



The emotion–facial expression link: evidence from human and automatic expression recognition

Anna Tcherkassof¹ · Damien Dupré²

Received: 11 March 2020 / Accepted: 6 November 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

While it has been taken for granted in the development of several automatic facial expression recognition tools, the question of the coherence between subjective feelings and facial expressions is still a subject of debate. On one hand, the “Basic Emotion View” conceives emotions as genetically hardwired and, therefore, being genuinely displayed through facial expressions. Consequently, emotion recognition is perceiver independent. On the other hand, the constructivist approach conceives emotions as socially constructed, the emotional meaning of a facial expression being inferred by the perceiver. Hence, emotion recognition is perceiver dependent. In order (1) to evaluate the coherence between the subjective feeling of emotions and their spontaneous facial displays, and (2) to compare the recognition of such displays by human perceivers and by an automatic facial expression classifier, 232 videos of expressers recruited to carry out an emotion elicitation task were annotated by 1383 human perceivers as well as by Affdex, an automatic classifier. Results show a weak consistency between self-reported emotional states by expressers and their facial emotional displays. They also show low accuracy both of human perceivers and of the automatic classifier to infer the subjective feeling from the spontaneous facial expressions displayed by expressers. However, the results are more in favor of a perceiver-dependent view. Based on these results, the hypothesis of genetically hardwired emotion genuinely displayed is difficult to support, whereas the idea of emotion and facial expression as being socially constructed appears to be more likely. Accordingly, automatic emotion recognition tools based on facial expressions should be questioned.

Introduction

With the development of commercial automatic facial expression recognition tools (Dupré et al., 2018, 2020), industries and governments are gradually implementing this technology to monitor humans’ emotions in various scenarios (e.g., in marketing, healthcare, and the automotive industry to name a few). This technology rests on the premise that emotional state is revealed by the measurement of facial features thanks to a one-to-one mapping between subjective feeling and facial expression. This premise comes

from the Basic Emotion View (Ekman et al., 1987; Ekman, 1992, 2007; Ekman & Heider, 1988). By contrast, for the constructivist approach (Barrett 2017 emotions; Gendron & Barrett, 2017), emotions being socially constructed, a making-meaning process takes part in the link between emotion and facial expression. Emotion recognition is considered as resulting from the perceiver making inferences about the expresser’s underlying emotions. The constructivist approach is, therefore, a perceiver-*dependent* view; whereas, the Basic Emotion View is a perceiver-*independent* view (Gendron & Barrett, 2017). In the emotion–facial expression link contention, the issue of facial expressions’ naturalness must also be considered. Since automatic recognition tools are meant to be used in real-life conditions, spontaneous displays of emotional feelings are the heart of the matter as well as their accurate recognition. The potential significance of the present study is that it aims at providing data, not only on the emotion–facial expression link, but also on the perceiver-independent vs. perceiver-dependent question of emotion recognition as regard to spontaneous facial expressions of self-reported emotions. It looks at the triangulation

✉ Anna Tcherkassof
anna.tcherkassof@univ-grenoble-alpes.fr
Damien Dupré
damien.dupre@dcu.ie

¹ Psychology Department, Université Grenoble Alpes, Bâtiment Michel Dubois, 1251 Avenue Centrale, Saint-Martin-d’Hères 38400, France

² Business School, Dublin City University, DCU Glasnevin Campus, Dublin D09, Ireland

emotional feeling—spontaneous facial display—emotion recognition within a single paradigm comparing the recognition accuracy between human and machines.

The basic emotion view

Since Darwin's book *The Expression of the Emotions in Man and Animals* (Darwin, 1872), countless studies have fortified the idea that emotional states are inherently coupled to a set of prototypic facial expressions (e.g., Ekman & Cordaro, 2011). The Basic Emotion View holds that facial expressions are genuine displays of an individual's inner emotional state. More specifically, a set of six so-called basic emotions (*fear, anger, surprise, disgust, sadness* and *happiness*) are universally displayed and are genetically hardwired not only in humans (Ekman, 1992), but also in different animal species (Waal, 2019). According to the Basic Emotion View, "when emotions are aroused by perception of a social event, a set of central commands produce patterned emotion-specific changes in multiple systems, including [...] facial expressions." (Ekman, 2007, p. 49). To respond to criticisms, several amendments have been made to the Basic Emotion View, increasing the number of basic emotions from six to seven (Ekman & Heider, 1988) as well as adding the concept of "display rules" to explain cultural differences in the management of facial expressions (Ekman et al., 1987). The most recent versions of the Basic Emotion Theory even distinguish some twenty expressions of emotions, including love, pride, embarrassment, shame, among others (Keltner et al., 2019).

Relatively few studies have actually tested Basic Emotion View's fundamental claim regarding the facial production (or "expression") of emotions. Rosenberg and Ekman (1994) reported the first evidence of coherence between self-report of emotion and displayed facial expressions. Participants were shown emotionally evocative films and reported their own emotions using a moment-by-moment reporting procedure. Analysis of participants' facial expressions and reports of emotions showed that there was a high degree of temporal linkage and emotional agreement between facial expressions and self-reports. Notwithstanding this study, few have used purely descriptive methods such as electromyography or objective face coding systems to identify and to measure the actual changes in the face when a given emotion is felt (see Wagner, 1997, for methodological points).

Much of the research has focused instead on the recognition of emotional facial expressions (EFE), that is on the issue of the perceiver's interpretation of facial displays. According to the Readout Hypothesis (Buck, 1985) which formalizes the basic Emotion View, facial expressions have evolved to provide "an external readout of those motivational-emotional processes that have had social implications during the course of evolution" (Buck, 1985, pp. 396–397).

Thus, as readouts of emotional states, spontaneous expressive displays are directly accessible to other organisms. Research on the inference of emotions from facial expressions has established that perceivers show considerable agreement that the so-called basic emotions are associated with specific facial displays (e.g., reviews from Elfenbein & Ambady, 2002; Russell, 1994).

Yet, the empirical data called upon by the Basic Emotion View remain unpersuasive. Notably, it fails to explain, in instances in which display rules cannot be called upon, how individuals can feel emotions without expressing them or how individuals can express emotions without feeling them, and why perceivers can make mistakes in recognizing facial expressions of basic emotions, among others. All these findings subverting the Basic Emotion viewpoint come from field observations and laboratory experiments on spontaneous expression of emotions (for reviews, Fernández-Dols & Crivelli, 2013; Reisenzein, Studtmann, & Horstmann, 2013; Russell, Bachorowski, & Fernández-Dols, 2003). The key findings are summarized below as well as the main issues regarding recognition studies.

Spontaneous facial expression in naturalistic studies

Naturalistic studies look at the ecological frequency of co-occurrence of certain emotions and facial displays (Fernández-Dols & Crivelli, 2013). Though they cannot be considered as a straightforward test of the triggering role of emotion on facial behavior, they have the advantage of considering situations that cannot be created in a laboratory. For instance, the ethological study of Kraut and Johnston (1979) contradicts the Basic Emotion View premise that a smile is only the major component of a facial display associated with and caused by feelings of happiness. Naturalistic observation at a bowling alley showed that bowlers do not necessarily smile after scoring a spare or a strike (a situation likely to elicit a positive emotion). Rather, they often smile when interacting with other people. More generally, examination of bowlers' facial display showed that they rarely smiled while facing the pins but often smiled when facing their friends. These findings were confirmed by Ruiz-Belda, Fernández-Dols, Carrera, and Barchard (2003) who analyzed the facial displays of bowlers after a strike and soccer fans after their team scored. Their results show a low probability of smiling when participants were not interacting with someone else. These findings have been supported in other realistic field settings such as the one of Fernández-Dols and Ruiz-Belda (1995) suggesting that happiness is not a sufficient cause of smiling. Fernández-Dols and Ruiz-Belda (1995) watched extremely happy gold medalist athletes displaying facial expressions of sadness (sometimes associated with tears) during their Olympics awards ceremonies. More

specifically, winners showed Duchenne smiles and other types of smiles when they were interacting with other people during the awards ceremony, but smiles were scarce or nonexistent when waiting behind the podium and/or when turning toward the flagpoles and focusing their attention on the flags and the national anthem. Other naturalistic studies have also found that happiness cannot be considered the strongest predictor for the occurrence of Duchenne smiles (Crivelli, Carrera, & Fernández-Dols, 2015).

Happiness is not the only emotion weakly associated with the predicted facial expression. A naturalistic study conducted by Scherer and Ceschi (1997) in an airport's baggage handling office showed that passengers claiming for their lost luggage displayed very few facial expressions of negative emotions while self-reporting subjective feeling states of anger or sadness among others. The covariation between passengers' self-ratings and the claims agents' attributions of the passengers' emotions was very low. Another refutation of the Basic Emotion View's predictions regarding the link between emotions and facial expressions is provided by naturalistic observations of infants' productions of facial expressions. It is often acknowledged that adults regulate their expressive behavior. In accordance with various display rules (personal, social, cultural ones), they exert a control over the supposed automatic readouts of their emotions. Developmental studies provide relevant settings in which display rules are inoperative. Camras et al. observed that facial expressions of negative emotions were displayed in circumstances that were unlikely to have elicited those emotions. For instance, infants displayed "fear" expressions in settings not related to fear (Camras et al., 1991). Bennett, Bendersky, and Lewis (2002) videotaped a sample of 4-month-old infants during tickle, sour taste, jack-in-the-box, arm restraint, and masked stranger situations. Infants displayed a variety of facial expressions in each eliciting situation. Yet, more infants exhibited positive than negative facial expressions across all situations—except sour taste. No evidence for emotion-specific facial expressions corresponding to anger, fear, and sadness was obtained. Camras et al. also observed that 11-month-old European-American, Chinese, and Japanese infants did not display distinct negative emotion-specific patterns of facial muscles in response to two elicitors meant to induce fear and anger (Camras et al., 2007). Thus, these findings in naturalistic settings provide little support for the one-to-one mapping of subjective feeling and facial expression.

Spontaneous facial expression in laboratory studies

Laboratory findings also support field studies. As stressed by Reisenzein et al. (2013), experimental studies permit both better control of various factors (e.g., emotion elicitors) and tests about likely moderators of the emotion–facial

expression link (e.g., the social context). As an example, a strong disconfirmation of the Basic Emotion View's premise is put forward by Reisenzein's studies on surprise (e.g., Reisenzein, 2000). In eight controlled laboratory situations, surprise was induced by Reisenzein, Bördgen, Holtbernd, and Matz (2006) by establishing and then invalidating a set of beliefs concerning the experimental events such as the unexpected appearance of a picture of one's own face as the last picture in a series of portraits that had to be rated. Visible or electromyography-detected facial displays of surprise occurred only in few participants. Yet, most participants reported subjective feelings of surprise and most believed that they had shown a strong surprise facial expression. Schützwohl and Reisenzein (2012) also observed similarly low frequencies of surprise facial expressions when their participants, after leaving the laboratory, unexpectedly found themselves not in the corridor but in a new room with green walls and a red office chair. Less than a quarter of the participants displayed an expression of surprise: only 5% showed widened eyes, raised eyebrows, and opened mouth which correspond to the complete expression of surprise according to the Basic Emotion View and 17% showed widened eyes and raised eyebrows. Again, participants overestimated their surprise expressivity.

Studies on happiness and related positive emotions such as sensory pleasantness also show a low coherence between emotion and facial display (Durán et al., 2017; Reisenzein et al., 2013). In fact, only experiments on amusement provide a fairly strong association between emotion and smiling. When confronted to humorous events (e.g., jokes, being tickled), the number of participants smiling and laughing is indeed quite fair, whereas when confronted to other positive emotions (e.g., happiness), few participants show the expected facial expressions (for instance, the Duchenne smile or any kind of expression related to happiness; see Mehu et al., 2007; Lee & Wagner, 2002). More largely, the meta-analysis conducted by Durán et al. (2017) on the degree of statistical covariation between emotions and facial expressions provides conclusive evidence. First of all, the basic emotions are weakly correlated with the specific configuration of facial muscles that the Basic Emotion View ascribes to them (the correlation drops when happiness/amusement studies are excluded). Very few participants who relived an experience of sadness were reactive, only a small number of them showing components of a sad facial expression with oblique eyebrows or lip corners pulled down (Tsai, Chentsova-Dutton, Freire-Bebeau, & Przymus, 2002). Regarding fear, about one third of spider phobic participants exposed to a live tarantula displayed some components of the prototypical fear expression such as eye widening, brow raising and knitting (Vernon & Berenbaum, 2002). Similar observation is made for anger. The proportion of facially reactive participants displaying at least one component

among frowning or lid/lip tightening and reporting anger does not exceed 35% (Johnson, Waugh, & Fredrickson, 2010; Tsai et al., 2002). The same proportion of reactive participants confronted to disgusting elicitors is observed (e.g., Ekman et al., 1980; Fernandez-Dols, Sanchez, Carrera, & Ruiz-Belda, 1997). In any case, neither an insufficient intensity of the emotion to cause a facial expression nor the intervention of display rules (nor measurement issues) can explain the low emotion–facial expression consistency.

Hence, laboratory studies show that facial expressions of emotion are often not displayed in situations in which the Basic Emotion View would predict them to be expressed. Moreover, when corresponding facial expressions are indeed displayed, they are only partially displayed. All in all, research on the spontaneous expression of emotions does not yield strong support for the Basic Emotion View. The available evidence steadily indicates weak links between emotions and their predicted facial expressions both in natural and in semi-naturalistic settings.

Emotional facial expression recognition studies

The Basic Emotion View postulates that, when triggered, each basic emotion is expressed by a prototypical face (non-basic emotions being blends of the basic ones). In return, the recognition of EFE is claimed to be based on the identification of specific patterns of facial movements associated with each emotion, as if expression and recognition were the two sides of the same coin. Seemingly compelling evidence supports this claim, sustaining the possibility of a clear readout of subjective feelings from facial expressions. Furthermore, recognition systems rely on such a principle and are considered as an objective coding tool because they are based on the identification of specific muscular changes in the face. Many people strongly agree that so-called basic emotions are associated with specific facial configurations (Ekman, 2017; Elfenbein & Ambady, 2002) and also argue that this is strong evidence for the Basic Emotion View. Moreover, it implies that EFE recognition of human perceivers should be as accurate as automatic classifiers.

Yet, some researchers have highlighted the limitations of Basic Emotion View empirical research. Among others, evidence has been questioned on methodological grounds (e.g., Russell, 1994). The response format usually used in recognition studies (i.e., forced choice: selection of one word from a pre-specified list of emotion labels), notably, leads to a biased consensus (Russell, 1993). Depending on the list of emotion labels at participants' disposal, EFE of sadness can easily be categorized as sad expressions as well as fear expressions, as one example. Russell and Fehr (1987) also shown that the same facial expression can be seen as expressing different types of emotions, depending on what other faces are seen. DiGirolamo and Russell

(2017) conducted seven experiments that establish that high agreement between participants can be an artifact of the standard method used by EFE recognition studies. Thus, results gathered with forced choice cannot demonstrate the unequivocal link between emotion and facial expressions claimed by the Basic Emotion View. Using alternative recognition methods (emotion satiation procedure, face-matching task, sorting task; e.g., Lindquist, Barrett, Bliss-Moreau, & Russell, 2006; Gendron, Roberson, Vyver, & Barrett, 2014), it has been shown on the contrary that facial muscle movements are not linked in a one-to-one manner to a specific discrete emotional experience. Instead, emotions are probably mentally constructed by the perceiver and mental categories of emotions are needed to accurately categorize facial movements among contextual information.

To the methodological limitations contaminating the hundreds of studies apparently supporting the Basic Emotion View, a stimulus bias must be added. Facial stimuli used in experiments also constitute a methodological bias because they are unrepresentative of ordinary facial expressions. Basic Emotion View empirical evidence is based for the most part on methods using a static and unnatural material, namely, still photographs of posed facial expressions of emotion (e.g., intentionally encoded by the expresser). This kind of methodology raises questions about its ecological validity and the generalizability of the results to real interpersonal emotional communication (e.g., Tcherkassof et al., 2007). Indeed, a number of pieces of evidence indicate that research cannot content itself with data collected with static and posed material. These data come from researches studying the case of dynamic and/or spontaneous facial expressions of emotion. They show that the dynamic aspects of facial movement are likely to be of importance (e.g., Kamachi et al., 2013). Cohn and Schmidt (2004) have shown that spontaneous smiles are of smaller amplitude and have a more consistent relation between amplitude and duration than deliberate smiles. Hess and Kleck (1990) have also pointed out the importance of the dynamics of facial movements, and particularly the irregularity, or phasic changes, of the expressions' unfolding. Thus, the motion of facial expression provides perceivers with other information than the one provided by static expressions. It may be that differences in the social information displayed by static and dynamic expressions leads to facial recognition differential effects. Regarding the issue of spontaneous vs. posed expressions (the latter are overused in experiments), as Meillon et al. (2010) conclude, EFE have been typically studied as static displays. As a consequence, even though the central role of the dynamics of facial expressions is endorsed, little is still known about the temporal course of facial expressions. Furthermore, studied EFE exhibit emotions simulated or posed by actors. Yet, the lack of spontaneity and

naturalness of this material constitutes a serious objection raised against such studies (Kanade, Cohn, & Tian, 2000).

Finally, as many doubts can be raised about the standard method, experiments conducted with such a method cannot be considered as providing solid empirical support to the Basic Emotion View. Based on the numerous methodological criticisms, but also theoretical, addressed to this view, alternative conceptions have emerged. Among them, the constructivist approach is gaining in importance. The constructivist approach represents a different way of understanding the emotion–facial expression link. It affirms that facial expressions do not provide a direct access to individuals' subjective feelings. Therefore, instead of considering that emotions can be “read” on facial displays, it claims that the emotion is “in the eye” of the perceiver (Barrett, Mesquita, & Gendron, 2019).

Constructivist approach

Starting from the empirical evidence suggesting that spontaneous facial expressions in ordinary life are equivocal, Fernández-Dols (2017) argues in favor of a pragmatic conception of natural facial displays. He makes a plea for the idea that natural facial displays, rather than “saying”—because facial expressions do not have a specific meaning—“make” things. Facial expressions are actions in a communicative interaction. They do not express emotions but they “prompt, on the receiver's side, important inferences about the context, the sender [expresser], and the course of the interaction between sender [expresser] and receiver” (Fernández-Dols, 2017, p. 466). As such, the fact that facial displays are able to signal emotions is a byproduct of one of their main functions; implementing actions performing practical ends. Therefore, a certain number of expressive displays (e.g., surprise, disgust, fear, anger) hold a motor intention (Pacherie, 2003). They implement their aim which is their motor intention. They are not primarily communicative signals, and even less so can they be considered the outlet of an internal state. Facial displays are parts of pragmatic actions aiming at orienting the person's relation to their environment. For instance, facial displays are maintaining, breaking or restoring the relationship between expresser and perceiver (Frijda, 2012). Facial displays are not recognized in semantic terms but are perceived as intentional actions. In the face of the continual flow of uninterrupted facial movements, perceivers see behaviors directed towards a goal. They translate the continuous flow of movements into coordinated sequences of actions holding a beginning and an end. Facial displays are not simple strings of action units, the morphological configuration of which would be the prototype of a given emotion, and consequently identified as such (Ekman et al., 1987). They are best conceived as a Gestalt, the same way as a string of musical notes establishes

a melody (Tcherkassof & Frijda, 2014). This is why even inauthentic facial displays can still be recognized as emotional expressions. Duchenne de Boulogne explains that the artist who has shaped the famous Laocoon antic sculpture, exhibited in the Vatican's museum, has made a modeling mistake since no face can display its emotional expression (Duchenne, 1876). Indeed, no muscular contraction can produce it. He even rectifies the “mistake” by presenting a statue which face is shaped according to the physiology of facial expressive movements. His demonstration gives food for thought. Even though no objective coding system can correctly code the discordant facial features of the Laocoon's face, anyone can easily recognize the suffering and despair he expresses. This example is aligned with the constructivist approach.

The constructivist approach claims that facial displays are behaviors whose meaning is inferred by perceivers. Findings support this perceiver dependence (Lindquist & Gendron, 2013; Niedenthal, Wood, Rychlowska, & Korb, 2017). They show that to make meaning of another person's facial behavior, the perceiver relies in particular on her/his knowledge about emotion categories. For instance, Gendron et al. used a face-sorting task allowing them to manipulate the influence of emotion concepts on how facial expressions were perceived. They conducted their experiment among U.S. participants and Himba participants from remote regions of Namibia (Gendron et al., 2014) and Hazda participants of the Eastern Rift Valley of Tanzania (Gendron et al., 2020) both groups with limited exposure to Western culture. Gendron et al. demonstrated that facial expressions were not universally recognized in discrete emotional terms. Indeed, when Himba and Hazda participants did not have emotion concepts at their disposal to structure perception, they perceived the facial expressions as behaviors, such as looking or smelling, that did not have a necessary relationship to emotions. They did not infer inner states, rather they proceeded with action identification that pointed out the functions of behaviors (see also Crivelli, Russell, Jarillo, & Fernández-Dols, 2017 for similar observations among a small-scale society of Papua New Guinea). The constructivist approach considers that specific emotion categories, as conceptualized by Western cultures' knowledge, are cast on the perceived face to make meaning of the expresser's facial displays. Following this approach, faces convey a range of information essential for social communication. They are best conceived as tools displaying signals in social interactions (Crivelli & Fridlund, 2018). These signals can convey individuals' motivations and readiness (Frijda & Tcherkassof, 1997) or social messages (Fridlund, 1994). As for emotional meaning, more specifically, this is shaped by the perceiver according to the specific context in which the facial displays are observed.

Having reached this point, one can assert that numerous questions regarding the link between emotions and facial

expressions remain unanswered. The two main competing approaches to facial expressiveness, the Basic Emotion View and the constructivist approach, entail completely opposite predictions regarding the decoding of facial expressions, as evidenced above. The present study aims to examine these predictions to provide empirical evidence to allow the discussion to evolve. To date, no systematic study has looked at facial expressions spontaneously displayed in reaction to emotional triggers and how they are decoded, both by human perceivers and by automatic emotion recognition tools based on the detection of facial muscular configurations. This study fills that gap. It intends to investigate the consistency between the subjective feeling of emotions and its recognition from facial expressions. Spontaneous and dynamic facial reactions to emotional elicitation are under consideration to ensure the generalizability of the results to emotional behaviors in ordinary life. More specifically, this study aims to examine the recognition of EFE produced by ordinary people during situations judged and/or self-reported to involve different emotions. It (a) examines consistency between ordinary people's self-reported emotional experience and perceivers' judgments of these ordinary people's EFE, and (b) examines consistency between ordinary people's self-reported emotional experience and an automatic classifier's analysis of these ordinary people's EFE. In other words, it is interested in how people actually move their faces to express self-reported emotions, in how human perceivers accurately infer the expresser's emotional state, and in how automatic recognition accurately codes the expresser's emotional state. Because of their superior ability to exactly recognize EFE, it is expected that human perceivers' will have a higher accuracy to identify expressers' subjective feeling than automatic EFE recognition tools.

Methods

To evaluate the consistency between subjective feeling of emotions and their recognition from facial expressions, expressers were first recruited to perform an emotion elicitation task, while their facial expression was video recorded. To reduce the likelihood of facial control, the expressers were alone in the room and were filmed by hidden cameras, so they had no reason to comply with social display rules. Then, the videos of the expressers' faces were shown to human perceivers and were also analyzed by an automatic classifier to identify which emotion was displayed.

Emotion elicitation

For the emotion elicitation experiment, 358 encoding participants (182 females, 176 males, $M_{\text{age}} = 47.9$, $SD_{\text{age}} = 9.2$) were recruited to perform one out of 11 emotion elicitation

tasks designed to trigger a positive, a specific negative or a neutral emotional state. These participants were recruited by a private company for a study supposedly devoted to an "ergonomic visual task" (cover story). After a description of the general aims, participants agreed and signed the experiment consent form. At the end of the elicitation task, they received an equivalent of €50 in voucher for their participation. A second consent form was signed by the participants to allow their video to be processed for research purposes.

Two different types of task were used to elicit emotions. Some tasks were passive and consisted in the participant watching videos selected to trigger emotions (e.g., television commercials). Other tasks were active; they required the participant to interact with a computer (e.g., to answer questions or to test a flawed software). Expressers' faces were recorded using a hidden camera resulting 358 front facing 768×576 videos varying from 1 to 1479 s. These recordings form the DynEmo database (Tcherkassof et al., 2013).

After the emotion elicitation task, the expressers rated their subjective feeling on Likert scales from 0 ("not at all") to 5 ("strongly") related to six "basic" emotion labels (i.e., *anger*, *disgust*, *fear*, *happiness*, *surprise* and *sadness*) as well as six "non-basic" emotion labels (i.e., *pride*, *curiosity*, *boredom*, *shame*, *humiliation*, and *disappointment*).

Finally, a debriefing session was performed to ensure that expressers were not durably affected by the emotion elicitation task. The debriefing was also used to check that expressers did not guess the real purpose of the experiment (e.g., being filmed while they were performing an emotional elicitation task) to guarantee facial expressions' genuineness. All expressers gave their agreement on their data and recordings being processed for research purposes only.

Human facial expression recognition

To analyze how the recorded facial expressions are perceived, 1383 participants were recruited among social science under- and post-graduates through advertising to watch and to annotate these recordings. They received a course credit for their participation. After a description of the general aims, participants agreed and signed the experiment consent form. An iterative procedure was set up to insure that each recording has been annotated by at least 20 participants: 10% of the video database were randomly selected and annotated until all the videos reached the amount of annotation required, then another section of the database is randomly selected among the remaining recordings to be annotated. To avoid potential decrease in the quality of the annotations due to cognitive fatigue, each participant only annotated recordings during a maximum of 30 min. As a result, only 232 out of the 358 videos have been annotated. Each video was annotated 29 times on average ($SD = 12$).

The annotation of facial expressions was performed on-site using *Oudjat*, a software for designing video annotation experiments (Dupré et al., 2015). For each video, the annotation procedure followed two steps. First, the participants had to identify the emotional sequences by pressing the space bar of their keyboard to indicate the beginning and the end of the emotional sequences while watching the video. Second, the participants watched each emotional sequence previously identified and labeled the sequence using one of the 12 emotions proposed including six basic emotion labels (i.e., *anger*, *disgust*, *fear*, *happiness*, *surprise* and *sadness*) and six “non-basic” emotion labels (i.e., *pride*, *curiosity*, *boredom*, *shame*, *humiliation*, and *disappointment*). They also had the possibility to indicate that the sequence was expressing none of the proposed emotion labels.

This annotation procedure resulted in a uni-dimensional time-series for each video per human perceiver, identifying for each second of the video which emotion was recognized. Then, the time-series corresponding to the same video was aggregated to calculate the proportion of human perceivers $x_{\text{video}_i, \text{label}_j, t_k}$ for each second of the video per emotional label shown in Eq. (1).

$$x_{\text{video}_i, \text{label}_j, t_k} = \frac{n_{\text{video}_i, \text{label}_j, t_k}}{n_{\text{video}_i}}, \quad (1)$$

where i is one of the 232 videos, j is one of the six “basic” emotion labels, k for each second of the video.

Automatic facial expression recognition

The 232 annotated video were processed with Affdex (SDK v3.4.1). Affdex is an automatic facial expression recognition classifier developed and distributed by Affectiva, which is a spin-off company resulting from the research activities of MIT media lab created in 2009 (McDuff et al., 2016). Affdex’s algorithm uses Histogram of Oriented Gradient (HOG) features and Support Vector Machine (SVM) classifiers to recognize facial expressions. For each video frame, Affdex identifies the probability $p_{\text{video}_i, \text{label}_j, t_k}$ of expressing each of the six basic emotion labels (i.e., *anger*, *disgust*, *fear*, *happiness*, *surprise* and *sadness*) as well as additional psychological states such as *valence*, *engagement* or *contempt*, and facial features such as *cheek raise*, *eye widen* or *jaw drop*. The result of Affdex is a multivariate time-series output which provides for each label a probability from 0 to 100 (rescaled from 0 to 1 for the analysis) of expressing the corresponding label. For this study, only the six basic emotion labels are analyzed to match with self-reports and human perceivers.

For both human and automatic recognition, to determine which of the six basic emotions can be used to identify each video, the recognition probability for each label by frame

was converted into odd ratio by label (Dente, Küster, Skora, & Krumhuber, 2017). The highest sum of each odd ratio time-series defines the label recognized (see Eq. (2) for human recognition and Eq. (3) for automatic recognition).

$$\text{video}_i.\text{label} = \max \left(\frac{\sum_{k=1}^n x_{\text{video}_i, \text{label}_j, t_k}}{\sum_{k=1}^n x_{\text{video}_i, t_k}} \right), \quad (2)$$

$$\text{video}_i.\text{label} = \max \left(\frac{\sum_{k=1}^f p_{\text{video}_i, \text{label}_j, t_k}}{\sum_{k=1}^f p_{\text{video}_i, t_k}} \right), \quad (3)$$

where i is one of the 232 videos, j is one of the six “basic” emotion labels, k for each second of the n second video or for each frame of the f frame video.

Since expressers’ self-reports, human annotations and the automatic recognition include data on “non-basic” emotion labels and features, the analysis is performed using only the six basic emotion labels to compare them. The maximum rating for self-reports, human annotations and automatic recognition is used to label the video. In case of more than one label obtaining the maximum rating, the video is labeled as undetermined.

The confusion matrices obtained by crossing expressers’ self-reports, human annotations and the automatic recognition are analyzed by calculating two-tailed t-tests of Pearson’s product-moment correlation between dichotomous variables, accuracy which corresponds to the overall agreement rate averaged over cross-validation iterations and Cohen’s (unweighted) Kappa.

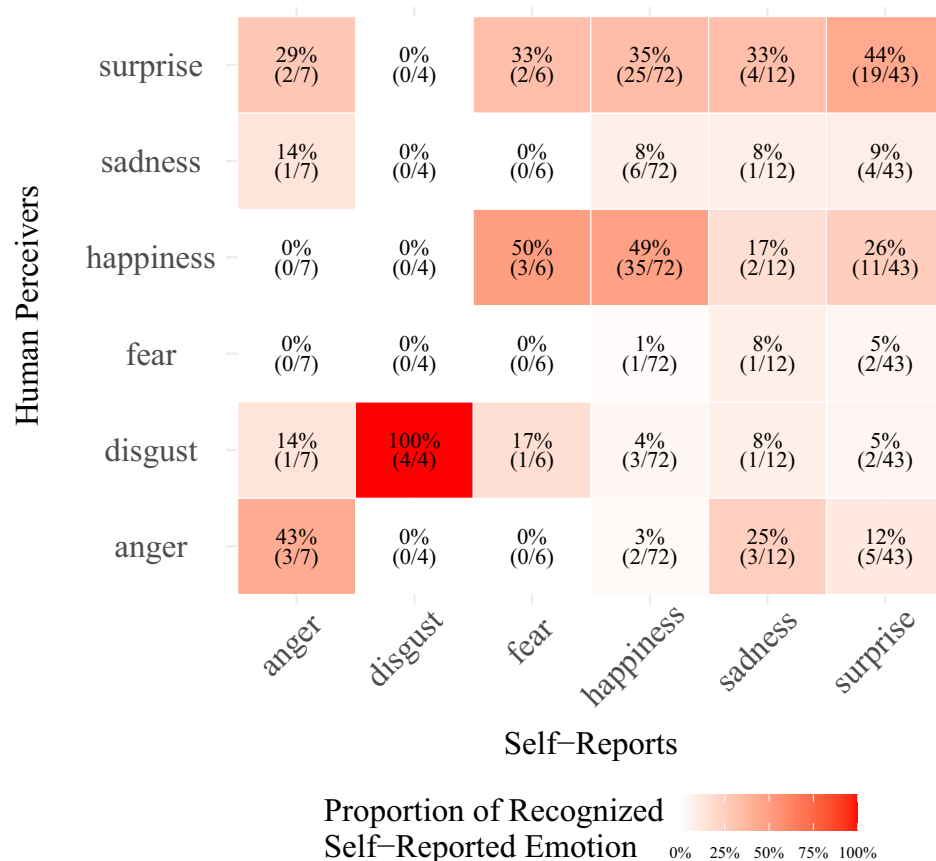
Results

Human perceivers’ accuracy

The overall correlation of recognition and non-recognition between self-reported emotions and human perceivers recognition is significant but low ($r = 0.24$, 95% CI [0.19, 0.29], $t(1384) = 9.21$, $p < 0.001$). To identify differences according to the emotional labels, expressers’ subjective feelings are compared with human perceivers’ recognition in a confusion matrix (Fig. 1).

Each emotion label used to describe expressers’ self-reported subjective feeling (i.e., the label rated with the highest value) is compared with the emotion labels which were rated with the highest score by human perceivers. Results of the confusion matrix show a low agreement between the emotion felt by the expresser during the elicitation and the emotion recognized by the human perceivers (Accuracy = 0.43, 95% CI [0.35, 0.52]; Kappa = 0.19) except for *disgust* (100% of the videos self-reported). These results

Fig. 1 Confusion matrix between the emotion self-reported as being characteristic of the elicitation with the emotion recognized by the human perceivers



are far from those classically obtained in the literature for emotional facial expression recognition (Krumhuber et al., 2020, report an average accuracy of 65%). However these results are mostly obtained with posed facial expressions (i.e., displayed by actors) in a forced-choice recognition paradigm.

Interestingly human perceivers seem to recognize surprise expressed in videos where anger, fear happiness and sadness was the highest self-reported emotion (respectively 29%, 33%, 35% and 33% of the videos self-reported), and in a lower instance *happiness* was recognized in videos where *fear* and *surprise* was the highest self-reported emotion (respectively 50% and 26% of the videos self-reported).

Sensitivity, specificity, precision and *F1* scores for each emotion reveal that *happiness* has the highest coherence ratio, whereas *sadness* has the lowest coherence ratio between true positives and false positives (Table 1).

Accuracy metrics by emotional labels indicate a discrepancy in the ratio of true/false positives. Whereas *happiness* and *disgust* obtain the highest scores, *anger*, *surprise* and *sadness* have the lowest recognition ratio. This difference between emotions is usually observed in the literature (e.g., Krumhuber, Kappas, & Manstead, 2020). Multiple reasons can explain why *happiness* and *disgust* are more easily recognized than *anger*, *surprise* and *sadness*. It is possible that

Table 1 Human recognition accuracy metrics for each emotion

Emotion	Sensitivity	Specificity	Precision	<i>F1</i>
Anger	0.43	0.93	0.23	0.3
Disgust	1.00	0.94	0.33	0.5
Fear	0.00	0.97	0.00	na
Happiness	0.49	0.78	0.69	0.57
Sadness	0.08	0.92	0.08	0.08
Surprise	0.44	0.67	0.37	0.4

na values are produced when not enough data are available to compute accuracy indicators

these expressions are usually more intense than others. In addition, *anger* and *sadness* as non-socially desirable emotions may have been felt but not expressed.

However, self-reports show a significant proportion of undetermined emotional states (35% of the 358 videos) which reveals the potential limit of using 6-points Likert scales to measure emotional self-reports. Indeed, Likert scales may not be able to discriminate between two dominant emotions felt with different intensities due to their reduced number of possibilities. To better discriminate between emotions, an alternative would be to use a continuous slider from 0 to 100 (see for example Lottridge & Chignell, 2009).

Automatic classifier’s accuracy

Similarly to the previous analysis, the overall correlation of recognition and non-recognition revealed a significant but very low coherence between self-reported emotions and automatic classifier’s recognition ($r = 0.12$, 95% CI [0.07, 0.17], $t(1384) = 4.50$, $p < 0.001$). A confusion matrix was used to compare expressers’ subjective feeling with the emotion label recognized by the automatic classifier (Fig. 2).

Results obtained for the comparison between emotions self-reported and recognized by the automatic classifier are somewhat similar to the ones with human perceivers (Table 2). Overall, a low agreement between emotion self-reported and emotion recognized by the automatic classifier (Accuracy = 0.30, 95% CI [0.22, 0.38]; Kappa = 0.07) except for *happiness* (44% of the video self-reported) is evident.

Surprisingly, the automatic classifier incorrectly recognized as *disgust* a significant proportion of videos in which *anger*, *happiness* and *surprise* was the highest self-reported emotion (respectively, 43%, 25% and 33% of the videos self-reported). In parallel, the automatic classifier recognized as *happiness* videos in which *fear* and *surprise*

Table 2 Automatic recognition accuracy metrics for each emotion

Emotion	Sensitivity	Specificity	Precision	F1
Anger	0.00	0.90	0.00	na
Disgust	0.25	0.72	0.02	0.04
Fear	0.00	0.94	0.00	na
Happiness	0.44	0.73	0.62	0.52
Sadness	0.15	0.95	0.25	0.19
Surprise	0.19	0.86	0.36	0.25

na values are produced when not enough data are available to compute accuracy indicators

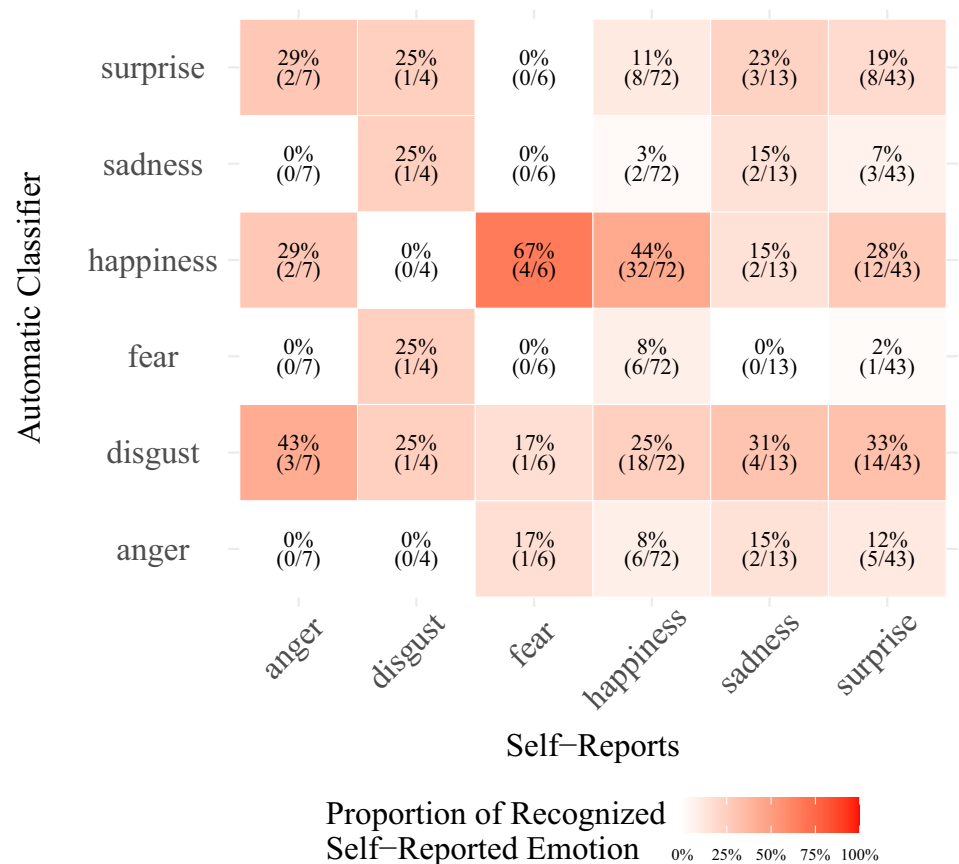
were the highest self-reported emotions (respectively, 67% and 33% of the videos self-reported).

A comparable explanation involving the amount of undetermined video based on self-reports can be provided, as the level of undetermined emotions are very high for the self-reports.

Comparison between human and automatic recognition

As previously mentioned, human perceivers appear to be more accurate than the automatic classifier to recognize an

Fig. 2 Confusion matrix of between the emotion self-reported as being characteristic of the elicitation with the emotion recognized by the automatic classifier



individual's subjective feeling (human perceivers Accuracy = 0.43; automatic classifier Accuracy = 0.30). However, both make similar mistakes as the two-tailed t-test of Pearson's product-moment correlation between human perceivers and automatic classifier recognition is significant ($r = 0.22$, 95% CI [0.17, 0.27], $t(1384) = 8.31$, $p < 0.001$).

A third confusion matrix is used to compare similarities (diagonal) and differences between human perceivers and automatic classifier in classifying the six emotion labels (Fig. 3).

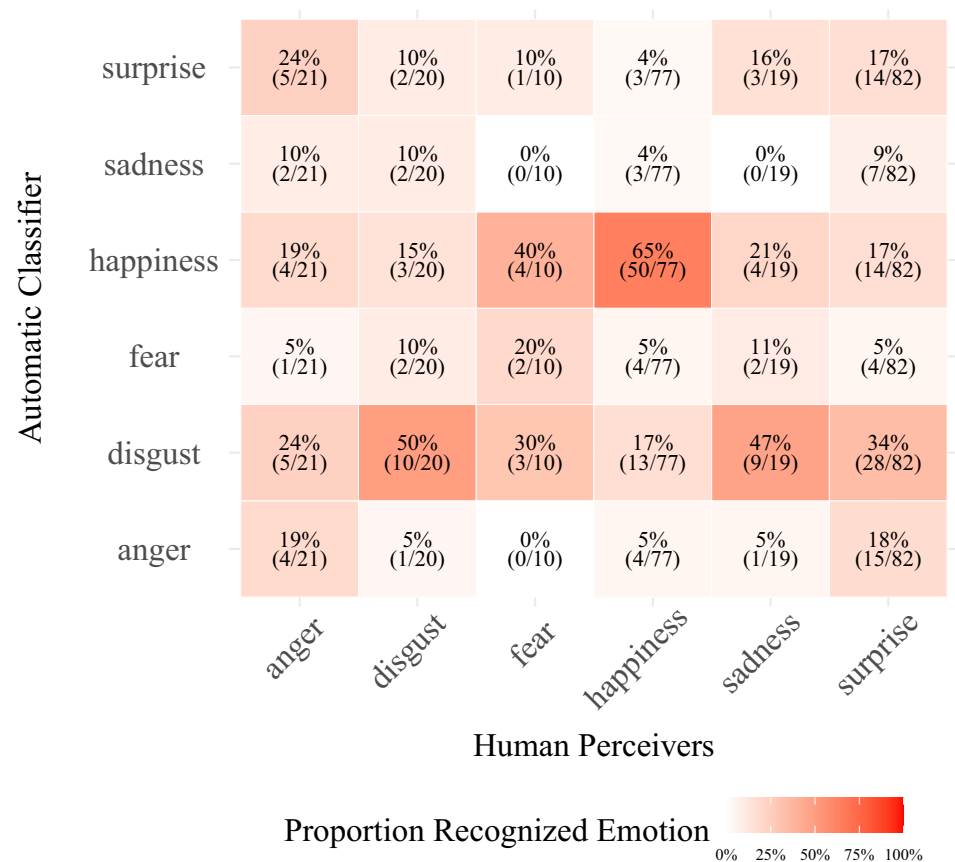
The overall agreement between human perceivers and the automatic classifier is in fact very low (Kappa = 0.18). Except for *happiness* and *disgust* (respectively, 65% and 50% of common labelling), there is no clear common pattern. Moreover, the automatic classifier has a tendency to label as *disgust* videos labeled as *sadness* by human perceivers, and as *happiness* videos labeled as *fear* by human perceivers.

Discussion and conclusion

Despite being one of the most investigated questions in affective science, the consistency between emotion felt and facially displayed on the one hand, and facial expression recognized on the other is a hot topic. To date, no clear

evidence has been found to definitively solve the questions raised. Yet, with the growing interest of industries and government in monitoring individual's psychological states, this issue is under intense scrutiny. The present research aims to provide some empirical data to answer some of the questions posed. The faces expressers spontaneously displayed when confronted with an emotional eliciting task were submitted both to human and to automatic recognition. The criterion for recognition accuracy was the subjective feeling self-reported by the expresser once the elicitation task was carried out. Results first reveal a low consistency between emotion felt and facial expression displayed. They show that facial expressions of emotion are often not displayed when the Basic Emotion View would predict them to be expressed. Second, results show low accuracy rates for both humans and the automatic classifier in identifying the inner emotional states of these expressers based on their facial expressions. Third, human perceivers prove to be better at recognizing the emotion facially expressed than the automatic recognition tool is. Such results support the hypothesis advanced by some authors of low emotion-expression coherence (Kappas, 2003). In many instances, facial displays are not associated with a concordant emotional state, even any emotional state at all (Bonanno & Keltner, 2004; Fernández-Dols & Crivelli, 2013). More and more evidence is showing

Fig. 3 Proportion of emotion labels classified by human perceivers which are recognized by the automatic classifier



that facial expressions are in reality not expressing emotions (McKeown, 2013). Increasing studies show that, for most emotions, the EFE elicited by emotional triggers are scarce and partial, even when micro-expressions are taken into account (Durán et al., 2017). These studies are conducted either in laboratory settings or in the field. This is not to say that facial expression is not an informative modality for understanding the emotional state of a person. There is indeed an affinity between emotion and facial display (Frijda & Tcherkassof, 1997). However, claiming that unique facial muscles configurations are used both to express and to infer the presence of a specific emotion is misleading. As well as other nonverbal behaviors, facial movements are not only assumed to be determined by emotion but also by various other causes, such as psychological states (e.g., motivations or pain), to say nothing of social context and sociocultural norms (Ekman et al., 1987). Hence, facial movements have causes and functions other than the expression of emotion. This multiple determination excludes any possibility of drawing a linear inference from facial activity on the underlying psychological state (emotional or other). Beyond the present observations showing a weak consistency between subjective feelings and spontaneous facial expressions, this study sheds some light on the controversy between the Basic Emotion View and the constructivist approach as to the facial recognition issue. The former (Ekman, 1992) assumes that expressions of emotion are brief and coherent patterns of facial muscle movements that co-vary with discrete subjective experiences. In return, this information displayed by the face corresponds to the one extracted by perceivers. Instead of viewing emotions as natural kinds (Barrett, 2006), the constructivist approach supposes that emotions are social constructions and that facial behaviors intrinsically situated. The emotions that are recognized by the perceiver are constructed in her/his mind. Therefore, facial movements do not express specific emotions because they do not carry intrinsic emotional signification. It is the perceiver that infers the emotional meaning of the facial expression, this interpretation depending on availability of different information such as context or linguistic categories. As a consequence, one can predict from the first line of thinking that individuals' emotional subjective feeling should be correlated with the recognition of facial expressions from both human perceivers and automatic classifiers, whereas if emotions are social constructs, as stated by the second line of thinking, human perceivers should be better at perceiving emotions expressed on the face than automatic classifiers.

The present findings speak against any strong version of the Basic Emotion View. The correlations between the self-measured emotions and the observed facial behaviors are low and the latter are weakly recognized. Results show that human perceivers are more accurate than the automatic recognition tool to identify an individual's subjective feeling on

the basis of their face. Moreover, mistakes made by human perceivers look less arbitrary than the ones made by the classifier. For instance, even if a mix-up between disgust and anger is sometimes reported in recognition studies, odd confusions such as the present ones produced by the classifier have never been noted for human perceivers. The latter obviously make sense of the facial behavior they are witnessing, as claimed by the constructivist stance. That said, these results do not unequivocally argue in favor of the latter. Human perceivers in the present study were far from being perfectly accurate. Of course, one could argue that they were presented with faces without any contextual information which could have helped them to shape more precisely their interpretation. Hassin, Aviezer, and Bentin (2013) reviewed evidence to demonstrate that faces are inherently ambiguous and that perceivers rely on situational cues when they process facial displays (see also Aviezer et al., 2008; Barrett et al., 2011). It happens that contextual information even shifts emotion perception. In the present case, the facial stimuli displayed to the decoders were as equivocal as real-life facial expressions (Aviezer, Ensenberg, & Hassin, 2017), yet perceivers having no cue at their disposal. So, without any possibility of integrating faces and context, their decoding accuracy is necessarily reduced (e.g., Wagner et al., 1986). Their superiority to the classifier tool is probably owing to their capacity to rely on their previous personal experience to invent a context in which a face could display such an expression. Once a credible context is retrieved, they can affix an emotional label to the facial behavior.

Several limitations should be stated, highlighting the need for further research. One of them is effectiveness of the emotion elicitation tasks. One can consider that an intensity threshold needs to be exceeded for a visible emotional expression to occur. In the case of the present study, it may be that insufficient emotion intensity accounts for the low number of visibly reactive participants. However, such line of reasoning fits badly with the Basic Emotion View according to which emotions and their related prototypical facial behavior are universal because they are considered as innate mechanisms allowing individuals to respond adaptively to evolutionary significant events (threats, opportunities, etc.) encountered in the environment, whatever their potency. Moreover, various studies have used fairly strong emotion inductions without obtaining any visible facial display of any kind of emotion (Durán et al., 2017). Moreover, the present results show a quite good correlation between reported emotion and facial behavior for happiness. Thus, there are no reasons to believe that an explanation in terms of too low an emotion threshold applies for the other emotions under consideration. The latter were triggered and measured with conceptually identical elicitation and assessment procedures (Reisenzein et al., 2013). The use of self-reports to evaluate expressers' subjective feelings is another limitation that

can also be put forward because of the numerous cognitive biases they entail. Well-known problems with the reliability of self-reports are, among others, the reconstructive nature of memory, the influence of attentional biases on reports, demand characteristics, distorting effects of implicit causal theories and personal motives, as stressed by Nielsen and Kaszniak (2007). It is clearly obvious that self-reports are not simple outlets of inner mental processes but personal constructions and they are affected by many factors (Kappas, 2003). The emotional feeling echoes motivational tendencies, bodily changes, and cognitive appraisals of events (Sander & Scherer, 2014). All these are encapsulated into semantic categories referred to by labels. As things stand currently, there is unfortunately no objective way for accessing and assessing inner subjective emotional feelings except for asking people to report their subjective experience in words. The procedure used for human recognition can also be open to dispute. First, one can criticize the decoders' expertise level since they were not FACS coders but untrained students. Of course one can expect a difference between skilled annotators and novice ones regarding the assessment of emotional facial displays. This said, in everyday life, few people are specialist coders yet the quantity of successful social interactions proves lay persons recognize quite well others' facial behaviors. Therefore, and especially in the interest of generalizability, asking inexpert people seems relevant. Second, instead of using a classic forced-choice procedure, a more subtle approach was chosen to mimic results provided by the automatic classifier. As explained above, annotators first delimited a temporal sequence during which they noticed an emotional display on the face, and then attributed an emotion label to this behavioral sequence in a second step. Whereas this paradigm is longer and more complicated, it can lead to more robust results in reducing the forced-choice bias (Russell, 1993). However this procedure can also reduce the human perceivers' accuracy. In this regard, the results of the human observation could have been more ambiguous because it is not the natural way that people are inferring meaning from facial expressions. An alternative explanation relies in reducing the recognition bias involved in the classic recognition paradigm. Classic forced-choice paradigms obtain artificially high results; thus, using a more evolved approach, perceivers' accuracy may have been lowered. Another flaw is the lack of comparison with various facial expression recognition methods. Human recognition has only been compared to the Affdex classifier. Future studies are needed to confront human assessments with different automatic recognition methods, both frame based methods and sequence based automatic ones. This latter issue is particularly decisive. Indeed, the issue of the recognition of dynamic expressive sequences is essential because ordinary facial behavior is made up of dynamically shifting morphological features

(Krumhuber et al., 2013). This temporal information is indisputably a key feature of facial activity. It is not only perceiver-based judgements of facial displays which must be compared to automated facial analysis, but also different kinds of human recognition measurements should be undertaken. The challenge researchers are especially confronted with is to find ways to appropriately collect data regarding the perception of spontaneous and dynamic facial behavior. Finally, our understanding of facial displays as they occur in everyday interactions requires a strong emphasis on ecological concerns. The present study is a laboratory experiment. It has the advantage of controlling different parameters of the emotions investigated, such as intensity, quality, and temporal (onset, duration) features. Moreover, as expressers are alone, facing an emotional trigger, it also controls for the social context, removing its possible influence on their facial behaviors. However, it is known that expressers' imagination can influence their expressiveness, for instance when they believe that their friend is doing the same (vs. a different) emotional task in another room (Jakobs, Manstead, & Fischer, 1999). Trying to exclude social influence by leaving expressers alone may be illusionary. From an ecological perspective, it is even a mistake to exonerate behavioral observations from social contexts. Facial activity measurement in dyadic interactions has shown that the facial behavior of the perceiver reflects sometimes more what the expresser is experiencing than what the perceiver is feeling. It is the case, for instance, of emotional mimicry in dynamic social interactions (Hess & Fischer, 2016). It is also the case of healthy patients interacting with schizophrenic patients whose facial activity is almost identical (Krause, Steimer-Krause, Merten, & Ullrich, 1998). Hence, to better comprehend emotional communication in human relationships, experimental research should be corroborated with more ecological protocols.

Nowadays, automatic recognition systems are based on the coding of the facial muscular activations from which they infer the expressed emotion. Such automatic classifier tools take for granted that, when experienced, an emotion is: firstly, displayed on the face; secondly, in the form of a configuration of facial muscles that is unique to a person; and thirdly, recognized by the perceiver (human being or automatic classifier). These are the Basic Emotion View assertions, jeopardized by field observations and laboratory experiments on spontaneous expression of emotions, such as the present study. All raise serious objections to the supposed close relation between emotion and face. They bring up several questions, in particular regarding the perceiver's role in the interpretation of the expresser's emotional face. Is EFE recognition a perceiver-independent or perceiver-dependent process? Although the present results tend to favor the second option, the finalization of operational and effective "reading emotional

faces” devices rests on the answer, if any, to the question raised. In any case, as stressed by Krumhuber et al. (2020), machines’ accurate emotion classification is predicted by the prototypicality of a facial display. More prototypical is the facial behavior, more accurate is the classification. However, ordinary emotional facial behaviors are seldom prototypical. As a result, despite being one of the most investigated questions in affective science, the growing interest of industries and governments in tracking individual’s psychological states is supported by controversial assumptions.

Considering the above, the present results add to the body of literature evidence that an individual’s subjective feelings are hardly inferred from spontaneous facial expressions in the absence of contextual cues, and they invalidate the hypothesis of hardwired emotions unambiguously displayed on the face. This study is the first one, to our knowledge, to look simultaneously at facial expressions spontaneously displayed in reaction to emotional triggers and their accurate decoding by human perceivers as compared to automatic emotion recognition tools. Results provide no support for the one-to-one mapping of subjective feelings and facial behaviors. Even if emotions were hardwired, in everyday life, one does not observe prototypical facial expressions because of their rarity, and therefore research should be focused on analyzing non-prototypical facial expressions. Indeed, this study provide additional corroboration that facial displays of emotion are not exhibited in circumstances in which the Basic Emotion View would predict them to be expressed (Durán et al., 2017). Advancements in identifying “non-basic” emotion labels as well as non-prototypical facial expression have indeed occurred with automatic facial expression recognition tools (McDuff, 2016). However, these present results contribute to a critical understanding of the issues at stake. They suggest that automatic facial expression recognition tools should merely evaluate facial morphology features such as action units (already evaluated in OpenFace, Baltrušaitis et al., 2016; Affectiva’s Affdex, McDuff et al., 2016; Vicar Vision’s FaceReader, Den Uyl & Van Kuilenburg, 2005, to name a few) rather than inferring supposedly emotional or affective states. Trying to interpret facial displays as a means of determining underlying emotional state (especially without any information about the context, nor person specifications of the expresser, etc.), in all likelihood, remains dubious methods.

Acknowledgements The authors would like to thank the developers of the following R packages used to process, analyze, display and report data: R (Version 4.0.2; R Core Team, 2019) and the R-packages *caret* (Version 6.0.86; Jed Wing et al., 2019), *here* (Version 0.1; Müller, 2017), *janitor* (Version 2.0.1; Firke, 2019), *kableExtra* (Version 1.2.1; Zhu, 2019), *papaja* (Version 0.1.0.9997; Aust & Barth, 2018), and *tidyverse* (Version 1.3.0; Wickham, 2017).

Author contributions The authors made the following contributions. AT: Conceptualization, Project administration, Validation, Writing—Original Draft Preparation, Writing—Review and Editing; DD: Investigation, Data curation, Formal Analysis, Software, Visualization, Writing—Original Draft Preparation, Writing—Review and Editing.

Data availability The R code and relevant data for statistical computing are available on the “The emotion–facial expression link (data and code)” repository from the Center for Open Science at <https://osf.io/erua5/> (Dupré & Tcherkassof, 2020). The data that support the findings of this study are available from the DynEmo database but restrictions apply to the availability of these data. DynEmo is devoted to academic research to the exclusion of any other purpose, and so the data are not publicly available. Data are however available from the authors upon reasonable request and with permission of DynEmo.

Compliance with ethical standards

Conflicts of interest No potential conflicts of interests regarding the research reported in the manuscript exist.

Ethical approval This is an observational study. The CERGA Research Ethics Committee has confirmed that no ethical approval is required.

Informed consent Written informed consent was obtained from all individual participants included in the original study.

References

- Aust, F., & Barth, M. (2018). Papaja: Create APA manuscripts with R Markdown. <https://github.com/crsh/papaja>. Accessed 6 Nov 2020.
- Aviezer, H., Ensenberg, N., & Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. *Current Opinion in Psychology*, 17, 47–54. <https://doi.org/10.1016/j.copsyc.2017.06.006>.
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, 19(7), 724–732. <https://doi.org/10.1111/j.1467-9280.2008.02148.x>.
- Baltrušaitis, T., Robinson, P., & Morency, L.-P. (2016). Openface: An open source facial behavior analysis toolkit. *Winter Conference on Applications of Computer Vision*. <https://doi.org/10.1109/WACV.2016.7477553>.
- Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, 1(1), 28–58. <https://doi.org/10.1111/j.1745-6916.2006.00003.x>.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68. <https://doi.org/10.1177/1529100619832930>.
- Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, 20(5), 286–290. <https://doi.org/10.1177/0963721411422522>.
- Bennett, D. S., Bendersky, M., & Lewis, M. (2002). Facial expressivity at 4 months: A context by expression analysis. *Infancy*, 3(1), 97–113. https://doi.org/10.1207/S15327078IN0301_5.
- Bonanno, G., & Keltner, D. (2004). The coherence of emotion systems: Comparing “on-line” measures of appraisal and facial expressions, and self-report. *Cognition and Emotion*, 18(3), 431–444. <https://doi.org/10.1080/02699930341000149>.

- Buck, R. (1985). Prime theory: An integrated view of motivation and emotion. *Psychological Review*, 92(3), 389–413. <https://doi.org/10.1037/0033-295X.92.3.389>.
- Camras, L. A., Malatesta, C., & Izard, C. E. (1991). The development of facial expressions in infancy. In R. S. Feldman & B. Rimé (Eds.), *Studies in emotion & social interaction. Fundamentals of nonverbal behavior* (pp. 31–54). New York: Cambridge University Press.
- Camras, L. A., Oster, H., Bakeman, R., Meng, Z., Ujiie, T., & Campos, J. J. (2007). Do infants show distinct negative facial expressions for fear and anger? Emotional expression in 11-month-old European American, Chinese, and Japanese infants. *Infancy*, 11(2), 131–155. <https://doi.org/10.1111/j.1532-7078.2007.tb00219.x>.
- Cohn, J. F., & Schmidt, K. (2004). The timing of facial motion in posed and spontaneous smiles. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(2), 121–132. <https://doi.org/10.1142/S021969130400041X>.
- Crivelli, C., Carrera, P., & Fernández-Dols, J.-M. (2015). Are smiles a sign of happiness? Spontaneous expressions of judo winners. *Evolution and Human Behavior*, 36(1), 52–58. <https://doi.org/10.1016/j.evolhumbehav.2014.08.009>.
- Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in Cognitive Sciences*, 22(5), 388–399. <https://doi.org/10.1016/j.tics.2018.02.006>.
- Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J.-M. (2017). Recognizing spontaneous facial expressions of emotion in a small-scale society of Papua New Guinea. *Emotion*, 17(2), 337–347. <https://doi.org/10.1037/emo0000236>.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- de Waal, F. (2019). *Mama's last hug: Animal emotions and what they teach us about ourselves*. London: Granta Books.
- Den Uyl, M., & Van Kuilenburg, H. (2005). The face reader: Online facial expression recognition. *Proceedings of Measuring Behavior*, 30, 589–590.
- Dente, P., Küster, D., Skora, L., & Krumhuber, E. (2017). Measures and metrics for automatic emotion classification via facet. In: *Proceedings of the Conference on the Study of Artificial Intelligence and Simulation of Behaviour*, pp 160–163.
- DiGirolamo, M. A., & Russell, J. A. (2017). The emotion seen in a face can be a methodological artifact: The process of elimination hypothesis. *Emotion*, 17(3), 538–546. <https://doi.org/10.1037/emo0000247>.
- Duchenne, G.-B. (1876). *Mécanisme de la physiologie humaine: Où, analyse électro-physiologique de l'expression des passions*. Paris: J.-B. Baillière.
- Dupré, D., Akpan, D., Elias, E., Adam, J.-M., Meillon, B., Bonnefond, N., et al. (2015). Oudjat: A configurable and usable annotation tool for the study of facial expressions of emotion. *International Journal of Human-Computer Studies*, 83, 51–61. <https://doi.org/10.1016/j.ijhcs.2015.05.010>.
- Dupré, D., Andelic, N., Morrison, G., & McKeown, G. (2018). Accuracy of three commercial automatic emotion recognition systems across different individuals and their facial expressions. *Proceedings of the International Conference on Pervasive Computing and Communications*. <https://doi.org/10.1109/PERCOMW.2018.8480127>.
- Dupré, D., & Tcherkassof, A. (2020). The emotion–facial expression link (data and code). *OSF*, <https://doi.org/10.17605/OSF.IO/ERUA5>.
- Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *PLoS ONE*, 15(4), e0231968. <https://doi.org/10.1371/journal.pone.0231968>.
- Durán, J. I., Reizenstein, R., & Fernández-Dols, J.-M. (2017). Coherence between emotions and facial expressions. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 107–129). New York: Oxford University Press.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>.
- Ekman, P. (2007). The directed facial action task. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 47–53). New York: Oxford University Press.
- Ekman, P. (2017). Facial expressions. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 39–56). New York: Oxford University Press.
- Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3(4), 364–370. <https://doi.org/10.1177/1754073911410740>.
- Ekman, P., Freisen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology*, 39(6), 1125–1134. <https://doi.org/10.1037/h0077722>.
- Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., et al. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53(4), 712–717. <https://doi.org/10.1037/0022-3514.53.4.712>.
- Ekman, P., & Heider, K. G. (1988). The universality of a contempt expression: A replication. *Motivation and Emotion*, 12(3), 303–308. <https://doi.org/10.1007/BF00993116>.
- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128(2), 203–235. <https://doi.org/10.1037/0033-2909.128.2.203>.
- Fernández-Dols, J. M. (2017). Natural facial expression: A view from psychological constructionism and pragmatics. In J. M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 457–478). New York: Oxford University Press.
- Fernández-Dols, J.-M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review*, 5(1), 24–29. <https://doi.org/10.1177/1754073912457229>.
- Fernández-Dols, J.-M., & Ruiz-Belda, M.-A. (1995). Are smiles a sign of happiness? Gold medal winners at the olympic games. *Journal of Personality and Social Psychology*, 69(6), 1113–1119. <https://doi.org/10.1037/0022-3514.69.6.1113>.
- Fernández-Dols, J.-M., Sanchez, F., Carrera, P., & Ruiz-Belda, M.-A. (1997). Are spontaneous expressions and emotions linked? An experimental test of coherence. *Journal of Nonverbal Behavior*, 21(3), 163–177. <https://doi.org/10.1023/A:1024917530100>.
- Firke, S. (2019). Janitor: Simple tools for examining and cleaning dirty data. <https://CRAN.R-project.org/package=janitor>. Accessed 6 Nov 2020.
- Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego: Academic Press.
- Frijda, N. H. (2012). Recognition of emotion in others. In F. Paglieri (Ed.), *Consciousness in interaction. The role of the natural and social context in shaping consciousness* (pp. 237–256). Amsterdam: J. Benjamins.
- Frijda, N. H., & Tcherkassof, A. (1997). Facial expressions as modes of action readiness. In J. A. Russell & J.-M. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 78–102). New York: Cambridge University Press.
- Gendron, M., & Barrett, L. F. (2017). Facing the past: A history of the face in psychological research on emotion perception. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 15–36). Oxford: Oxford University Press.
- Gendron, M., Hoemann, K., Crittenden, A. N., Mangola, S. M., Ruark, G. A., & Barrett, L. F. (2020). Emotion perception in Hadza Hunter-Gatherers. *Scientific Reports*, 10(1), 1–17. <https://doi.org/10.1038/s41598-020-60257-2>.

- Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion, 14*(2), 251–262. <https://doi.org/10.1037/a0036052>.
- Hassin, R. R., Aviezer, H., & Bentin, S. (2013). Inherently ambiguous: Facial expressions of emotions, in context. *Emotion Review, 5*(1), 60–65. <https://doi.org/10.1177/1754073912451331>.
- Hess, U., & Fischer, A. H. (2016). *Emotional mimicry in social context*. New York: Cambridge University Press.
- Hess, U., & Kleck, R. E. (1990). Differentiating emotion elicited and deliberate emotional facial expressions. *European Journal of Social Psychology, 20*(5), 369–385. <https://doi.org/10.1002/ejsp.2420200502>.
- Jakobs, E., Manstead, A. S., & Fischer, A. H. (1999). Social motives and emotional feelings as determinants of facial displays: The case of smiling. *Personality and Social Psychology Bulletin, 25*(4), 424–435. <https://doi.org/10.1177/0146167299025004003>.
- Jed Wing, M. K. C., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., & Hunt, T. (2019). Caret: Classification and regression training. <https://CRAN.R-project.org/package=caret>. Accessed 6 Nov 2020.
- Johnson, K. J., Waugh, C. E., & Fredrickson, B. L. (2010). Smile to see the forest: Facially expressed positive emotions broaden cognition. *Cognition and Emotion, 24*(2), 299–321. <https://doi.org/10.1080/02699930903384667>.
- Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2013). Dynamic properties influence the perception of facial expressions. *Perception, 42*(11), 1266–1278. <https://doi.org/10.1068/p3131n>.
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. In: *Proceedings Fourth International Conference on Automatic Face and Gesture Recognition*, pp 46–53.
- Kappas, A. (2003). What facial activity can and cannot tell us about emotions. In M. Katsikitis (Ed.), *The human face* (pp. 215–234). Boston: Springer.
- Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior, 43*, 133–160. <https://doi.org/10.1007/s10919-019-00293-3>.
- Krause, R., Steimer-Krause, E., Merten, J., & Ullrich, B. (1998). Dyadic interaction regulation, emotion, and psychopathology. In W. F. J. Flack & J. D. Laird (Eds.), *Emotions and psychopathology: Theory and research* (pp. 70–80). New York: Oxford University Press.
- Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology, 37*(9), 1539–1553. <https://doi.org/10.1037/0022-3514.37.9.1539>.
- Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5*(1), 41–46. <https://doi.org/10.1177/1754073912451349>.
- Krumhuber, E. G., Küster, D., Namba, S., & Skora, L. (2020). Human and machine validation of 14 databases of dynamic facial expressions. *Behavior Research Methods, 52*(1), 8–20. <https://doi.org/10.3758/s13428-020-01443-y>.
- Lee, V., & Wagner, H. (2002). The effect of social presence on the facial and verbal expression of emotion and the interrelationships among emotion components. *Journal of Nonverbal Behavior, 26*(1), 3–25. <https://doi.org/10.1023/A:1014479919684>.
- Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J. A. (2006). Language and the perception of emotion. *Emotion, 6*(1), 125–138. <https://doi.org/10.1037/1528-3542.6.1.125>.
- Lindquist, K. A., & Gendron, M. (2013). What's in a word? *Language constructs emotion perception. Emotion Review, 5*(1), 66–71. <https://doi.org/10.1177/1754073912451351>.
- Lottridge, D., & Chignell, M. (2009). Emotional bandwidth: Information theory analysis of affective response ratings using a continuous slider. In: *IFIP Conference on Human-Computer Interaction*, 111–114.
- McDuff, D. (2016). Discovering facial expressions for states of amused, persuaded, informed, sentimental and inspired. *Proceedings of the International Conference on Multimodal Interaction*. <https://doi.org/10.1145/2993148.2993192>.
- McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., & el Kaliouby, R. (2016). AFFDEX sdk: A cross-platform real-time multi-face expression recognition toolkit. *Proceedings of the Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/2851581.2890247>.
- McKeown, G. J. (2013). The analogical peacock hypothesis: The sexual selection of mind-reading and relational cognition in human communication. *Review of General Psychology, 17*(3), 267–287. <https://doi.org/10.1037/a0032631>.
- Mehu, M., Grammer, K., & Dunbar, R. I. (2007). Smiles when sharing. *Evolution and Human Behavior, 28*(6), 415–422. <https://doi.org/10.1016/j.evolhumbehav.2007.05.010>.
- Meillon, B., Tcherkassof, A., Mandran, N., Adam, J.-M., Dubois, M., Dupré, D., Benoit, A.-M., Guérin-Dugué, A., & Caplier, A. (2010). DynEmo: A corpus of dynamic and spontaneous emotional facial expressions. In: *Proceedings of International Workshop Series on Multimodal Corpora, Tools and Resources. Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality*, pp 31–36.
- Müller, K. (2017). Here: A simpler way to find your files. <https://CRAN.R-project.org/package=here>. Accessed 6 Nov 2020.
- Niedenthal, P. M., Wood, A., Rychlowska, M., & Korb, S. (2017). Embodied simulation in decoding facial expression. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 397–414). New York: Oxford University Press.
- Nielsen, L., & Kaszniak, A. W. (2007). Conceptual, theoretical, and methodological issues in inferring subjective emotion experience. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 361–375). New York: Oxford University Press.
- Pacherie, E. (2003). Modes de structuration des contenus perceptifs visuels. In J. Bouveresse & J.-J. Rosat (Eds.), *Philosophies de la perception. Phénoménologie, grammairie et sciences cognitives* (pp. 263–289). Paris: Odile Jacob.
- R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>. Accessed 6 Nov 2020.
- Reisenzein, R. (2000). Exploring the strength of association between the components of emotion syndromes: The case of surprise. *Cognition and Emotion, 14*(1), 1–38. <https://doi.org/10.1080/026999300378978>.
- Reisenzein, R., Bördgen, S., Holtbernd, T., & Matz, D. (2006). Evidence for strong dissociation between emotion and facial displays: The case of surprise. *Journal of Personality and Social Psychology, 91*(2), 295–315. <https://doi.org/10.1037/0022-3514.91.2.295>.
- Reisenzein, R., Studtmann, M., & Horstmann, G. (2013). Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emotion Review, 5*(1), 16–23. <https://doi.org/10.1177/1754073912457228>.
- Rosenberg, E. L., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition and Emotion, 8*(3), 201–229. <https://doi.org/10.1080/02699939408408938>.
- Ruiz-Belda, M.-A., Fernández-Dols, J.-M., Carrera, P., & Barchard, K. (2003). Spontaneous facial expressions of happy bowlers and

- soccer fans. *Cognition and Emotion*, 17(2), 315–326. <https://doi.org/10.1080/02699930302288>.
- Russell, J. A. (1993). Forced-choice response format in the study of facial expression. *Motivation and Emotion*, 17(1), 41–51. <https://doi.org/10.1007/BF00995206>.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, 115(1), 102–141. <https://doi.org/10.1037/0033-2909.115.1.102>.
- Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J.-M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology*, 54(1), 329–349. <https://doi.org/10.1146/annurev.psych.54.101601.145102>.
- Russell, J. A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *Journal of Experimental Psychology: General*, 116(3), 223–237. <https://doi.org/10.1037/0096-3445.116.3.223>.
- Sander, D., & Scherer, K. (2014). *Oxford companion to emotion and the affective sciences*. New York: Oxford University Press.
- Scherer, K. R., & Ceschi, G. (1997). Lost luggage: A field study of emotion–antecedent appraisal. *Motivation and Emotion*, 21(3), 211–235. <https://doi.org/10.1023/A:1024498629430>.
- Schützwohl, A., & Reisenzein, R. (2012). Facial expressions in response to a highly surprising event exceeding the field of vision: A test of darwin’s theory of surprise. *Evolution and Human Behavior*, 33(6), 657–664. <https://doi.org/10.1016/j.evolhumbeh.av.2012.04.003>.
- Tcherkassof, A., Bollon, T., Dubois, M., Pansu, P., & Adam, J.-M. (2007). Facial expressions of emotions: A methodological contribution to the study of spontaneous and dynamic emotional faces. *European Journal of Social Psychology*, 37(6), 1325–1345. <https://doi.org/10.1002/ejsp.427>.
- Tcherkassof, A., Dupré, D., Meillon, B., Mandran, N., Dubois, M., & Adam, J.-M. (2013). DynEmo: A video database of natural facial expressions of emotions. *The International Journal of Multimedia & Its Applications*, 5(5), 61–80. <https://doi.org/10.5121/ijma.2013.5505>.
- Tcherkassof, A., & Frijda, N. H. (2014). Les émotions: Une conception relationnelle. *L’Année Psychologique*, 114(3), 501–535.
- Tsai, J. L., Chentsova-Dutton, Y., Freire-Bebeau, L., & Przymus, D. E. (2002). Emotional expression and physiology in european americans and hmong americans. *Emotion*, 2(4), 380–397. <https://doi.org/10.1037/1528-3542.2.4.380>.
- Vernon, L. L., & Berenbaum, H. (2002). Disgust and fear in response to spiders. *Cognition and Emotion*, 16(6), 809–830. <https://doi.org/10.1080/02699930143000464>.
- Wagner, H. L. (1997). Methods for the study of facial behavior. In J. A. Russell & J.-M. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 31–54). Cambridge: Cambridge University Press.
- Wagner, H. L., MacDonald, C. J., & Manstead, A. S. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of Personality and Social Psychology*, 50(4), 737–743. <https://doi.org/10.1037/0022-3514.50.4.737>.
- Wickham, H. (2017). Tidyverse: Easily install and load the ‘tidyverse’. <https://CRAN.R-project.org/package=tidyverse>. Accessed 6 Nov 2020.
- Zhu, H. (2019). KableExtra: Construct complex table with ‘kable’ and pipe syntax. <https://CRAN.R-project.org/package=kableExtra>. Accessed 6 Nov 2020.

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