Acting Surprised: Comparing Perceptions of Different Dynamic Deliberate Expressions

Mircea Zloteanu[1] [ID] **· Eva G. Krumhuber[2] · Daniel C. Richardson[2]**

Abstract

People are accurate at classifying emotions from facial expressions but much poorer at determining if such expressions are spontaneously felt or deliberately posed. We explored if the method used by senders to produce an expression influences the decoder's ability to discriminate authenticity, drawing inspiration from two well-known acting techniques: the Stanislavski (internal) and Mimic method (external). We compared spontaneous surprise expressions in response to a jack-in-the-box (genuine condition), to posed displays of senders who either focused on their past affective state (internal condition) or the outward expression (external condition). Although decoders performed better than chance at discriminating the authenticity of all expressions, their accuracy was lower in classifying external surprise compared to internal surprise. Decoders also found it harder to discriminate external surprise from spontaneous surprise and were less confident in their decisions, perceiving these to be similarly intense but less genuine-looking. The findings suggest that senders are capable of voluntarily producing genuine-looking expressions of emotions with minimal effort, especially by mimicking a genuine expression. Implications for research on emotion recognition are discussed.

Keywords Facial expression · Emotion · Genuineness · Authenticity discrimination

Introduction

Emotions are integral to human interaction, having a strong impact on how we perceive others. However, emotional expressions can be either genuine or non-genuine. Spontaneously felt emotional expressions genuinely reflect the underlying affect of a person, whereas deliberately posed expressions reflect the communicative intent of the sender. Posed expressions are a strategic tool, with uses ranging from general social compliance to outright deception (Ekman and Rosenberg 2005).

Emotion recognition refers broadly to the processing and recognition of facial displays, as well as ascribing an expression to an emotional category. Researchers typically

✉ Mircea Zloteanu
    m.zloteanu@kingston.ac.uk

1   Department of Criminology and Sociology, Kingston University London,
    Kingston Upon Thames KT1 2EE, UK

2   University College London, London, UK

ask individuals (i.e., decoders) to assign pre-defined emotion labels to specific facial configurations (i.e., Action Units; Ekman et al. 2002) using a forced-choice procedure (for a review, see Calvo and Nummenmaa 2015), intending to compare classification scores among various populations (e.g., Carroll and Russell 1997; Marsh and Blair 2008).

This research has demonstrated that people can recognize with high accuracy (> 70%) what emotions are indicated by particular facial expressions (Matsumoto et al. 2008; Nelson and Russell 2013). However, when it comes to determining the affective authenticity of such expressions, their ability is lower (e.g., Ekman and Friesen 1974; Ekman and O'Sullivan 1991; Hess and Kleck 1994; Porter and ten Brinke 2008; Porter et al. 2012) and more variable (e.g., Boraston et al. 2008; Manera et al. 2011).

People's poor ability to discriminate emotion authenticity has been argued to stem from two (complementary) factors. Using a decoder-centric perspective, decoders rely on incorrect cues when attempting to determine authenticity (Ekman and O'Sullivan 1991) or lack the perceptual, attentional, and/or cognitive ability to discriminate affective authenticity (Perron and Roy-Charland 2013). Using a sender-centric perspective, behavioral differences (i.e., morphologic or dynamic) between genuine and non-genuine emotional displays may be nonexistent or imperceptible to the human eye (Namba et al. 2016; Porter et al. 2012). As a result, senders can utilize posed expressions to communicate false affective states (i.e., displays which are [mis-]interpreted by decoders as reflecting genuine affect, when the sender's underlying affect does not match; Gosselin et al. 2010; Gunnery et al. 2013; Krumhuber et al. 2014) or to strengthen affective signaling (e.g., exaggerating genuine emotional states in appropriate social settings; Fridlund 1991).

The literature on authenticity discrimination ability generally suffers from two limitations. First, human authenticity discrimination ability assumes the existence of behavioral differences between spontaneous and posed expressions of emotion (originating from research on the Duchenne smile; Duchenne 1862/1990), which has spurred a belief in facial markers of emotion authenticity, called reliable muscles (Ekman 2003). This has produced an overreliance on stimuli pre-selected based on specific muscle activations in investigations of affective authenticity discrimination (see Gunnery and Ruben 2016), and imposing an appearance-based dichotomy that may not exist (Kappas et al. 2013).

If the research question relates to authenticity discrimination, a comparison based on the elicitation method is more appropriate (Krumhuber et al. 2016). Ironically, studies exploring perceptual differences between alleged genuine and non-genuine smiles often rely on stimuli produced by actors posing both types of expressions (Bernstein et al. 2008; Calvo et al. 2013; Gosselin et al. 1995), undermining the notion of authenticity. Research has demonstrated that spontaneous and posed smiles, as measured by facial markers, can be easily produced in the absence of an underlying affect (see Gunnery et al. 2013). As a result, some scholars have argued that internal affective states and external displays can and should be treated as two different phenomena (Gunnery and Hall 2014).

When using an appearance-based approach for selecting stimuli, binarizing facial expressions into spontaneous versus posed transforms an authenticity discrimination task into a categorization task, where "accuracy" reflects the decoder's ability to group similar items. Under this perspective, authenticity discrimination is a measure of the decoder's ability to perceive specific behavioral cues and judge them correctly, assuming such cues reflect genuine affect. This conflates judgments of facial appearance with judgments of sender intent. The spontaneous-posed terminology should refer to the method used to produce the facial expression, while the genuine-non-genuine terminology should refer to decoders' judgments and inferences.

Second, the umbrella term "posed" obfuscates the various methods available to voluntarily produce facial expressions of emotions, overlooking differences in their appearance and perception. Typically, deliberately posed expressions are produced by providing senders with specific instructions. For instance, specifying which facial muscles to activate (see Directed Facial Action Task; Ekman et al. 1983), by asking senders to imitate a photograph of an expression (Stöckli et al. 2018), by providing a verbal prompt (Lewis et al. 1987), or through a combination of such approaches (Gosselin et al. 2002). They can also be produced through specific acting techniques, such as asking senders (typically, professional actors) to recall a congruent past affective experience or by asking them to act out a specific scenario (Gur et al. 2002; Scherer and Bänziger 2010). Deliberate expressions can also occur in response to specific social contexts even spontaneously and without explicit instructions, such as smiling for a photograph (Vazire et al. 2009), or to enable communication and learning (e.g., mother-infant interactions; Chong et al. 2003).

Decoders may respond differently to a posed expression based on the underlying elicitation method (e.g., Douglas et al. 2012; McLellan et al. 2010; Sauter and Fischer 2018; Soppe 1988). Treating all posed displays uniformly will result in perceptual differences from the elicitation method being lost, leading to incorrect, contradictory, and misleading inferences. Moreover, genuine, spontaneously felt emotional displays can also be elicited using various methods that may produce different outcomes (see Siedlecka and Denson 2019). For instance, one could use emotion-evoking imagery (Krumhuber and Manstead 2009) or social interaction (Vallverdu 2015). For a review of elicitation techniques, see Coan and Allen (2007).

The genuine-non-genuine terminology can produce confusion as it often represents different underlying constructs (Scherer and Bänziger 2010). A conceptual distinction should be made between stimulus features (e.g., facial markers) and sender veracity/intent. Appearance-based approaches make assumptions regarding genuine and non-genuine displays (e.g., the presence/absence of specific markers) and impose constraints on which exemplars are representative of the spontaneous-posed dimension (e.g., even excluding genuinely felt spontaneous displays if they lack specific markers). Indeed, decoders' perceptions of the genuineness of specific expressions do not always match their elicitation or production method, even showing the opposite perceptual patterns (Dawel et al. 2017).

In emotion research, it is crucial to consider the elicitation method for the spontaneous-posed dimension. Spontaneous emotional displays, aimed to reflect genuinely felt affect, must be elicited by engineering circumstances so that the target emotion is evoked, while deliberately posed emotional displays must reflect attempts to produce a genuine-looking emotional display in the absence of genuine underlying affect.

Thus, researchers should treat their explorations as a two-fold process: (1) classification accuracy-the ability to correctly categorize (i.e., label) facial expressions into emotion categories, (2) authenticity discrimination-the ability to determine if an emotional display reflects the genuine underlying affect of the sender or a non-genuine, deliberately posed display absent of underlying affect.

Presently, we aim to illustrate the importance of this operationalization on decoder judgments, by providing a comparison of emotion perception as a factor of the type of production method used to generate posed expressions. In a recent study, Zloteanu et al. (2018) found that the method used to produce emotional expressions impacted decoders' inferences. For this, they compared perceptions of spontaneous surprise (i.e., an emotional reaction to an affect-evoking stimulus) with two deliberately posed expressions: rehearsed (i.e., by reproducing a recent genuine affective event) and improvised (i.e., by relying on one's own beliefs to produce the display). While spontaneous surprise was correctly detected as

having occurred in the presence of a genuinely surprising event (a jack-in-the-box) fairly often (and rated the most genuine-looking and intense), rehearsed surprise was the more difficult to detect. Rehearsed surprise was also rated more genuine-looking in appearance, but improvised surprise was perceived as more intense. Hence, the method used to produce a posed expression affected perceptions of emotional authenticity and several other dimensions.

The present study extends the work of Zloteanu et al. (2018) by exploring additional methods for producing posed expressions. Two new methods were investigated. Facial expressions were produced (a) while the sender relies on a past affective memory, which we term the internal condition, and (b) by mimicking a genuine facial display of another individual, which we term the external condition. We drew inspiration from two well-known acting methods: the Mimic and the Stanislavski method (see Hull 1985). We note that the current experiment does not concern acting or trained actors per se (for an example of such work, see, Conson et al. 2013). We use this terminology only to conceptualize the difference between external and internal components of the deliberate emotional display.[1]

According to the Stanislavski method, emotions can be posed convincingly by recalling a previous affectively-congruent episode (Hull 1985). Using this perspective, a convincing sender only needs to draw on a memory that is affectively congruent with the emotion they wish to portray. It has been argued that relying on internal affective simulations may allow individuals to recreate the genuine outward expression more reliably, due to the congruent underlying affect, but may be insufficient to produce intense facial displays (Ekman et al. 1983) or activate all the expected facial muscles (Reisenzein et al. 2006).

Alternatively, according to the Mimic method, emotional expressions can be produced by mimicking the behavior of individuals who are experiencing genuine emotion. In this approach, seeing a genuine expression provides sufficient information for the sender to produce a genuine-looking expression at will, without the need for a direct experience of the underlying emotion (Hull 1985). However, research on mimicry argues such expressions may be perceived as less genuine, due to their caricatured appearance (Mehu et al. 2012), and lack of underlying affect needed to produce an emotionally appropriate   display (Hess et al. 1995).

The Present Study

In this paper, the focus is on a single emotion: surprise. Surprise is considered a basic emotion, having a distinctive facial configuration that is well recognized cross-culturally (Nelson and Russell 2013). To elicit it reliably, we used the startle response, a sudden defensive response to an external aversive stimulus. We employed a jack-in-the-box toy, an approach that has been successful in the past for inducing surprise (Bennett et al. 2002; Reissland et al. 2002).

---

[1] The use of actors for stimulus production is a debated topic. It is argued that actors generate more reliable and standardized displays (for details, see Scherer and Bänziger 2010), which are more intense and unambiguous (Cohn et al. 2007). Acting approaches focus on "recognizability" (see Cowie et al. 2005). However, the use of actors has been criticized due to the intentional communicative nature of these performances (Sauter &and Fischer 2018). Research using actors asks decoders "what emotion is being portrayed?" but not "what emotion is the sender feeling?".

Emotions were judged by decoders from dynamic video stimuli, as they more closely resemble real-life expressions (Arsalidou et al. 2011), thereby offering better ecological validity (Krumhuber et al. 2013). Videos are also beneficial to discriminating expression authenticity (Zloteanu et al. 2018), allowing for more subtle elements of an expression to be incorporated into the decoding process (e.g., timing, duration, fluidity; Tobin et al. 2016).

Decoders were presented with surprise expressions elicited in several ways: genuine, external, and internal. The genuine condition expressions reflect senders' spontaneous response to seeing the jack-in-the-box for the first time (see Zloteanu et al. 2018). The external condition expressions reflect senders' attempts to mimic a spontaneous expression after seeing an example from another person, while the internal condition expressions reflect senders' use of their affective memory of being surprised previously.

Emotion perception was measured on several dimensions: authenticity discrimination, perceived expression genuineness, perceived expression intensity, and confidence in judgment. Authenticity discrimination reflects the accuracy with which decoders can detect that the expression was produced in response to an emotion-evoking stimulus—spontaneous—or the absence of such a stimulus—posed.[2] Genuineness is the degree to which the emotional display is perceived as having occurred in response to an emotion-evoking stimulus, regardless of veracity. Considering judgments in this fashion allows for a separation of discrimination ability and response bias (Gosselin et al. 1995). Intensity reflects how strongly the display is perceived. Finally, confidence is used to measure decoders' perceptions and awareness of their ability to assess authenticity.

We predicted that decoders would show differences in how they perceive the three expression conditions. Specifically, authenticity discrimination and judgment confidence should be affected by the production method used, with spontaneous expressions (genuine condition) being rated more accurately and confidently than the two posed expression types (internal and external conditions). External surprise was hypothesized to be harder to detect as a posed expression (compared to internal surprise) due to the more intense nature of its appearance, as the sender focuses on the outward expression for mimicking a believable emotional display. In contrast, internal surprise should appear more genuine, as the sender relies on a past affectively congruent experience.

Method

Stimuli

The stimulus set used in this research comprised thirty-nine videos of university students (13 male, 26 females; age range: 18-36 years; $M_{age} = 21.32$, $SD = 4.45$) who were video recorded in one of three expression conditions: internal simulation, external simulation, or spontaneous expressions. The internal and external videos were created specifically for this study, while for the spontaneous expressions, the videos from Zloteanu et al. (2018)

---

[2] This operationalization of authenticity discrimination can also be referred to as "congruency" (see Dawel et al. 2017) where discrimination occurs if (a) a display is judged as "genuine" and the sender was experiencing the emotion or (b) the display is judged as "non-genuine" and the sender was not experiencing the emotion.
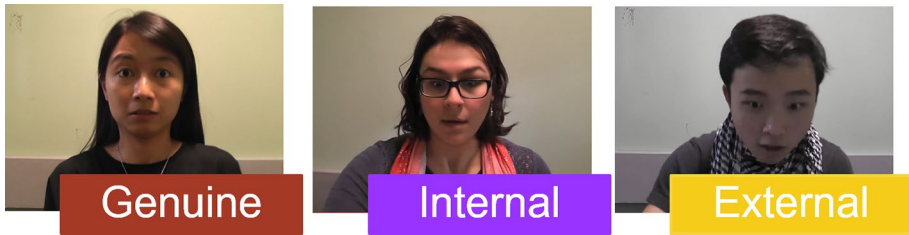
Fig. 1 Stimuli used in the study, illustrating the three types of surprise expressions

genuine condition were utilized. A Panasonic SDR-T50 camcorder was used to record the facial reactions at a resolution of 1920×1080 pixels and 25 fps.

In the genuine condition, encoders were seated in front of the jack-in-the-box and turned the crank until the toy "popped out". A melody played as the crank was turned, prompting the action from the toy. Their reaction was video recorded from the start of the winding action until the end of their behavioral response.

For the internal condition, encoders were instructed to focus during the task on their internal feeling of the emotion over the outward behavior they generated. They then experienced the jack-in-the-box, as in the genuine condition. After a short break, the jack-in-the-box crank was disconnected from the releasing mechanism and a tablet with a video countdown was placed in front of the toy. Participants were instructed to recollect the internal state they had experienced and react when prompted. When the countdown finished the word "NOW" appeared on screen, prompting the response from the participants, which was video recorded. The countdown video had the same melody and timing as the original toy.

In the external condition, participants first viewed a randomly selected video from one of the persons in the genuine condition and were told to study their behavioral reaction. Afterward, the inoperable toy and the tablet were placed in front of the participants. Participants were recorded while reproducing the expression they had seen when the word "NOW" appeared.

In total, 44 videos were recorded for the current study. Of these, excluding videos that contained recording issues (n=2) and senders not following instructions (n=4), 26 were selected. 13 videos were used for the internal (5 men, 8 women) and external (4 men, 9 women) conditions, alongside 13 genuine (4 men, 9 women) videos which were re-used from Zloteanu et al. (2018). The videos start from the moment the sender cranks the toy wheel until the end of their behavioral response. All videos are presented in color, without sound, and last approximately 10 s (see Fig. 1).

Participants and Design

The study employed a within-subjects design with three levels of the independent variable (Genuine, Internal, External). A total of 102 participants were recruited online through Amazon Mechanical Turk in exchange for $0.75. After deleting incomplete cases (n=52) the final data encompassed 50 participants (14 males, 36 females), between the ages of 18

to 50 years, (M = 25.0, SD = 7.2).[3] An a priori power analysis (G*Power 3.1.9.2; Faul et al. 2007), assuming a medium-sized effect of condition (Cohen's f = 0.25), determined that this sample size is sufficient to achieve 80% power. Informed consent was received from all participants. Ethical approval for the present study was granted by the Department of Psychology ethics committee.

Procedure

The study was conducted using Qualtrics (Provo, UT), a web-based platform. Participants were told that they would watch a series of videos of facial expressions of emotions. They were instructed to watch each video carefully and rate the behavioral response of the person in the video. It was made clear that some senders were genuinely reacting to a jack-in-the-box while others never saw the toy popping out and were merely attempting to appear surprised. Because mood can affect classification accuracy (Schmid et al. 2011), it was necessary to control for this factor by asking the following question: "How do you feel at this moment?" with a Likert-type scale response from 1 (extremely sad) to 5 (extremely happy).

Participants saw all 39 videos in randomized order. They first rated the extent to which they believed that the expression was produced while the person saw a jack-in-the-box; the scale ranged from -2 (certain NO jack-in-the-box), with a midpoint of 0 (not sure) to 2 (certain WITH jack-in-the-box). This question served to measure both authenticity discrimination (i.e., decoder accuracy) and perceptions of expression genuineness.[4] This was followed by ratings of decision confidence and perceived intensity of the expression, using a 5-point Likert scale anchored at 1 (not at all) and 5 (very much). At the end, all participants were fully debriefed.

Results

Preliminary analyses revealed no effects of participant gender (Fs < 1.24, ps > 0.30, JZS $BF_{01} > 4$), or mood (Fs < 2.00, ps > 0.10, JZS $BF_{01} > 2$); hence, these two factors were excluded from further analysis.

Genuineness To investigate how genuine the three expression types appeared, the responses decoders provided for the jack-in-the-box question for each expression were summed across the videos within each condition (range -26–26), and then averaged across decoders. Scores > 0 represent that the expressions within that condition were rated as appearing more genuine, and scores < 0 representing a more deliberate rating. This

---

[3] The attrition rate reported reflects the liberal definition of a "participant" as any individual who progressed past the online consent form screen of the study, even if they terminated the experiment before seeing the first stimulus exemplar.

[4] Operationalizing the scale in this fashion offers a separation akin to sensitivity and decision criterion in Signal Detection Theory. It allows for the exploration of scenarios where accuracy rates between two conditions are equitable (e.g., 40% in both) yet genuineness ratings are more polarized or consistent in one condition (e.g., "-2" ratings in condition A and "-1" in condition B), and for scenarios where genuineness ratings in two conditions are equitable (e.g., "-2" for every correct detection) yet differing accuracy scores (e.g., "60%" in condition A and "33%" in condition B).

Fig. 2 Violin plots for decoder accuracy in each expression condition. The shaded areas detail the distribution of the data in each condition. The dot inside each plot represents the mean authenticity discrimination score (error bars ± 1 SE)
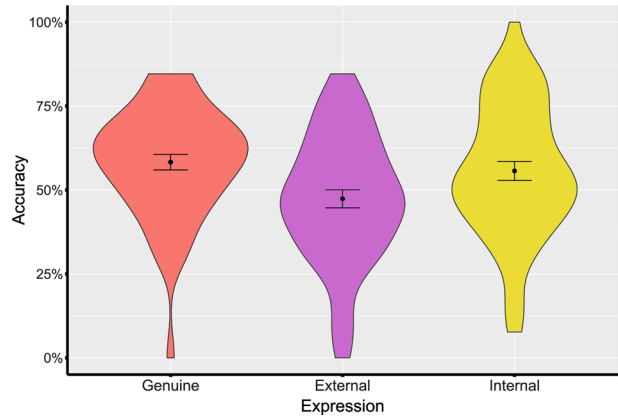
provides a metric for the magnitude in the perception of genuineness, separate from overall discrimination performance.

The analysis revealed a main effect of expression condition, $F(2, 98) = 54.75$, $p < 0.001$, $\eta^2 = 0.53$, 90% CI [0.41, 0.61], JZS $BF_{10} = 8.05e^{13}$ (decisive evidence for H_A; Wetzels et al. 2011). Genuine condition expressions were perceived as the most genuine ($M = 4.58$, $SD = 5.77$), followed by external condition expressions ($M = -2.16$, $SD = 7.10$), and lastly by internal condition expressions, which were rated as the most non-genuine ($M = -5.64$, $SD = 7.29$). Post-hoc tests (Bonferroni-corrected alpha-level) revealed that external condition expressions were seen as more genuine than internal condition expressions, $t(49) = 4.14$, $p < 0.001$, 95% CI [1.79, 5.17], $d_z = 0.59$, JZS $BF_{10} = 170.7$ (decisive evidence for H_A). Both posed conditions, however, were rated lower than genuine condition expressions, $t(49) = 6.48$, $p < 0.001$, 95% CI [4.65, 8.83], $d_z = 0.92$, JZS $BF_{10} = 3.24e^5$ (external condition; decisive evidence for H_A), and $t(49) = 9.45$, $p < 0.001$, 95% CI [8.05, 12.39], $d_z = 1.34$, JZS $BF_{10} = 6.58e^9$ (internal condition; decisive evidence for H_A).

Accuracy For authenticity discrimination, ratings were collapsed to form three possible states: -2 and -1 were coded as "posed", 0 was coded as "not sure", while 1 and 2 were coded as "spontaneous". These were matched to the experimental condition, such that if a decoder saw a genuine condition expression and responded with "spontaneous" it was considered "accurate" (score = 1). If there was a mismatch, it was treated as "inaccurate" (score = 0). The reverse was true for the internal and external conditions. Responses of "not sure" were treated as incorrect (score = 0). This yielded an accuracy score of correct detections out of the 13 exemplars per condition, which was converted to a percentage value.

On average 58.31% ($SD = 16.31$) of the genuine condition expressions were correctly classified, 55.70% ($SD = 20.00$) of the internal condition expressions, and 47.40% ($SD = 19.00$) of the external condition expressions. A repeated-measures ANOVA yielded an overall effect of expression condition on accuracy, $F(1.47, 71.81) = 5.14$, $p = 0.015$, $\eta^2 = 0.10$, 90% CI [0.01, 0.20] (Greenhouse–Geisser corrected), JZS $BF_{10} = 7.33$ (substantial evidence for H_A). Subsequent repeated-measures t-tests (Bonferroni-corrected alpha) revealed a significant difference between genuine and external condition expressions, $t(49) = 2.69$, $p = 0.01$, 95% CI [0.36, 2.48], $d_z = 0.38$, JZS $BF_{10} = 3.82$ (substantial evidence for H_A), and between internal and external condition expressions, $t(49) = 3.71$, $p < 0.001$, 95% CI [0.50, 1.67], $d_z = 0.52$, JZS $BF_{10} = 50.38$ (very strong evidence for H_A), indicating that external condition expressions were harder to accurately identify as posed. The

difference between genuine and internal condition expressions was non-significant, $t < 1$, $p = 0.522$, JZS $BF_{10} = 0.19$ (see Fig. 2).

When comparing accuracy rates to chance level (33%), genuine condition expressions were discriminated with above chance performance, $t(49) = 10.83$, $p < 0.001$, 95% CI [2.57, 3.81], $d_z = 1.53$, JZS $BF_{10} = 5.33e^{11}$ (decisive evidence for H_A), as were external, $t(49) = 5.23$, $p < 0.001$, 95% CI [1.13, 2.51], $d_z = 0.74$, JZS $BF_{10} = 5.11e^3$ (decisive evidence for H_A), and internal condition expressions, $t(49) = 7.91$, $p < 0.001$, 95% CI [2.15, 3.67], $d_z = 1.12$, JZS $BF_{10} = 3.97e^7$ (decisive evidence for H_A).

Confidence For ratings of confidence, analyses revealed a main effect of expression condition, $F(2, 98) = 21.02$, $p < 0.001$, $\eta^2 = 0.30$, 90% CI [0.17, 0.40], JZS $BF_{10} = 4.12e^5$ (decisive evidence for H_A). Decoders had reduced confidence in their discrimination ability for expressions from the internal ($M = 47.62$, $SD = 7.07$) and external ($M = 46.90$, $SD = 7.28$) conditions compared to the genuine condition ($M = 50.50$, $SD = 7.59$), $t(49) = 5.07$, $p < 0.001$, 95% CI [1.74, 4.02], $d_z = 0.71$, JZS $BF_{10} = 3.02e^3$ (decisive evidence for H_A), $t(49) = 5.80$, $p < 0.001$, 95% CI [2.35, 4.85], $d_z = 0.82$, JZS $BF_{10} = 3.29e^4$ (decisive evidence for H_A), but showed no difference in confidence between the two posed expressions, $t(49) = 1.26$, $p = 0.214$, 95% CI [-0.43, 1.87], JZS $BF_{10} = 0.32$.

Intensity Finally, a main effect of expression condition was found for ratings of intensity, $F(2, 98) = 35.09$, $p < 0.001$, $\eta^2 = 0.42$, 90% CI [0.29, 0.51], JZS $BF_{10} = 2.04e^9$ (decisive evidence for H_A). This revealed that external condition expressions ($M = 36.70$, $SD = 6.49$) were rated equally intense to genuine condition expressions ($M = 38.12$, $SD = 6.03$), $t(49) = 2.26$, $p = 0.028$, 95% CI [0.16, 2.68] (non-significant after Bonferroni corrections, $p = 0.017$), JZS $BF_{10} = 1.57$ (anecdotal evidence for H_A). Additionally, both external and genuine condition expressions received higher intensity ratings than internal condition expressions ($M = 32.72$, $SD = 7.13$), $t(49) = 6.46$, $p < 0.001$, 95% CI [2.74, 5.22], $d_z = 0.91$, JZS $BF_{10} = 3.03e^5$ (decisive evidence for H_A), and, $t(49) = 7.17$, $p < 0.001$, 95% CI [3.89, 6.91], $d_z = 1.01$, JZS $BF_{10} = 3.29e^6$ (decisive evidence for H_A).

Relationship Between Measures To explore the relationship between the dependent measures, correlations were conducted for each expression condition (see Table 1). This revealed an expected pattern of results for accuracy and genuineness, with the direction of the relationship being correlated to the veracity of the expression. A moderate relationship between accuracy and confidence was observed, but only in the genuine condition. Considering statistical significance and Bayes factors, there was no strong evidence for any other relationship.

Discussion

People can quickly and accurately recognize facial expressions of emotions but find it harder to determine their authenticity. Here, we show that the method used to produce deliberately posed emotional expressions impacts authenticity discrimination and perception. Overall, the genuine condition expressions were rated as the most genuine-looking and intense, yielding the highest authenticity discrimination accuracy and judgment confidence when compared to the two posed expression conditions, corroborating findings on genuine and non-genuine dynamic expressions (Zloteanu et al. 2018). The way posed expressions were produced resulted in substantial differences in perception. External condition surprise was harder to accurately classify as posed and was perceived as more

Table 1 Correlations between dependent measures for each expression condition

Condition		Spearman's rho	BF_{10}	p	95% CI Lower	Upper
External						
Accuracy	Confidence	0.197	0.644	0.171	−0.086	0.450
Accuracy	Intensity	−0.245	0.836	0.087	−0.490	0.036
Accuracy	Genuineness	−0.917[***,+++]	>999	<.001	−0.952	−0.858
Confidence	Intensity	0.329[*]	0.440	0.020	0.056	0.556
Confidence	Genuineness	−0.206	0.373	0.152	−0.458	0.077
Intensity	Genuineness	0.146	0.656	0.312	−0.138	0.408
Internal						
Accuracy	Confidence	0.175	1.205	0.224	−0.108	0.432
Accuracy	Intensity	−0.126	0.317	0.385	−0.390	0.158
Accuracy	Genuineness	−0.955[***,+++]	>999	<.001	−0.975	−0.922
Confidence	Intensity	0.173	0.296	0.230	−0.111	0.430
Confidence	Genuineness	−0.227	1.655	0.112	−0.476	0.054
Intensity	Genuineness	0.216	0.709	0.132	−0.066	0.466
Genuine						
Accuracy	Confidence	0.315[*,+]	9.911	0.026	0.040	0.545
Accuracy	Intensity	0.276	0.461	0.052	−0.002	0.515
Accuracy	Genuineness	0.869[***,+++]	>999	<.001	0.780	0.924
Confidence	Intensity	0.324[*]	1.275	0.022	0.051	0.553
Confidence	Genuineness	0.172	0.674	0.233	−0.112	0.430
Intensity	Genuineness	0.152	0.273	0.294	−0.132	0.412

[*] $p < 0.05$, [**] $p < 0.01$, [***] $p < 0.001$; [+] $BF_{10} > 10$, [++] $BF_{10} > 30$, [+++] $BF_{10} > 100$

genuine and intense than internal surprise. Internal condition surprise was rated the least genuine-looking, had the lowest ratings of intensity, and was easily detected as posed.

The data also suggest that people possess some ability to discriminate spontaneous from posed expressions of emotion as accuracy was consistently above chance level. Considering the correlations, genuineness ratings were strongly positively correlated with accuracy rates for spontaneous expressions, but strongly negatively correlated with accuracy rates for posed expressions.

External condition expressions were perceived equal in intensity to genuine condition expressions, supporting claims that posed expressions should appear intense as senders want their message to be clear (Conson et al. 2013; Sauter and Fischer 2018) and contradicting claims that they should appear less intense due to the lack of underlying affect (Hess et al. 1995; Hess et al. 1997). By contrast, internal condition expressions were rated low in intensity, which converges with claims that the affective memory of an emotion is insufficient for an intense reproduction (Ekman et al. 1983).

These results provide insight into the contentious issue of expression intensity and emotional authenticity (Dawel et al. 2015; Ekman et al. 1983; Hess et al. 1995; Thibault et al. 2009). Here, the internal condition expressions were perceived as less intense than the genuine condition expressions, indicating a positive relationship between discrimination and intensity. In comparison, the external condition expressions were perceived as equally intense to the genuine condition expressions, indicating no relationship between discrimination and intensity. Considering all the results, expression intensity does not appear to be a marker of emotional authenticity, but more a product of the elicitation method employed (see also Zloteanu et al. 2018).

Judgmental confidence also varied as a function of expression type. Specifically, it was lower for the two posed expressions conditions than the spontaneous expressions condition. However, this did not translate into improved authenticity discrimination. Looking at the correlations between accuracy and confidence indicate that when a decoder made a confidence judgment on a spontaneous expression, they tended to give a higher rating, but no reliable pattern emerged for judgments of posed displays. This may indicate that decoders possess an innate perceptual ability to detect the underlying veracity of expressions that is not captured by their overt judgment, paralleling research on unconscious lie detection (see DePaulo et al. 1997).

The "deceptive" superiority of the external condition expressions may have resulted from senders using a spontaneous, felt reaction as their target expression. This is supported by research on facial mimicry, where the reference expression—spontaneous or posed—has been found to affect the mimicked display (Gunnery et al. 2013; Lundqvist and Dimberg 1995). Conversely, the higher discriminability for the internal condition expressions may be due to the added complexity of the task the senders had to perform, minding both presentation and timing while controlling their nonverbal channels (see Gunnery et al. 2013; Zuckerman et al. 1981).

Implications

For the emotion expression literature, the present results have pertinent methodological implications. It is evident that there are not only perceptual differences between posed and spontaneous expressions but also between different types of posed expressions. This supports our argument that specificity regarding the production method employed is important. Relying on a simple spontaneous-posed dichotomy would not have provided a complete explanation of the findings, ignoring perceptual differences between the internal and external conditions.

We hope to have argued convincingly that assessing emotion recognition ability should be a two-fold process. First, the ability to categorize an expression based on emotional content (i.e., classification accuracy). Second, the ability to determine if an emotion reflects true affective content as felt by the sender (i.e., authenticity discrimination). There is value in distinguishing between the two when investigating emotion recognition ability, and we caution that aggregating the two abilities can obscure relevant effects and produce incorrect conclusions regarding human emotion recognition.

Approaches focusing on matching expressions to emotional categories reflect the process of agreeing that a sender accurately depicted the emotion they were supposed to display (e.g., frowning person at a funeral) without considering the underlying affect (i.e., are they actually sad?). If the aim is to test people's affective authenticity discrimination, then the veracity (or intent) of the sender should be the operationalizing factor. If the aim is to understand differences in the ability to categorize facial displays, an appearance-based approach may be suitable, with the caveat that the findings only speak to overt categorization ability.

Limitations

A limitation of the current approach, especially regarding real-world authenticity discrimination, is the rarity of such isolated, intense, and recognizable expressions occurring in day-to-day interactions (Scherer and Bänziger 2010). Our results represent a "best-case scenario" for human performance.

A methodological limitation is the presence of the camera and senders' knowledge that they were being recorded. Research has indicated that anticipation, context, social desirability, and display rules can impact expression presentation (Ekman and Friesen 1982; Ekman et al. 2005; Scherer and Bänziger 2010). We cannot know the impact this may have had on reactions and performances.

For our external condition expressions, the target expression used must be considered. Although randomly selected from the spontaneous expression videos in Zloteanu et al. (2018), the exemplar used may have impacted the produced expressions, and by extension decoders' ratings. The use of multiple exemplars could have improved the reliability of our inferences, but at the cost of added variability and heterogeneity in performances, creating two sources of noise in the data (first from individual differences in sender ability and second from differences between exemplars; Coan and Allen 2007). Considering our aims, the ratings for the genuine and external condition expressions speak favorably towards the ability to successfully mimic a genuine-looking expression of surprise.

The current design did not allow for an exploration of gender-specific effects between decoders and senders. However, future expansions should consider this interaction, given

gender differences in expression production (Brody and Hall 2008) and judgment (Gunnery and Ruben 2016).

Future Directions

Future expansions should focus on the methods through which posed expressions are produced in real-world interactions, comparing successful and unsuccessful performances. For example, convicts with psychopathic traits (e.g., flat affect) are better at deceiving others about being remorseful (Porter et al. 2009), supporting our assertion that knowledge of an emotional display is more important than the affect corresponding to said emotion. The present superiority of the external condition expressions supports this view.

Close consideration must also be given to the emotion being investigated. Decoders show variability in recognizing different emotions (e.g., surprise being highly recognizable; Gosselin et al. 1995), while senders show differences in their ability to voluntarily produce different emotional expressions (Gosselin et al. 2010).

Here, we focused on exploring the human perception of posed and spontaneous expressions. It falls to future research to analyze how such dynamic expressions differ (i.e., behaviorally) and which objective markers of affective authenticity, if any, separate genuine from non-genuine emotional displays. A machine-learning approach may reveal quantifiable and diagnostic differences between expressions under different elicitation conditions (even if such differences are not perceivable by humans). Subsequently, a lens model (Brunswik 1956; Scherer and Bänziger 2010) may be used, exploring how subjective judgment and objective markers combine in emotion recognition.

Conclusion

The approach described presently illustrates the importance of being explicit with the operationalization of emotional stimuli in studies of affective authenticity discrimination. The method used to produce posed expressions affected decoders' ability to distinguish them from spontaneous surprise, resulting in differences in perceived intensity, genuineness, confidence, and accuracy. For successful "deceptive" expressions, having information on the physiognomic features of a spontaneous display (i.e., external expression) was more important than the affective experience (i.e., internal sensation). By these criteria, the Mimic method appears to be superior to Stanislavski for our senders. Our findings demonstrate the importance of the specific technique used to elicit emotional displays, and the need to treat authenticity discrimination separately from classification accuracy. These considerations may reduce inconsistencies regarding posed expressions and authenticity, whilst further improving the methodological rigor in the field of emotion recognition.

References

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topography, 24*(2), 149–163.https://doi.org/10.1007/s10548-011-0171-4.

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.

Bernstein, M. J., Young, S. G., Brown, C. M., Sacco, D. F., & Claypool, H. M. (2008). Adaptive responses to social exclusion: Social rejection improves detection of real and fake smiles. *Psychological Science, 19*(10), 981–983. https://doi.org/10.1111/j.1467-9280.2008.02187.x.

Boraston, Z. L., Corden, B., Miles, L. K., Skuse, D. H., & Blakemore, S.-J. (2008). Brief report: Perception of genuine and posed smiles by individuals with autism. *Journal of Autism and Developmental Disorders, 38*(3), 574–580. https://doi.org/10.1007/s10803-007-0421-1.

Brody, L. R., & Hall, J. A. (2008). Gender and emotion in context. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 395–408). New York: Guilford Press.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed., p. 154). Berkeley: University of California Press.

Calvo, M. G., Gutiérrez-García, A., Avero, P., & Lundqvist, D. (2013). Attentional mechanisms in judging genuine and fake smiles: Eye-movement patterns. *Emotion, 13*(4), 792–802. https://doi.org/10.1037/a0032317.

Calvo, M. G., & Nummenmaa, L. (2015). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion, 30*(6), 1081–1106. https://doi.org/10.1080/02699931.2015.1049124.

Carroll, J. M., & Russell, J. A. (1997). Facial expressions in Hollywood's portrayal of emotion. *Journal of Personality and Social Psychology, 72*(1), 164–176. https://doi.org/10.1037/0022-3514.72.1.164.

Chong, S. C. F., Werker, J. F., Russell, J. A., & Carroll, J. M. (2003). Three facial expressions mothers direct to their infants. *Infant and Child Development, 12*(3), 211–232. https://doi.org/10.1002/icd.286.

Coan, J. A., & Allen, J. J. B. (Eds.). (2007). *Handbook of emotion elicitation and assessment*. Oxford: Oxford University Press.

Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Observer-based measurement of facial expression with the facial action coding system. In J. A. Coan & J. J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment* (3rd ed., pp. 203–221). Oxford: Oxford University Press.

Conson, M., Ponari, M., Monteforte, E., Ricciato, G., Sarà, M., Grossi, D., et al. (2013). Explicit recognition of emotional facial expressions is shaped by expertise: Evidence from professional actors. *Frontiers in Psychology, 4,* 382. https://doi.org/10.3389/fpsyg.2013.00382.

Cowie, R., Douglas-Cowie, E., & Cox, C. (2005). Beyond emotion archetypes: Databases for emotion modelling using neural networks. *Neural Networks, 18*(4), 371–388. https://doi.org/10.1016/j.neunet.2005.03.002.

Dawel, A., Palermo, R., O'Kearney, R., & McKone, E. (2015). Children can discriminate the authenticity of happy but not sad or fearful facial expressions, and use an immature intensity-only strategy. *Frontiers in Psychology, 6,* 462. https://doi.org/10.3389/fpsyg.2015.00462.

Dawel, A., Wright, L., Irons, J., Dumbleton, R., Palermo, R., O'Kearney, R., et al. (2017). Perceived emotion genuineness: Normative ratings for popular facial expression stimuli and the development of perceived-as-genuine and perceived-as-fake sets. *Behavior Research Methods, 49*(4), 1539–1562. https://doi.org/10.3758/s13428-016-0813-2.

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review, 1*(4), 346–357. https://doi.org/10.1207/s15327957pspr0104_5.

Douglas, K. M., Porter, R. J., & Johnston, L. (2012). Sensitivity to posed and genuine facial expressions of emotion in severe depression. *Psychiatry Research, 196*(1), 72–78. https://doi.org/10.1016/j.psychres.2011.10.019.

Duchenne, D. B. (1862). The mechanism of human facial expression or an electro-physiological analysis of the expression of the emotions (A. Cuthbertson, Trans.). New York: Cambridge University Press.

Ekman, P. (2003). Darwin, deception, and facial expression. *Annals of the New York Academy of Sciences, 1000*(1), 205–221. https://doi.org/10.1196/annals.1280.010.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology, 29*(3), 288–298. https://doi.org/10.1037/h0036006.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior, 6*(4), 238–252. https://doi.org/10.1007/BF00987191.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system. Manual and investigator's guide*. Salt Lake City: Research Nexus.

Ekman, P., Friesen, W. V., & Simons, R. C. (2005). Is the startle reaction an emotion? In What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS) (pp. 21–38). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195179644.003.0002.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science, 221*(4616), 1208–1210. https://doi.org/10.1126/science.6612338.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*(9), 913–920. https://doi.org/10.1037/0003-066X.46.9.913.

Ekman, P., & Rosenberg, E. L. (2005). What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146.

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology, 60*(2), 229–240. https://doi.org/10.1037/0022-3514.60.2.229.

Gosselin, P., Beaupré, M., & Boissonneault, A. (2002). Perception of genuine and masking smiles in children and adults: Sensitivity to traces of anger. *The Journal of Genetic Psychology, 163*(1), 58–71. https://doi.org/10.1080/00221320209597968.

Gosselin, P., Kirouac, G., & Doré, F. Y. (1995). Components and recognition of facial expression in the communication of emotion by actors. *Journal of Personality and Social Psychology, 68*(1), 83–96. https://doi.org/10.1037/0022-3514.68.1.83.

Gosselin, P., Perron, M., & Beaupré, M. (2010). The voluntary control of facial action units in adults. *Emotion, 10*(2), 266–271. https://doi.org/10.1037/a0017748.

Gunnery, S. D., & Hall, J. A. (2014). The Duchenne smile and persuasion. *Journal of Nonverbal Behavior, 38*(2), 181–194. https://doi.org/10.1007/s10919-014-0177-1.

Gunnery, S. D., Hall, J. A., & Ruben, M. A. (2013). The deliberate Duchenne smile: Individual differences in expressive control. *Journal of Nonverbal Behavior, 37*(1), 29–41. https://doi.org/10.1007/s10919-012-0139-4.

Gunnery, S. D., & Ruben, M. A. (2016). Perceptions of Duchenne and non-Duchenne smiles: A meta-analysis. *Cognition and Emotion, 30*(3), 501–515. https://doi.org/10.1080/02699931.2015.1018817.

Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., et al. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods, 115*(2), 137–143. https://doi.org/10.1016/s0165-0270(02)00006-7.

Hess, U., Banse, R., & Kappas, A. (1995). The intensity of facial expression is determined by underlying affective state and social situation. *Journal of Personality and Social Psychology, 69*(2), 280–288. https://doi.org/10.1037/0022-3514.69.2.280.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior, 21*(4), 241–257. https://doi.org/10.1023/A:1024952730333.

Hess, U., & Kleck, R. E. (1994). The cues decoders use in attempting to differentiate emotion-elicited and posed facial expressions. *European Journal of Social Psychology, 24*(3), 367–381. https://doi.org/10.1002/ejsp.2420240306.

Hull, S. L. (1985). *Strasberg's method as taught by Lorrie Hull: A practical guide for actors, teachers, and directors*. London: Ox Bow Press.

Kappas, A., Krumhuber, E., & Küster, D. (2013). Facial behavior. In M. L. Knapp & J. A. Hall (Eds.), *Handbook of communication science: Nonverbal communication* (pp. 131–166). Berlin: de Gruyter. https://doi.org/10.1515/9783110238150.131.

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., Likowski, K. U., & Weyers, P. (2014). Facial mimicry of spontaneous and deliberate Duchenne and non-Duchenne smiles. *Journal of Nonverbal Behavior, 38*(1), 1–11.https://doi.org/10.1007/s10919-013-0167-8.

Krumhuber, E. G., & Manstead, A. S. R. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion, 9*(6), 807–820.https://doi.org/10.1037/a0017844.

Krumhuber, E. G., Skora, L., Küster, D., & Fou, L. (2016). A review of dynamic datasets for facial expression research. *Emotion Review, 9*(3), 280–292.https://doi.org/10.1177/1754073916670022.

Lewis, M., Sullivan, M. W., & Vasen, A. (1987). Making faces: Age and emotion differences in the posing of emotional expressions. *Developmental Psychology, 23*(5), 690–697.https://doi.org/10.1037/0012-1649.23.5.690.

Lundqvist, L.-O., & Dimberg, U. (1995). Facial expressions are contagious. *Journal of Psychophysiology, 9*(3), 203–211.https://doi.org/10.1002/ejsp.2420240306.

Manera, V., Del Giudice, M., Grandi, E., & Colle, L. (2011). Individual differences in the recognition of enjoyment smiles: No role for perceptual–attentional factors and autistic-like traits. *Frontiers in Psychology, 2,* 143.https://doi.org/10.3389/fpsyg.2011.00143.

Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 32*(3), 454–465.https://doi.org/10.1016/j.neubiorev.2007.08.003.

Matsumoto, D., Keltner, D., Shiota, M., Frank, M. G., & O'Sullivan, M. (2008). What's in a face? Facial expressions as signals of discrete emotions. In M. Lewis, J. M. Haviland, & L. F. Barrett (Eds.), *Handbook of emotion* (pp. 211–234). New York: Springer.

McLellan, T., Johnston, L., Dalrymple-Alford, J., & Porter, R. J. (2010). Sensitivity to genuine versus posed emotion specified in facial displays. *Cognition and Emotion, 24*(8), 1277–1292. https://doi.org/10.1080/02699930903306181.

Mehu, M., Mortillaro, M., Bänziger, T., & Scherer, K. R. (2012). Reliable facial muscle activation enhances recognizability and credibility of emotional expression. *Emotion, 12*(4), 701–715. https://doi.org/10.1037/a0026717.

Namba, S., Makihara, S., Kabir, R. S., Miyatani, M., & Nakao, T. (2016). Spontaneous facial expressions are different from posed facial expressions: Morphological properties and dynamic sequences. *Current Psychology, 6,* 1–13. https://doi.org/10.1007/s12144-016-9448-9.

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review, 5*(1), 8–15. https://doi.org/10.1177/1754073912457227.

Perron, M., & Roy-Charland, A. (2013). Analysis of eye movements in the judgment of enjoyment and non-enjoyment smiles. *Frontiers in Psychology, 4,* 659. https://doi.org/10.3389/fpsyg.2013.00659.

Porter, S., & ten Brinke, L. M. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science, 19*(5), 508–514. https://doi.org/10.1111/j.1467-9280.2008.02116.x.

Porter, S., ten Brinke, L. M., & Wallace, B. (2012). Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior, 36*(1), 23–37. https://doi.org/10.1007/s10919-011-0120-7.

Porter, S., ten Brinke, L. M., & Wilson, K. (2009). Crime profiles and conditional release performance of psychopathic and non-psychopathic sexual offenders. *Legal and Criminological Psychology, 14*(1), 109–118. https://doi.org/10.1348/135532508X284310.

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.

Reissland, N., Shepherd, J., & Cowie, L. (2002). The melody of surprise: Maternal surprise vocalizations during play with her infant. *Infant and Child Development, 11*(3), 271–278. https://doi.org/10.1002/icd.258.

Sauter, D. A., & Fischer, A. H. (2018). Can perceivers recognise emotions from spontaneous expressions? *Cognition and Emotion, 32*(3), 504–515. https://doi.org/10.1080/02699931.2017.1320978.

Scherer, K. R., & Bänziger, T. (2010). On the use of actor portrayals in research on emotional expression. In K. R. Scherer, T. Bänziger, & E. Roesch (Eds.), *A blueprint for affective computing: A sourcebook and manual* (pp. 166–176). Oxford: Oxford University Press.

Schmid, P. C., Schmid Mast, M., Bombari, D., Mast, F. W., & Lobmaier, J. S. (2011). How mood states affect information processing during facial emotion recognition: An eye tracking study. *Swiss Journal of Psychology, 70*(4), 223–231. https://doi.org/10.1024/1421-0185/a000060.

Siedlecka, E., & Denson, T. F. (2019). Experimental methods for inducing basic emotions: A qualitative review. *Emotion Review, 11*(1), 87–97. https://doi.org/10.1177/1754073917749016.

Soppe, H. J. G. (1988). Age differences in the decoding of affect authenticity and intensity. *Journal of Nonverbal Behavior, 12*(2), 107–119. https://doi.org/10.1007/BF00986929.

Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. C. (2018). Facial expression analysis with AFFDEX and FACET: A validation study. *Behavior Research Methods, 50*(4), 1446–1460. https://doi.org/10.3758/s13428-017-0996-1.

Thibault, P., Gosselin, P., Brunel, M.-L., & Hess, U. (2009). Children's and adolescents' perception of the authenticity of smiles. *Journal of Experimental Child Psychology, 102*(3), 360–367. https://doi.org/10.1016/j.jecp.2008.08.005.

Tobin, A., Favelle, S., & Palermo, R. (2016). Dynamic facial expressions are processed holistically, but not more holistically than static facial expressions. *Cognition and Emotion, 30*(6), 1208–1221. https://doi.org/10.1080/02699931.2015.1049936.

Vallverdu, J. (Ed.). (2015). Handbook of research on synthesizing human emotion in intelligent systems and robotics. Hershey: Information Science Reference, an imprint of IGI Global.

Vazire, S., Naumann, L. P., Rentfrow, P. J., & Gosling, S. D. (2009). Smiling reflects different emotions in men and women. *Behavioral and Brain Sciences, 32*(5), 403–405. https://doi.org/10.1017/S0140 525X09991026.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using t-tests. *Perspectives on Psychological Science, 6*(3), 291–298. https://doi.org/10.3389/fpsyg.2018.01184.

Zloteanu, M., Krumhuber, E. G., & Richardson, D. C. (2018). Detecting genuine and deliberate displays of surprise in static and dynamic faces. *Frontiers in Psychology, 9,* 1184. https://doi.org/10.3389/fpsyg.2018.01184.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). New York: Academic Press. https://doi.org/10.1016/S0065-2601(08)60369-X.