Emotional expression in psychiatric conditions: New technology for clinicians

Karol Grabowski, MD, PhD,1 Agnieszka Rynkiewicz, MD, MAT, PhD2,3 Amandine Lassalle, PhD4,5 Simon Baron-Cohen, MPhil, PhD5 Björn Schuller, PhD,6 Nicholas Cummins, PhD,6 Alice Baird, MFA,7 Justyna Podgórska-Bednarz, PhD,6,8 Agata Pieniążek, MA6,10,11 and Izabela Łucka, MD, PhD12

Aim: Emotional expressions are one of the most widely studied topics in neuroscience, from both clinical and nonclinical perspectives. Atypical emotional expressions are seen in various psychiatric conditions, including schizophrenia, depression, and autism spectrum conditions. Understanding the basics of emotional expressions and recognition can be crucial for diagnostic and therapeutic procedures. Emotions can be expressed in the face, gesture, posture, voice, and behavior and affect physiological parameters, such as the heart rate or body temperature. With modern technology, clinicians can use a variety of tools ranging from sophisticated laboratory equipment to smartphones and web cameras. The aim of this paper is to review the currently used tools using modern technology and discuss their usefulness as well as possible future directions in emotional expression research and treatment strategies.

Methods: The authors conducted a literature review in the PubMed, EBSCO, and SCOPUS databases, using the following key words: ‘emotions,’ ‘emotional expression,’ ‘affective computing,’ and ‘autism.’ The most relevant and up-to-date publications were identified and discussed. Search results were supplemented by the authors’ own research in the field of emotional expression.

Results: We present a critical review of the currently available technical diagnostic and therapeutic methods. The most important studies are summarized in a table.

Conclusion: Most of the currently available methods have not been adequately validated in clinical settings. They may be a great help in everyday practice; however, they need further testing. Future directions in this field include more virtual-reality-based and interactive interventions, as well as development and improvement of humanoid robots.

Keywords: affective computing, autism, emotions, expressed emotion, nonverbal communication.


Emotions are a universal aspect of human behavior. They reveal what we are feeling and the ways in which we might act as a result of those feelings. For instance, somebody exhibiting anger might be prone to entering a fight. Emotions also inform decision-making in a critical way. Indeed, given that an angry person is likely to fight, an observer should prepare to either fight back or to escape in order to survive. Our emotional state is apparent in our facial expression, tone of voice, gestures, and physiological parameters. The purpose of this article is to investigate the physical characteristics of emotion expressiveness across these modalities in several psychiatric conditions, and to examine how new technology can support clinical practice. A summary of the most important methods and studies on emotion expression assessment methods is presented in Table 1.1,42 The idea that emotions are innate, unlearned responses associated with a complex set of movements affecting the face, body, and speech traces back to Darwin,43 who suggested that the way in which we express emotions has evolved from emotional expression in other animals.

Different Aspects of Emotional Expression

Facial expression

The face has been the most studied of the emotion expression modalities. Darwin’s ideas about a set of universal emotions were refined by Ekman,44,45 who presented work indicating that people from New Guinea had no difficulty recognizing facial expressions in Westerners and vice versa, even though the two cultures had had no contact with each other. In addition, it has been shown that the expression of basic facial emotions is not affected by congenital blindness.46–48
<table>
<thead>
<tr>
<th>Reference</th>
<th>Method of emotion expression</th>
<th>Technology used</th>
<th>Potential target population</th>
<th>Validation study</th>
<th>Sample size</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Du et al. (2014)</td>
<td>Facial expression</td>
<td>Video analysis</td>
<td>No specific target population</td>
<td>No</td>
<td>230 subjects</td>
<td>Identification of compound emotion. Accuracy up to 76.91% in emotion identification.</td>
</tr>
<tr>
<td>Chickerur and Joshi (2015)</td>
<td>Facial expression</td>
<td>3-D video analysis</td>
<td>Students, general population</td>
<td>No</td>
<td>95 healthy students</td>
<td>Creating a standardized 3-D database of facial expressions of emotions. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Saneiro et al. (2014)</td>
<td>Facial expression, body movement</td>
<td>Video analysis</td>
<td>No specific target population</td>
<td>No</td>
<td>75 healthy subjects</td>
<td>Automated method using Kinect camera, integrating 2-D and 3-D data. Allows capturing dynamic emotions evoked by a specific task. No data on statistical power of the results. Study designed to improve teaching based on students' reactions.</td>
</tr>
<tr>
<td>Agarwal et al. (2016)</td>
<td>Facial expression</td>
<td>Video analysis</td>
<td>ASC, schizophrenia, depression</td>
<td>No</td>
<td>54 healthy subjects</td>
<td>Emotion recognition with accuracy between 40% for anger and 100% for disgust, joy, sorrow, and surprise. Study designed to create a system to recognize emotions from streaming videos. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Bartlett et al. (2006)</td>
<td>Facial expression</td>
<td>Video analysis based on FACS</td>
<td>ASC, schizophrenia</td>
<td>No</td>
<td>119 healthy subjects</td>
<td>Automated system showing correlation of 0.83 between automatic scoring and human scorer. Designed to capture spontaneous facial expression. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Littlewort et al. (2011)</td>
<td>Facial expression</td>
<td>Video analysis based on FACS</td>
<td>ASC, schizophrenia</td>
<td>No</td>
<td>26 healthy subjects and dataset of 100 healthy subjects</td>
<td>90.1% recognition performance for facial actions and 80% accuracy for spontaneous facial expression. Study aimed at creating a software for automatic real-time face expression recognition.</td>
</tr>
<tr>
<td>Huang et al. (2012)</td>
<td>Facial expression and EEG</td>
<td>Video analysis, EEG analysis</td>
<td>ASC, schizophrenia</td>
<td>Yes, in non-clinical setting</td>
<td>100 healthy subjects (9% men)</td>
<td>Real-time analysis, validation of Chinese face images database. Sensitivity index for all emotions was 0.94. No data on statistical power of the results.</td>
</tr>
</tbody>
</table>
Table 1. (Continued)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Method of emotion expression</th>
<th>Technology used</th>
<th>Potential target population</th>
<th>Validation study</th>
<th>Sample size</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golan et al. (2006)&lt;sup&gt;9&lt;/sup&gt;</td>
<td>Facial expression, voice analysis</td>
<td>Video and audio analysis</td>
<td>Adults with ASC</td>
<td>Yes</td>
<td>21 adults with Asperger’s syndrome (15 males, 6 females) and 17 healthy adults (12 males, 5 females)</td>
<td>CAM test battery assessing complex emotion recognition. Shows strong correlation with standard tests used in ASC.</td>
</tr>
<tr>
<td>Golan et al. (2015)&lt;sup&gt;10&lt;/sup&gt;</td>
<td>Facial expression, voice analysis</td>
<td>Video and audio analysis</td>
<td>Children with ASC</td>
<td>Yes</td>
<td>30 ASC children and 25 healthy controls</td>
<td>CAM test battery assessing complex emotions, version for children</td>
</tr>
<tr>
<td>Ghimire et al. (2013)&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Facial expression</td>
<td>Video analysis</td>
<td>ASC, depression, schizophrenia</td>
<td>No</td>
<td>Video database of 123 healthy subjects</td>
<td>Up to 97.35% accuracy in emotion recognition using tested automated method. Not tested clinically. Study aimed at creating emotion recognition in still images.</td>
</tr>
<tr>
<td>Grzadzinski et al. (2016)&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Facial expression, gestures, behavior analysis</td>
<td>Video analysis</td>
<td>Children with ASC</td>
<td>Yes</td>
<td>56 children</td>
<td>Tool to measure social communication change, requires rater training</td>
</tr>
<tr>
<td>Lord et al. (2012)&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Facial expression, gestures, analysis, behavior analysis</td>
<td>Video analysis, real time observation and interaction</td>
<td>Children with ASC</td>
<td>Yes</td>
<td>98 ASC children in original validation study, 1574 and 1282 children in subsequent validation studies</td>
<td>Gold standard in ASC diagnosis with 89.6% sensitivity and 72.2% specificity for ASC diagnosis.</td>
</tr>
<tr>
<td>Kohler et al. (2008)&lt;sup&gt;14&lt;/sup&gt;</td>
<td>Facial expression</td>
<td>Video analysis with FACS</td>
<td>Adults with schizophrenia</td>
<td>Yes</td>
<td>12 schizophrenic patients and 12 healthy controls</td>
<td>Study designed for assessment of severity of flattened and inappropriate affect. Proved to be clinically useful in this purpose. No data on statistical power of the results (small sample size).</td>
</tr>
<tr>
<td>Hamm et al. (2014)&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Facial expression</td>
<td>Video analysis with FACS</td>
<td>Adults with schizophrenia</td>
<td>Yes</td>
<td>28 schizophrenic patients, 26 healthy controls</td>
<td>Automated FACS, gives more repeatable results, good in assessment of affect flattening but does not measure valence. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Alvino et al. (2007)&lt;sup&gt;16&lt;/sup&gt;</td>
<td>Facial expression</td>
<td>Photo analysis, regional volumetric differences</td>
<td>Adults with schizophrenia</td>
<td>Yes</td>
<td>32 actors as a model database. Validated with 12 schizophrenic patients and 12 healthy controls</td>
<td>Efficient in detecting changes in emotion expression between healthy people and schizophrenia patients. Significant correlation with SANS scale scores.</td>
</tr>
<tr>
<td>Wang et al. (2008)&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Facial expression</td>
<td>Video analysis</td>
<td>Adults with schizophrenia and ASC</td>
<td>Yes, in non-clinical setting</td>
<td>3 Asperger’s syndrome patients, 3 schizophrenia patients, 3 healthy controls</td>
<td>Automated video analysis using face and landmarks detection. No data on statistical power of the results due to small sample size. Requires further clinical validation.</td>
</tr>
<tr>
<td>Reference</td>
<td>Method of emotional expression</td>
<td>Technology used</td>
<td>Potential target population</td>
<td>Validation study</td>
<td>Sample size</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------------------------</td>
<td>------------------------------------------------------</td>
<td>----------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Hamm et al. (2011)</td>
<td>Facial expression</td>
<td>Video analysis with FACS</td>
<td>Adults with schizophrenia</td>
<td>Yes</td>
<td>28 patients and 26 healthy controls</td>
<td>Automatic analysis validated by trained observer/scorer. Show differences between patients and healthy controls. Designed for more automated assessment of blunted/ inappropriate affect.</td>
</tr>
<tr>
<td>Kring and Sloan (1995)</td>
<td>Facial expression</td>
<td>Video analysis with FACES</td>
<td>ASC, schizophrenia, depression, PTSD</td>
<td>Yes</td>
<td>196 students, 56 schizophrenia patients, 32 healthy adults</td>
<td>Does not identify specific emotions but their frequency, intensity, and duration. Reference paper summarizes results from different studies.</td>
</tr>
<tr>
<td>Scherer et al. (2015)</td>
<td>Voice</td>
<td>Audio analysis</td>
<td>No specific target population</td>
<td>No</td>
<td>20 healthy people</td>
<td>Identifies acoustic features of emotions in speaking and singing voice. Not tested in clinical setting. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Rutherford et al. (2002)</td>
<td>Voice</td>
<td>Audio analysis</td>
<td>Adults with ASC</td>
<td>Yes</td>
<td>19 adults with ASC (17 males, 2 females)</td>
<td>Test results may help to distinguish ASC adults from healthy controls.</td>
</tr>
<tr>
<td>Meeren et al. (2005)</td>
<td>Posture and gestures, facial expression, EEG</td>
<td>Video analysis, EEG signal analysis-evoked potentials</td>
<td>ASC, depression, schizophrenia</td>
<td>No</td>
<td>12 healthy adults</td>
<td>Method for specific research, not tested clinically. Shows integration of emotional processing. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Dael et al. (2012)</td>
<td>Body posture and gestures</td>
<td>Video analysis</td>
<td>No specific target population</td>
<td>No</td>
<td>10 actors (5 male, 5 female)</td>
<td>Requires trained scorer. Method can be incorporated in other systems. Not tested clinically.</td>
</tr>
<tr>
<td>Piana et al. (2014)</td>
<td>Gestures</td>
<td>Video analysis, made with Kinect</td>
<td>ASC patients</td>
<td>No</td>
<td>Healthy subjects, actors (4 females, 8 males)</td>
<td>Emotion recognition accuracy 61.3%. System included games to train emotion expression with gestures. Background for further clinical testing.</td>
</tr>
<tr>
<td>Castellano et al. (2007)</td>
<td>Gestures</td>
<td>Video analysis</td>
<td>No specific target population</td>
<td>No</td>
<td>Healthy subjects: 6 males, 4 females</td>
<td>Not suitable for distinguishing specific emotions. Possible part of complex systems. No data on statistical power of the results (small sample size).</td>
</tr>
<tr>
<td>Reference</td>
<td>Method of emotion expression</td>
<td>Technology used</td>
<td>Potential target population</td>
<td>Validation study</td>
<td>Sample size</td>
<td>Comments</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------------</td>
<td>----------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Choppin (2000)</td>
<td>EEG</td>
<td>EEG</td>
<td>Intellectually disabled</td>
<td>No</td>
<td>22 subjects (13 male, 17 female)</td>
<td>64% accuracy in emotion recognition in valence–arousal spectrum. Study designed to create EEG-based human interface. Early development phase of this method.</td>
</tr>
<tr>
<td>Zhang et al. (2016)</td>
<td>EEG</td>
<td>EEG signal analysis (2 channel)</td>
<td>ASC, schizophrenia, amyotrophic lateral sclerosis, depression</td>
<td>No</td>
<td>32 subjects</td>
<td>Identifies valence and arousal with accuracy up to 94.98%. Easy to use, with only 2 EEG channels, which is possible with data extraction method created by the authors. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Chai et al. (2016)</td>
<td>EEG</td>
<td>EEG subspace alignment auto encoder</td>
<td>No specific target population</td>
<td>Comparison with other methods of EEG signal processing</td>
<td>15 subjects</td>
<td>77.88% accuracy in valence identification. Uses 64 EEG channels, thus difficult to implement in clinical setting.</td>
</tr>
<tr>
<td>Aydin et al. (2016)</td>
<td>EEG</td>
<td>Wavelet-based feature extraction from EEG signal</td>
<td>No specific target population</td>
<td>No</td>
<td>32 subjects</td>
<td>Only for valence–arousal spectrum, does not distinguish specific emotions. Shows use of LabVIEW software. Not tested clinically. No data on statistical power of the results.</td>
</tr>
<tr>
<td>Jirayucharoensak et al. (2014)</td>
<td>EEG</td>
<td>Deep learning networks</td>
<td>No specific target population</td>
<td>No</td>
<td>32 subjects</td>
<td>55.07% accuracy for valence and 52.56% for arousal classification. Study designed to create new method of EEG signal extraction and processing. Uses 32 EEG channels, thus difficult to implement in clinical setting.</td>
</tr>
<tr>
<td>Tuck et al. (2016)</td>
<td>HRV</td>
<td>HRV measured with watch HR monitor</td>
<td>ASC, schizophrenia</td>
<td>No</td>
<td>80 females</td>
<td>Designed to establish a correlation between resting HRV and emotion expression. Statistically significant for overall expressive skills and anger. Not tested clinically.</td>
</tr>
<tr>
<td>Reference</td>
<td>Method of emotion expression</td>
<td>Technology used</td>
<td>Potential target population</td>
<td>Validation study</td>
<td>Sample size</td>
<td>Comments</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------</td>
<td>-----------------------------------------</td>
<td>-----------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Wolf et al. (2006)34</td>
<td>EMG</td>
<td>EMG of 3 facial muscles</td>
<td>Schizophrenia</td>
<td>Yes</td>
<td>32 schizophrenic patients, 21 healthy controls</td>
<td>Study designed to establish a new method of EMG and establish its correlation with psychopathology. It measures valence and arousal spectrum, shows a good correlation with clinical assessment scales.</td>
</tr>
<tr>
<td>Valenza et al. (2016)35</td>
<td>EEG and HR</td>
<td>EEG signal analysis, HR</td>
<td>ASC, schizophrenia</td>
<td>No</td>
<td>22 healthy subjects</td>
<td>Exploring brain–heart dynamics modulated by emotions. It measures valence–arousal spectrum.</td>
</tr>
<tr>
<td>Huang et al. (2016)36</td>
<td>EEG and facial expression</td>
<td>Video analysis, EEG analysis</td>
<td>ASC, schizophrenia</td>
<td>No</td>
<td>30 healthy subjects (17 females, 13 males)</td>
<td>Study designed to test a new method of automated emotion recognition. It shows 66.28% accuracy for valence and 63.22% for arousal identification. Clinical validation is lacking.</td>
</tr>
<tr>
<td>Vos et al. (2012)37 and (2013)38</td>
<td>Multimodal system</td>
<td>Behavioral analysis, GSR, SKT, HP</td>
<td>Intellectually disabled patients</td>
<td>Yes</td>
<td>27 intellectually disabled patients</td>
<td>Designed specifically for intellectually disabled patients. It measures valence and arousal spectrum and cannot be used as a single method.</td>
</tr>
<tr>
<td>Jackson et al. (2015)39</td>
<td>Multimodal and interactive</td>
<td>Video analysis based on FACS, use of virtual avatars, HR, RR, GSR, eye movements</td>
<td>ASC, schizophrenia</td>
<td>Yes (in non-clinical setting)</td>
<td>19 healthy subjects (10 females, 9 males)</td>
<td>Platform to study and train empathy. Clinical validation is lacking.</td>
</tr>
<tr>
<td>Verma (2014)40</td>
<td>Multimodal system</td>
<td>EEG, GSR, BVP, respiration pattern, SKT, EMG, electrooculogram, video analysis</td>
<td>No specific population</td>
<td>No</td>
<td>32 healthy subjects</td>
<td>Emotion identification accuracy between 57.74% and 85.46% in valence–arousal spectrum. Study designed to find signals enabling to predict emotions based on physiological signals. Difficult to implement in clinical practice due to large channel number.</td>
</tr>
<tr>
<td>Bekele et al. (2013)41</td>
<td>Virtual reality-based multimodal system</td>
<td>Eye tracking, ECG, PPG, SKT, GSR</td>
<td>ASC</td>
<td>Yes</td>
<td>10 ASC patients and 10 healthy controls</td>
<td>Study designed to create emotional expression method with a possibility to monitor emotional reaction of a user. Method effective in research, not tested in clinical purposes. Limited statistical power due to small sample size.</td>
</tr>
</tbody>
</table>
Table 1. (Continued)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Method of emotion expression</th>
<th>Technology used</th>
<th>Potential target population</th>
<th>Validation study</th>
<th>Sample size</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown et al.</td>
<td>Multimodal system</td>
<td>FACS, gaze, eye tracking, sound detection</td>
<td>ASC</td>
<td>No</td>
<td>None</td>
<td>Interactive mobile application for therapy augmentation. Easy to use, low hardware requirements. Proof-of-concept study, not tested in clinical setting.</td>
</tr>
</tbody>
</table>

ASC, autism spectrum conditions; BP, blood pressure; BVP, blood volume pressure; CAM, Cambridge Mindreading Face-Voice Battery; ECG, electrocardiogram; EEG, electroencephalography; EMG, electromyography; FACES, Facial Expression Coding System; FACS, Facial Action Coding System; GSR, galvanic skin response; HP, time in milliseconds between two consecutive heartbeats; HR, heart rate; HRV, heart rate variability; PPG, pulse plethysmograph; RR, respiration rate; SKT, skin temperature.

...facial expressions. In addition, the FACS is typically applied to the emotions: calm (sadness, disgust, neutral), medium arousal (joy and happiness, amusement), and excited/activated (surprise, fear, anger, anxiety). In the valence spectrum, we can also divide emotions into three classes: unpleasant (fear, anger, disgust, sadness, anxiety), neutral (surprise, neutral), and pleasant (joy and happiness, amusement).

Apart from the FACS and FACES, video analysis of facial emotion expression can also use vector analysis, in which feature vectors are built by a change of position of different points (landmarks) in specific face regions. Sets of vectors characteristic for specific emotions can be used as a comparator and allow us to achieve 97.35% accuracy in recognizing emotions.

When analyzing facial emotion expression, we have to acknowledge confounding factors such as: (i) the emotional context (previous emotional state); (ii) the temperament of the person expressing the emotion; and (iii) individual expression features, which may include type and number of repetitions of specific movements (e.g., blinking three times with left eye when nervous). A further problem with emotion expression is cognitive processes omnipresent during these expressions. The content of thinking may be emotionally different than a stimulus triggering the emotional reaction. This may affect the expression of complex emotions, thereby acting as a confounding factor in emotion analysis.

Practical and clinical applications of automatic facial emotion recognition have been widely tested and validated. For example, schizophrenia is another condition with emotion expression impairment. Emotional expression dysfunction is connected with emotional blunting and is considered a core negative symptom of schizophrenia. Automated video analysis of emotion expression can be used as a screening tool or a measure to assess the severity of symptoms or...
progression connected with therapeutic effects of medications. The associated changes in emotional facial expressions may be subtle and difficult to notice for an observer, especially when they do not have sufficient clinical experience or are unfamiliar with a patient. Automatic recognition of expressed emotions can also serve as a communication aid that could be used for psychotherapy enhancement.

Voice
Apart from facial expressions, emotions are expressed in the voice and this expression can be both verbal and non-verbal. For example, in Sauter et al., European English-speaking participants were presented with recordings of sounds of nonverbal vocalizations from natives of an isolated northern Namibian village and participants had no difficulty matching the emotional vocalization to a situation (anger, disgust, fear, sadness, surprise, or amusement). This suggests that intonation of voice can also be categorized into basic and complex emotions. Research presented by Scherer and colleagues indicates that emotions can be expressed at the same level in a singing or a speaking voice, regardless of the meaning of the words. Although we can identify acoustic features characteristic for specific emotions, such as sadness or anger, automatic emotion recognition from the voice is mostly limited to the valence/arousal spectrum or emotional intentions. This may be sufficient for clinical application of voice analysis because according to Douglas-Cowie and colleagues the distinction between specific basic emotions is not relevant for emotional prosody (i.e., sound features of speech reflecting the emotional state of the speaker). They argue that the cultural influences and the display rules that color our emotional lives affect emotional prosody more than they affect facial expressions. They speculate that this is due to the cultural nature of speech, suggesting that cultural influences and display rules could even affect the prosody associated with basic emotions. They also argue that, in real life, we express rather subtle emotions with our voices and usually not the ‘pure’ basic emotions that are studied in emotion research.

The sound features of speech responsible for emotional prosody are: pitch, accent, energy, loudness, and speaking rate. These characteristics are crucial to determine, given the importance of emotional prosody in social interaction. For example, stroke patients who cannot identify prosody that accompanies speech, or who cannot generate prosody, have been found to have severe interpersonal difficulties.

There are systems and applications using voice as a single source of emotion recognition, such as ComParE and GeMAPS, which can be clinically useful. Possible clinical implications may include diagnosis and treatment of ASC and schizophrenia, patients with intellectual disabilities, and stroke patients. Voice analysis may help teach patients prosody and improve social interactions. Sound analysis does not require verbal communications, which makes it useful in nonverbal patients (children, ASC, dementia patients). One of the limitations of these methods is sensitivity for external sounds and noise and that is why it may require standardized conditions. Another limitation is that voice analysis is better in the valence and arousal spectrum than in identifying specific emotions. However, emotion expression and recognition in the voice are more valuable as elements of many multimodal systems that we will describe below.

Gestures and movement
In addition to being conveyed by facial expressions, speech acoustics and prosody, emotions are also reflected by body language. For example, the sight of clenched fists can allow the detection of anger in another, as the perception of others’ facial expressions is facilitated when we have access to their body postures. In addition, body postures can activate the brain region involved in emotion processing to the same extent as facial expressions. For instance, pictures of a fearful body activate the amygdala as much as pictures of a fearful face. Despite its importance in affect communication and its potential for conveying emotions (substantially more muscles can move in the body than in the face), there has been less research on body language and emotions than on facial and vocal expressions of emotion. This has been attributed to the absence of a precise measurement system similar to the FACS coding system for body postures. However, recently the Body Action and Posture (BAP) system, which includes a list 141 body action units, was developed to describe the principle characteristics of each emotion, expressed in terms of body gestures. Emotion expression can be assessed based on expressivity, dynamics, and direction of movement. The identification of specific gestures characteristic for basic or complex emotions can be difficult, as it has significant interpersonal variability. Implementing references (i.e., 3-D models of emotional gestures) can help to overcome these issues and has been shown to achieve 61.3% accuracy in emotion recognition in the valence and arousal spectrum. Gestures can be recorded and measured by simple systems; however, they are better in describing the valence and arousal spectrum rather than identifying specific emotions. That is why gesture and movement analysis is a part of the complex system of emotion analysis rather than a single diagnostic tool.

Physiological parameters
Each emotional state is connected with physiological changes consistent with meaning and the role of the specific emotion. The most classical example of this is fear that triggers arousal, which prepares the organism for the ‘fight-or-flight’ reaction. Since these physiological changes involve the whole organism, there are many parameters that can be measured and connected with specific emotions. One of the useful physiological parameters is the electric activity of the cerebral cortex measured with electroencephalography (EEG; for a review, see Alarcão and Fonseca). It is debatable whether specific emotions can be robustly identified with this measure, but analysis of EEG spectral power can provide information on valence and arousal as well as on the strength of expressed emotions. A common issue when working with EEG is that there is no gold standard in the number of channels and their selection for analysis, nor in the data extraction and analysis methods. Practical issues, such as placing the electrodes, removing artifacts and noise (e.g., caused by muscle tension or movement), and real-time recording and analysis may also limit the use of EEG-based emotion expression in everyday practice.

Physiological parameters, like galvanic skin reaction, muscle tension heart rate, and heart rate variability, are also used to measure arousal and valence. Physiological parameter analysis can also help to predict emotional expression. People with higher resting vagally mediated heart rate variability are able to deliberately express anger and interest. The single physiological parameter that may also be useful is respiration, which allows 73.06% accuracy for valence and 80.78% for arousal identification. This method has not been clinically validated so far but its simplicity and potential low costs make it promising.

Another physiological parameter connected with emotional expression is body temperature. As a single feature, it is generally considered not sufficient for emotion recognition, even in the arousal and valence spectrum. However, studies conducted on macaque monkeys show that during expression of negative emotional states, the skin temperature in the nasal area decreases. In studies in people with severe and profound intellectual disabilities, changes in skin temperature during expressions of emotions were detected: during the first 6 s of low-intensity negative emotions expression, skin temperature was higher compared to during low-intensity positive emotions.

Measuring physiological parameters is valuable due to its simplicity and easy practical application, especially when assessment of the valence–arousal spectrum is enough. A good example of such use is biofeedback based on heart rate variability or EEG. Analysis of physiological parameters in emotion expression assessment can have great clinical value in populations with limited verbal and non-verbal communication, like amyotrophic lateral sclerosis or dementia patients. It can also be used in forensic psychiatry, especially in combination with other measures.
Multimodal systems
Considering the low specificity and accuracy of a single physiological parameter in emotion expression, development of multi-modal systems seems the obvious direction. At the moment, a range of different systems have been proposed in the relevant literature, combining different sets of parameters. For instance, combining gesture and facial expression is natural as they can be measured with similar video equipment. Combining facial expression with EEG increases accuracy in emotion detection, especially in arousal, where it can be more accurate than a human observer. This combination has also been shown to improve valence assessment. It is a valuable method even acknowledging that the artifacts from facial muscles may contaminate the EEG signal and affect the results. Multimodal systems based on physiological parameters can also complement or validate behavioral observations in patients with compromised emotional expression (e.g., patients with intellectual disabilities).

The combination of facial expressions, gestures, and social communication analysis together with implementation of specific diagnostic algorithms can create new possibilities in ASC assessment. An example of such a diagnostic tool is the Brief Observation of Social Communication Change, which is based on the Autism Diagnostic Observation Schedule, Second Edition and measures subtle changes in social communication behavior. It captures even minor changes in behavior and thus reduces subjective observer bias. Such an accurate assessment is possible thanks to the use of audiovisual systems capable of recording and subsequent replaying of diagnostic sessions. The recording allows researchers to save every gesture or emotion and to evaluate it both qualitatively and quantitatively, which is almost impossible in the case of direct observation and simultaneous evaluation of a patient. Moreover, the recording of diagnostic sessions enables the verification of agreement in case of two independent evaluators, and provides reliable training for clinicians. Incorporating new technology in diagnostic and therapeutic procedures not only can improve the accuracy of diagnosis and the efficiency of therapy, but can also help in redefining diagnostic criteria.

Modern technology can also be of considerable value to diagnosticians in terms of exploring the way ASC is expressed differently in girls than in boys, and for further investigation of the camouflage in females, which may pose a risk of under-diagnosis or misdiagnosis for this population. It is important to highlight that systems using different sets of physiological parameters without facial expression are in many cases limited to the valence and arousal spectrum and fail to accurately identify specific emotions. Examples of such systems include combining: EEG with heart rate; heart rate variability with respiration; heart rate, skin conductance, and body temperature; and electrocardiogram (ECG), photoplethysmography, and galvanic skin response. Validation of these systems is still insufficient and there is very scarce data to reliably assess their clinical application. One new promising system is the JAKE Multimodal Data Capture System, which integrates EEG, ECG, sleep monitoring, eye tracking, caregivers’ observations, and clinicians’ conducted procedures. This system has been tested in autistic children in order to improve therapy and to aid the identification of biomarkers in specific ASC subpopulations.

Interactive systems
Emotion expressions are being used in clinical settings, not only to improve diagnostics but also in the enhancement of therapy. In an interactive system, emotion expression is a tool used to play a game, control an avatar, or complete a task. Prototypes of these systems are biofeedback systems based on galvanic skin reaction, heart rate, and heart rate variability or EEG. Interactive systems with automated recognition of emotions are also based on facial expression and could be used during therapy. An example of such a system is FACE, developed by Pioggia and colleagues. An issue with FACE is that it requires the assistance of a therapist and thus is not widely used. Computer-assisted systems are still a valuable option in improving therapy. By focusing on emotion expression and recognition, children can learn social and mentalizing skills, as demonstrated by Rice and colleagues with the use of their FaceSay program, designed to work with autistic children.

Children and adults with difficulties in recognizing emotions can learn emotion expression with the use of modern technology. Simple tools can be based on a website (e.g., Micro Expression Training Tool 3.0 and Subtle Expression Training Tool 3.0), on a smartphone application, via a DVD or from computer and console video games. For example, after training for the recognition of anger and joy expressions with use of a video game, children with ASC were able to express these emotions at the same level as other children (typically developed, IQ-matched children from a control group). Facial emotion expression can be used as a brain–computer interface and in this case, computers can be controlled by expression of certain emotion. This can be used in developing games based on real-time interaction with a computer that helps children to learn to identify and express emotions. There is a lot of independently developed software to train emotion recognition and expression but they are not clinically validated. However, there are international projects integrating scientists and clinicians and one of the biggest is ASC-Inclusion, which is aimed at development of diagnostic tools and interactive games used in therapy.

At present, there are more sophisticated tools becoming available for more exact measures of emotion expression. One of the new possibilities is augmented reality. It is a real-world environment with its elements augmented by computer effects. Objects can be transformed, gain new features (e.g., visual or auditory), or be created from scratch. In prototypical augmented reality, emotions experienced and expressed by characters of a book or a video are augmented in order to show children emotions they are learning.

Another rapidly developing field is virtual reality. There is growing data on applying virtual reality in cognitive behavioral therapy, especially exposure therapy in phobias and post-traumatic stress disorder. In simpler forms, it is used to interact with patients on the basis of their expressed emotions; for example, to control their avatar in virtual reality (VR) or to interact with other created avatars. Facial emotion expression using VR systems is often based on FACS, but the software used in a VR environment can produce clinically reliable and valuable facial emotion expressions. VR systems may be designed to teach and improve social skills, enhance facial affect recognition, or to promote emotion expression or empathy learning. Further, it has been observed that there are no notable differences in reaction to a VR environment between typically developed individuals and children and adolescents with ASC. VR environments enable the creation of individually tailored therapeutic programs for different indications (anxiety disorders, schizophrenia, ASC) and for individual patients with specific needs. Because of the great development of VR and its integration with game consoles and smartphones, we may expect more rapid growth in its use in therapy for mental disorders. An interactive system may be integrated with simple home devices, such as tablets, smartphones, or game consoles, or may be a separate machine. These systems can be based on speech analysis and help users to identify their current emotion in the valence and arousal spectrum and adjust their behavior or may be used in non-clinical settings. An excellent example of an interactive system is an application developed as an eBook reader on a tablet where the story line is dependent on children's emotional reaction expressed with movement, gesture, or facial expression. It enhances expression of emotions, and helps in the learning of their context and meaning. It also promotes improvement of a child’s motor skills as it requires moves like clapping, snapping, and so forth. Apart from creating an artificial environment or augmenting an existing one, we can develop devices that help with the expression, recognition, and understanding of emotions in real life. Measurement and identification of a currently experienced emotion may help in
adjusting an appropriate behavior in social situations. Devices used for this purpose may be based on simple wearable sensors placed in a wrist band signaling arousal,121 or involve more complex sets of wireless sensors connected with a smartphone, providing information about experienced and expressed emotions in real time, for example the ‘Capture My Emotion’ platform.122,123 The invention of separate devices dedicated to helping includes the development of robots that work with people experiencing communication impairment. Giannopulu et al. created a minimalistic robot, shaped as a small plant for improving communication in autistic children.124 The robot simply nodded along while a person spoke to it. This helped children to sustain their attention and encouraged them to speak. The minimalistic character of this intervention is of great importance, as it was easier for children to cope with it and the probability of the robot’s reaction seems to facilitate better vocal and emotion expression in autistic children. Another example of such a device is a mobile robot created by Goulart et al. that interacts with a child based on emotion expression EEG data together with laser sensor-based data on the child’s localization and position. The drawback of this device is the necessity for the child to wear EEG sensors while using the robot. It was created to work with autistic children; however, to date, there are no data available on the clinical application of this device.125

Humanoid robots are a well-known occupational therapy strategy for the education of autistic children.100 Robotics were first proposed as a method for enhancing autistic education with the robotic turtle ‘LOGO,’ which was found to be suitable for eliciting a range of verbal and nonverbal social responses from ASC children.126 Since then, a large variety of interactive educational programs have been investigated with a particular focus on enhancing the ability of children with ASC to recognize emotions.127–129 However, these therapies are generally based on a ‘Wizard-of-Oz’ paradigm in which the robot is remotely controlled by a therapist. With advances in affective computing technologies in the study of emotions in human–computer interaction, research attention has started to focus on developing fully automated humanoid systems that monitor behavioral signals, such as speech, facial expression, and body movements, to adapt to meet the specific needs of an autistic child.130,131 Furthermore, with the integration of these technologies, the next generation of ASC therapy robots will be able to automatically, robustly, and objectively detect a child’s emotion and thus be able to assess the appropriateness of a child’s emotional response,132 a skill not yet fully targeted in current conventional robot therapies.

Summary

Emotion expressions are a ‘hot’ topic in neuroscience research. Most of the work is performed for purposes other than clinical practice; however, the use of new technology in everyday work with patients is a very promising approach. In this article, we could not cover all possible directions of technical development. Many methods are still in the early stages of development and are lacking clinical validation. Some of them are designed strictly for research purposes and do not have any practical value for clinicians. On the other hand, many newly developed tools available on the market have not been adequately tested and their clinical value is questionable. The aim of this paper is to summarize the available methods of emotion expression analysis that can be clinically useful and point to possible future directions. We have tried to include the diagnostic and therapeutic methods that have been published in scientific journals, but the data, including quantitative and qualitative information, are limited, which causes an important drawback of this paper. For the summary of these methods see Table 1.1,121 What we can expect in the future is a rapid growth of use of modern technologies, such as VR, augmented reality, and interactive systems based on game consoles, in therapy enhancement. Hardware for these systems has become more available and new applications are easy to use in everyday life and tailored for the specific, individual needs of a patient. Another dynamically growing direction is the development of humanoid robots, and we may expect them to become more present as our assistants both in everyday life and clinical practice. Currently, most of the new technology methods in emotion expression are designed for ASC patients; however, they can be adapted and applied in patients with depression, schizophrenia, amyotrophic lateral sclerosis, locked-in syndrome, and in forensic psychiatry. The goal of future research is first to establish a set of parameters to obtain the best accuracy in emotion recognition; second, to make emotion-based systems as simple as possible; and finally, to create a method or multimodal system that could be applied in a real-life setting, not just in a neuroscience laboratory.

Disclosure statement

Preparation of this paper was not supported by any external financing source. The authors declare no conflicts of interest.

Author contributions

K.G.: literature analysis, manuscript preparation and editing; A.R.: literature analysis, drafting manuscript; A.L.: literature analysis, drafting manuscript; S.B.-C.: literature analysis, manuscript editing; B.S.: literature analysis, manuscript editing; N.C.: literature analysis, manuscript editing; J.P.-B.: literature analysis, manuscript preparation; A.P.: literature analysis, manuscript preparation; I.L.: literature analysis, manuscript preparation. All authors revised, developed, read, and approved the final manuscript.

References


Cowrie R, Cornelius RR. Describing the emotional states that are expressed in speech. Speech Commun. 2003; 40: 5–32.


132. DE-ENIGMA. Playfully Empowering Autistic Children. [Cited 4 December 2018.] Available from URL: www.de-enigma.eu