doi: 10.5821/conference-9788419184849.14

CHARACTERISTIC ANALYSIS OF FACIAL STIFFNESS USING AVERAGE FACES OF SCHIZOPHRENIA

Naoko KANEKON[ab] , Yoshimasa TAWATSUJI[a] , Tatsunori MATSUI [a]

[a] *Waseda University, Japan, kanekon.n@akane.waseda.jp*
[b] *Takatsuki Hospital, Japan, tk-ot@takatuki.or.jp*

ABSTRACT

There is a significant need for objective assessment methods for schizophrenia. Therefore, this study examined the emotional characteristics of facial stiffness—one of the indicators for estimating patients' flat affect—using average faces. First, we conducted an experiment for psychological evaluation of facial stiffness in 16 Japanese patients with schizophrenia. Twelve medical professionals rated the patients' facial stiffness in 147 videos. Their gaze data were also collected. To extract the physical characteristics of facial stiffness, 11 average faces of 14 male patients and one average face of a group of healthy subjects were generated. The average faces were subjected to an emotion evaluation test by the existing application and 33 medical experts. The results showed that most of the average faces differed in the proportions of the eight emotions between two tests. The most common emotions were anger for the stiffest average face and calm for the healthy average face. In the application analysis, the percentage of emotions other than calm, such as anger and sadness, ranged from 15–65%. In contrast, in the judgment by medical experts ranged from 55–97% for emotions other than calm. The results suggest that “the perception of anger” or “the perception of complex emotions with a mix of anger, confusion, disgust, fear or sadness” is related to the judgment of facial stiffness by medical professionals.

Keywords: *Schizophrenia, Stiffness of facial expression, Average face, Emotion analysis*

1 INTRODUCTION

Evaluation of psychiatric patients' medical conditions requires both interview and observation. However, observation is generally subjective, and formalization of the professional evaluation process would facilitate the development of an evaluation system and be effective in responding to changes in illness. Lack of insight is a problem in schizophrenia, which is a common psychiatric disorder. As it is difficult for these patients to correctly recognize their illness, it is desirable to

develop an objective assessment method. The facial stiffness characteristic of schizophrenia is thought to be related to the patient's reactions to hallucinations and flat affect—the latter is a negative symptom of schizophrenia. Specifically, it is a state of decreased emotional reactivity, including facial stiffness, gestures, and lack of appropriate emotional expression. Gur et al. (2006) reported that flat affect is an important clinical feature of schizophrenia and is associated with prognosis because it adversely affects the course of the illness. To assess flat affect, an objective evaluation of facial stiffness, which is one of its indicators, is required. Although facial stiffness has been evaluated based on the subjective judgment of the evaluator, it is necessary to formalize the judgment. A unique feature of this study is that average faces were created and used as the subject of analysis, instead of using the face images of individual patients. An average face is a representation of the positional information, such as eyes and nose, averaged from multiple face images that constitute the average face. It is mainly used in the field of face research because it is suitable for representing the characteristics of faces as a group. It is generally believed that averaging creates more attractive faces, and factors such as facial symmetry (Jones, DeBruine, & Little, 2007) and improved skin texture are cited as contributing factors. Additionally, an attempt to reveal the cognitive component of facial attractiveness (DeBruine, Jones, Unger, Little, & Feinberg, 2007) by processing average faces using morphing techniques has also been reported. Thus, it is believed that the average face shows the characteristics of the faces as a group. Therefore, it is expected that the average faces of schizophrenia, which is the subject of this study, will also show the characteristics of stiffness. In this study, average faces of schizophrenic patients were graded based on their rated values of facial stiffness, and an attempt was made to identify the emotional characteristics of facial stiffness specific to patients. The structure of this study is shown below. In this study, five experiments were conducted. 1) We conducted an experiment in which 12 medical professionals evaluated the facial stiffness of 147 images of patients, and ranked the images according to the evaluation values to confirm whether there was a common evaluation tendency. 2) To clarify the knowledge of medical professionals when evaluating facial stiffness, we showed groups of images created according to the degree of stiffness and interviewed medical professionals. 3) To extract facial stiffness characteristics, average faces were created for each level of facial stiffness. 4) Emotional analysis of the average faces was performed using an application. 5) Emotional judgment experiments by medical professionals were conducted to clarify the emotional characteristics related to the judgment of facial stiffness by each method. Additionally, the agreement and differences between the two methods were examined. Through these experiments, the following findings were obtained: 1) There are two types of criteria for judging stiffness—the perception of anger and the perception of complex emotions with a mix of multiple emotions. 2) The proportions of mixed emotions differed between applications and medical professionals' evaluation. Judgments by medical professionals showed a mix of anger, confusion, disgust, and fear or sadness, suggesting that the perception of more complex emotions is related to the evaluation of stiffness. This study was conducted with the approval of the Ethics Review Committee on Research with Human Subjects of Waseda University and the Ethical Review Committee of Takatsuki Hospital.

2 EXPERIMENT 1：EXPERIMENTAL EVALUATION OF FACIAL STIFFNESS BY MEDICAL PROFESSIONALS

The evaluation experiment of patients' facial stiffness was used to obtain evaluation values and rank the images, and to confirm whether there was a common evaluation tendency.

2.1 Methods

The participants were 12 (6 males and 6 females) medical professionals with at least 10 years of clinical experience in psychiatry, and included five nurses, two occupational therapists, two clinical psychologists, and three mental health workers. They studied the facial expressions of 16 Japanese schizophrenic patients (14 males and 2 females) admitted to a psychiatric hospital. The stimuli for the facial stiffness evaluation experiment were 147 videos of the patients' faces when they entered the occupational therapy room. The time required to recognize facial expressions is less than 400 ms (Tobimatsu, 2012). Additionally, considering the time required to evaluate facial stiffness, the average length of the stimulus videos was 1.96 seconds (standard deviation 0.39). An evaluation application was originally developed to playback the stimulus videos and obtain evaluation values. Figure 1 shows the evaluation experiment. The evaluation was conducted by moving the knob of the evaluation slider. The evaluation values were obtained by a method in which the position information of the knobs was recorded in a text file. At the same time, the gaze position during the evaluation was recorded at 60 fps by an eye-tracking device (Tobii ProX3-120 Eye Tracker) attached to the bottom of the monitor.

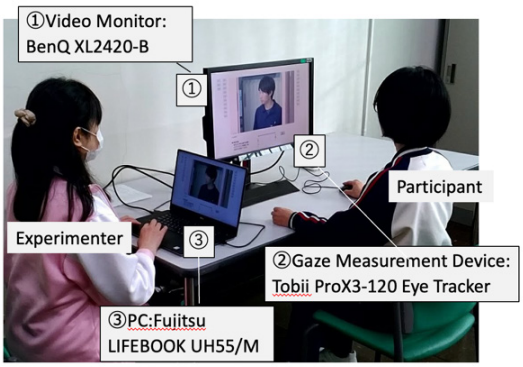

Figure 1. Experiment Environment

2.2 Results

The facial stiffness ratings obtained from the facial stiffness evaluation experiment were standardized. The box-and-whisker plots in Figure 2 show the ranking of video stiffness based on the mean of the standardized scores of the 12 participants, divided into groups of 30 videos each. Figure 2 shows that the variation in scores was lowest in the middle group of stiffness ratings. The variation of the scores was highest in the middle group of the stiffness evaluation, and it was larger in the lower group than in the higher group. Therefore, the F-test was used to compare the variance of each group. The results showed that the following six combinations (1–30 and 61–90, 1–30 and 121–147, 31–60 and 91–120, 31–60 and 121–147, 61–90 and 91–120, and 61–90 and 121–147) had $p<.01$, indicating significant differences in the variability of ratings. In particular, the lowest group (121–120) showed the largest variation in scores, indicating a significant

difference from the top-three groups. In other words, the higher scores showed less variation in the evaluation of stiffness, suggesting the possibility that there is some index that medical professionals have in common regarding the evaluation of stiffness.

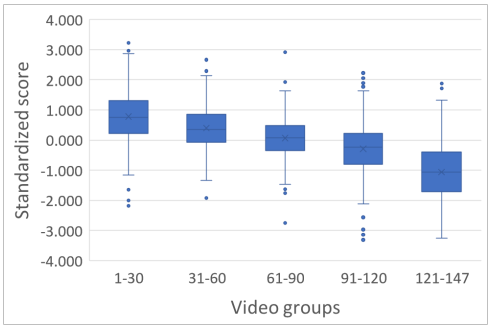

Figure 2. Box-and-whisker diagram of standardized evaluated values of stiffness of facial expression collected per 30 ranks

3 EXPERIMENT 2：INTERVIEWS ABOUT STIFFNESS RATINGS

Interviews were conducted to determine the knowledge of medical professionals in assessing facial stiffness.

3.1 Methods

For each rater, three groups of videos were created from the 147 videos based on the evaluation values. The top-30 videos in terms of stiffness were grouped together as the "group with stiff facial expressions"; the middle-30 videos as the "middle group"; and the bottom-30 videos as the "group with soft facial expressions." After presenting these group videos to the raters, we obtained their comments on the facial characteristics of each group through semi-structured interviews.

3.2 Results

Eight of the 12 participants in the "stiff facial expression" group responded that the videos were characterized by the state of their facial muscles. Six of the 12 respondents indicated that there was no movement of the muscles. Four respondents answered that they had no facial expression, and one said that their muscles were relaxed. Ten of the 12 respondents related to the eyes. Five of them mentioned the eye direction, especially downward gaze. Five of the 12 respondents mentioned eye movement, and four mentioned sharpness of the eyes. These results suggest that two pieces of information: the state of facial muscles and the state of the eyes, are used as cues in the evaluation of facial stiffness.

4 EXPERIMENT 3： ANALYSIS OF GAZE DATA DURING EVALUATION OF STIFFNESS

To extract facial stiffness features, average faces were created for each stage of facial stiffness. Furthermore, we mapped the gaze data during the evaluation of stiffness onto the average face, and clarified the clues for evaluation based on the characteristics of the distribution of gazing points at different levels of stiffness.

4.1 Methods

4.1.1 Creating average faces

The average faces of the patients were created using 127 videos of 14 male patients out of 147 videos. The most front-facing image was extracted from each video. Average Face PRO (version 4.7.3) was used to create the average faces. Average face 1 was constructed using the images ranked from 1 to 30 in terms of stiffness. Average face 2 is composed of images ranked from 11 to 40. Eleven faces were created by shifting the ranking of the stiffness of the images used in the same manner by 10 (Figure 3). The average number of included patients in the average face was 9.7. Images from the Photo AC website were used to create the average faces of the healthy subjects. After cropping the parts of the faces that were straight and facing forward, and after collecting the 20 faces, 14 were randomly selected to create the average face of healthy subjects (Figure 3).

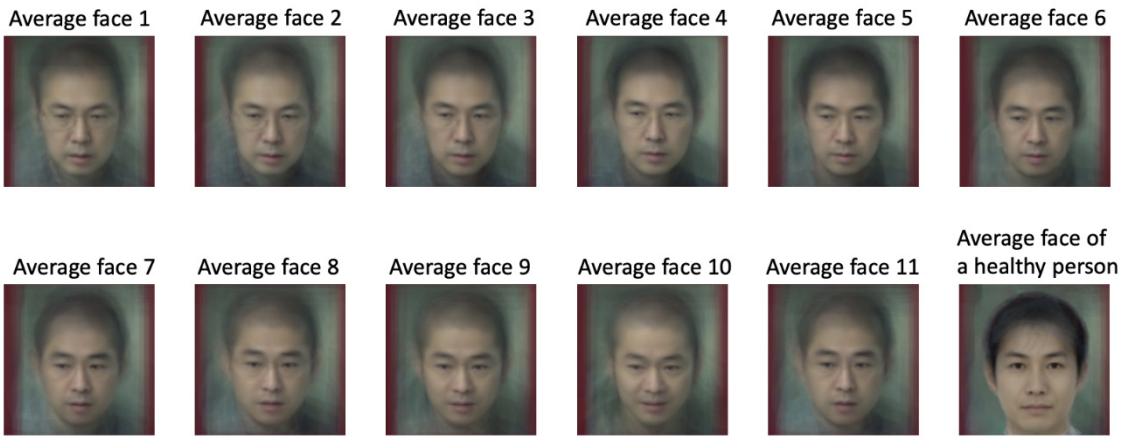

Figure 3. Eleven average faces of schizophrenia by stiffness of facial expression and one average face of a healthy person

4.1.2 Mapping gaze data to average face

For the analysis of the gaze data, the data recorded by the eye-tracking device was output as videos in Tobii Pro Studio. The videos were then divided into frames using a video editing software (Free Video to JPG Converter v.5.0.101 build 201). The position of the gazing point on the patient's face shown in the segmented images was mapped to the relative position of the average face.

4.2 Results

The gaze data of 12 medical professionals during the evaluation of the videos with stiff facial expressions are superimposed in Figure 4. The left figure shows the gaze data of the video with the stiffest facial expression mapped to average face 1. In contrast, the right figure shows the results of mapping the gaze data of the video rated as the seventh stiff expression to average face 1. There were two types of distribution of gazing points in the video group evaluated as having stiff facial expressions: the first was that the gazing points were concentrated around the eyes; the second was that the gazing points were distributed over the entire face. In the videos with the stiffest facial expressions, the gazing points were concentrated around the eyes, inside the eyebrows, and on the upper part of the nose. This is consistent with the "anger" feature of Ekman's six basic emotions (Ekman & Friesen, 1971). In contrast, in the seventh-ranked video, the

gazing points were distributed not only around the eyes but also around the entire face, including the cheeks and mouth. Figure 5 shows a mapping of the gaze data during the evaluation of the videos with the softest facial expressions to the average face 11. In the videos with the softest expressions, the gazing points tended to be distributed to the eyes and their extensions at the corners of the eyes. These results suggest that the range of gazing points during evaluation is narrow when emotions can be judged only around the eyes and nose, such as when "anger" or " happy" emotions are expressed, and that the gazing points are widely distributed when no specific emotion is expressed.

Figure 4. Results of mapping gaze data to the average face of the video images rated as stiffest

Figure 5. Results of mapping gaze data to the average face in the video image rated as soft

5 EXPERIMENT 4：EMOTIONAL ANALYSIS OF AVERAGE FACES BY APPLICATION

Eight types of emotional analyses were conducted using the application with the aim of clarifying the emotional characteristics of the average faces.

5.1 Methods

Emotion analysis was performed on each average face using a web-based application (Amazon Rekognition Face Analysis). Emotion prediction with Amazon Rekognition is based on the physical appearance of the person's face in the image, and each emotion is assigned a confidence score. The eight emotions analyzed were calm, angry, sad, confused, disgusted, surprised, happy, and fear.

5.2 Results

The results of the analysis are shown in Figure 6. Average face 1 on the left-side of the figure is composed of the images ranked from 1 to 30 with the stiffest facial expressions. Moving to the right, ranking of the stiffness of facial expressions of the images decreases, and average face 11 is the face with the softest facial stiffness. The rightmost image is the result of the emotion analysis of the average face of a healthy subjects. Average face 1 had the highest confidence in the emotion of anger at 53.6%. All other average faces had the highest confidence in the emotion of calmness, indicating that the characteristics of the average faces are different from average face 1. Average faces 2–5 and 9 contained about 10–20% anger emotion. Additionally, average faces 2–5 contained about 10–15% sadness. Average faces 3 and 5–10 contained about 5–10% confusion. Average face 11 had the highest degree of certainty of being calm (85.4%) among patients' average faces. The average face of a healthy subjects was 99.5% certain of calmness, and contained few other emotions, which was different from the average faces of the patients. In summary, there were two types of average faces that were judged to have stiff facial expressions: those that showed anger, and those that had a mix of anger, sadness, and mild emotion facial features. The average faces judged to be intermediate to soft tended to show a decrease in the proportion of anger and an increase in the proportion of confusion.

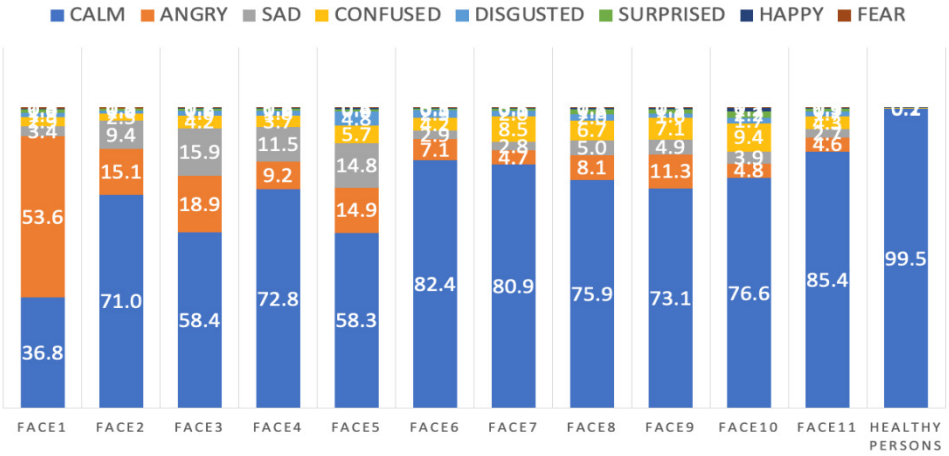

Figure 6. Results of emotion analysis of average faces using the application

6 EXPERIMENT 5：AN EXPERIMENT IN JUDGING THE EMOTIONS OF AVERAGE FACES BY MEDICAL PROFESSIONALS

An emotion judgment experiment was conducted to characterize medical professionals' emotion judgments about average faces, and the results were compared with the results of the application in Experiment 4.

6.1 Methods

The experiment included 33 participants (13 males, 19 females, and one non-binary): four physicians, 16 nurses, four licensed psychologists, five occupational therapists, and four mental health workers employed at a psychiatric hospital. Their ages ranged from 25 to 60 years (mean 40.0 years, standard deviation 9.38). The program for the emotion judgment experiment was

created using a questionnaire tool (qualtrics). Emotion judgments were made by selecting up to three possible emotions from the displayed average face images in order of likelihood. After one practice session using the practice images, the participants responded to the 12 average faces that were displayed at random. Eight emotions were selected—same as in the application of Experiment 3. The second and third place selections could also be made by selecting "not applicable."

6.2 Results

Figure 7 shows the percentages of emotions that the raters selected as first place. Average face 1 had the highest percentage of anger, judged at 36.4%, followed by disgust at 30.3%. Average faces 2-6, 8, and 11 had the highest percentage of confusion, ranging from 30 to 50%. In particular, half of the raters (51.5%) were confused about average face 6. The average faces 1 to 6 were rated as disgusting by 12 to 30% of the raters. Average face 7 was rated as mild by the highest percentage of raters (45.5%). For average face 10, 69.7% of the patients judged it as joyful. The average face of healthy subjects was judged as moderate by 78.8% of the total respondents, which was different from the tendency of the average faces of the patients. The application and medical professionals shared the same judgments about the average face 1. The highest percentage of both judgments was anger. In the application's analysis of emotions, the highest confidence was found in the eight emotions, ranging from 60–85%, with the exception of average face 1, which had the highest confidence in the emotion of calm for all the average faces. In contrast, in the judgment by medical professionals, the highest rate of selecting "calm" was 45% for average face 7, followed by 30% for average faces 4 and 9, indicating that emotions other than "calm" were selected in a high percentage of cases. Additionally, a significant difference was that confusion was selected more frequently by the human judges, with more than 30% of the judges selecting this emotion for average faces 2–6, 8, and 11. In the application, the facial features rated as stiffness were a mix of anger, sadness, and calm, but in the judgment by medical professionals, a mix of anger, confusion, disgust, and fear, and a mix of anger, confusion, disgust, and sadness were indicated, suggesting that a more complex perception of emotion is associated

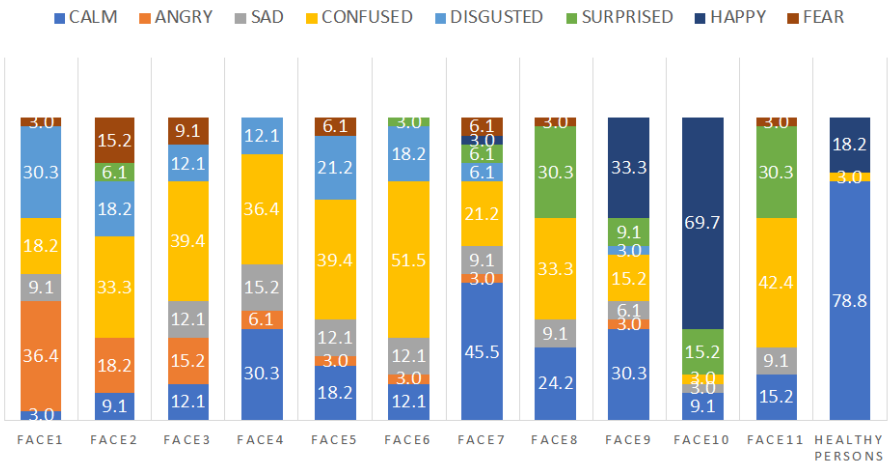

Figure 7. Results of emotion judgment of average faces by medical professionals

with the evaluation of stiffness. The results suggest that a more complex perception of emotions is associated with the evaluation of stiffness.

7 DISCUSSION

With the aim of developing an objective evaluation method for schizophrenia, this study attempted to clarify facial stiffness in patients in terms of emotional characteristics. After interviews with medical professionals, two characteristics were identified in the evaluation of stiffness: state of facial muscles and sharpness of eye contact. This was supported by the fact that the mapping of gaze data showed that the average face with stiffness was distributed to the area that expressed anger emotion. Additionally, we conducted an emotion analysis application and emotion judgments by medical professionals on the average faces, and confirmed the consistency and differences between the two judgments. As a result, the only commonality extracted between the two was that the detection of anger emotion for the stiffest average face 1 was related to facial stiffness and that the average face of healthy subjects was judged as calm. In the emotion judgments for many of the other average faces, the composition of emotions differed significantly between the applications and the medical professionals. Faces rated as stiff by the application showed a mix of facial features of anger, sadness, and calm emotions, whereas medical professionals' judgments showed a mix of anger, confusion, disgust, and fear, as well as anger, confusion, disgust, and sadness, suggesting that a more complex perception of emotions was associated with the stiffness ratings. These results suggest that facial stiffness was evaluated using two criteria. Criterion 1 was the perception of anger and criterion 2 was the perception of complex emotions. A possible reason for the difference between the two groups could be that the facial features of schizophrenics are different from those of healthy people, which had a significant impact on the differences between the two groups. Most emotion analysis applications are based on Ekman et al.'s (2002) Facial Action Coding System (FACS), and emotion judgments are mainly based on positional information of facial parts. Therefore, it is considered that both judgments were consistent for the average face of a healthy person. However, it can be inferred that the application found it difficult to judge the emotions of patients' faces because of the difficulty in expressing typical emotions that can be handled by FACS because of flat affect and parkinsonism. Additionally, confusion was the most frequently selected feature of emotion judgments by medical professionals for many average faces. This may be due to the possibility that patients' psychological states are influenced by psychological symptoms such as auditory hallucinations, which may be harmful to them. The criteria used by medical professionals to judge facial stiffness were a mix of anger, embarrassment, disgust, and fear, or a mix of anger, embarrassment, disgust, and sadness, which were difficult to distinguish from specific emotions. However, it is possible that this ambiguity in facial expressions is related to the Praecox Feeling proposed by Rümke, and this will require further study. Although the accuracy of emotion analysis tools has been improved by the use of big data and deep learning technologies in recent years, there was a difference between the judgment of average faces of schizophrenia by the application and that by medical professionals, suggesting that a system for detecting facial stiffness in schizophrenia requires a different algorithm than conventional emotion recognition.

8 CONCLUSION AND FUTURE WORKS

In this study, we examined the emotional characteristics of facial stiffness—one of the indicators for estimating flat affect in patients—using average faces. There were two types of criteria for judging stiffness—Criterion 1 “the perception of anger” and criterion 2 “the perception of complex emotions”. The extraction of the emotion of anger with a stiff face was common to both the application and medical professionals. However, the proportions of mixed emotions differed between applications and medical professionals' evaluation. Judgments by medical professionals showed a mix of anger, confusion, disgust, and fear or sadness, suggesting that the perception of more complex emotions is related to the evaluation of stiffness. Future work is needed to develop a new sophisticated emotional evaluation system for schizophrenia based on Criterion 2, which is a characteristic of a patient's facial stiffness.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.com) for English language editing.

REFERENCES

DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating Averageness and Attractiveness: Attractive Faces Are Not. Journal of Experimental Psychology: Human Perception and Performance, 33(6), 1420-1430.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of Personality and Social Psycology, 17(2), 124-129.

Ekman, P., Friesen, W. V, & Hager, J.C. (2002). Facial Action Coding System. Retrieved from http://face-and-emotion.com/dataface/facs/manual/TitlePage.html

Grube, M. (2006). Towards an Empirically Based Validation of Intuitive Diagnostic: Rümke’s ‘Praecox Feeling’ across the Schizophrenia Spectrum: Preliminary Results. Psychopathology, 39(5), 209-17.

Gur, R. E., Kohler, C. G., Ragland, J. D., Siegel, S. J., Lesko, K., Bilker, W. B., & Gur, R. C. (2006). Flat Affect in Schizophrenia: Relation to Emotion Processing and Neurocognitive Measures. Schizophrenia Bulletin, 32(2), 279-287.

Jones, B. C., DeBruine, L. M., & Little, A. C. (2007). The role of symmetry in attraction to average faces. Perception & Psychophysics, 69(8), 1273-1277.

Moskalewicz, M., Kordel, P., Brejwo, A., Schwartz, M. A., & Goze, T. (2021). Psychiatrists Report Praecox Feeling and Find It Reliable. A Cross-Cultural Comparison. Frontiers in Psychiatry | www.frontiersin.org, 12(642322).

Tobimatsu, S. (2012). Neural Mechanisms of Face Recognition: An Event-related Potential Study. BRAIN and NERVE, 64(7), 717-726.