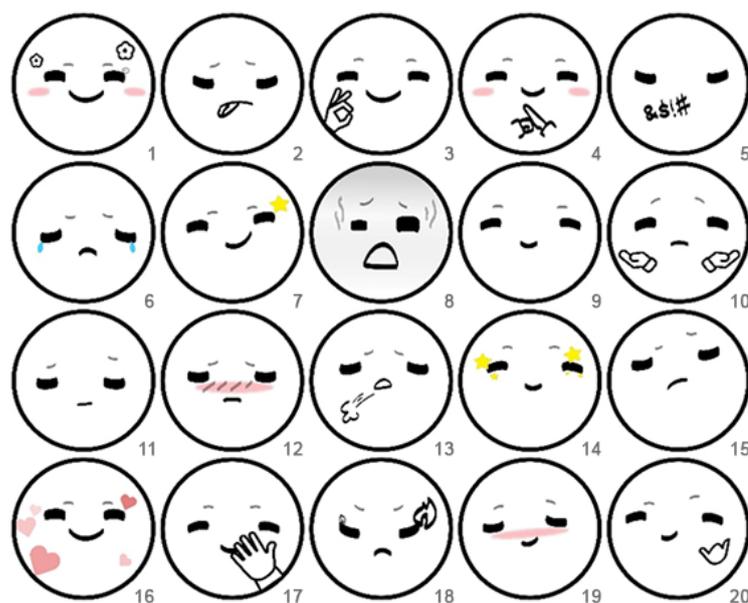
Figure 4. Iterative expression design. (The numbers in the figure mark the expressions, as described in paragraph 2 of 5.2).

5.3. Procedure

Before the start of the experiment, the experimenter introduced the purpose and tasks of the experiment and handed out the basic information form and the experiment informed consent form to the participants. Participants studied the meaning of the three emotional dimensions of PAD, the distribution in the three-dimensional space, and the scoring method of the scale and then officially start the experiment after understanding and adapting.

At the beginning of the experiment, after the participants heard the "ding" prompt, the emotion words and expression designs were presented on the screen in front of the participants in random order. In the normative study of psychology, it takes 21 seconds for the participants to complete the evaluation of a picture. Thus each emotion word or expression design was displayed for 6 seconds. Participants had to watch the screen during the entire display process. After 6 seconds, the screen was switched to a white screen. Participants responded to the display on the screen at a constant 15-second interval. The expression design or emotion words were scored on three dimensions, and the subjective feelings of the heart when they saw these emotion words were evaluated. After scoring all expression designs or emotion words, the researcher conducted in-depth interviews with the participants and selected specific ambiguous expression designs or emotion words for investigation.

5.4. Results

In this study, 20 emotion words and emotion expressions were scored on the SAM scale, and a total of 137 valid results were recovered in the first task of the first experiment. Thirty-three valid results were recovered in the second task of the first experiment, and 50 valid results were recovered in the second task of the second experiment.

According to the PAD nine-point scale, the emotional grade deviation was divided into 0~1 emotional value ranges which were light < 0.25, $0.25 \leq$ mild < 0.5, $0.5 \leq$ moderate < 0.75, and $0.75 \leq$ severe < 1. The PAD value was analyzed correspondingly with the expression elements. The following Table 2 (below) shows the data with emotional offset in the first round of the relationship between the emotion words and the emotion expressions corresponding to the minimum distance which is the PAD value of the emotion words displayed horizontally and the PAD value of the emotion expressions displayed vertically. The yellow marked part refers to the minimum distance between the emotion expression and the corresponding emotion word, while the green one indicates that the distance between other emotion expressions and the given emotion word is less than the distance between the corresponding emotion expressions. The high frequency of green annotation indicates that the emotion expression needs to be modified (slight deviation is normal).

Most expressions in the first edition of facial design in the second task have emotion offset. For example, the minimum distance between the emotion expression "Pride" and its corresponding emotion word is 0.51. However, the minimum distance between the emotion expressions "Admiration", "Gratitude", "Hope", and "Happiness" and the emotion word "Pride" is 0.48, 0.38, 0.50, and 0.38, which are all less than 0.51, indicating that emotion expression "Pride" needs to be modified. During the interview, some participants mentioned that they could not understand the meaning of the expression; thus we considered the iterative design of these expressions. Having explored the PAD value of the "Pride" emotion expression, the arousal degree of the expression is not high. Combining with the interview content, we made some adjustments to enlarge the eyes in iterative design. Thus adjusting the inclination of the expression is helpful to emphasize the expression of the "Pride" emotion. Other emotion expressions listed in the table were also analyzed according to the process above, then their optimization direction was obtained, which is not described here.

Table 2. The minimum distance between the first round of emotion words and emotion expressions.

	Pride	Love	Admiration	Gratitude	Resentment	Hope	Relief	Satisfaction	Mildness	Shame	Happiness
Reproach	1.29	1.22	0.89	0.91	0.40	1.16	0.74	1.13	0.62	0.62	0.40
Pride	0.51	0.46	0.52	0.49	1.21	0.42	0.31	0.28	0.51	1.40	0.38
Love	0.57	0.43	0.23	0.20	1.13	0.40	0.46	0.40	0.55	1.30	0.42
Admiration	0.48	0.40	0.43	0.41	1.18	0.35	0.38	0.27	0.55	1.37	0.34
Gratitude	0.38	0.30	0.46	0.43	1.28	0.25	0.47	0.18	0.65	1.46	0.23
Resentment	1.58	1.49	1.05	1.07	0.48	1.44	0.83	1.37	0.61	0.59	1.44
Hope	0.50	0.39	0.34	0.31	1.17	0.35	0.42	0.31	0.55	1.35	0.35
Relief	1.02	0.92	0.54	0.55	0.85	0.89	0.30	0.90	0.11	0.75	0.92
Satisfaction	0.67	0.55	0.29	0.28	1.03	0.52	0.32	0.46	0.40	1.21	0.52
Fear	1.46	1.34	0.90	0.92	0.37	1.64	1.06	1.34	0.92	0.44	1.33
Disappointment	1.79	1.69	1.19	1.22	0.29	0.45	1.13	1.61	0.92	0.25	1.65
Mildness	0.61	0.50	0.29	0.27	1.05	0.45	0.38	0.43	0.48	1.22	0.46
Pity	1.63	1.55	1.10	1.13	0.39	1.49	0.91	1.43	0.70	0.49	1.50
Distress	1.78	1.68	1.20	1.23	0.20	1.63	1.17	1.61	0.97	0.12	1.65
Anger	1.47	1.44	1.21	1.22	0.67	1.36	1.17	1.40	1.10	0.82	1.39
Shame	0.79	0.66	0.26	0.27	0.95	0.63	0.35	0.58	0.35	1.11	0.64
Happiness	0.38	0.27	0.41	0.38	1.28	0.23	0.51	0.21	0.66	1.46	0.23
Hate	1.40	1.34	1.01	1.02	0.40	1.28	0.79	1.23	0.65	0.61	1.28
Boredom	1.61	1.52	1.08	1.07	0.58	1.47	0.83	1.38	0.61	0.68	1.46
Gloating	0.77	0.69	0.44	0.43	0.88	0.63	0.30	0.58	0.34	1.07	0.65

In the second experiment of this study, the eyebrow elements were added based on the first round of experiments to further strengthen the embodiment of different emotions in the virtual image. We adopted and expanded the PAD emotional dimension method to construct the emotions expressed by facial expressions and analyzed and summarized the features. The summary table (Table 1) concludes emotions in the eight emotional dimensions including eyebrows, eyes, mouth, and auxiliary elements. At the same time, the three visual information dimensions proposed in the previous research [11], namely the inclination of the mouth, the opening of the face, and the inclination of the eyes, are used to iterate the facial expression design combined with the teaching value of PAD. For instance, in the positive emotional space, the eyebrows of the avatar are in a horizontal state; the eyes are slightly narrowed or widened; the shape of the eye notch is downward; and the corners of the mouth are turned upward. The auxiliary elements such as blush and love are integrated. In the boring emotional space, the eyebrows of the virtual image are lifted at both ends; the whole eyebrows are turned down; the distance between the eyebrows is reduced; the eyes are narrowed or squinted; the notch is up; and the corners of the mouth are slightly curled down. The auxiliary elements and symbols such as tears can help convey emotions better.

Comparing the results of the iterative expression experiment with those of the first edition, it is found that the emotional offset of the iterative emotion expression design is significantly reduced compared with that of the first edition. However, there is still a slight emotional offset between some emotion expressions, as shown in Table 3.

The element of eyebrows was added to the iterative expression design. Two raised eyebrows are common facial expressions. When the eyebrows are raised, the virtual images have higher arousal and dominance such as "Satisfaction" and "Pride". When the eyebrows are down, it means that the pleasure and arousal are low and the mood is in a neutral or negative state such as "Pity" and "Relief". When the eyebrows are high and low, the virtual image is between positive emotion and negative emotion, and the meaning of "Love" and "Resentment" is very different. When the distance between two eyebrows becomes smaller, most of them represent negative emotions. In addition, some of the hand gestures were changed in the second experiment. For example, the hand gesture "Satisfaction" was modified, and the minimum distance value in the second experiment was significantly smaller than the minimum distance value in the first experiment.

Table 3. The minimum distance between second-round emotion words and emotion expression.

	Admiration	Resentment	Hope	Satisfaction	Disappointment	Mildness	Pity
Reproach	1.29	0.85	1.34	1.35	1.19	1.09	1.14
Pride	0.78	1.61	0.21	0.31	1.95	1.04	1.85
Love	0.62	1.51	0.06	0.26	1.85	0.95	1.75
Admiration	0.43	1.18	0.32	0.27	1.49	0.58	1.39
Gratitude	0.31	1.14	0.35	0.35	1.45	0.58	1.35
Resentment	1.21	0.39	1.64	1.59	0.18	0.86	0.09
Hope	0.59	1.31	0.24	0.24	1.64	0.75	1.54
Relief	0.52	0.87	0.83	0.82	0.97	0.09	0.88
Satisfaction	0.59	1.31	0.24	0.24	1.64	0.75	1.54
Fear	1.14	0.55	1.67	1.58	0.89	1.21	0.86
Disappointment	1.23	0.43	0.93	1.62	0.17	0.88	0.10
Mildness	0.58	0.68	0.93	0.86	0.90	0.19	0.80
Pity	1.21	0.39	1.64	1.59	0.18	0.86	0.09
Distress	1.47	0.50	1.91	1.87	0.13	1.17	0.23
Anger	1.41	0.69	1.57	1.56	0.96	1.15	0.93
Shame	1.05	0.19	1.50	1.51	0.43	0.92	0.38
Happyiness	0.53	1.40	0.07	0.20	1.73	0.83	1.63
Hate	1.01	0.40	1.28	1.23	0.67	0.65	0.51
Boredom	1.08	0.58	1.47	1.38	0.50	0.61	0.42
Gloating	0.44	0.88	0.63	0.58	1.18	0.34	1.08

The experimental results show that it is normal for a small number of emotions to have a slight emotional deviation. For example, the emotional offset value of "Admiration" and "Gratitude" is less than the minimum distance. The difference between these two emotion expressions lies in the different hand gestures and the subtle changes of eyes and eyebrows. After analysis and comparison of Q (0.48, 0.34, −0.13) of "Admiration" and Q (0.50, 0.35, −0.10) of "Gratitude", the PAD values of these two emotions are in the same emotional space and very close. Thus the emotional offset difference between "Admiration" and "Gratitude" is 0.12.

The results of the task in the second experiment show that, after the second task of design iteration optimization, the number of offsets of emotional expressions and emotion words is significantly smaller than that of the first task. In addition, the iterative version of the design is significantly optimized based on the first task. The calculation method of PAD space combined with the minimum distance can effectively measure the numerical difference between expressions in this research to obtain the expressions that are most in line with human emotional cognition. The larger the inclination of eyes, eyebrows, and mouth, the greater the degree of awakening. Incorporating blush and emojis can effectively enhance pleasure and arousal, which is because the non-single color of the face can improve pleasure and dominance. Eyebrows are an indispensable element of facial expressions that can improve the arousal of users. Thus adding appropriate semantic emoticons is very helpful to improve the dominance. In addition, the adjustment of hand gestures can increase the user's perception of emotion expressions.

6. Conclusions and Future Work

In this study, the virtual image was used as a carrier to design the facial expressions of the emotion, and a quantitative evaluation method for virtual images based on the PAD 3D emotion model was introduced. The focus of this study was whether it can guide designers to effectively enhance human–robot communication. The results show that

using the quantitative evaluation method, namely the PAD emotion value and emotion tendency obtained by evaluating the emotion words and emotion expressions, the PAD emotion value and the image design are mapped symmetrically and can determine whether the expression design conforms to human emotion cognition. The regular features and expression design principles of the VA's facial design expressions can be used to improve the design. Based on the features of facial elements, the forms of visual presentation can produce many variations. Both experiments found that auxiliary elements such as hand gestures can effectively increase the emotional recognition of virtual images.

The research on the abstract expression design of the virtual image is part of our work on HRI interaction. Multi-channel interactions such as voice interaction and hand gesture interaction will be added to test the overall emotion interaction. Next, we will improve the HRI and UX by continuously iterating the emotion expression design of virtual images.

Author Contributions: Conceptualization, F.Y., J.W. and Y.W.; methodology, Y.W. and Y.L.; software, F.Y. and C.W.; validation, Y.W., F.Y. and T.Y.; formal analysis, Y.W.; investigation, Y.W.; resources, J.W.; data curation, Y.W. and Y.L.; writing—original draft preparation, Y.W. and J.W.; writing—review and editing, Y.W., J.W., T.Y. and P.H.; visualization, Y.W. and W.Y.; supervision, J.W. and F.Y.; project administration, J.W. and F.Y.; funding acquisition, J.W. and F.Y. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Iannizzotto, G.; Bello, L.L.; Nucita, A.; Grasso, G.M. A Vision and Speech Enabled, Customizable, Virtual Assistant for Smart Environments. In Proceedings of the 2018 11th International Conference on Human System Interaction (HSI), Gdansk, Poland, 4–6 July 2018; pp. 50–56.
2. Blairy, S.; Herrera, P.; Hess, U. Mimicry and the Judgment of Emotional Facial Expressions. *J. Nonverbal Behav.* **1999**, *23*, 5–41. [CrossRef]
3. Itoh, K.; Miwa, H.; Nukariya, Y.; Imanishi, K.; Takeda, D.; Saito, M.; Hayashi, K.; Shoji, M.; Takanishi, A. Development of face robot to express the facial features. In Proceedings of the RO-MAN 2004, 13th IEEE International Workshop on Robot and Human Interactive Communication, Kurashiki, Japan, 22 September 2004; pp. 347–352.
4. Liu, Z.; Wu, M.; Cao, W.; Chen, L.; Xu, J.; Zhang, R.; Zhou, M.; Mao, J. A facial expression emotion recognition based human-robot interaction system. *IEEE/CAA J. Autom. Sin.* **2017**, *4*, 668–676. [CrossRef]
5. Tiago, F.; Alves, S.; Miranda, J.; Queirós, C.; Orvalho, V. Lifeisgame: A facial character animation system to help recognize facial expressions. In ENTERprise Information Systems, Proceedings of the International Conference on ENTERprise Information Systems, CENTERIS 2011, Vilamoura, Portugal, 5–7 October 2011; Springer: Berlin/Heidelberg, Germany, 2011; Volume 221, pp. 423–432.
6. Park, S.; Yu, W.; Cho, J.; Cho, J. A user reactivity research for improving performance of service robot. In Proceedings of the 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Incheon, Korea, 23–26 November 2011; pp. 753–755.
7. Koda, T.; Sano, T.; Ruttkay, Z. From cartoons to robots part 2: Facial regions as cues to recognize emotions. In Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), New York, NY, USA, 6 March 2011; pp. 169–170.
8. Steptoe, W.; Steed, A. High-fidelity avatar eye-representation. In Proceedings of the 2008 IEEE Virtual Reality Conference, Reno, NV, USA, 8–12 March 2008; pp. 111–114.
9. Han, J.; Kang, S.; Song, S. The design of monitor-based faces for robot-assisted language learning. In Proceedings of the 2013 IEEE RO-MAN, Gyeongju, Korea, 26–29 August 2013; pp. 356–357.

10. Koda, T.; Nakagawa, Y.; Tabuchi, K.; Ruttkay, Z. From cartoons to robots: Facial regions as cues to recognize emotions. In Proceedings of the 5th ACM/IEEE International Conference on Human Robot Interaction, Osaka, Japan, 2 March 2010; pp. 111–112.

11. Tawaki, M.; Kanaya, I.; Yamamoto, K. Cross-cultural design of facial expressions for humanoid robots. In Proceedings of the 2020 Nicograph International (NicoInt), Tokyo, Japan, 5–6 June 2020; p. 98.

12. Albrecht, I.; Haber, J.; Kähler, K.; Schröder, M.; Seidel, H.-P. May I talk to you?—Facial animation from text. In Proceedings of the 10th Pacific Conference on Computer Graphics and Applications, Beijing, China, 9–11 October 2002; pp. 77–86.

13. Williams, K.J.; Peters, J.C.; Breazeal, C.L. Towards leveraging the driver's mobile device for an intelligent, sociable in-car robotic assistant. In Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV), Gold Coast, QLD, Australia, 23–26 June 2013; pp. 369–376.

14. Foen, N. *Exploring the Human-Car Bond Through an Affective Intelligent Driving Agent (AIDA)*; Massachusetts Institute of Technology: Cambridge, MA, USA, 2012.

15. Ofli, F.; Erzin, E.; Yemez, Y.; Tekalp, A.M. Estimation and Analysis of Facial Animation Parameter Patterns. In Proceedings of the 2007 IEEE International Conference on Image Processing; Institute of Electrical and Electronics Engineers (IEEE), San Antonio, TX, USA, 16–19 September 2007; Volume 4, pp. 4–293.

16. Zhang, T. *Research on Affective Speech Based on 3D Affective Model of PAD*; Taiyuan University of Technology: Taiyuan, China, 2018. (In Chinese)

17. Ekman, P.; Friesen, W.V. Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.* **1971**, *17*, 124–129. [CrossRef] [PubMed]

18. Liu, Y.; Jang, S.S. The effects of dining atmospherics: An extended Mehrabian–Russell model. *Int. J. Hosp. Manag.* **2009**, *28*, 494–503. [CrossRef]

19. Russell, J.A.; Lewicka, M.; Niit, T. A cross-cultural study of a circumplex model of affect. *J. Pers. Soc. Psychol.* **1989**, *57*, 848–856. [CrossRef]

20. Qi, X.; Wang, W.; Guo, L.; Li, M.; Zhang, X.; Wei, R. Building a Plutchik's Wheel Inspired Affective Model for Social Robots. *J. Bionic Eng.* **2019**, *16*, 209–221. [CrossRef]

21. Wundt, W. Outline of psychology. *J. Neurol. Psychopathol.* **1924**, *1–5*, 184. [CrossRef]

22. Osgood, C.E. Dimensionality of the semantic space for communication via facial expressions. *Scand. J. Psychol.* **1966**, *7*, 1–30. [CrossRef] [PubMed]

23. Mehrabian, A. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament. *Curr. Psychol.* **1996**, *14*, 261–292. [CrossRef]

24. Becker-Asano, C.; Wachsmuth, I. Affective computing with primary and secondary emotions in a virtual human. *Auton. Agents Multi-Agent Syst.* **2009**, *20*, 32–49. [CrossRef]

25. Davis, M.; Davis, K.J.; Dunagan, M.M. Communication without Words. In *Scientific Papers and Presentations*, 3rd ed.; Elsevier Inc.: Cham, Switzerland, 2012; Volume 14, pp. 149–159. ISBN 978-0-12-384727-0.

26. DiSalvo, C.; Gemperle, F.; Forlizzi, J.; Kiesler, S. All robots are not created equal: The design and perception of hu-manoid robot heads. In Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, London, UK, 25 June 2002; pp. 321–326.

27. Manning, A. Understanding Comics: The Invisible Art. *IEEE Trans. Dependable Secur. Comput.* **1998**, *41*, 66–69. [CrossRef]

28. Mori, M. Bukimi no Tani [The Uncanny Valley]. *Energy* **1970**, *7*, 33–35. [CrossRef]

29. Schneider, E.; Wang, Y.; Yang, S. Exploring the Uncanny Valley with Japanese video game characters. In Proceedings of the DiGRA Conference, Tokyo, Japan, 28 September 2007.

30. House, D.; Beskow, J.; Granström, B. Timing and interaction of visual cues for prominence in audiovisual speech perception. In Proceedings of the INTERSPEECH, Aalborg, Denmark, 3–7 September 2001; pp. 387–390.

31. Ekman, P. About brows: Emotional and conversational signals. In *Human Ethology*; Routledge: London, UK, 1979.

32. Young, J.E.; Xin, M.; Sharlin, E. Robot expressionism through cartooning. In Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, Arlington, VA, USA, 10–12 March 2007; pp. 309–316. [CrossRef]

33. Gebhard, P. ALMA: A layered model of affect. In Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, Utrecht, The Netherlands, 25 July 2005; pp. 29–36.

34. Shao, Z. Visual Humor Representation of Cartoon Character Design—Taking (Tom and Jerry) as an Example. *Decorate* **2016**, *4*, 138–189. (In Chinese) [CrossRef]

35. Laurans, G.; Desmet, P. Introducing PrEmo2: New directions for the non-verbal measurement of emotion in design. In Proceedings of the 8th International Conference on Design and Emotion, London, UK, 11–14 September 2012.

36. Bradley, M.M.; Lang, P.J. Measuring emotion: The self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **1994**, *25*, 49–59. [CrossRef]

37. Li, X.; Fu, X.; Deng, G. Preliminary Trial of Chinese Simplified Pad Emotion Scale in Beijing University Students. *Chin. J. Ment. Health* **2008**, *22*, 327–329. (In Chinese)

38. Jiang, N.; Li, R.; Liu, C. Application of PAD Emotion Model in User Emotion Experience Assessment. *Packag. Eng.* **2020**, 1–9. (In Chinese)

39. Ortony, A.; Clore, G.L.; Collins, A.J. *The Cognitive Structure of Emotions*; Cambridge University Press (CUP): Cambridge, UK, 1988; ISBN 9780511571299.
40. Elfenbein, H.A.; Ambady, N. Universals and Cultural Differences in Recognizing Emotions. *Curr. Dir. Psychol. Sci.* **2003**, *12*, 159–164. [CrossRef]