



Universidad Autónoma
de Madrid

Biblos-e Archivo
Repositorio Institucional UAM

Repositorio Institucional de la Universidad Autónoma de Madrid

<https://repositorio.uam.es>

Esta es la **versión de autor** del artículo publicado en:

This is an **author produced version** of a paper published in:

Emotion 21.7 (2021): 1550 – 1569

DOI: <https://doi.org/10.1037/emo0001015>

Copyright: © 2021, American Psychological Association

El acceso a la versión del editor puede requerir la suscripción del recurso

Access to the published version may require subscription

EXPRESSION AND EMOTION

©American Psychological Association, 2021. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. The final article is available, upon publication, at: <https://doi.org/10.1037/emo0001015>

Do Emotions Result in their Predicted Facial Expressions? A Meta-Analysis of Studies on the Co-Occurrence of Expression and Emotion

Juan I. Durán¹, José-Miguel Fernández-Dols²

¹Universidad a Distancia de Madrid, Spain

²Universidad Autónoma de Madrid, Spain

Author Note

Juan I. Durán  <https://orcid.org/0000-0001-8562-0919>

José-Miguel Fernández-Dols  <https://orcid.org/0000-0002-9755-4338>

This research was funded by the Spanish Government's grant PSI2017-88776-P.

Correspondence concerning this article should be addressed to either José-Miguel Fernández-Dols, Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain. (E-mail: jose.dols@uam.es) or Juan I. Durán, Universidad a Distancia de Madrid, 28400, Collado Villalba, Spain (E-mail: juanignacio.duran@udima.es).

Abstract

That basic emotions produce a facial signal would – if true – provide a foundation for a science of emotion. Here, random-effects meta-analyses tested whether happiness, sadness, anger, disgust, fear, and surprise each co-occurs with its predicted facial signal. The first meta-analysis examined only those studies that measured full expressions through FACS. Average co-occurrence effect size was .13. The second meta-analysis included both full and partial expressions, as measured by FACS or another system. Average co-occurrence effect size rose to .23. A third meta-analysis estimated the Pearson correlation between intensity of the reported emotion and intensity of the predicted facial expression. Average correlation was .30. Overall, co-occurrence and correlation were greatest for disgust, least for surprise. What are commonly known as the six classic basic emotions do not reliably co-occur with their predicted facial signal. Heterogeneity between samples was found, suggesting a more complex account of facial expressions.

Keywords: facial expression; basic emotion theory; coherence; meta-analysis.

Do Emotions Result in their Predicted Facial Expressions? A Meta-Analysis of Studies on the Co-Occurrence of Expression and Emotion

A central thesis of the classic Basic Emotion Theory is “that the universal in facial expressions of emotion is the connection between particular facial configurations and specific emotions” (Ekman, 2017, p. 51). This thesis, in turn, provides a foundation for a science of emotion: “emotional states are signaled in distinct patterns of expressive behavior (...) Evidence establishing which psychological states are signaled in expressive behavior (and which states are not) informs taxonomic claims about which states might be considered emotions and the boundaries between them” (Cordaro et al., 2020, p. 1292). On the other hand, this thesis remains in need of empirical testing (for a recent review, see Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019).

This view might not be the core of all modern Basic Emotion theories (e.g., Oatley & Johnson-Laird, 2011), but it has become part of the psychological zeitgeist. One revealing example is Ekman’s (2016) survey in which 80% of a large sample of emotion researchers agreed with the assumption that some emotions are expressed with universal facial configurations and vocalizations. A second illustrative example is the large number of studies whose explicit purpose is the study of “facial emotions”, taking as granted an essential identity of emotion and facial expression (a search for “facial emotion” in PsycInfo retrieved 600 documents).

In the applied realm, researchers and professionals have assumed that the categorization of posed facial displays in terms of basic emotions is a core feature of emotional competence (e.g., Matsumoto et al. 2000, Scherer and Scherer, 2011, Passarelli, Masini, Bracco, Petrosino, & Chiorri, 2018). Teaching of classical facial expressions of emotion is promoted as an accurate way of improving the emotional competence of children (e.g., Lierheimer & Stichter, 2011), managers (e.g., Caruso

&Salovey, 2004) or even customs officers (United States Government Accountability Office [GAO], 2013). Companies and governments are producing algorithms to detect basic emotions from facial expressions through IA systems (e.g., Schwartz, 2019).

A direct test of the thesis that facial expressions are emotion signals is difficult, but two predictions from this thesis have been tested. One such prediction is universal “*recognition*” of an emotion from its facial expression. Reviews of extensive evidence on this prediction have resulted in controversy, with different scientists coming to different conclusions (Cowen, Sauter, Tracy, & Keltner, 2019; Ekman, 2017; Keltner & Cordaro, 2017; Barrett, et al., 2019; Fridlund, 2017; Nelson & Russell, 2013; Sauter & Russell, in press). Even if the evidence overwhelmingly supported the prediction of universal “recognition,” however, this type of evidence would not be definitive, for universal agreement might be a universal mistake, as when people everywhere once agreed that the sun circled the earth, rising in the morning, circling overhead, and setting in the evening.

A second prediction from the assumption of facial signals of emotion is that each basic emotion produces the predicted facial expression; therefore, an emotion and its purported facial expression *co-occur*. This second prediction is the topic of the present study. The issue of the co-occurrence of emotion and its predicted facial expression has been labelled “coherence” (Rosenberg & Ekman, 1994) and “synchronization” (e.g., Scherer, 2013). In this article, we use the theoretically less charged term “co-occurrence.” The opposite of coherence, “incoherence,” would suggest a nonsensical relation of emotion to face; and lack of “synchronization” would be implausible because all behavior must be synchronized to be functional. Less than strong emotion-face occurrence (or the absence of any significant occurrence) would describe a relative independence between basic emotions and their facial expressions as currently predicted

by classic Basic Emotion Theory, and it would also raise the question of what, if not emotion, accounts for what are called facial expressions.

Method

General overview

Reviewing evidence gathered over decades of research, Ekman (2017) concluded: “The evidence [about universal facial expressions] is strongest for happiness, anger, disgust, sadness, and fear/surprise”. These “classic” (e.g., Jack & Schyns, 2015) or “basic six” (Cowen et al., 2019) emotion categories are the only ones with an appreciable literature on their co-occurrence with their predicted expressions, and they were the target of our analyses. A PsycInfo search combining each of these six terms with “facial expression” returned 713 entries, whereas “love”, “shame”, “contempt”, “hate”, “interest”, “distress”, and “guilt” combined with “facial expression” only returned 65 entries; practically all of them were not on the actual co-occurrence of these emotion categories with their predicted, if any, facial expression.

Most research on this topic scored co-occurrence in one of two ways. (a) An instance of co-occurrence was scored only when the emotion co-occurred with the *whole* predicted facial expression –i.e., a facial display that included all the facial movements predicted by Ekman and Friesen (e.g., Ekman & Friesen, 1978; Ekman, Friesen & Hager, 2002). For example, the whole expression of sadness would include pulled up eyebrows, downturned mouth, and drooped eyelids (technically, action units 1+4+15+54). (b) An instance of co-occurrence was scored when the emotion co-occurred with a *part* of the predicted facial pattern. For example, a partial expression of sadness might be a raised inner brow (action unit 1). It is questionable whether this second method is a strict test of the prediction because, for example, raised inner brow is

also predicted to co-occur with surprise and fear. We addressed the problem of whole versus part by conducting two meta-analyses.

The first meta-analysis (Whole Predicted Facial Expressions Only) was the most theoretically relevant. It included only co-occurrence studies in which whole predicted facial expressions were assessed with a precise and systematic coding system (FACS).

The second meta-analysis (Whole and Partial Predicted Facial Expression) expanded the set of studies analyzed. It included all the studies from the first meta-analysis (with only full expressions coded by FACS) plus additional studies that included other alternative coding systems and either whole or *partial* occurrences of the predicted facial expression. The inclusion of less strict, arguably overly inclusive coding of the facial expression could also help to determine the upper limit on the amount of its co-occurrence with emotion.

A third meta-analysis (Intensity) analyzed a subset of studies that reported the intensity of the emotion and the intensity of the facial expression. Whereas the first two meta-analyses assessed the presence or absence of the predicted facial expression, this third analysis concerned the hypothesis that the intensity of the emotion predicts the intensity of the facial expression.

Besides these three main meta-analyses, we report two additional meta-analyses on amusement. Some researchers included amusement within the category of happiness (e.g., Ekman, Friesen & Ancoli 1980), but there are important conceptual and empirical reasons not to do so. Keltner and Cordaro (2017) has theorized that amusement and happiness are distinct emotions. Ruch (1995) found evidence of a high co-occurrence between amusement and laughter, but laughter has not been found to co-occur with happiness lacking amusement. Cordaro et al. (2020) offered evidence that happiness and amusement have distinct facial expressions. The co-occurrence between amusement and

laughter is therefore important in itself (see Ruch, 1995). For these reasons, in the present study, amusement data havenot been included into the meta-analyses as part of happiness or as one of the six basic emotions. Instead, we include two independent meta-analyses on amusement alone.

Literature Search

The relevant studies were located through a number of independent searches:

1. A predecessor of these meta-analyses is Duran, Reisenzein, and Fernández-Dols's (2017) quantitative review in which they concluded that co-occurrence between emotion and facial expression was less frequent than would be expected by Basic Emotion Theory. The departure point for the literature search was the sample of studies included in Duran et al.'s study (2017). However, some of these studies were excluded for the current analyses because of not meeting the inclusion criteria (see below).
2. Additional studies located by a PsycInfo search (June 2018, renewed in June 2019) using the term "facial expression" as the major subject, and "spontaneous expression" or "coherence" or "correlation" or "production" or "display" or "Duchenne smile" as keywords.
3. Wescreeend 74 articles obtained through the previous steps andadding the onesreported in the reviews on co-occurrence by Matsumoto, Keltner, Shiota, O'Sullivan, and Frank (2008), Reisenzein, Studtmann, and Horstmann (2013) and Fernández-Dols and Crivelli (2013). Thosesources yielded 40 articles with 86 effect sizes.
4. In November 2016, we posted an announcement on the listserve of the International Society for Research on Emotion, asking for published or

unpublished works relevant to our meta-analysis. We obtained 22 documents, 5 of them with testable data from which we obtained 12 effect sizes.

5. In October 2017, Lisa Feldman-Barrett allowed us to share an additional database of facial EMG measurements compiled and coded by Erika Siegel, Madeleine Devlin, Ludger Hartley and Tori Tavormina for a related meta-analysis. Although the articles did not report typical co-occurrence studies, a portion of them (10 articles out of 123) included usable data on co-occurrence, allowing us to include 12 new effect sizes in the meta-analysis.

A flowchart describing the process of screening and selecting the articles is shown in Figure 1.

--- Figure 1 ---

All the entries used in the meta-analyses for each of the considered emotion categories were based on data from different samples. However, the same sample could contribute to the meta-analysis with co-occurrence indexes for more than one emotion category. The meta-analyses were based on 55 articles that reported 69 studies from which were extracted 110 effect sizes. The studies included a total of 3847 participants with a mean of 55 and a median of 34.5 participants per study. The full database with all the descriptive information of each study is provided as supplementary materials.

Inclusion and exclusion criteria

For the first two meta-analyses, we included all studies that, for a given emotion category, reported or allowed us to calculate an effect-size index and confidence interval (CI) of the proportion of participants who displayed the predicted expression. We excluded those studies (a) with no available data about proportions or correlations, (b) with dependent variables that conflated multiple nonverbal channels, or (c) with an

independent variable not directly related to one of the six classic categories of basic emotion (e.g., manipulating affective or a motivational state).

For the third meta-analysis, we included all studies that included or allowed us to calculate a Pearson correlation between the intensity of observed expression and the intensity of reported emotion. Therefore, correlation indices different from inter-individual Pearson or other Pearson correlations were excluded. If the authors did not report an exact correlation but reported that it was below a cutoff value (e.g., $r < .20$), the correlation was estimated as being .05 below the cutoff (e.g., $.20 - .05 = .15$).

If the authors did not report an effect size but the article included sufficient data to compute it (see “effect size” section below), the effect size was computed from these data.

If the authors reported several usable co-occurrence indices (i.e., one for each action unit predicted for a particular emotion category), the analysis always included the most favorable to the co-occurrence hypothesis, namely the one with the highest magnitude, excluding the others in order to avoid interdependence problems.

Clinical studies in which co-occurrence was not reported for a group of neurotypical individuals were excluded. Studies with infants and young children (age below 5; e.g., Bennett, Bendersky & Lewis, 2002; Camras et al., 2002; Izard & Abe, 2004; Scherer, Zentner, & Stern, 2004) raise complex problems both in measurement of emotion and of facial behavior, which could be an important source of heterogeneity; for this reason, they were also excluded. For a review on the relatively low co-occurrence between expression and emotion in infants and young children, see Camras, Castro, Halberstadt, & Shuster (2017).

References List S1 (see supplemental materials) reports the studies on co-occurrence that were not included and the reason for their exclusion.

Coding Procedure

Juan I. Durán coded all the moderators, and J.M. Fernández-Dols coded a subsample of 20% of the articles to calculate the coding reliability for categorical and quantitative moderating variables. The variables were: size of the sample, emotions studied (six categories: happiness/enjoyment, surprise, disgust, sadness, anger, fear), method of emotion induction (films or texts, imagery or reminiscence, situations or sensorial inputs), whole expression vs. specific facial movements, registering system (EMG, FACS, EMFACS, other observational systems and subjective estimates), intensity of the reported emotion, social interaction in the experimental setting, participant's awareness of being observed, and age group (children vs. adults).

Given the relatively low number of studies and the factual nature of the coded data, there were no significant disparities in the coding. The agreement for the categorical variables according to Cohen's Kappa was .97 for emotion category, 1.0 for Vividness of the Emotion Elicitors, .97 for Whole Expression vs Specific Facial Movements, 1 for Observation or Registering System, 1 for Age Group, .96 for Sociality, and .83 for Self-Awareness. The intraclass correlation coefficient for the quantitative moderator Intensity was .96. Disagreements were resolved by discussion between the authors.

Effect Size Indices

Proportions and correlations are the most frequently reported indices in co-occurrence studies (Reisenzein, Studtmann, & Horstmann, 2013). The effect sizes summarized in the first and second meta-analyses were the proportion p_{rp} of participants in the experimental sample who displayed the predicted expression. The estimate of intensity co-occurrence (third meta-analysis) was based on the Pearson correlation (r)

between the intensity of the observed expression and the intensity of the reported emotion.

Proportions. To avoid the potential problems associated with the estimation of CIs from raw proportions (Newcombe, 2012), we used the *logit* transformation of the proportions of participants who displayed the predicted expression p_{rp} :

$$\text{logit}(p_{rp}) = \log_e \left(\frac{p_{rp}}{(1 - p_{rp})} \right), \quad (1)$$

computing their corresponding variances through

$$\hat{\sigma}_{\text{logit}(p_{rp})}^2 = \frac{1}{n \cdot p_{rp} \cdot (1 - p_{rp})}, \quad (2)$$

where n is the study sample size.

For studies in which the reported p_{rp} was lower than .03, their p_{rp} was included as .03 in order to obtain suitable values for their variances and CIs.

Pearson correlations. Fisher's Z index, obtained from the Pearson correlation r , was used to avoid the problems associated with asymmetries in the distribution of r when its population value is 0 (Botella & Sánchez-Meca, 2015):

$$Z_r = \frac{1}{2} \cdot \log_e \frac{1+r}{1-r}. \quad (3)$$

The Z_r index variance was calculated as:

$$\hat{\sigma}_{Z_r}^2 = \frac{1}{n-3}. \quad (4)$$

We also included studies that did not report a correlation value but provided enough information about intensity of the expression elicited by an emotional condition (E) versus a non-emotional or baseline condition (BL). The r value from the standardized difference of means d was obtained through

$$r = \frac{d}{\sqrt{d^2 + a}}; a = \frac{(n_E + n_{BL})^2}{n_E \cdot n_{BL}}, \quad (5)$$

where n_E and n_{BL} are sample sizes of E and BL conditions, and $a = 4$ when $n_E = n_{BL}$. For these studies, r variance was obtained through

$$\hat{\sigma}_{d-r}^2 = \frac{a^2 \cdot v_d}{(d^2 + a)^3}, \quad (6)$$

where v_d is the variance of the difference of means d .

When the study reported the average intensity of the expression in an emotional (\bar{X}_E) versus a non-emotional baseline condition (baseline average intensity or \bar{X}_{BL}) across independent groups of participants, the standardized mean difference $d_{between-groups}$ and its variance v_d was obtained through:

$$d_{between-groups} = \frac{\bar{X}_E - \bar{X}_{BL}}{\sqrt{\frac{(n_E - 1) \cdot S_E^2 + (n_{BL} - 1) \cdot S_{BL}^2}{n_E + n_{BL} - 2}}} \cdot c(df), \quad (7)$$

$$v_{d_{between-groups}} = \frac{n_E + n_{BL}}{n_E \cdot n_{BL}} + \frac{d^2}{2 \cdot (n_E + n_{BL})},$$

(8)

where S_E^2 and S_{BL}^2 are E and BL groups variances in expression intensity, respectively.

The term $c(df)$ in the equation is a correction factor (Hedges, 1981) defined as:

$$c(df) = 1 - \frac{3}{4 \cdot (n_E + n_{BL} - 2) - 1}. \quad (9)$$

When the study reported a within-subjects comparison between the average intensity of the expression in an emotional condition and a non-emotional baseline condition, the $d_{within-groups}$ and its variance $v_{d_{within-groups}}$ were computed through:

$$d_{within-groups} = \frac{\bar{X}_{BL} - \bar{X}_E}{\sqrt{S_{BL}^2 + S_E^2 - 2 \cdot r_{BL-E} \cdot S_{BL} \cdot S_E}} \quad (10)$$

$$v_{d_{within-groups}} = [c(df)]^2 \cdot \left(\frac{1}{n}\right) \cdot \left(\frac{n-1}{n-3}\right) \cdot (1 + n + d_{within-groups}^2) - d_{within-groups}^2 \quad (11)$$

where r_{BL-E} is the correlation between BL and E measures, and the correction factor $c(df)$ was:

$$c(df) = 1 - \frac{3}{4 \cdot (n-1) - 1} \quad (12)$$

Given that the r_{BL-E} value is not reported in any of the included studies, we set it to .7 --following Rosenthal's (1991) recommendation-- as representative of pre-post test studies. In some studies, the only information available about the within-subject comparison was the t -statistic value. When this was the case, d was obtained as:

$$d = \frac{t}{\sqrt{n}} \cdot c(df) \quad (13)$$

Combined Estimation of Effect Size and Heterogeneity

Combined estimations of effect sizes were computed weighting each effect size by the inverse of its variance (Hedges and Olkin, 1985). Restricted Maximum Likelihood (RML) was used as the estimation method, using the Metafor package (Viechtbauer, 2010) for R software (R Core Team, 2019).

A random-effects model (see Borenstein, Hedges, Higgins, & Rothstein, 2009) for variance estimation allowed a broader generalization of results than those obtained through a fixed-effects model (Borenstein, Hedges, Higgins & Rothstein, 2010). We used cluster robust tests and confidence intervals for the model coefficients (Hedges, Tipton, & Johnson, 2010) to account for the interdependence of data. This problem of interdependence arises from cases in which more than one effect size from the same sample were included. In particular, 42% of the proportions and 37% of correlations of

the present sample of studies are interdependent, due to the inclusion of studies that yielded effect sizes corresponding to different emotion categories. For that reason, studies that elicited the same emotion with the same elicitation method were clustered together before obtaining the overall estimates, their confidence intervals, and the tests for the moderating variables. On the other hand, estimates for each category of the moderating variables were obtained by clustering the effect sizes coming from the same study.

Statistics Q and I^2 (Borenstein, Hedges, Higgins, & Rothstein, 2009) assessed the heterogeneity among the effect sizes of the studies. The null hypothesis of the Q -test corresponds to the assumption of effect size homogeneity. The homogeneity hypotheses are tested by Q_w and cluster robust tests. Rejection of the null hypothesis tested by Q_w means that there is more heterogeneity than the expected by mere sampling error across studies. When this is the case, cluster robust tests are used to check if this heterogeneity can be explained by different moderating variables (i.e. there are differences between the moderator categories or a linear relationship between them and the effect size magnitude). In addition, the statistic I^2 assessed the amount of overall heterogeneity (Huedo-Medina, Sánchez-Meca, Marin-Martinez, & Botella, 2006).

Moderating Variables

Next is a list of all the potentially moderating variables that were included in the analyses. The goal was to be as comprehensive as the available data allowed.

Categories of Emotion. An obvious factor that might play an important role in the degree of co-occurrence between experience and expression is the specific category or categories of basic emotion addressed by the study. Emotions might have a different degree of “expressive acuity” for a number of reasons: duration or frequency of the

experience, typical intensity, etc. (e.g., Verduyn, van Mechelen, Tuerlinckx, Meers, & Van Coillie, 2009).

Vividness of the Emotion Elicitors. Critics of the co-occurrence studies have pointed out that the emotions elicited in the laboratory cannot reach the level of vividness that characterizes “emotions in the wild”. The emotion elicitors reported in the studies on co-occurrence were therefore classified in terms of the psychological distance between the eliciting situation and the participant. We classified elicitors into the three categories: less vivid, when the emotional experience is mediated by films or texts; moderately vivid, when the emotional experience is elicited through imagery of or reminiscence about emotional events; and most vivid, when the emotional experience is directly elicited by sensorial inputs (e.g., odors, see Soussignan & Schaal, 1996) and real or simulated situations (e.g., to open a room to an unexpected place, see Schutzwahl & Reisenzein, 2021; to play against someone who is openly cheating, see Hubbard et al., 2002).

Whole Expression vs. Specific Facial Movements. A possible source of heterogeneity between the studies was the absence of a distinction between cases in which the whole predicted facial expression –i.e., an expression that included all the facial movements predicted by Basic Emotion Theory-- had been observed and cases in which only a part of that facial expression had been. Therefore, in the second and third meta-analyses we registered if the study targeted the whole predicted expression for the elicited emotion or specific facial movements. In a substantial number of studies (37 out of 64) the authors, either for theoretical or practical reasons, limited their analyses to some specific muscles (typically zygomaticus major and/or corrugator supercilii) or a portion of the face (e.g., upper face vs. lower face; e.g., Reisenzein, 2000; Tomarken & Davidson, 1992 or Van der Graaf et al., 2016).

Observation or Registering System. The most important coding system is the Facial Actions Coding System (FACS) developed by Ekman and Friesen (1978) and based on the coding system of Hjortsjö, a Swedish professor of anatomy who developed a complex system for numerically coding the movements of all the facial muscles (Hjortsjö, 1969). Another systematic observational coding system is the Maximally Discriminative Facial Movement Coding System (MAX, Izard, 1979), but this was mostly restricted to the observation of infant's facial expression (Bennet, Bendersky, & Lewis, 2002; Izard & Abe, 2004) which were not included in these meta-analyses.

The FACS is based on visual appraisals, requires trained observers and is highly time-consuming. Although researchers are developing computerized versions of FACS, they do not reach the level of description, muscle by muscle, of the original system. Many researchers have resorted to coding systems based on FACS (e.g., EMFACS) or to other coding systems with a more holistic or subjective basis. Finally, researchers have also used electromyography, an objective recording system, but limited to a few muscles.

The studies were classified according to the expression measurement system. Insufficient data exist to analyze each measurement system separately. The resulting four categories included 1) observational coding systems with an analytic approach - such as FACS, 2) observational coding systems with a global, less analytic approach, such as EMFACS, "Emotional Behavior Coding System" (e.g. Gross, 1998) or coding systems based on Ekman, Friesen, and Tomkins' FAST (1971), 3) coding systems that rely on subjective criteria of coders like the "Global cultural informant approach" (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005), subjective coding (e.g. Vazire

et al., 2009; Visser, Kramer & Swerts, 2015) or non-specified methods (e.g. Underwood & Bjornstad, 2001), and 4) electromyography studies.

Intensity of the Reported Emotion. A potential objection to the co-occurrence studies and by extension to all the experimental studies on emotion is that the emotions elicited in the laboratory are of low intensity and, for this reason, they do not produce observable instances of their corresponding predicted facial expressions. For this reason, the intensity of the participants' emotion, as a predictor of the expression, is a moderating variable in the first and second meta-analyses (the third meta-analysis is focused on the assumption behind this objection, the lineal relationship between intensity of the experience and intensity of the predicted expression).

Studies included in the first and second meta-analyses used different metrics and many of them did not report the standard deviations of their measurements. For these reasons, we used two approaches to transform mean intensity ratings (\bar{X}_k) to a common metric: (1) by $\bar{X}_k / S_{X_k} \cdot \sqrt{n}$, where S_x is the ratings standard deviation and n is the k study sample size, and (2) $(\bar{X}_k - X_{k(\min)}) / (X_{k(\max)} - X_{k(\min)})$, where $X_{k(\min)}$ and $X_{k(\max)}$ are respectively the minimum and maximum possible values of the intensity scale presented in the study.

Age Group. Age plays an important role in the recognition and production of expressions. following the authors' self-description of the study we differentiated between studies with children (all of, at least, elementary school age) and studies with adults as participants.

Sociality. The sociality of the situation is a variable with theoretical relevance (e.g., Fridlund, 1991). We classified the studies in three groups. Studies in which the participant was alone, studies in which there was a mere presence of another person in

the experimental setting but no interaction (e.g., the experimenter was sitting behind the participant), and studies in which the experimental procedure required interaction (e.g., in a clinical interview, in a game with two players).

Self-Awareness. In order to control the effect of potential display rules (Ekman, 1972), we classified the studies in two groups; studies with a procedure in which the participants could infer that they were observed (e.g. because there was a visible camera or they had electrodes on their face), and studies in which the participants were not aware of being observed (e.g., they were recorded with a hidden camera).

Publication Date. Conceivably, as methods and information improved, more recent studies were more likely to capture the co-occurrence of emotion with facial pattern. This potential moderator was assessed as the number of years since publication date. The range was 1 to 39.

Publication bias

Several methods for assessing publication bias were used: Egger's regression test (Egger, Smith, Schneider, & Minder, 1997), Trim-and-Fill (Duval & Tweedie, 2000) and Rosenthal's (1979) and Rosenberg's (2005) fail-safe n .

Egger's regression test assesses the asymmetries in the funnel plots, which represent both the effect size estimates and the standard error for each included study. If there were a publication bias, studies that report null results have a lower probability of being published, and this would produce asymmetries in the funnel plot. Using the same rationale, the Trim-and-Fill method iteratively adds the effect sizes of the potentially censored studies in order to obtain summary estimates that take this problem into account. Finally, fail-safe n is an estimation of how many studies with null effect would be necessary to add to the sample to reduce the combined estimation to a non-significant magnitude according to a standard alpha level of .05.

From now on, the results of the three meta-analyses are presented according to the following structure. First, a forest plot, i.e., a graphic, descriptive representation of the CIs of the effect sizes for the studies included in the meta-analysis. Second, the weighted mean of the effect sizes. Third, the estimation of the effect sizes' heterogeneity and the analysis of moderators that could explain it. And finally, the tests of publication bias.

First Analysis: Whole Predicted Facial Expression Only

Arguably, the clearest test of the assumption of the co-occurrence between a classic basic emotion and its corresponding facial expression requires that the facial expression match that predicted by theories of emotion signals such as that of Ekman (1972, 1992, 2017). In this first analysis, researchers tested if the whole expression specified in and measured with the standardized observational coding system FACS (Ekman & Friesen, 1978; Ekman, Friesen, & Hager, 2002) was taken to be one predicted for the elicited basic emotion. Thirteen such studies have been reported. They yielded fifteen estimates of p_{rp} , i.e., the proportions of participants showing the predicted facial expression.

The forest plot is a graphic representation of the CIs of the compared studies, represented by horizontal bars, and, at the bottom, equally represented by a horizontal bar, the CI that results of combining the CI of all the studies. The forest plot with global and by-emotion combined estimations of p_{rp} for those studies is shown in Figure 2.

--- Figure 2---

Weighted Estimates

The weighted mean estimate of proportion of participants who displayed the predicted expression p_{rp} was .13, and its CI at 95% was [.06 - .27]. Whilst this

proportion of matches is significant (the corresponding interval does not include 0), its symmetrical value (the proportion of misses) is also significant and its magnitude is much larger. Thus, on average, .87 of participants failed to display the predicted whole facial expression.

Analysis of Moderators

The statistical test of heterogeneity for the reported effect sizes on which this estimation was based was significant ($Q_w(12) = 45.82, p < .001, I^2 = 81\%$).

The examination of the five moderators that could explain this heterogeneity requires keeping in mind that, given the small number of studies, any test about the effects of any of moderating variables could have low statistical power. According to the results of the analysis of moderating variables shown in Table 1, there are significant differences between the emotion categories for which co-occurrence was computed. The studies included in this meta-analysis reported effect sizes for only three emotion categories. The highest degree of co-occurrence was obtained for disgust ($p_{rp} = .30 [.20 - .41]$). The co-occurrence estimates for happiness and surprise were lower, but statistically different from zero ($p_{rp} = .12 [.02 - .51]$, and $p_{rp} = .06 [.02 - .15]$). The remaining heterogeneity within the emotion categories was significantly higher than expected by sampling for happiness ($Q_w(3) = 12.99, p = .005, I^2 = 80\%$), but not for disgust ($Q_w(3) = 4.77, p = .190, I^2 = 7\%$) and surprise ($Q_w(4) = 2.51, p = .642, I^2 = 12\%$).

Publication Date also explains differences in the magnitude of the effect size for proportions of participants p_{rp} . Contrary to our anticipation—that new improved methods would have been more likely to capture the co-occurrence of emotion with facial expression--, the tendency is for older studies to report greater proportions of participants who displayed the predicted expression.

Lastly, the Intensity of the Reported Emotion also shows an inverse relation with p_{rp} , although the result does not replicate when using the standardized value.

The remaining testable moderating variables (Vividness of Emotion Elicitor, Age Group, Sociality and Self-Awareness) did not have a significant effect on the magnitude of p_{rp} .

--- Table 1 ---

Publication Bias

There was no evidence that the significant effect of the weighted mean estimate of p_{rp} was due to publication bias. The funnel plot is shown in Figure 3's upper right panel. The observed asymmetry was the opposite of what would be expected in the case of publication bias according to Egger's test ($z = -3.30, p = .001$). Publication bias would artificially increase the effect size, so an estimate corrected by its presence would be lower than the original mean estimate. However, the bias-corrected p_{rp} estimate obtained with the Trim-and-Fill method was .19 [.11 - .31]. There was no likely risk that results were a consequence of publication bias according to fail-safe n : it would be necessary to add 625 or 397 studies with null effects according to Rosenthal's and Rosenberg's fail-safe n , respectively, to reduce the combined estimation to a non-significant effect for an alpha level of .05.

--- Figure 3---

Results Summary

In summary, the estimate of the proportion of participants who *failed* to display the predicted facial expression was within a CI that ranged from .73 to .94. Two moderators, Publication Date and Category of Emotion explained a substantial portion of the observed heterogeneity among the effect sizes of the studies included in the

analysis. Disgust showed the greatest amount of co-occurrence (0.30 [0.20, 0.41]) with its predicted facial expression, although still not a large amount.

Second Analysis: Whole and Partial Predicted Facial Expression

It might be argued that the first analysis just reported set the bar too high by requiring the *whole* predicted facial expression. To establish the assumption of co-occurrence of emotion and facial expression, perhaps it would suffice to show that *part* of the expression occurs, and not just as measured by FACS but also by any other alternative coding method. Indeed, in 35 studies, the authors defined a successful prediction of co-occurrence either for a whole (i.e., the studies already included in the first analysis) or a partial expression. In this second meta-analysis we coded whether the original study reported co-occurrence indices for both complete and partial displays of the predicted expression. If it was available, we used the co-occurrence index for the complete expression; if the authors reported only the co-occurrence index for a partial display, that index was used in the second meta-analysis (in the figures, the partial-display indices are marked with “*”).

The forest plot with the global and by-emotion combined estimations is shown in Figure 4. The CIs of the forest plot are not symmetrical due to the back-transformation of effect sizes from *logit* (p_{rp}) to their original raw proportion metrics.

Weighted Estimates

These 35 studies yielded 50 estimates of p_{rp} , i.e., the proportion of participants who produced the whole or partial version of the predicted expression. The weighted mean of p_{rp} was .23 (CI = .13 - .36). Put differently, .77 of participants failed to display even a partial version of the predicted expression.

--- Figure 4 ---

Analysis of Moderators

As expected, significant heterogeneity was found among the effect sizes ($Q_w(49) = 353.22, p < .001, I^2 = 91\%$).

The results of the analysis of moderators that explored the potential sources of that heterogeneity are shown in Table 2. The emotion categories for which co-occurrence was computed showed different mean effect size magnitudes. The effect sizes ranged from $p_{rp} = .42$ for disgust to $p_{rp} = .13$ for surprise. Proportions were significantly different from zero for all emotions.

All categories but sadness shown an amount of heterogeneity above from what can be explained by sampling error (happiness: $Q_w(6) = 45.60, p < .001, I^2 = 88\%$; surprise: $Q_w(18) = 64.58, p < .001, I^2 = 75\%$; disgust: $Q_w(9) = 95.45, p < .001, I^2 = 96\%$; sadness: $Q_w(4) = 8.43, p = .077, I^2 = 55\%$; anger: $Q_w(4) = 47.12, p < .001, I^2 = 88\%$; fear: $Q_w(3) = 43.60, p < .001, I^2 = 96\%$).

--- Table 2 ---

There were differences between different degrees of Sociality, finding the greatest proportions in studies with social interactions, followed by non-social settings and the mere presence of other people. There were also differences between the effect sizes of the studies reporting Whole Expressions and Specific Facial Movements, being higher for the second group of studies.

In line with the first meta-analysis, although marginally significant, there was an inverse relationship between the effect size magnitude and the Intensity of the Reported Emotion.

The remaining moderating variables (Vividness of the Emotion Elicitor, Observation or Registering System, Age Group, Self-Awareness, Intensity of Reported

Emotion and Publication Date) did not significantly affect the observed proportions of participants who displayed the predicted expression.

Publication Bias

There was no evidence of publication bias. The funnel plot is shown in the upper left panel of Figure 3. For the case of co-occurrence studies, a publication bias would produce a censure of studies located at the left side of the effect size distribution. However, as in the analysis on whole expressions only, Egger's test ($z = -3.79$, $p < .001$) indicated the opposite asymmetry. The p_{rp} estimate corrected with the Trim-and-Fill method (Duval & Tweedie, 2000) was .27 [.20 - .36]. There is no threat to the validity of the effect found in this meta-analysis due to publication bias either according to Rosenthal's (1979) or Rosenberg's (2005) fail-safe n : it would be necessary to add 5395 or 2964 studies with null effects, respectively, to reduce the combined estimation to a non-significant magnitude according to a standard alpha level of .05.

Results Summary

In summary, the estimate of the proportion of participants who *failed* to display the predicted facial expression was within a CI that ranged from .64 to .87. The observed heterogeneity among the effect sizes in these studies was largely explained by three moderators: Category of Emotion, Whole vs Partial Expression, and Sociality. Disgust showed the greatest amount of co-occurrence (0.42 [0.14, 0.75]) with its predicted facial expression. Whole plus partial facial expressions showed more agreement with prediction than did Whole alone. Finally, greater Sociality showed more agreement with prediction while in the first meta-analysis no significant differences were found. In short, despite this last difference, this second meta-analysis yielded results highly similar to those of our first meta-analysis. Importantly, the majority of

cases of a basic emotion failed to co-occur with even a part of the predicted facial expression.

Third Analysis: Intensity

Classic basic emotion theory (e.g., Ekman, 1993) predicts that more intense emotions trigger more intense facial movements. The effect sizes found in our first two meta-analyses did not take into account the intensity of the predicted expression. The third meta-analysis examined a possible linear relationship between the intensity of the emotional experience and the intensity of the predicted expression.

Collected data were Pearson correlations (r) between the *reported intensity* of the emotional experience and the observed *intensity* of the predicted facial reaction (typically the activity of a muscle measured by EMG). If the authors reported correlations for both whole and partial displays of the predicted expression, this meta-analysis used the correlation index for the intensity of the whole expression.

The forest plot with the global and by-emotion combined Fisher Z estimations, back-transformed to Pearson correlations, is shown in Figure 5. CIs can be non-symmetrical due to this back-transformation.

--- Figure 5---

Weighted Estimates

Thirty-five estimates of r from 28 different studies yielded a combined estimate of $r = .30$, with a CI at 95% = [0.18 - 0.41].

Analysis of Moderators

The results of the statistical test for heterogeneity between studies reporting correlations was significant ($Q_w(34) = 71.77, p < .001, I^2 = 55\%$).

The results of the analysis of moderating variables for r index are shown in Table 3. The analysis with the moderating variable Category of Emotion (i.e., the category of

basic emotion addressed by the study) showed that there were differences between emotion categories. The highest correlation was obtained for disgust ($r = .45$ [.31 - .57]) and sadness ($r = .36$ [.30 - .42]), followed by happiness ($r = .32$ [.24 - .39]) and anger ($r = .18$ [.07 - .28]). The co-occurrence estimates for surprise and fear were not statistically different from zero, although they were based on only two and one study, respectively. Studies within every emotion category showed a non-significant amount of heterogeneity (Happiness: $Q_w(6) = 6.58$, $p = .361$, $I^2 = 0\%$; Surprise: $Q_w(1) = 0.18$, $p = .665$, $I^2 = 0\%$; Disgust: $Q_w(9) = 15.09$, $p = .088$, $I^2 = 44\%$; Sadness: $Q_w(4) = 2.16$, $p = .706$, $I^2 = 0\%$; Anger: $Q_w(9) = 13.74$, $p = .131$, $I^2 = 39\%$).

--- Table 3 ---

The moderating variable Vividness of the Emotion Elicitors also had an impact on co-occurrence. The highest co-occurrence estimates were obtained by showing films, images or texts. Lower estimates were obtained when the elicitor was an emotional situation or imagining or remembering events. Significant differences also arise between the age groups studied (Age Group), being the effect size of studies with adults larger than with children. Lastly, there also are differences between outcomes obtained with registering systems (Observation or Registering Systems), obtaining a greater effect size with an analytic approach, followed by EMG and a Global one.

No evidence was found in support of an effect of the remaining moderating variables (Whole Expression vs Specific Facial Movements, Sociality, and Self-Awareness) on the magnitude of observed correlation.

Publication Bias

Egger's test did not provide evidence of a relevant publication bias ($z = 0.75$, $p = .451$). No evidence of a relevant publication bias was found according to Rosenthal's and Rosenberg's fail-safe n either: it would be necessary to add 2142 or 1806 studies

with null effects, respectively, to reduce the correlation combined estimation to a non-significant magnitude with an alpha level of .05 (see funnel plot in the lower panel of Figure 3). However, we found a negative asymmetry compatible with the effect of publication bias according to the Trim-and-Fill method, that, when corrected, yielded an estimate of $r = .26$ [.19 - .34].

Results summary

In summary, the estimate of the correlation between the intensity of the reported emotion and the intensity of the observed expression was within a CI that ranged from .18 to .41. Four moderators, Category of Emotion, Vividness of the eliciting stimulus, participants' Age, and Registering System, explained a substantial portion of the observed heterogeneity among the correlations found in the studies included in the analysis. The highest correlation was obtained for disgust.

The case of amusement

Durán et al. (2017) found a greater frequency of co-occurrence of emotion and expression for amusement– greater than that found for the six basic-emotion categories included in our three previous analyses.

Figures 6a and 6b show the effect size estimates for expressions of amusement, taken from studies that used humor elicitors. The method and procedures of these studies were essentially identical to the co-occurrence studies on classic basic emotions reviewed in this article, but effect sizes coming from the same study were clustered for obtaining the robust estimations. The weighted mean estimate of the proportions of participants who displayed the predicted expression was .60, and its CI at 95% was [.38-.79]; for correlation between intensity of reported emotion and expression the weighted mean estimate was .41 [.34-.49].

Results summary

Amusement co-occurred with smile or laughter, on average, in 60% of the cases. This index of co-occurrence is in stark contrast with estimate for the six basic emotion categories: The estimated proportion of cases that fail to show the amusement expression range from .21 to .62 according to the corresponding confidence interval.

--- Figure 6a---

--- Figure 6b---

Discussion and Conclusions

The first analysis examined the most rigorous tests of the co-occurrence hypothesis: that the full predicted facial expression, as measured by the dominant assessment method, FACS, co-occurs with a basic emotion. The hypothesis failed in a majority of cases.

In the second analysis the assessment of facial behavior was expanded into not just whole but also part of the predicted facial movements. The point estimation of co-occurrence obtained from Whole + Partial Predicted Expressions meta-analysis should be the same or higher than the one obtained in the Whole Predicted Expressions Only meta-analysis, because it included data from more liberal tests of co-occurrence. And indeed the proportion of co-occurrences increased from .13 (for Whole Expressions only) to .23, but, according to the CIs of these estimations, the difference between the mean estimates of the two meta-analyses was not statistically significant. Surprisingly, no reliable evidence supports the intuition that looser criteria lead to significantly higher co-occurrence effect sizes. A barplot with a comparison of the summary results of the first and second analysis by emotion category is shown in Figure 7.

--- Figure 7---

In the third meta-analysis, the correlation between intensity of emotion and that of facial behaviors, the CI for the weighted estimate indicates a moderate effect size (.30, [0.18 - 0.41]). An interpretation, based on the results of the third analysis is that research should turn to contextual factors other than determine intensity of facial movement (e.g., Wenzler, Levine, van Dick, Oertel-Knöchel, & Aviezer, 2016).

Confidence in the Findings

A first source of concern about these meta-analyses is the striking shortage of effect size estimates from strict tests of co-occurrence (i.e., co-occurrence between reported emotion and whole predicted facial expressions assessed with a precise and systematic coding system). Given this limited sample size, we carried out a second and third meta-analyses with less strict requirements or a different dependent variable. Their results were not much more encouraging and confirmed the reliability of the first meta-analysis. Additionally, the case of amusement shows that a substantial degree of co-occurrence between psychological processes and facial behavior exists beyond the six classic categories of basic emotion.

A second obvious problem for the reliability of the theoretical interpretation of the results is the unexplained heterogeneity of the reported effect sizes. The statistical test of heterogeneity for the reported effect sizes was significant in the two meta-analyses but there was an important difference in the results of the analyses of moderating variables.

In the first meta-analysis, two moderators (Category of Emotion, i.e., the specific category of emotion addressed in the study, and Publication Date) explained 53% of the observed heterogeneity. Publication Date explained differences in the magnitude of the effect size, such that the tendency is for older studies to report greater proportions of participants who displayed the predicted expressions. There are several

potential explanations of this finding: new ethical constraints, more careful experimental procedures, more demanding statistical requirements for significance, etc.

In the second meta-analysis three moderators (Category of Emotion, Whole Expression vs. Specific Facial Movements, and Sociality) explained 22% of the observed heterogeneity. Some heterogeneity was simply due to the increased variability in the set of facial expressions that counted as predicted in each study. Other heterogeneity was due to differences in co-occurrence during social interaction versus non-interaction. This last finding is consistent with the sociality effect found for expressions (Fridlund, 1991).

Differences in the amounts of co-occurrence between emotion categories were found in both the first and second meta-analyses. What is more, additional heterogeneity among studies was found *within some emotion categories*. Our list of moderators did not fully account for this remaining heterogeneity in all cases.

With respect to the third meta-analysis (correlation between intensities of emotions and of facial expressions) the moderating variable Category of Emotion (i.e., the category of basic emotion addressed by the study) showed, as in the two other meta-analyses, that there were differences between emotion categories. Additionally, Vividness of the Emotion Elicitors and Age Group also had an impact on co-occurrence. The highest co-occurrence estimates were obtained by showing films, images or texts, whereas lower estimates were obtained when the elicitor was an actual emotional situation. This finding is counter-intuitive given that vividness of the stimuli seems to have an inverse effect in coherence between emotion and expression, and it opens the way to a broader question on what precisely is the meaning of an emotion category when it is reported about stimuli with different degree of vividness.

The effect size of studies with adults was larger than with children, which is also surprising, given that children are typically expected to be more expressive than adults. These findings suggest that experimental procedures play a significant role in the way in which participants categorize and display their emotional reactions.

Our interpretation is that most of this heterogeneity is due to the lack of a unitary measurement of the predicted facial expressions for testing co-occurrence, and the unavoidable subjective character of self-reports, the most common measure of emotion in the studies on co-occurrence. Although future studies could examine additional moderators, we are pessimistic about the possibility of attaining a significant reduction of the reported heterogeneity. Our hypothesis, in line with Barrett et al. (2019), is that most of the heterogeneity comes not only from the limitations of the measures used in these studies, but from a confound of the observed behavior with its conceptual interpretation.

In any case, the observed heterogeneity points to variables that can be usefully studied. An illustrative example, taken from studies sampled in this meta-analysis, are the absence of systematic information about unpredicted facial movements observed in the experiments, which raises the question about the potential existence of a significant co-occurrence of some categories of basic emotion with unpredicted facial movements.

Implications for Research

Do these data reveal an empirical impossibility of finding a higher degree of co-occurrence in the current state of the art? It could be argued that methodological problems such as weak experimental elicitors or poor ecological validity of the situations studied led researchers to underestimate the true co-occurrence of emotion and expression. However, our findings on co-occurrence are consistent with other meta-analytic reviews that found that classic emotions are not concurrent with distinct

autonomic nervous system “fingerprints” (Siegel et al., 2018) or brain regions (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012).

Additionally, this argument is weakened by a special case in our own review, the expression of amusement, which yielded a higher degree of co-occurrence between reports of amusement and smiling/laughter through studies methodologically identical to most of the co-occurrence studies for the six basic-emotion categories. It could be argued that is easier to elicit amusement in the laboratory than to elicit emotions like fear or sadness, but some categories of basic emotion that are relatively unproblematic to elicit in the laboratory, as disgust or surprise, had much lower co-occurrence indexes than amusement. This finding shows that it is possible to detect a higher degree of co-occurrence between psychological processes and facial configurations if it existed. Studies on co-occurrence can be feasible and fruitful if researchers focus on other not necessarily emotional facial behaviors.

The relatively high co-occurrence reported in the studies on amusement suggests that the expression of amusement is probably a facial behavior that should be considered a potential candidate for a universal, uniform signal. Laughter and smiles in humorous contexts are combined with social vocalizations and they reflect a complex cognitive and affective process with multiple functions, such as for example facilitation of communication, induction of positive affect in others, and self-regulation (e.g., Gervais & Wilson, 2005). Interestingly, laughter is a human facial display clearly homologous to a facial display (the “play face”) observed in other primates (e.g., Owren & Bachorowski, 2003).

In the light of these findings, we believe that researchers should approach the study of facial behavior from a fresh perspective. In the cumulative data on co-occurrence gathered in our analyses, the CI shows that between 6% and 27% of the

participants who reported one of the six basic emotions produced the whole predicted facial expression. The theory that, for the last fifty years, has inspired the tests of co-occurrence of facial expression and emotion has not provided conclusive evidence that specific facial expressions are a constitutive “characteristic which distinguish basic emotions from one another and from other affective phenomena” (Ekman, 1992, p. 175).

Moving Beyond

The results of these meta-analyses can be read from more than one single fresh perspective, and it can be useful for all of them.

From an applied point of view, these meta-analyses question a simplistic psychological axiom ingrained in introductory psychology lectures, clinical and educational practice, and algorithms produced by multinational information-technology companies and governments. Our data reinforces previous warnings (e.g. Fernández-Dols & Ruiz-Belda, 1997; Barrett et al., 2019) that the extremely widespread assumption of a universal, uniform and stable production of six facial expressions of basic emotion is not valid (for an exhaustive analysis of this problem see Barrett et al. 2019).

From a theoretical point of view, our data constitute a necessary “ground truth” for the search of a new, more dynamic view of the function, morphology, and signal value of facial behavior.

On the one hand, this work is an unavoidable empirical reference for those authors who have developed new arguments in favor of the existence of dynamic, multimodal expressions of basic emotion (e.g., Cowen et al., 2019). The results of our meta-analyses can be read as a support for the assumption that emotional experience “does not manifest itself in prototypical face-muscle configurations alone but rather in

multimodal expressions, and varieties of expressions within a given modality” (Cowen et al., 2019, p. 83). Our findings could even be read as an invitation to explore the hypothesis of a multimodal expression of emotion, but they are also a warning of the formidable methodological challenges that testing the hypothesis of a strong co-occurrence between multimodal expressions and basic emotions will have to contemplate. Our summary of a fifty-year search for a strong co-occurrence between a single-modal expression and a limited set of basic emotions are plenty of important lessons. For example, as we pointed out above, the unexplained heterogeneity of the reported effect sizes of co-occurrence in our meta-analyses is most probably the consequence of a lack of conceptual precision. This problem might be aggravated by future multimodal co-occurrence studies if they include a large number of insufficiently precise variables (cf. Cowen & Keltner, 2017; Cowen et al., 2021; Barrett, 2021).

On the other hand, our meta-analyses can also be a useful empirical reference for other approaches with differences of degree in the support of the existence of basic emotions, and the existence of universal expressions, because our data could also be read as a disconfirmation of one of the central features of the concept of expression (its essential co-occurrence with the experience of emotion), and a call for alternative approaches to the role of facial behavior in affective and emotional processes.

In the framework of alternative theoretical approaches to the Basic Emotion Theory some research teams are exploring new forms of co-occurrence between facial movements and appraisals (e.g., Scherer, Ellgring, Dieckmann, Unfried & Mortillaro, 2019), while others hypothesize that facial movements are a powerful source of social influence (e.g., Crivelli & Fridlund, 2019). Meanwhile new models are renewing the classic assumptions on the functions and evolutionary origins of facial expressions (e.g., from optical needs; e.g., Lee, Mirza, Flanagan & Anderson, 2014), and the

communicative functions of facial behavior (e.g., Fernández Dols, 2017, Parkinson, 2021).

Whatever the theoretical framework, the findings of these meta-analyses strongly suggest that the field needs more proactive and precise methods for studying the two terms of the causal relationship. We need to widen our research by de-coupling emotion and facial behavior. Emotion may have correlates in more than the face. Facial behavior may have correlates in more than emotion.

Author contributions

JMFD and JD developed the study concept. All authors contributed to the study design. JMFD and JD performed data collection. JD performed the data analysis. All authors contributed to the interpretation of the results. JMFD drafted the paper, and JD provided critical revisions and wrote parts of the main text. The two authors approved the final version of the paper for submission.

References

References with an asterisk () are the studies included in the meta-analysis*

- *Ajaya, Y., Peckham, A.D., & Johnson, S.L. (2016). Emotion regulation and mania risk: Differential responses to implicit and explicit cues to regulate. *Journal of Behavior Therapy and Experimental Psychiatry*, 50, 283-288.
<https://doi.org/10.1016/j.jbtep.2015.10.004>
- * Armstrong, J.E., Hutchinson, I., Laing, D.G., & Jinks, A.L. (2007). Facial electromyography: responses of children to odor and taste stimuli. *Chemical senses*, 32(6), 611-621. <https://doi.org/10.1093/chemse/bjm029>
- *Beerman, U., & Ruch, W. (2011). Can People Really “Laugh at Themselves?”-Experimental and Correlational Evidence. *Emotion*, 11(3), 492-501.
<https://doi.org/10.1037/a0023444>
- Barrett, L.F. (2021). AI weighs in on debate about universal facial expressions. *Nature*, 589(7841), 202–203. doi:10.1038/d41586-020-03509-5
- 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, 1-68.
<https://doi.org/10.1177/1529100619832930>
- 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
- *Boecker, L., Likowski, K. U., Pauli, P., & Weyers, P. (2015). The face of schadenfreude: Differentiation of joy and schadenfreude by electromyography. *Cognition and Emotion*, 29(6), 1117-1125.
<https://doi.org/10.1080/02699931.2014.966063>

- *Bonnano, G. A., & Keltner, D. (2004). The coherence of emotion systems: comparing “on line” measures of appraisal and facial expressions, and self-report. *Cognition and Emotion*, 18, 431-444. <https://doi.org/10.1080/02699930341000149>
- Borenstein, M., Hedges, L.V., Higgins, J.P., & Rothstein, R.H. (2009). *Introduction to meta-analysis*. London: Wiley. <https://doi.org/10.1002/9780470743386>
- Borenstein M., Hedges L.V., Higgins J.P.T., & Rothstein H.R. (2010). A basic introduction to fixed-effects and random effects models for meta-analysis. *Research Synthesis Methods*, 1, 97–111. <https://doi.org/10.1002/jrsm.12>
- Botella, J., & Sánchez-Meca, J. (2015). *Meta-análisis en ciencias sociales y de la salud*. Madrid: Síntesis.
- *Brown, S.-L., & Schwartz, G. E. (1980). Relationships between facial electromyography and subjective experience during affective imagery. *Biological Psychology*, 11, 49-62. [https://doi.org/10.1016/0301-0511\(80\)90026-5](https://doi.org/10.1016/0301-0511(80)90026-5)
- Camras, L.A., Castro, V.L., Halberstadt, A.G., & Shuster, M.M. (2017). Spontaneously produced facial expressions in infants and children. In J.M. Fernández-Dols, & J.A. Russell (Eds.), *The Science of Facial Expression* (pp. 279-296). New York: Oxford University Press.
- <https://doi.org/10.1093/acprof:oso/9780190613501.003.0015>
- Camras, L.A., Meng, Z., Ujie, T., Dharamsi, S., Miyake, K., Oster, H., Wang, L., Cruz, J., Murdoch, A., & Campos, J. (2002). Observing emotion in infants: Facial Expression, body behavior, and rater judgments of responses to an expanticy-violating event. *Emotion*, 2(2), 179-193. <https://doi.org/10.1037/1528-3542.2.2.179>
- *Cannon, P.R., Schnall, S., & White, M. (2011). Transgressions and expressions: Affective facial muscle activity predicts moral judgments. *Social Psychological*

and Personality Science, 2(3), 325-331.

<https://doi.org/10.1177/1948550610390525>

Caruso, D.R., & Salovey, P. (2004). *The Emotionally Intelligent Manager: How to Develop and Use Four Key Emotional Skills of Leadership*. San Francisco, CA: Jossey-Bass.

*Castro, V., Camras, L.A., Halberstadt, A.G., & Shuster, M. (2018). Children's prototypic facial expression during emotion-eliciting conversations with their mothers. *Emotion*, 18, 260-276. <https://doi.org/10.1037/emo0000354>

*Ceschi, G., & Scherer, K.R. (2003). Children's ability to control the facial expression of laughter and smiling: Knowledge and behavior. *Cognition and Emotion*, 17(3), 385-411. <https://doi.org/10.1080/02699930143000725>

*Chapman, H.A., Kim, D.A., Susskind, J.M., & Anderson, A.K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, 323, 1222-1226. <https://doi.org/10.1126/science.1165565>

Cordaro, D.T., Sun, R., Kamble, S., Hodder, N., Monroy, M., Cowen, A., Bai, Y., & Keltner, D. (2020). The recognition of 18 facial-bodily expressions across nine cultures. *Emotion*, 20, 1292-1300. <https://doi.org/10.1037/emo0000576>

Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, 114(38), E7900–E7909. doi:10.1073/pnas.1702247114

Cowen, A., Sauter, D., Tracy, J.L., & Keltner, D. (2019). Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychological Science in the Public Interest*, 20, 69-90. <https://doi.org/10.1177/1529100619850176>

- Cowen, A.S., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2021). Sixteen facial expressions occur in similar contexts worldwide. *Nature* 589, 251–257.
<https://doi.org/10.1038/s41586-020-3037-7>
- *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, 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 Science*, 22, 388-399.
<https://doi.org/10.1016/j.tics.2018.02.006>
- *Davidson, R.J. (1992). Prolegomena to the structure of emotion: Gleanings from neuropsychology. *Cognition & Emotion*, 6, 245–268.
<https://doi.org/10.1080/02699939208411071>
- *Davydov, D.M., Zech, E., & Luminet, O. (2011). Affective context of sadness and physiological response patterns. *Journal of Psychophysiology*, 25(2), 67-80.
<https://doi.org/10.1027/0269-8803/a000031>
- Durán, J. I., Reisenzein, R., & Fernández-Dols, J. M. (2017). Coherence between emotions and facial expressions: A research synthesis. In J.M. Fernández-Dols, & J.A. Russell (Eds.), *The Science of Facial Expression* (pp. 107-129). New York: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780190613501.003.0007>
- Duval, S., & Tweedie, R. (2000). A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95(449), 89-98. <https://doi.org/10.2307/2669529>

- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 629-634.
<https://doi.org/10.1136/bmj.315.7109.629>
- Ekman, P. (1972). Universal and cultural differences in facial expressions of emotion. In J. R. Cole (Ed.), *Nebraska Symposium on Motivation, 1971*, (pp. 207-283). Lincoln, NE: Nebraska University Press.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.
<https://doi.org/10.1080/02699939208411068>
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48, 376-379.
<https://doi.org/10.1037/0003-066x.48.4.384>
- 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. <https://doi.org/10.1093/acprof:oso/9780190613501.003.0003>
- Ekman, P. (2016). What scientists who study emotion agree about. *Perspectives on Psychological Science*, 11 (1), 31-34. <https://doi.org/10.1177/1745691615596992>
- Ekman, P., & Friesen, W.V. (1978). *Facial Action Coding System (FACS). A technique for the measurement of facial action*. Palo Alto, CA: Consulting Psychologists Press.
- *Ekman, P., Friesen, W.V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology*, 39, 1125-1134.
<https://doi.org/10.1037/h0077722>
- Ekman, P., Friesen, W. V., & Hager, J. V. (2002). *Facial Action Coding System (2nd edition)*. Salt Lake City, UT: Research Nexus eBook
- Ekman, P., Friesen, W.V., & Tomkins, S.S. (1971). Facial Affect Scoring Technique: A first validity study. *Semiotica*, 3, 37-58. <https://doi.org/10.1515/semi.1971.3.1.37>

- Fernández-Dols, J-M. (2017). Natural facial expression: A view from social constructionism and pragmatics. In J.M. Fernández-Dols, & J.A. Russell (Eds.) *The Science of Facial Expression* (pp. 457-475). New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190613501.003.0024>
- Fernández-Dols, J-M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review*, 5, 24-29. <https://doi.org/10.1177/1754073912457229>
- *Fernández-Dols, J-M., Sánchez, F., Carrera, P., & Ruiz-Belda, M.A. (1997). Are spontaneous expressions and emotions linked? An experimental test of coherence. *Journal of Nonverbal Behavior*, 21, 163-177. <https://doi.org/10.1023/a:1024917530100>
- *Fiacconi, C.M., & Owen, A.M. (2015). Using psychophysiological measures to examine the temporal profile of verbal humor elicitation. *PLoS ONE* 10(9): e0135902. <https://doi.org/10.1371/journal.pone.0135902>
- 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>
- Fridlund, A. (2017). The behavioral ecology view of facial displays, 25 years later. In J.M. Fernández-Dols, & J.A. Russell (Eds.), *The Science of Facial Expression* (pp. 77-92). New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190613501.003.0005>
- Gervais, M., & Wilson, D.S. (2005). The evolution and functions of laughter and humor: A synthetic approach. *The Quarterly Review of Biology*, 80, 395-430. <https://doi.org/10.1086/498281>
- *Gross, J.J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression and physiology. *Journal of Personality*

and Social Psychology, 74(1), 224-237. <https://doi.org/10.1037/0022-3514.74.1.224>

*Gross, J.J., John, O.P., & Richards, J.M. (2000). The dissociation of emotion expression from emotion experience: A personality perspective. *Personality and Social Psychology Bulletin*, 26, 712-726.
<https://doi.org/10.1177/0146167200268006>

*Gruber, J., Hay A.C., & Gross, J.J. (2014). Rethinking emotion: Cognitive reappraisal is an effective positive and negative emotion regulation strategy in bipolar disorder. *Emotion*, 14(2), 388-396. <https://doi.org/10.1037/a0035249>

Hall, J.A., & Horgan, T.G. (2003). Happy affect and smiling: is their relation moderated by interpersonal power? *Emotion*, 3(3), 303-309. <https://doi.org/10.1037/1528-3542.3.3.303>

*Harrigan, J.A., & O'Connell, D.M. (1996). How do you look when feeling anxious? Facial displays of anxiety. *Personality and Individual Differences*, 21, 205-212.
[https://doi.org/10.1016/0191-8869\(96\)00050-5](https://doi.org/10.1016/0191-8869(96)00050-5)

Harris, C.R., & Alvarado, N. (2005). Facial expressions, smile types, and self-report during humour, tickle, and pain. *Cognition and Emotion*, 19, 655-669.
<https://doi.org/10.1080/02699930441000472>

Hedges, L.V. (1981). Distribution theory for Glass's estimator effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107-128.
<https://doi.org/10.2307/1164588>

Hedges L.V., & Olkin I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press. <https://doi.org/10.1016/c2009-0-03396-0>

- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*, 39–65. <https://doi.org/10.1002/jrsm.5>
- *Herring, D.R., Burleson, M.H., Roberts, N.A., & Devine, M.J. (2011). Coherent with laughter: Subjective experience, behavior, and physiological responses during amusement and joy. *International Journal of Psychophysiology, 79*, 211-218. <https://doi.org/10.1016/j.ijpsycho.2010.10.007>
- Hjortsjö, C.H. (1969). *Man's face and mimic language*. Lund, Sweden: Studentlitteratur.
- *Hubbard, J.A., Smithmyer, C.M., Ramsden, S.R., Parker, E.H., Flanagan, K.D., Dearing, K.F., Relyea, N., & Simons, R.F. (2002). Observational, physiological, and self-report measures of children's anger: Relations to reactive versus proactive aggression. *Child Development, 73*(4), 1101-1118. <https://doi.org/10.1111/1467-8624.00460>
- Huedo-Medina T.B., Sánchez-Meca J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I² index? *Psychological Methods, 11*, 193–206. <https://doi.org/10.1037/1082-989x.11.2.193>
- Izard, C.E. (1979). *The maximally discriminative facial movement coding system (Max)*. Newark, DE: University of Delaware, Office of Academic Computing and Instructional Technology.
- Izard, C. E., & Abe, J. A. A. (2004). Developmental changes in facial expressions of emotions in the strange situation during the second year of life. *Emotion, 4*(3), 251-265. <https://doi.org/10.1037/1528-3542.4.3.251>
- Jack, R.E., & Schyns, Ph. (2015). The human face as a dynamic tool for social communication. *Current Biology, 25*, R621-R634. <https://doi.org/10.1016/j.cub.2015.05.052>

- *Jäncke, L., & Kaufmann, N. (1994). Facial EMG responses to odors in solitude and with an audience. *Chemical Senses*, 19, 99-111.
<https://doi.org/10.1093/chemse/19.2.99>
- *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, 299-321. <https://doi.org/10.1080/02699930903384667>
- *Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, 68, 441-454. <https://doi.org/10.1037/0022-3514.68.3.441>
- *Keltner, D., & Bonanno, G.A. (1997). A study of laughter and dissociation: Distinct correlates of laughter and smiling during bereavement. *Journal of Personality and Social Psychology*, 73, 687-702. <https://doi.org/10.1037/0022-3514.73.4.687>
- Keltner, D., & Cordaro, D.T. (2017). Understanding multimodal emotional expressions: Recent advances in Basic Emotion Theory. In J.M. Fernández-Dols, & J.A. Russell (Eds.), *The Science of Facial Expression* (pp. 57-75). New York, NY: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780190613501.003.0004>
- *Kreibig, S.D., Samson, A.C., & Gross, J.J. (2013). The psychophysiology of mixed emotional states. *Psychophysiology*, 50, 799-811.
<https://doi.org/10.1111/psyp.12064>
- Lee, D.H., Mirza, R., Flanagan, J.G., & Anderson, A.K. (2014). Optical origins of opposing facial expression actions. *Psychological Science*, 25, 745-752.
<https://doi.org/10.1177/0956797613514451>

- *Lerner, J.S., Dahl, R., Hariri, A.R., & Taylor, S.E. (2007). Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses. *Biological Psychiatry*, 61, 253-260. <https://doi.org/10.1016/j.biopsych.2006.08.016>
- Lierheimer, K., & Stichter, J. (2011). Teaching facial expressions of emotion. *Beyond Behavior*, 21, 20-27.
- Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35 (3), 121-143. <https://doi.org/10.1017/S0140525X11000446>
- *Ludden, G.D.S., Schifferstein, H.N.J., & Hekkert, P. (2009). Visual - tactual incongruities in products as sources of surprise. *Empirical Studies of the Arts*, 27, 63-89. <https://doi.org/10.2190/em.27.1.d>
- * Lumley, M. A., & Melamed, B. G. (1992). Blood phobics and nonphobics: psychological differences and affect during exposure. *Behaviour research and therapy*, 30(5), 425-434. [https://doi.org/10.1016/0005-7967\(92\)90026-d](https://doi.org/10.1016/0005-7967(92)90026-d)
- Matsumoto, D., & Kupperbusch, C. (2001). Idiocentric and allocentric differences in emotional expression, experience, and the coherence between expression and experience. *Asian Journal of Social Psychology*, 4, 113-131. <https://doi.org/10.1111/j.1467-839x.2001.00080.x>
- Matsumoto, D., Keltner, D., Shiota, M.N., O'Sullivan, M., & Frank, M. (2008). Facial expressions of emotion. In: M. Lewis, J.M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions (3rd edition)* (pp. 211-234). New York: The Guilford Press.
- *Matsumoto, D., & Willingham, B. (2006). The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners of the 2004 Athens Olympic

Games. *Journal of Personality and Social Psychology*, 91(3), 568-581.

<https://doi.org/10.1037/0022-3514.91.3.568>

- Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kookan, K., Ekman, P., Yrizarry, N., Loewinger, S., Uchida, H., Yee, A., Amo, L., & Goh, A. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBERT). *Journal of Nonverbal Behavior*, 24(3), 179–209. <https://doi.org/10.1023/A:1006668120583>
- *Matsumoto D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology*, 96(1), 1-10. <https://doi.org/10.1037/a0014037>
- *Mauss, I.B., Levenson, R.W., McCarter, L., Wilhelm, F.H., & Gross, J.J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, 5, 175-190. <https://doi.org/10.1037/1528-3542.5.2.175>
- *Mui, P.H.C., Goudbeek, M.B., Swerts, M.G.J., & Hovasapian, A. (2017). Children's nonverbal displays of winning and losing: Effects of social and cultural contexts on smiles. *Journal of Nonverbal Behavior*, 41, 67-82. <https://doi.org/10.1007/s10919-016-0241-0>
- *Namba, S., Kagamihara, T., Miyatani, M., & Nakao, T. (2017). Spontaneous facial expressions reveal new action units for the sad experiences. *Journal of Nonverbal Behavior*, 41, 203-220. <https://doi.org/10.1007/s10919-017-0251-6>
- Nelson, N.L., & Russell, J.A. (2013). Universality revisited. *Emotion Review*, 5, 8-15. <https://doi.org/10.1177/1754073912457227>
- Newcombe, R.G. (2012). *Confidence intervals for proportions and related measures of effect size*. Boca Raton, FL: CRC Press. <https://doi.org/10.1201/b12670>

- Oatley, K., & Johnson-Laird, P. N. (2011). Basic Emotions in Social Relationships, Reasoning, and Psychological Illnesses. *Emotion Review*, 3(4), 424–433.
doi:10.1177/1754073911410738
- Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315–331. <https://doi.org/10.1037/0033-295X.97.3.315>
- Owren, M.J., & Bachorowski, J.A. (2003). Reconsidering the evolution of nonlinguistic communication: The case of laughter. *Journal of Nonverbal Behavior*, 27, 183-200. <https://doi.org/10.1023/a:1025394015198>
- *Park S., & Kim, K. (2011). Physiological reactivity and facial expression to emotion-inducing films in patients with schizophrenia. *Archives of Psychiatric Nursing*, 25(6), 37-47. <https://doi.org/10.1016/j.apnu.2011.08.001>
- Passarelli, M., Masini, M., Bracco, F., Petrosino, M., & Chiorri, C. (2018). Development and validation of the Facial Expression Recognition Test (FERT). *Psychological Assessment*, 30(11), 1479–1490.
<https://doi.org/10.1037/pas0000595>
- R Core Team (2016). R: A language and environment for statistical computing. R. Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Reisenzein, R. (2000). Exploring the strength of association between the components of emotion syndromes: The case of surprise. *Cognition and Emotion*, 14, 1-38.
<https://doi.org/10.1080/026999300378978>
- Reisenzein, R., Studtmann, M. & Horstmann, G. (2013). Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emotion Review*, 5, 16-23. <https://doi.org/10.1177/1754073912457228>

- *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, 295-315. <https://doi.org/10.1037/0022-3514.91.2.295>
- Rosenberg, M. S. (2005). The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution*, 59, 464–468. <https://doi.org/10.1111/j.0014-3820.2005.tb01004.x>
- Rosenberg, E.L., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition and Emotion*, 8, 201-229. <https://doi.org/10.1080/02699939408408938>
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, 86, 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (revised edition). Newbury Park, CA: Sage. <https://doi.org/10.4135/9781412984997>
- Ruch, W. (1995). Will the real relationship between facial expression and affective experience please stand up: The case of exhilaration. *Cognition and Emotion*, 9, 33-58. <https://doi.org/10.1080/02699939508408964>
- *Ruch, W. (1997). State and trait cheerfulness and the induction of exhilaration: A FACS study. *European Psychologist*, 2, 328-341. <https://doi.org/10.1027/1016-9040.2.4.328>
- Ruch, W., Platt, T., Proyer, & Hsueh-Chih, Ch. (2019). Editorial: Humor and laughter, playfulness and cheerfulness: Upsides and downsides to a life of lightness. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00730>

- *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, 315-326. <https://doi.org/10.1080/02699930302288>
- Sauter, D. & Russell, J. A. (in press). Nonverbal expressions of emotion. In A. Scarantino (Ed.). *Routledge Handbook of Emotion Theory*. London: Routledge.
- Scarantino, A. (2017). How to do things with emotional expressions: The theory of affective pragmatics. *Psychological Inquiry*, 28(2-3), 165–185. <https://doi.org/10.1080/1047840x.2017.1328951>
- Scherer, K. R. (2013). The evolutionary origin of multimodal synchronization in emotional expression. *Journal of Anthropological Sciences*, 91, 185-200. <https://doi.org/10.1093/acprof:oso/9780199583560.003.0001>
- Scherer K.R., Ellgring H., Dieckmann A., Unfried M., & Mortillaro M. (2019). Dynamic Facial Expression of Emotion and Observer Inference. *Frontiers in Psychology*, 10:508. <https://doi.org/10.3389/fpsyg.2019.00508>
- Scherer, K.R., Scherer, U. (2011). Assessing the Ability to Recognize Facial and Vocal Expressions of Emotion: Construction and Validation of the Emotion Recognition Index. *Journal of Nonverbal Behavior*, 35, 305-326. <https://doi.org/10.1007/s10919-011-0115-4>
- Scherer, K. R., Zentner, M. R., & Stern, D. (2004). Beyond Surprise: The puzzle of infants' expressive reactions to expectancy violation. *Emotion*, 4(4), 389-402. <https://doi.org/10.1037/1528-3542.4.4.389>
- *Schmidt, K.L., Cohn, J.F., & Tian, Y. (2003). Signal characteristics of spontaneous facial expressions: Automatic movement in solitary and social smiles. *Biological Psychology*, 65(1), 49-66. [https://doi.org/10.1016/s0301-0511\(03\)00098-x](https://doi.org/10.1016/s0301-0511(03)00098-x)

- *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, 657–664.
<https://doi.org/10.1016/j.evolhumbehav.2012.04.003>
- Schwartz, O. (2019, March 6). Don't look now: why you should be worried about machines reading your emotions. *The Guardian*. Retrieved from <https://www.theguardian.com/>
- *Shenhav, A., & Mendes, W.B. (2014). Aiming for the stomach and hitting the heart: Dissociable triggers and sources for disgust reactions. *Emotion*, 14(2), 301-309.
<https://doi.org/10.1037/a0034644>
- Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K. S., & Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin*, 144(4), 343–393.
<https://doi.org/10.1037/bul0000128>
- *Slomine, B.S., & Greene, A.F. (1993). Anger imagery and corrugator electromyography. *Journal of Psychosomatic Research*, 37(6), 671-676.
[https://doi.org/10.1016/0022-3999\(93\)90062-k](https://doi.org/10.1016/0022-3999(93)90062-k)
- *Soussignan, R., & Schaal, B. (1996). Children's facial responsiveness to odors: Influences to hedonic valence of odor, gender, age, and social presence. *Developmental Psychology*, 32(2), 367-379. <https://doi.org/10.1037/0012-1649.32.2.367>
- Stark, R., Bertram, W., Schienle, A., & Vaitl, D. (2005). Psychophysiological correlates of disgust and disgust sensitivity. *Journal of Psychophysiology*, 19, 50-60.
<https://doi.org/10.1027/0269-8803.19.1.50>

- *Tomarken, A.J., & Davidson, R.J. (1993) Fear and the face: Differential accessibility of fear and disgust to the face [Manuscript in preparation].
- *Topolinski, S., & Strack, F. (2015). Corrugator activity confirms immediate negative affect in surprise. *Frontiers in Psychology*, 6, 134.
<https://doi.org/10.3389/fpsyg.2015.00134>
- *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, 380-397. <https://doi.org/10.1037/1528-3542.2.4.380>
- *Underwood M.K., & Bjornstad, G.J. (2001). Children's emotional experience of peer provocation. The relation between observed behavior and self-reports of emotions, expressions, and social goals. *International Journal of Behavioral Development*, 25(4), 320-330. <https://doi.org/10.1080/01650250143000085>
- United States Government Accountability Office (2013). *Aviation Security: TSA Should Limit Future Funding for Behavior Detection Activities*. Washington DC: Author.
- *Van der Graaff, J., Meeus, W., de Wied, M., van Boxtel, A., van Lier, P.A., Koot, H. M., & Branje, S. (2016). Motor, affective and cognitive empathy in adolescence: Interrelations between facial electromyography and self-reported trait and state measures. *Cognition and Emotion*, 30(4), 745-761.
<https://doi.org/10.1080/02699931.2015.1027665>
- *Vanhamme, J. (2000). The link between surprise and satisfaction: An exploratory research on how to best measure surprise. *Journal of Marketing Management*, 16, 565-582. <https://doi.org/10.1362/026725700785045949>
- 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, 403-405. <https://doi.org/10.1017/s0140525x09991026>

- *Vanhamme, J. (2003). Surprise...surprise. An empirical investigation on how surprise is connected to consumer satisfaction. In *ERIM Report Series Research in Management ERS-2003-005-MKT*. Rotterdam: Erasmus Research Institute of Management. <https://ssrn.com/abstract=411645>
- Verduyn, P., Van Mechelen, I., Tuerlinckx, F., Meers, K., & Van Collie H. (2009). Intensity profiles of emotional experience over time. *Cognition and Emotion*, 23, 1427-1443. <https://doi.org/10.1080/02699930902949031>
- *Vernon, L.L., & Berenbaum, H. (2002). Disgust and fear in response to spiders. *Cognition and Emotion*, 16, 809-830. <https://doi.org/10.1080/02699930143000464>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48. <https://doi.org/10.18637/jss.v036.i03>
- *Visser, M., Krahmer, E., & Swerts, M. (2014). Contextual effects on surprise expressions: A developmental study. *Journal of Nonverbal Behavior*, 38, 523-547. <https://doi.org/10.1007/s10919-014-0188-y>
- *Visser, M., Krahmer, E., & Swerts, M. (2015). Children's spontaneous emotional expressions while receiving (un)wanted prizes in the presence of peers. *Frontiers in Psychology*, 6, 1401. <https://doi.org/10.3389/fpsyg.2015.01401>
- *Wang, N., Marsella, S., & Hawkins, T. (2008). Individual differences in expressive response: a challenge for ECA design. *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, 3, 1289-1292. <http://www.ccs.neu.edu/~marsella/publications/pdf/WangAAMAS08.pdf>
- Wenzler, S., Levine, S., van Dick, R., Oertel-Knöchel, V., & Aviezer, H. (2016). Beyond Pleasure and Pain: Facial Expression Ambiguity in Adults and Children

During Intense Situations. *Emotion*, 16, 807-814.

<https://doi.org/10.1037/emo0000185>

- *Whitton, A.E., Henry, J.D., Rendell, P.G., & Grisham, J.R. (2014). Disgust, but not anger provocation, enhances levator labii superioris activity during exposure to moral transgressions. *Biological Psychology*, 96, 48-56.

<https://doi.org/10.1016/j.biopsycho.2013.11.012>

- *Wu, P., Gonzalez, I., Patsis, G., Jiang, D., Sahli, H., Kerckhofs, E., & Vandekerckhove, M. (2014). Objectifying facial expressivity assessment of Parkinson's patients: preliminary study. *Computational and mathematical methods in medicine*, 2014, 1–12.

<https://doi.org/10.1155/2014/427826>

- *Zaalberg, R., Manstead, A., & Fischer, A. (2004). Relations between emotions, display rules, social motives, and facial behaviour. *Cognition and Emotion*, 18, 183-207.

<https://doi.org/10.1080/02699930341000040>

Table 1

Results of moderating variables analysis for Whole Predicted Facial Expression Only

Categorical Moderator	Category	<i>F</i> (df)	<i>p</i>	<i>R</i> ² *	<i>p_{rp}</i> **	CI (95%)
<i>Emotion Category</i>		70.75 (2, 4)	.001	.56		
	Happiness (<i>k</i> = 4)				.12	[-.02 - .51]
	Surprise (<i>k</i> = 5)				.06	[-.02 - .15]
	Disgust (<i>k</i> = 4)				.30	[-.20 - .41]
<i>Vividness of the Emotion Elicitor</i>		0.42 (2, 4)	.424	.00		
	Film-imag-text (<i>k</i> = 3) ^a				.15	[-.03 - .52]
	Imagery-remembrance (<i>k</i> = 2) ^a				.18	[-.08 - .35]
	Situation (<i>k</i> = 8)				.10	[-.03 - .27]
<i>Age Group</i>		1.02 (1, 5)	.358	.00		
	Adults (<i>k</i> = 11)				.12	[-.05 - .26]
	Children (<i>k</i> = 2) ^a				.19	[-.05 - .53]
<i>Sociality</i>		0.45 (1, 5)	0.532	.00		
	Social (<i>k</i> = 5)				.17	[-.06 - .38]
	Non-Social (<i>k</i> = 8)				.10	[-.03 - .29]
<i>Self-awareness</i>		1.73 (1, 5)	.245	.00		
	Aware (<i>k</i> = 5)				.10	[-.02 - .36]
	Notaware (<i>k</i> = 8)				.16	[-.07 - .35]
Quantitative Moderator		<i>F</i> (df)	<i>p</i>	<i>R</i> ² *	Slope	CI (95%)
<i>Intensity of the Reported Emotion</i>		173.32 (1, 4)	<.001	1	-12.70	[-15.38 - -10.02]
<i>Standardized Intensity</i>		3.14 (1, 4)	.151	.24	-1.63	[-4.19 - 0.93]
<i>Publication Date</i>		16.37 (1, 5)	.001	.61	.070	[-.03 - .11]

^a CI bounds could not be obtained through cluster robust estimates

* The amount of explained variance was obtained with tradicional (not robust) tests of moderating variables

** Reported *p_{rp}* and confidence interval bound values for categories of categorical moderators were the result of returning logit(*p_{rp}*) estimated values to raw proportion metric after analysis

Table 2

Results of moderating variables analysis for Whole Expressions and Specific Facial Movements

Categorical Moderator	Category	<i>F</i> (df)	<i>p</i>	<i>R</i> ² *	<i>p_{rp}</i> **	CI (95%)
<i>Emotion</i>		20.58 (5, 10)	<.001	.07		
	Happiness (<i>k</i> = 7)				.20	[.07 - .45]
	Surprise (<i>k</i> = 19)				.13	[.08 - .21]
	Disgust (<i>k</i> = 9)				.42	[.14 - .75]
	Sadness (<i>k</i> = 5)				.28	[.18 - .41]
	Anger (<i>k</i> = 5)				.35	[.14 - .64]
	Fear (<i>k</i> = 4)				.25	[.02 - .84]
<i>Vividness</i>		0.58 (2, 13)	.576	.00		
	Film-imag-text (<i>k</i> = 9)				.31	[.09 - .67]
	Imagery-reminiscence (<i>k</i> = 8)				.27	[.12 - .49]
	Situation (<i>k</i> = 31)				.19	[.12 - .28]
<i>Whole vs. Specific</i>		10.74 (1, 14)	.006	.10		
	Whole (<i>k</i> = 19)				.15	[.09 - .23]
	Specific (<i>k</i> = 31)				.29	[.21 - .39]
<i>Registering system</i>		1.13 (3, 12)	.377	.00		
	Analytic (<i>k</i> = 34)				.20	[.14 - .27]
	Global (<i>k</i> = 12)				.27	[.11 - .53]
	EMG (<i>k</i> = 4)				.49	[.01 - .99]
<i>Age group</i>		3.81 (1, 14)	.071	.00		
	Adults (<i>k</i> = 41)				.21	[.15 - .28]
	Children (<i>k</i> = 9)				.33	[.10 - .69]
<i>Sociality</i>		9.75 (2, 13)	.003	.00		
	Social (<i>k</i> = 23)				.27	[.20 - .36]
	Mere presence (<i>k</i> = 5)				.12	[.03 - .37]
	Non-social (<i>k</i> = 21)				.20	[.10 - .37]
<i>Self-awareness</i>		0.27 (1, 12)	.610	.00		
	Aware (<i>k</i> = 8)				.26	[.04 - .74]
	Not aware (<i>k</i> = 38)				.21	[.16 - .27]
Quantitative Moderator		<i>F</i> (df)	<i>p</i>	<i>R</i> ² *	Slope	CI (95%)
<i>Intensity</i>		4.40 (1, 12)	.058	.08	-2.86	[-5.83- 0.11]
<i>Standardized Intensity</i>		2.16 (1, 8)	.180	.10	-0.99	[-2.55 - 0.57]
<i>Publication Date</i>		1.67 (1, 14)	.218	.01	0.03	[-0.02 - 0.09]

* The amount of explained variance was obtained with tradicional (not robust) tests of moderating variables

** Reported *p_{rp}* and confidence interval bound values for categories of categorical moderators were the result of returning logit(*p_{rp}*) estimated values to raw proportion metric after analysis

Table 3

Results of moderating variables analysis for correlation between reported intensities of emotions and of facial behaviors

Categorical Moderator	Category	<i>F</i> (df)	<i>p</i>	<i>R</i> ^{2*}	<i>r</i> ^{**}	CI (95%)
<i>Emotion</i>		6.73 (5, 22) ^a	< .001	.61		
	Happiness (<i>k</i> = 7)				.32	[.24 - .39]
	Surprise (<i>k</i> = 2)				.06	[-.62 - .69]
	Disgust (<i>k</i> = 10)				.45	[.31 - .57]
	Sadness (<i>k</i> = 5)				.36	[.28 - .43]
	Anger (<i>k</i> = 10)				.18	[.07 - .28]
	Fear (<i>k</i> = 1)				.27	[-.20 - .64]
<i>Vividness</i>		5.24 (2, 9)	.031	.71		
	Film-imag-text (<i>k</i> = 21)				.38	[.30 - .45]
	Imagery-reminiscence (<i>k</i> = 4)				.22	[-.90 - .96]
	Situation (<i>k</i> = 10)				.18	[.08 - .29]
<i>Whole vs. Specific</i>		0.22 (1, 9)	0.647	.00		
	Whole (<i>k</i> = 5)				.24	[-.01 - .46]
	Specific (<i>k</i> = 25)				.31	[.22 - .40]
<i>Registering system</i>		25.77 (4, 7)	<.001	.24		
	Analytic (<i>k</i> = 2) ^a				.51	[.25 - .71]
	Global (<i>k</i> = 10)				.27	[.16 - .38]
	EMG (<i>k</i> = 21)				.33	[.22 - .42]
<i>Age group</i>		10.41 (1, 10)	.009	.46		
	Adults (<i>k</i> = 31)				.32	[.26 - .39]
	Children (<i>k</i> = 3)				.14	[-.06 - .34]
<i>Sociality</i>		2.47 (1, 10)	.147	.40		
	Social (<i>k</i> = 7)				.21	[.04 - .37]
	Non-social (<i>k</i> = 27)				.34	[.27 - .40]
<i>Self-awareness</i>		0.82 (1, 8)	.391	.08		
	Aware (<i>k</i> = 7)				.32	[.13 - .48]
	Not aware (<i>k</i> = 13)				.24	[.13 - .34]
Quantitative Moderator		<i>F</i> (df)	<i>p</i>	<i>R</i> ^{2*}	Slope	CI (95%)
<i>Publication Date</i>		0.18 (1, 10)	.682	.00	0.00	[-0.02 - 0.01]

^a Result obtained clustering outcomes from the same study

* The amount of explained variance was obtained with tradicional (not robust) tests of moderating variables

** Reported *r*. and confidence interval bound values for categories of categorical moderators were the result of returning Fisher's *Z* estimated values to Pearson correlation metric after analysis

Figure 1

Flowchart with the process of screening and selection of the articles included in the meta-analyses

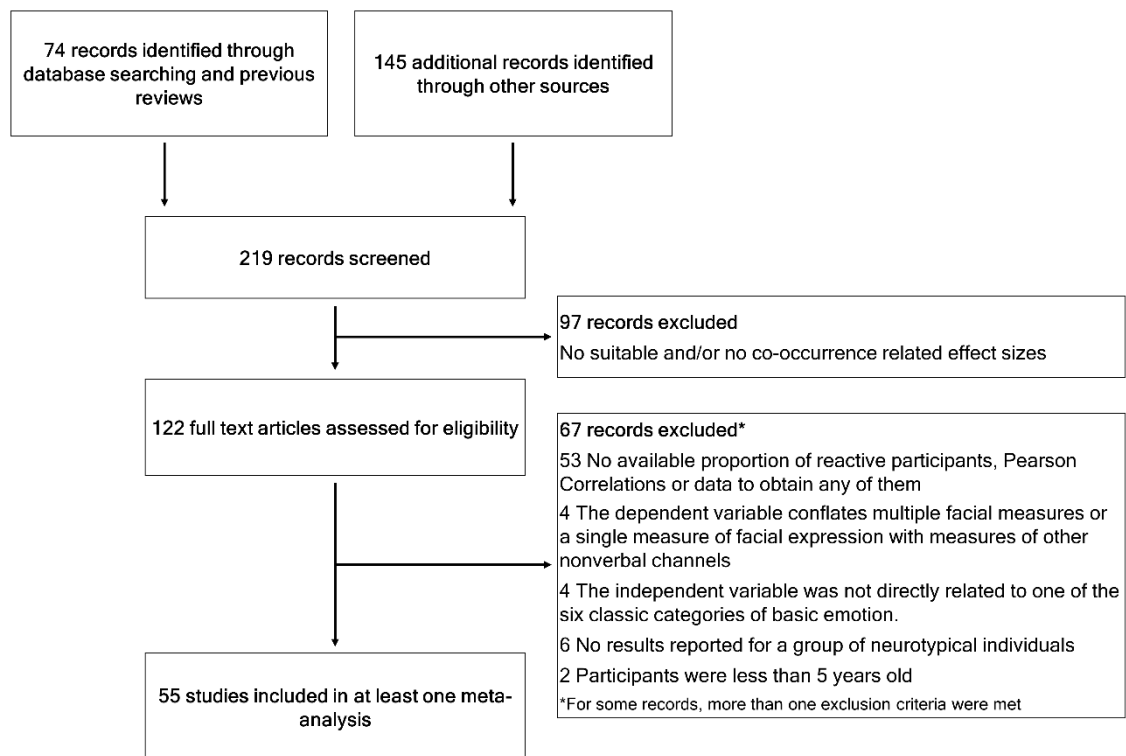


Figure 2

Forest plot for the sample of studies included in the Whole Expressions analysis

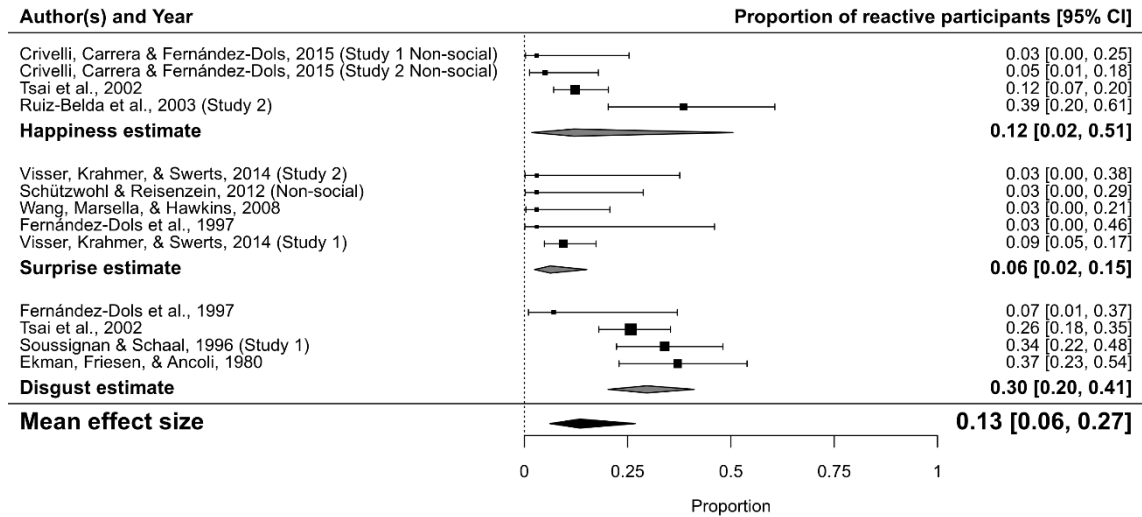


Figure 3

Funnel plots for Whole Expressions & Specific Facial Movements (upper left panel), Whole Facial Expression Only (upper right panel) and Intensity (lower panel) primary studies

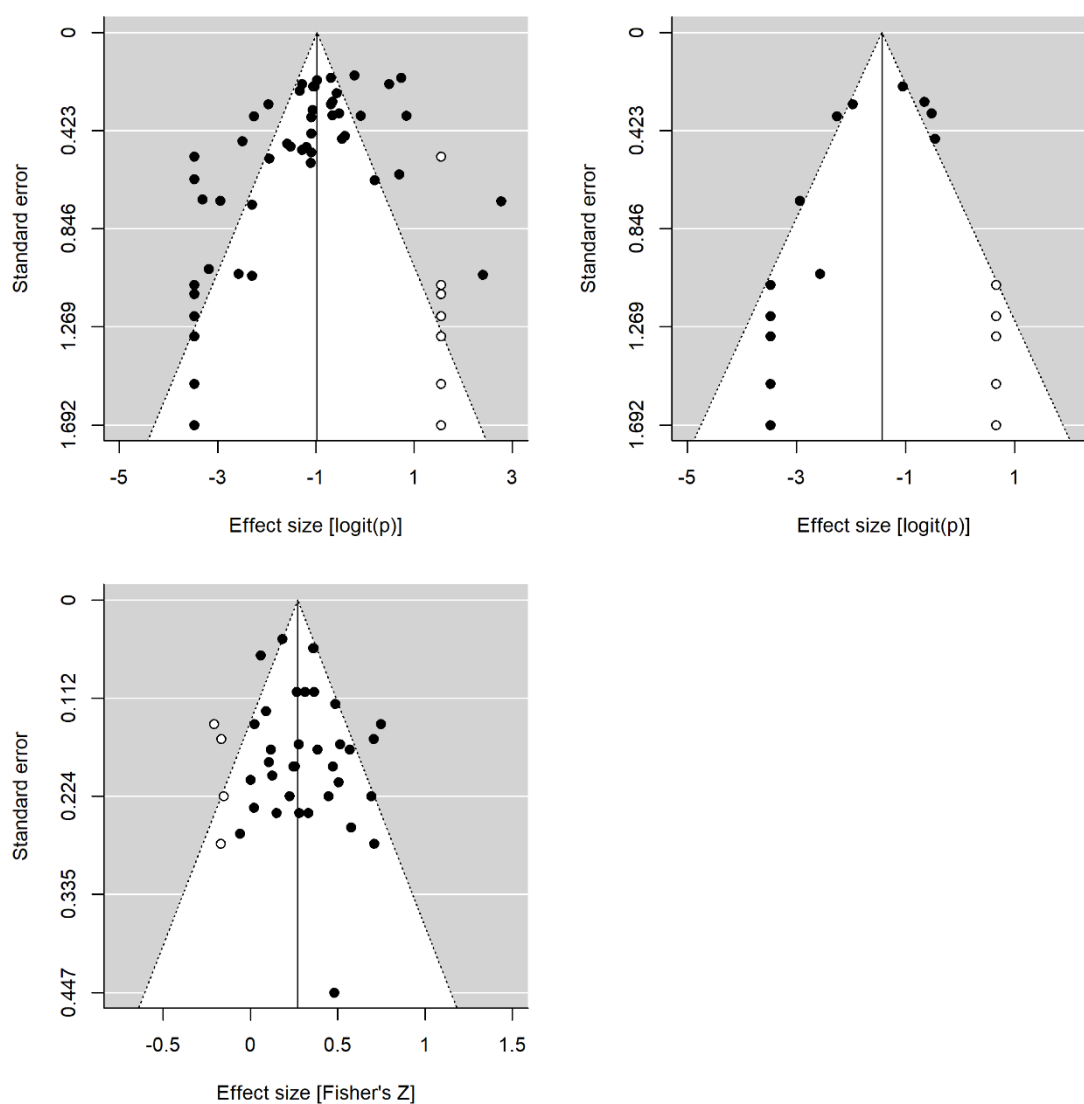


Figure 4

Forest plot for the sample of studies included in the Whole Expressions & Specific Facial Movements analysis

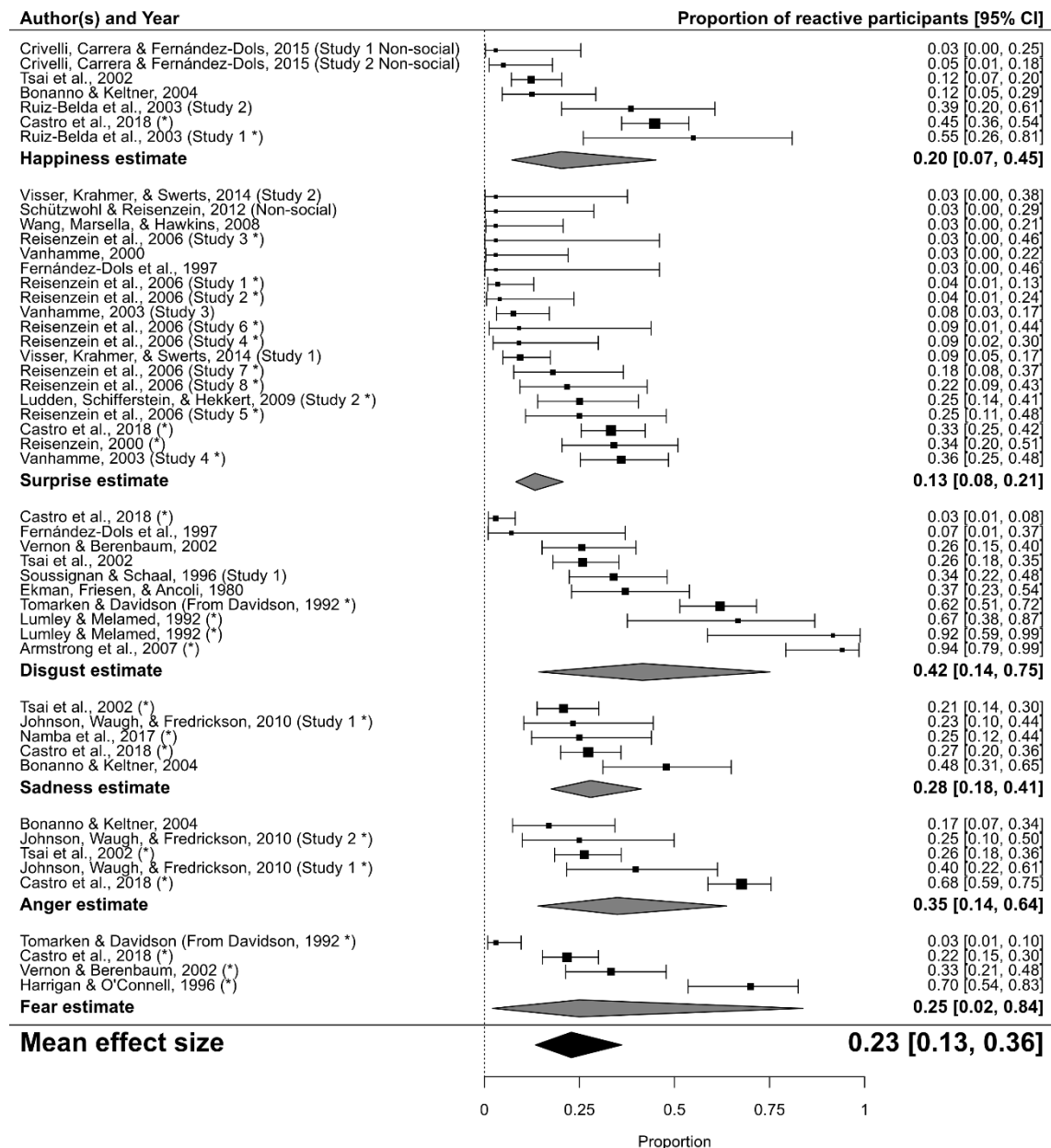


Figure 5

Forest plot for the sample of studies included in the Intensity analysis

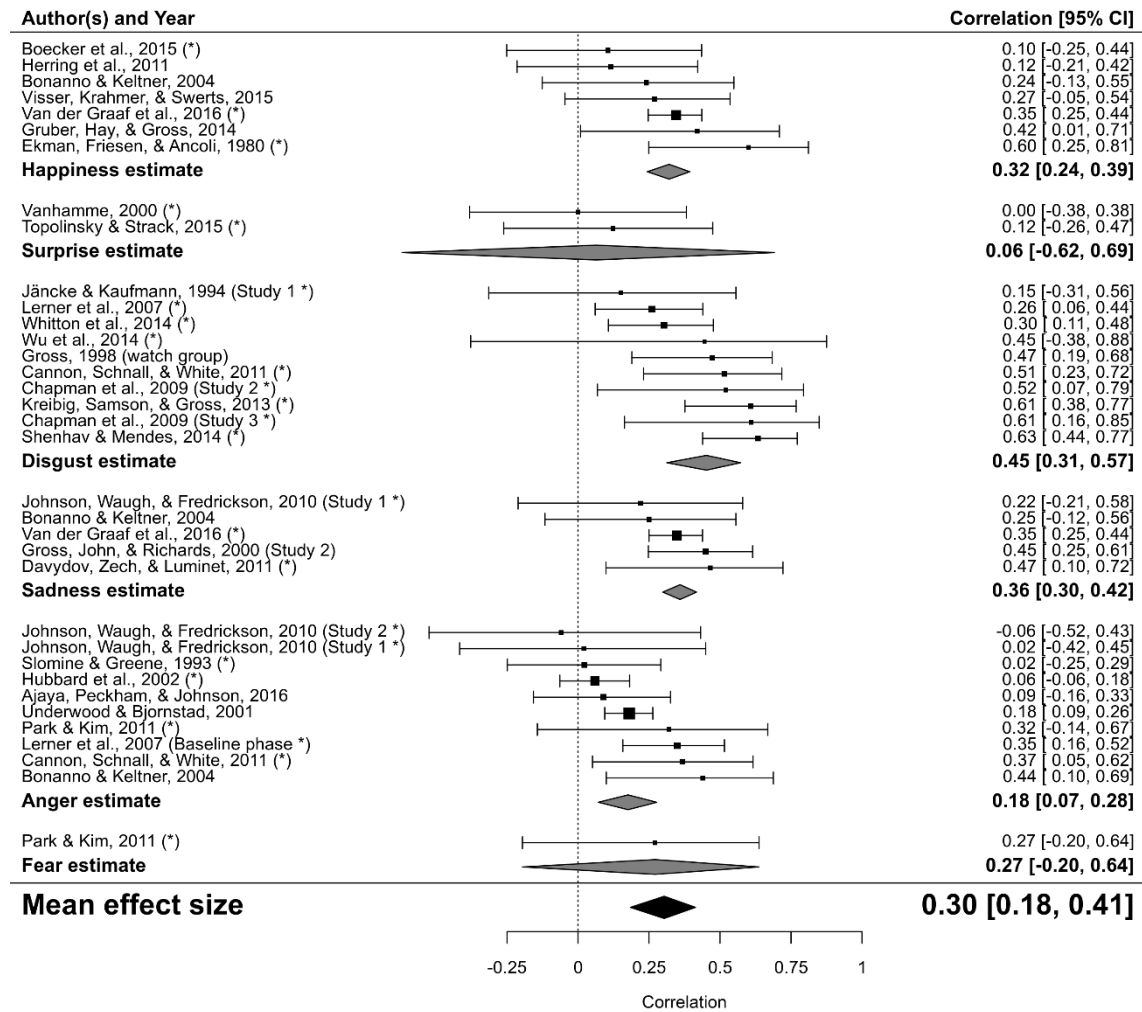


Figure 6a

Forest plot for the amusement studies reporting proportions

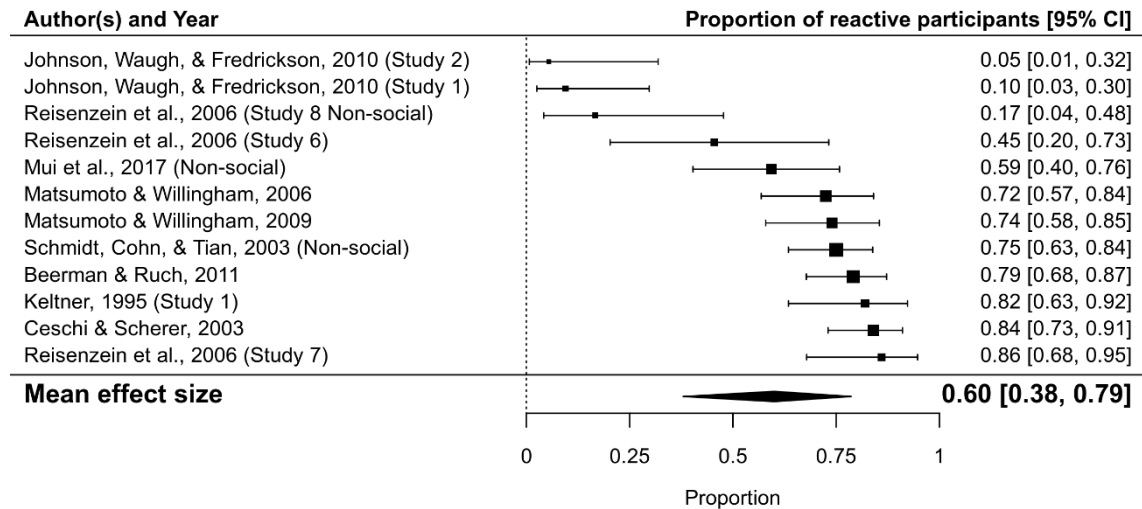


Figure 6b

Forest plot for the amusement studies reporting intensity correlations

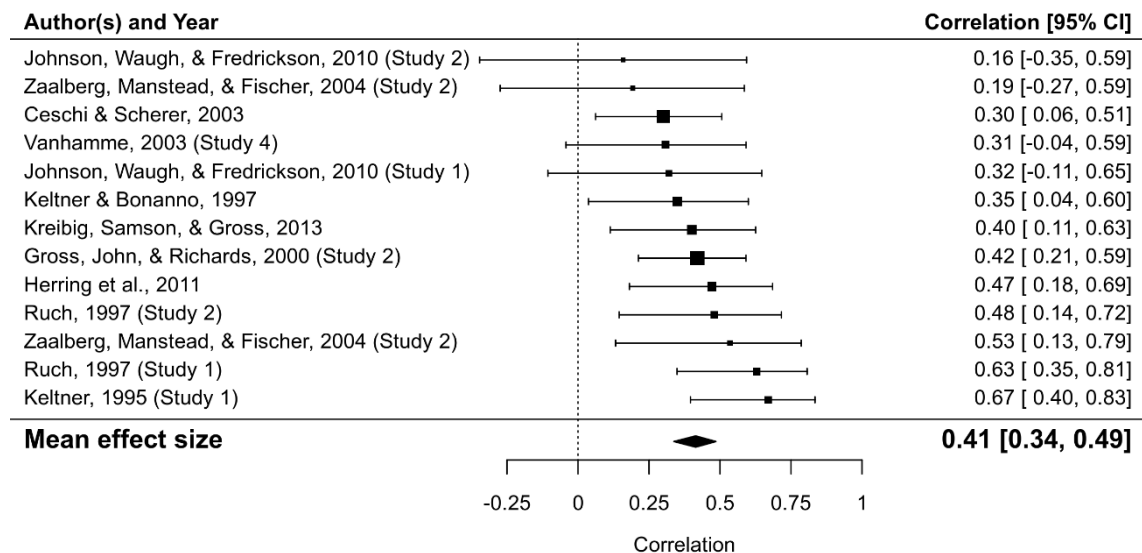
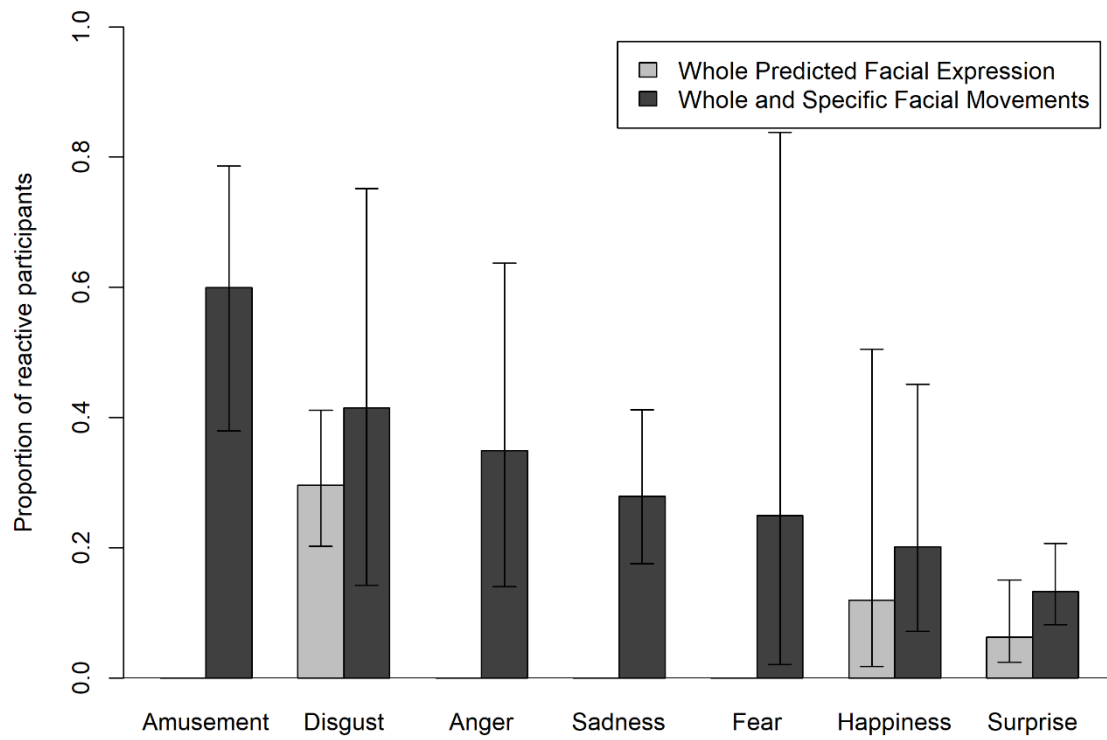


Figure 7

Barplot with the meta-analytic summaries by emotion obtained in the Whole and Whole & Specific Facial Movements analysis



Supplemental Materials

S1: Excluded Studies.

Reasons for exclusion:

1. Studies in which there is not enough data to compute co-occurrence either in terms of proportions of reactive participants or as an inter-individual Pearson correlation between the intensity of the reported emotion and the intensity of the expression.
2. Studies in which the dependent variable (proportion or reactive participants or intensity of the expression) conflates facial expression with other nonverbal channels (e.g., face plus voice) or it conflates a number of facial measures (e.g., frequency, intensity and duration) into a sole index.
3. Studies with an independent variable not directly related to one of the six classic categories of basic emotion (e.g., manipulating an affective or a motivational state).
4. Clinical studies in which co-occurrence is not reported for a group of neurotypical individuals.
5. Studies in which participants were babies or children less than 5 years old (elementary school).

¹Ahmed, W., van der Werf, G., & Minnaert, A. (2010). Emotional Experiences of Students in the Classroom. *European Psychologist*, 15(2), 142-151.

<https://doi.org/10.1027/1016-9040/a000014>

⁵Bennett, 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

- ⁴Borod, J. C., Rogers, K., Spielman, J., Halfacre, D., Flanagan, T., & Ramig, L. O. (2008). *Emotional experience and expression in Parkinson's disease*. Poster presented at the meeting of the Movement Disorders Society, Chicago, IL.
- ¹Brown, S.-L., & Schwartz, G. E. (1980). Relationships between facial electromyography and subjective experience during affective imagery. *Biological Psychology*, 11, 49-62. [https://doi.org/10.1016/0301-0511\(80\)90026-5](https://doi.org/10.1016/0301-0511(80)90026-5)
- ⁵Camras, L.A., Meng, Z., Ujie, T., Dharamsi, S., Miyake, K., Oster, H., Wang, L., Cruz, J., Murdoch, A., & Campos, J. (2002). Observing emotion in infants: Facial Expression, body behavior, and rater judgments of responses to an expanticy-violating event. *Emotion*, 2(2), 179-193. <https://doi.org/10.1037/1528-3542.2.2.179>
- ⁴Constantinou, E., Panayiotou, G., & Theodorou, M. (2014). Emotion processing deficits in alexithymia and response to a depth of processing intervention. *Biological Psychology*, 103, 212-222. <https://doi.org/10.1016/j.biopsycho.2014.09.011>
- ¹Davydov, D. M., Luminet, O., & Zech, E. (2013). An externally oriented style of thinking as a moderator of responses to affective films in women. *International Journal of Psychophysiology*, 87, 152-164. <https://doi.org/10.1016/j.ijpsycho.2012.12.003>
- ¹de Jong, P. J., Peters, M., & Vanderhallen, I. (2002). Disgust and disgust sensitivity in spider phobia: Facial emg in response to spider and oral disgust imagery. *Journal of Anxiety Disorders*, 16(5), 477-493. [https://doi.org/10.1016/s0887-6185\(02\)00167-6](https://doi.org/10.1016/s0887-6185(02)00167-6)

- ²Deckers, L., Kuhlhorst, L., & Freeland, L. (1987). The effects of spontaneous and voluntary facial reactions on surprise and humor. *Motivation and Emotion*, 11, 403-412. <https://doi.org/10.1007/bf00992852>
- ²Devereux, P. G., & Ginsburg, G. P. (2001). Sociality effects on the production of laughter. *The Journal of General Psychology*, 128(2), 227-240. <https://doi.org/10.1080/00221300109598910>
- ¹Dimberg, U., Hansson, G., & Thunberg, M. (1998). Fear of snakes and facial reactions: A case of rapid emotional responding. *Scandinavian Journal of Psychology*, 39, 75-80. <https://doi.org/10.1111/1467-9450.00059>
- ¹Dimberg, U., & Karlsson, B. (1997). Facial reactions to different emotionally relevant stimuli. *Scandinavian Journal of Psychology*, 38, 297-303. <https://doi.org/10.1111/1467-9450.00039>
- ¹Edelman, R. I. (1972). Vicarious fear induction and avowed autonomic stereotypy. *Behaviour Research and Therapy*, 10, 105-110. [https://doi.org/10.1016/s0005-7967\(72\)80002-0](https://doi.org/10.1016/s0005-7967(72)80002-0)
- ³Eisenberg, N., Schaller, M., Fabes, R. A., Bustamante, D., Mathy, R. M., Shell, R., & Rhodes, K. (1988). Differentiation of personal distress and sympathy in children and adults. *Developmental Psychology*, 24, 766-775. <https://doi.org/10.1037/0012-1649.24.6.766>
- ¹Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: emotional expression and brain physiology II. *Journal of Personality and Social Psychology*, 58(2), 342-353. <https://doi.org/10.1037/0022-3514.58.2.342>

- ¹Ekman, P., Friesen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology*, 39, 1125-1134.
<https://doi.org/10.1037/h0077722>
- ¹Falkenberg, I. Kohn, N., Schoepker, R. Habel, U. (2012). Mood induction in depressive patients: A comparative Multidimensional Approach. *PLoS ONE*, 7(1).
<https://doi.org/10.1371/journal.pone.0030016>
- ¹Fiacconi, C.M., & Owen, A.M. (2015). Using psychophysiological measures to examine the temporal profile of verbal humor elicitation. *PLoS ONE* 10(9).
<https://doi.org/10.1371/journal.pone.0135902>
- ¹Fridlund, 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>
- ¹Fridlund, A. J., Schwartz, G. E., & Fowler, S. C. (1984). Pattern recognition of self-reported emotional state from multiple-site facial EMG activity during affective imagery. *Psychophysiology*, 21(6), 622-637. <https://doi.org/10.1111/j.1469-8986.1984.tb00249.x>
- ⁴Gehricke, J-G, Shapiro, D. (2000). Reduced facial expression and social context in major depression: discrepancies between facial muscle activity and self-reported emotion. *Psychiatry Research*, 95, 157-167. [https://doi.org/10.1016/s0165-1781\(00\)00168-2](https://doi.org/10.1016/s0165-1781(00)00168-2)
- ²Gross, J. J., & Levenson, R. W. (1997). Hiding Feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, 106(1), 95-103.
<https://doi.org/10.1037/0021-843x.106.1.95>

- ^{1,4}Gruber, J., Dutra, S., Eidelman, P., Johnson, S. L., & Harvey, A. G. (2011). Emotional and physiological responses to normative and idiographic positive stimuli in bipolar disorder. *Journal of Affective Disorders*, 133, 437-442. <https://doi.org/10.1016/j.jad.2011.04.045>
- ³Gruber, J., Hay, A.C., & Gross, J.J. (2014). Rethinking emotion: Cognitive reappraisal is an effective positive and negative emotion regulation strategy in bipolar disorder. *Emotion*, 14(2), 388-396. <https://doi.org/10.1037/a0035249>
- ¹Hall, J. A., & Horgan, T. G. (2003). Happy affect and smiling: Is their relation moderated by interpersonal power? *Emotion*, 3, 303-309. <https://doi.org/10.1037/1528-3542.3.3.303>
- ¹Harris. C. R. (2001). Cardiovascular responses of embarrassment and effects of emotional suppression in a social setting. *Journal of Personality and Social Psychology*, 81(5), 886-897. <https://doi.org/10.1037/0022-3514.81.5.886>
- ¹Harris, C. R., & Alvarado, N. (2005). Facial expressions, smile types, and self-report during humour, tickle, and pain. *Cognition and Emotion*, 19, 655-669. <https://doi.org/10.1080/02699930441000472>
- ¹Hess, 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>
- ³Hoefling, A., Likowski, K. U., Deutsch, R., Häfner, M., Seibt, B., Mühlberger, A., Weyers, P., & Strack, F. (2009). When hunger finds no fault with moldy corn: food deprivation reduces food-related disgust. *Emotion*, 9(1), 50. <https://doi.org/10.1037/a0014449>

- ¹Horton, R. E., & Riddell, R. R. P. (2010). Mothers' facial expressions of pain and fear and infants' pain response during immunization. *Infant Mental Health Journal*, 31(4), 397-411. <https://doi.org/10.1002/imhj.20262>
- ⁴Izard, C. E., & Abe, J. A. A. (2004). Developmental changes in facial expressions of emotions in the strange situation during the second year of life. *Emotion*, 4(3), 251-265. <https://doi.org/10.1037/1528-3542.4.3.251>
- ¹Jänke, L. (1996). Facial EMG in an anger provoking situation: individual differences in directing anger outwards and inwards. *International Journal of Psychophysiology*, 23, 207-214. [https://doi.org/10.1016/s0167-8760\(96\)00062-1](https://doi.org/10.1016/s0167-8760(96)00062-1)
- ¹Jakobs, E., Manstead, A. S., & Fischer, A. H. (2001). Social context effects on facial activity in a negative emotional setting. *Emotion*, 1, 51-69. <https://doi.org/10.1037/1528-3542.1.1.51>
- ¹Juslin, P. N., Harmat, L. (2014) What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, 42(4), 599-623. <https://doi.org/10.1177/0305735613484548>
- ¹Kätsyri, J., Kinnunen, T., Kusumoto, K., Oittinen, P., & Ravaja, N. (2016). Negativity bias in media multitasking: the effects of negative social media messages on attention to television news broadcasts. *PLOS ONE*, 11(5). <https://doi.org/10.1371/journal.pone.0153712>
- ¹Kendler, K. S., Halberstadt, L. J., Butera, F., Myers, J., Bouchard, T., & Ekman, P. (2008). The similarity of facial expressions in response to emotion-inducing films in reared-apart twins. *Psychological Medicine*, 38(10), 1475-1483. <https://doi.org/10.1017/s0033291707001535>

- ¹Krippel, M. (2002). *Induction of disgust, anger, happiness, relaxation, hunger and curiosity by odors* [Unpublished manuscript]. Department Psychology III, University of Erlangen-Nürnberg.
- ^{1,3}Leersnyder J. D., & Mesquita, B. (2017). *Emotional frame switching in biculturals: How salient cultural concerns shape emotion* [Manuscript submitted for publication]. Center for Social and Cultural Psychology, University of Leuven.
- ¹Leersnyder, J. D., & Mesquita, B. (2014, April 24-26) *Concerned about autonomy, relatedness or both? How cultural concerns shape emotions during social interactions* [Poster presentation]. Conference of the Society for Affective Sciences. Bethesda, DC, US.
- ¹Lerner, J.S., Dahl, R., Hariri, A.R., & Taylor, S.E. (2007). Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses. *Biological Psychiatry*, 61, 253-260. <https://doi.org/10.1016/j.biopsych.2006.08.016>
- ¹Leutgeb, V., Übel, S., & Schienle, A. (2013). Can you read my pokerface? A study on sex differences in dentophobia. *European Journal of Oral Sciences*, 121, 465-470. <https://doi.org/10.1111/eos.12079>
- ¹Levenson, R. W., & Carstensen, L. L. (1991). Emotion, Physiology, and Expression in Old Age. *Psychology and Aging*, 6(1), 28-35. <https://doi.org/10.1037/0882-7974.6.1.28>
- ¹Lohani, M., & Isaacowitz, D. M. (2014). Age differences in managing response to sadness elicitors using attentional deployment, positive reappraisal and suppression. *Cognition and Emotion*, 28(4), 678-697. <https://doi.org/10.1080/02699931.2013.853648>

- ¹Matsumoto, D., & Kupperbusch, C. (2001). Idiocentric and allocentric differences in emotional expression, experience, and the coherence between expression and experience. *Asian Journal of Social Psychology*, 4, 113-131.
<https://doi.org/10.1111/j.1467-839x.2001.00080.x>
- ¹Mauss, I.B., Levenson, R.W., McCarter, L., Wilhelm, F.H., & Gross, J.J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, 5, 175-190. <https://doi.org/10.1037/1528-3542.5.2.175>
- ¹Mauss, I. B., Butler, E. A., Roberts, N. A., & Chu, A. (2010). Emotion control values and responding to an anger provocation in Asian-American and European-American individuals. *Cognition and Emotion*, 24(6), 1026-1043.
<https://doi.org/10.1080/02699930903122273>
- ¹Mauss, I. B., Shallcross, A. J., Troy, A. S., John, O. P., Ferrer, E., Wilhelm, F. H., & Gross, J. J. (2011). Don't hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning. *Journal of Personality and Social Psychology*, 100(4), 738-748. <https://doi.org/10.1037/a0022410>
- ²Mehu, 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>
- ¹Reisenzein, 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, 295-315. (Study 7).
<https://doi.org/10.1037/0022-3514.91.2.295>
- ¹Reissland, N., & Shepherd, J. (2006). The Effect of Maternal Depressed Mood on Infant Emotional Reaction in a Surprise-Eliciting Situation. *Infant Mental Health*

Journal: Official Publication of The World Association for Infant Mental Health,
27(2), 173-187. <https://doi.org/10.1002/imhj.20087>

- ¹Ritz, T., Claussen, C., & Dahme, B. (2001). Experimentally induced emotions, facial muscle activity, and respiratory resistance in asthmatic and non-asthmatic individuals. *British Journal of Medical Psychology*, 74, 167-182.
<https://doi.org/10.1348/000711201160894>
- ¹Rosenberg, E. L. & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition and Emotion*, 8, 201-229.
<https://doi.org/10.1080/02699939408408938>
- ¹Ruch, W. (1994). Extraversion, alcohol, and enjoyment. *Personality and Individual Differences*, 16(1), 89-102. [https://doi.org/10.1016/0191-8869\(94\)90113-9](https://doi.org/10.1016/0191-8869(94)90113-9)
- ¹Ruch, W. (1995). Will the real relationship between facial expression and affective experience please stand up: The case of exhilaration. *Cognition and Emotion*, 9, 33-58. <https://doi.org/10.1080/02699939508408964>
- ⁴Scherer, K. R., Zentner, M. R., & Stern, D. (2004). Beyond Surprise: The puzzle of infants' expressive reactions to expectancy violation. *Emotion*, 4(4), 389-402.
<https://doi.org/10.1037/1528-3542.4.4.389>
- ¹Schienze, A., Schäfer, A., Walter, B., Stark, R., & Vaitl, D. (2005). Elevated disgust sensitivity in blood phobia. *Cognition and Emotion*, 19(8), 1229-1241.
<https://doi.org/10.1080/02699930500286364>
- ¹Seider, B. H., Shiota, M. N., Whalen, P., & Levenson, R. W. (2011). Greater sadness reactivity in late life. *Social Cognitive and Affective Neuroscience*, 6, 186-194.
<https://doi.org/10.1093/scan/nsq069>

- ¹Sloan, D. M. (2004). Emotion regulation in action: Emotional reactivity in experiential avoidance. *Behaviour Research and Therapy*, 42(11), 1257-1270.
<https://doi.org/10.1016/j.brat.2003.08.006>
- ¹Stark, R., Walter, B., Schienle, A., & Vaitl, D. (2005). Psychophysiological correlates of disgust and disgust sensitivity. *Journal of Psychophysiology*, 19(1), 50-60.
<https://doi.org/10.1027/0269-8803.19.1.50>
- ¹Tsai, J. L., Pole, N., Levenson, R. W., & Muñoz, R. F. (2003). The effects of depression on the emotional responses of Spanish-speaking Latinas. *Cultural Diversity and Ethnic Minority Psychology*, 9(1), 49-63. <https://doi.org/10.1037/1099-9809.9.1.49>
- ¹Van der Graaff, J., Meeus, W., de Wied, M., van Boxtel, A., van Lier, P. A., Koot, H. M., & Branje, S. (2016). Motor, affective and cognitive empathy in adolescence: Interrelations between facial electromyography and self-reported trait and state measures. *Cognition and Emotion*, 30(4), 745-761.
<https://doi.org/10.1080/02699931.2015.1027665>
- ¹Vazire, S., Naumann, L. P., Rentfrow, P. J., & Gosling, S. D. (2009). Smiling reflects different emotions in men and women. *Behavioral and Brain Sciences*, 32, 403-405. <https://doi.org/10.1017/s0140525x09991026>
- ¹Vrana, S. R. (1993). The psychophysiology of disgust: Differentiating negative emotional contexts with facial EMG. *Psychophysiology*, 30, 279-286.
<https://doi.org/10.1111/j.1469-8986.1993.tb03354.x>
- ¹Vrana, S. R., & Rollock, D. (2002). The role of ethnicity, gender, emotional content, and contextual differences in physiological, expressive, and self-reported

emotional responses to imagery. *Cognition and Emotion*, 16(1), 165-192.

<https://doi.org/10.1080/02699930143000185>

- ¹Whitton, A. E., Henry, J. D., & Grisham, J. R. (2014). Cognitive and psychophysiological correlates of disgust in obsessive-compulsive disorder. *British Journal of Clinical Psychology*, 54, 16-33.

<https://doi.org/10.1111/bjc.12058>

- ¹Whitton, A. E., Henry, J. D., Rendell, P. G., & Grisham, J. R. (2014). Disgust, but not anger provocation, enhances levato labii superioris activity during exposure to moral transgressions. *Biological Psychology*, 96, 48-56.

<https://doi.org/10.1016/j.biopsycho.2013.11.012>

- ¹Wolf, K., Mass, R., Kiefer, F., Eckert, K., Weinhold, N., Wiedemann, K., & Naber, D. (2004). The influence of olanzapine on facial expression of emotions in schizophrenia-an improved facial EMG study. *German Journal of Psychiatry*, 7, 14-19. <https://doi.org/10.1055/s-2003-825564>