Emotional processing and the autonomic nervous system: a comprehensive meta-analytic investigation

Pedro Silva Moreira[1,2,3], Pedro Chaves[3,4], Nuno Dias[3,5], Patrício Costa[1,2], Pedro R Almeida[3,6]

[1]Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal;

[2]ICVS/3B's, PT Government Associate Laboratory, Braga/Guimarães, Portugal;

[3]MindProber Labs, Porto, Portugal;

[4]Department of Experimental Biology, Faculty of Medicine, University of Porto, Portugal

[5]2AI, Instituto Politécnico of Cávado and Ave, Barcelos, Portugal

[6]School of Criminology, Faculty of Law, University of Porto, Portugal

Correspondence:

Pedro Silva Moreira

E-mail: pedromsmoreira@gmail.com

Word count:

Abstract: 213 words

Main text: 5010 words

Abstract

Background: The search for autonomic correlates of emotional processing has been a matter of interest for the scientific community with the goal of identifying the physiological basis of emotion. Despite an extensive state-of-the-art exploring the correlates of emotion, there is no absolute consensus regarding how the body processes an affective state.

Objectives: In this work, we aimed to aggregate the literature of psychophysiological studies in the context of emotional induction.

Methods: For this purpose, we conducted a systematic review of the literature and a meta-analytic investigation, comparing different measures from the electrodermal, cardiovascular, respiratory and facial systems across emotional categories/dimensions. Two-hundred and ninety-one studies met the inclusion criteria and were quantitatively pooled in random-effects meta-analytic modelling.

Results: Heart rate and skin conductance level were the most reported psychophysiological measures. Overall, there was a negligible differentiation between emotional categories with respect to the pooled estimates. Of note, considerable amount of between-studies' heterogeneity was found in the meta-analytic aggregation. Self-reported ratings of emotional arousal were found to be associated with specific autonomic-nervous system (ANS) indices, particularly with the variation of the skin conductance level.

Conclusions: Despite this clear association, there is still a considerable amount of unexplained variability that raises the need for more fine-grained analysis to be implemented in future research in this field.

Keywords: emotion; autonomic nervous system; sympathetic nervous system; parasympathetic nervous system; psychophysiology; emotional arousal; emotional valence; films

1. Background

Current conceptualizations of emotion have been consistently influenced by William James' description of "what is an emotion?" in the 19th century (James, 1884). However, this has been a matter of relevance since the historical periods of ancient Greece and Rome, where emotion was widely perceived as a threat to reason and to philosophical thinking (Solomon, 1993). On the reason-emotion dualism (where these two entities are viewed as antagonist natural kinds), emotion typically receives a primitive, less intelligent and dangerous role, which needs to be controlled by the wisdom of reason (Solomon, 1993).

Despite the longstanding interest on the study of emotion and the accumulating state-of-the-art tackling this topic, a lot is still left to unravel. A matter of active discussion pertains to the exact nature of the physiological correlates of emotional processing, as well as the exact role of bodily changes in emotion. For instance, even though it is widely established that there are rich reciprocal connections between states in the central and peripheral nervous systems and what can be broadly described as emotional events, the direction of this association (*i.e.,* the cause-effect) is a matter of discordance (Larsen, Berntson, Poehlmann, Ito, & Cacioppo, 2008). One hypothesis – the Cannon-Bard theory (Cannon, 1927) – states that the efferent connections from the brain to periphery causes the peripheral variations in response to the subjective processing of emotions (*i.e.,* feelings). An opposite argument suggests that bodily changes follow directly the sensorial perception, which will not only precede, but will also generate the emotional experience (James, 1884). The third argument – the two-factor theory of emotion (Schachter & Singer, 1962) – shares a similar periphery-to-brain perspective, but suggesting that there is an undifferentiated peripheric response to different emotional states. Instead, the emotional experience will be cognitively interpreted, by conscientiously linking the experienced arousal with the situational context.

Psychophysiological correlates of emotional processing

The relevance of studying the brain/body correlates of emotional processing originally arises from the James-Lange theory of emotions, which highlights the close link between emotions and behavior. "Without the bodily states following on the perception, the latter would be purely cognitive in form, pale, colorless, destitute of emotional warmth. We might then see the bear, and judge it best to run, receive the insult and deem it right to strike, but we could not actually feel afraid or angry." (James, 1884). An extensive body of literature has examined the biological underpinnings underlying emotional processing, through the lens of central (central nervous system level; CNS) or peripheral measurement (autonomic nervous system level; ANS). Altogether, these findings provide evidence for a reciprocal relationship between the bodily expression of emotion and how emotional information is attended. This underlines the central tenets in the theories of embodied cognition – the perspective that the perception and conceptualization of emotion involves perceptual, somato-visceral and motoric reexperiencing of the relevant emotion in one's self (Niedenthal, 2007).

At the central level, most studies have mainly approached the brain correlates of emotional processing, based on electrophysiological (namely, electroencephalography; EEG) or hemodynamic (namely, functional magnetic resonance imaging; fMRI) measurements. The results from the meta-analytic aggregation of CNS measures are not consensual regarding the nature of the emotional processing on the brain: one branch of research suggests that different emotional categories have distinct signatures on brain correlates – supporting the view of the classical basic theories of emotion (Lench, Flores, & Bench, 2011); contrasting evidence points to non-specific patterns of brain response to the processing of discrete emotional categories (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012).

Many studies have characterized ANS responses to emotional processing by measuring the two main branches of this system: the sympathetic (which primes the body for action by, for instance, increasing the heart-rate) and the parasympathetic divisions (which aids in restorative functions; *e.g.,* simulation digestion). In general terms, ANS measures can be grouped according to different systems: electrodermal, cardiovascular, respiratory and facial. A summary of these systems and a description of their constituting measures is presented on Table 1. Cacioppo and colleagues (2000) performed the first aggregation of ANS correlates of emotional processing. The authors observed that there was some variation of between-studies' effect-sizes for different emotional categories, however with no evidence for discrete emotions' specificity. More recently, Siegel and colleagues (2018) reported a meta-analytic investigation of 202 studies, and demonstrated that the pattern of effect sizes from multiple systems did not allow the disentanglement between emotional categories. The authors suggested that these results provided evidence for the constructionist view of emotion (or the population hypothesis) (Barrett, 2017b) – which perceives emotional categories as conceptual categories (*i.e.,* the brain uses emotion concepts to categorize sensations to construct an instance of emotion) (Barrett, 2017a), in detriment of the classical view of emotion (referred as the fingerprint hypothesis) – by which, each emotional category is independent of the others in its behavioral, psychological, and physiological manifestations (Posner, Russell, & Peterson, 2005).

The current work

The goal this work was to aggregate the results of experimental studies investigating the ANS (cardiovascular, electrodermal and respiratory) correlates of emotion elicitation. For this purpose, due to their increased complexity and dynamic nature, audiovisual stimuli (*i.e.,* videos) are thought to provide a richer and ecologically valid methodology to induce affective sates (Baumgartner, Esslen, & Jäncke, 2006). In addition, video clips typically induce a more sustained affective state compared to presentation of static stimuli that elicit only short-lived affective responses (Bos, Jentgens, Beckers, & Kindt, 2013; Gross & Levenson, 1995). As such, we restricted our meta-analytic investigation to experiments using videos to elicit any affective response. It was also our goal to assess these correlates, using different perspectives of emotional processing – through the lens of a classic view – looking at the ANS correlates of main emotional categories (including: sadness, disgust, fear, anger and happiness), but also from a dimensional perspective – namely the valence–arousal model (Russell, 1980) – which postulates a bipolar valence dimension ranging from positive to negative, and an orthogonal arousal dimension ranging from low arousal to high arousal. For the latter aim, we associated the pooled effect-size for each emotional contrast with the variance of self-reported levels of arousal, for emotional content with positive and negative valence.

2. Methods

2.1. Data sources

The systematic review was implemented following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, Altman, & Group, 2009). The literature search was performed in multiple online databases, on August 2017, including PubMed, PsycInfo and Google Scholar, to identify relevant studies in the context of autonomic nervous system correlates of emotional processing. The following keywords and logical aggregations were used: (autonomic OR peripher*) AND (emotion OR arousal OR valence) AND (films OR movies OR clips OR videos). In addition, studies citing validated sets of emotional films were also included (Carvalho, Leite, Galdo-Álvarez, & Gonçalves, 2012; Gilman et al., 2017; Gross & Levenson, 1995; Hewig et al., 2005; Jenkins & Andrewes, 2012; Kaviani, Gray, Checkley, Kumari, & Wilson, 1999; Maffei et al., 2014; Philippot, 1993; Ray, 2007; Samson, Kreibig, Soderstrom, Wade, & Gross, 2016; Schaefer, Nils, Sanchez, & Philippot, 2010). Last, we also screened the reference list from relevant reviews on the field. Studies obtained from more than one database were identified as duplicates. References from relevant manuscripts were also included. The selection of individual studies was conducted in two consecutive phases: screening and full-text assessment. The inclusion criteria for each phase is described below.

2.2. Inclusion criteria

For the screening phase, the following criteria was established to determine study selection: (1) the study was published as an original research article in a peer-reviewed journal – *i.e.,* reviews, commentaries, protocols, publications in book chapters or conferences were not considered; (2) the study was published in English language; (3) the study involved human subjects; (4) the study implemented the visualization of films with emotional content; (5) one or more measures of autonomous nervous system correlates were obtained. The studies meeting these criteria were comprehensively assessed during the full-text assessment phase, in which the following inclusion criteria were defined: (1) the study presents results for healthy individuals; (2) films were classified according to one specific emotional category or one specific emotional valence – *i.e.,* studies implementing mixed emotional content were not included; (3) the study summarizes results for individual peripheral measures (*i.e.,* not composites of two or more measures) and in response to one single film or one single category (*e.g.,* average of peripheral response to films from the same emotional category); (4) the study presents the results of peripheral measures, contrasting emotional films to either an emotional film with neutral content or to a baseline period.

2.3. Data extraction

A structured database was constructed to aggregate the characteristics retrieved from individual studies, including sample characteristics (sample size, participants' mean age, proportion of male/female participants), description of ANS measures, stimuli-related variables (length of film clips, emotional category/valence of each stimulus). In addition, we also extracted valence and arousal ratings, assessed with the self-assessment manikin (SAM) (Bradley & Lang, 1994) or, alternatively, by assessing the intensity of the target emotions, as measured through the Positive and Negative Affect Schedule (PANAS) (Watson, Clark, & Tellegen, 1988) or similar structured or non-structured questionnaire forms. Self-reported scores were normalized, considering the lower and upper limits of the measurement scale to allow for between-studies' comparison.

When the relevant statistics were not available from the main text, tables or supplementary information, but was represented in plots, we used the GetData graph digitizer tool (Fedorov, 2008) to extract mean and dispersion measures, based on the manual definition of axes scaling. Similar strategies have been previously described and revealed the validity of this approach to estimate real values (*e.g.,* Kalluri, Zhang, Caritis, & Venkataramanan, 2017).

2.4. Data analysis

Separate meta-analyses were conducted for each emotional category/dimension and peripheral measure. For each individual study, average scores and dispersion measures were used to compute effect sizes. To estimate within-subjects' standardized effect sizes, the mean difference between an emotional condition and neutral/baseline scores was calculated. If not directly reported, the standard deviation (SD) for difference was estimated, considering the individual SDs for each measure, *i.e.,* based on the individual dispersion values for the emotional category and the baseline/neutral condition, according to:

$$SD_{diff} = \sqrt{SD_{cond1}^{2} + SD_{cond2}^{2}},$$

where SD_{diff} corresponds to the standard deviation of the difference. Cohen's d and confidence intervals from individual studies were aggregated using random-effects models (restricted maximum-likelihood), which constitutes a more conservative approach, to account for significant between-studies' heterogeneity. For each emotional contrast, the effect size Cohen's d was computed as:

$$d=\frac{\acute{x}_{emotion}-\acute{x}_{baseline/neutral}}{SD_{diff}},$$

Effect sizes were interpreted as small (≥0.20), medium (≥0.50) and large (≥0.80) (Cohen, 1988). The variance of Cohen's d (var_d) was calculated to establish the 95% confidence interval (CI), according to the formula:

$$var_d=\frac{1}{n}+\frac{d^2}{2n}$$

Between-studies heterogeneity was estimated based on the significance of the Cochran Q test (X^2 statistic) and I^2 statistic. I^2 was calculated as

$$I^2=\frac{Q-degrees\,of\ freedom}{Q}\times 100$$

where Q is the Cochran's statistic. Leave-one-out sensitivity analyses were conducted to assess the impact of individual studies on the overall estimated. Several strategies were implemented for assessing publication bias, including the Begg's rank correlation statistic for funnel plot asymmetry. In addition, contour-enhanced funnel plots were used which enables the consideration of the statistical significance of study estimates. Cluster-robust meta-analytic procedures were implemented to account for the statistical dependence of multiple effect sizes obtained from the same study, as these are likely to produce clusters of internally correlated effect size estimates (Pustejovsky & Tipton, 2014). Meta-regression analyses were conducted to assess the impact of sample characteristics (mean age, proportion of male/female individuals), stimuli-related [duration of the stimuli and the nature of comparison (*i.e.,* comparison of the emotional category against neutral stimuli or against a baseline period)] type self-reported) and stimuli-related variables (valence and arousal ratings) on the individual meta-analytic estimates. With the goal of assessing the adequacy of a dimensional perspective, correlation analyses were performed between the standardized mean differences (i.e., Cohen's *d* values) and scaled measures of self-reported arousal and valence, independently of the emotional category.

Statistical analysis was performed in RStudio (v3.31.1, RStudio, Boston, MA, USA). The meta-analytic pipeline was implemented using the metafor (Viechtbauer, 2010) and clubSandwich (Pustejovsky & Tipton, 2014) packages. The dataset used for the meta-analytic investigation, and the code for the analysis is available at the Open Science Framework (osf.io/).

3. Results

Fig. 1 represents the process of article selection, as a PRISMA flowchart. The literature search yielded 3445 articles. After the combination of the datasets, 636 duplicated articles were removed, resulting in a total of 2809 articles being included in the screening phase. During the screening phase, 2299 articles were excluded for not meeting the pre-defined inclusion criteria. Of the 591 articles assessed for eligibility, 288 articles were further excluded, mainly for not presenting psychophysiological data, not having emotional contrasts, or because data were not reported for individual psychophysiological measures.

3.1. Characterization of the included studies

The characteristics of the included studies are summarized on Table 2. Most studies reported two ANS measures. Among the studies included in this review (considering all the different emotional categories/dimensions), mean HR (186 studies) was the most described measure, followed by mean SCL (135 studies), HF-HRV (98 studies), RR (53 studies), EMG (47 studies), SCR (40 studies), and FT (26 studies). The most reported emotional category was sadness (123 studies), followed by fear (67 studies), happiness (58 studies), disgust (51 studies) and anger (36 studies). Most studies were conducted in Europe (48% of the studies), and North America (42% of the studies). Sample size varied from 4 to 408 participants. Most studies recruited university samples (approximately, 54%), followed by community samples (approximately, 28%) and by mixed community/university samples (approximately, 11%), while only about 6% of the studies assessed pediatric samples. For the studies comprised by clinical and/or psychiatric patients, only the data for healthy individuals were considered. Most studies assessed subjective measures of emotional induction, including Self-Assessment Manikin (SAM) or similar Visual-Analog Scales (VAS), Positive and Negative Affect Schedule (PANAS) or the intensity of target emotions.

3.2. Individual meta-analyses of ANS correlates

The global estimates obtained for the different meta-analyses (which are described below) are summarized on Table 3.

3.2.1. Electrodermal system

Mean SCL was significantly (with a small magnitude) increased in response to sad content, in comparison to neutral/baseline values (d=.27, p=.004, k=31) – nevertheless, the significance of these effects was lost when a cluster-robust model was conducted (d=.23, p=.077). Disgust and fear were characterized by moderate/large, significant, increases of mean SCL [d=.80 (p<.001; k=16), d=.78, (p<.001; k=14)]. For happiness, there were significant, moderate [d=.70 (p=.027; k=10) overall effects across studies. Combining all the categories with negative valence together, there were moderate increases of mean SCL (d=.46, p<.001, k=87). The SCL correlates for the different emotional categories are summarized as forest plots on Fig. 2. The results of the meta-regression analysis indicated that the use of baseline vs neutral stimuli as the control category did not produce significant effects on the overall estimates (b=-0.17, SE=.194, p=.378) self-reported arousal had a significant impact on the pooled estimates for sadness (b=3.33, SE=.98, p<.001). In addition, there were no effects of age on the pooled estimates (b=-0.003, SE=.007, p=.636). Lastly, there was no evidence for a statistical impact of the proportion of male individuals on the meta-analytic effects for sadness (b=-.168, SE=.696, p=.711). For the remaining pooled estimates, it was only possible to retrieve the information for self-reported arousal for seven or less studies, which precluded us to compute reliable meta-regression estimates for these emotional categories.

The number of SCRs was significantly increased for sadness (d=.42, p=.040, k=7) and happiness (d_{Ha}=1.00, p_{Ha}=.044) with a moderate magnitude. The remaining emotional categories were associated with a trend (although, non-significant) for increases in the number of SCRs (d_{Di}=3.21, p_{Di}=.330; d_{Fe}=3.11, p_{Fe}=.180; d_{An}=.24, p_{An}: .340). Nevertheless, it is important to highlight that these results should be interpreted with caution, since the number of studies reporting results for this metric ranged from four (Anger and Happiness) to seven (Sadness) studies. The combined effect of categories with negative valence did not produce significant overall effects (d=1.14, p=.157). The reduced number of studies also precluded the use of meta-regression analyses for this ANS measure (Fig. 3).

3.2.2. Cardiovascular system

Regarding the mean HR, no significant effects were noted for sadness (d=-.05, k=32). Disgust was associated with a small decrease, although non-significant, of the mean HR (d=-0.39, k=26). Anger was characterized by marginally significant increases of the mean HR (d=0.46, p=.100, k=16). Comparing with neutral/baseline stimuli, happiness was associated with small reductions of the mean HR (d=-.25, p=.090, k=28) (Fig. 4). None of the tested moderators significantly affected the pooled estimates for this metric. The combined effects of categories with negative emotional valence yielded significant, although with small magnitude, overall estimates (d=-.19, p=.038).

For high-frequency HRV, there were no significant overall effects for sadness (*d*=1.48, *p*=.440*; k*=9*),* disgust (*d*=-.08, *p*=.660*; k*=6*)*, fear (*d*=-.36, *p*=.207*; k*=8*)*, anger (*d*=-59, *p*=.200*; k*=2) or happiness (*d*=-.16, *p*=.420*; k*=6) (Figure 4). The combined set of negative emotional categories did not produce significant overall effects (*d*=-.22, *p*=.120). The low number of studies reporting other HRV measures, either in the time (*e.g.,* square-root of the mean squared differences, or RMSDD) or frequency domains (*e.g.,* low-frequency HRV) did not enable a meta-analytic pooling of individual studies (Fig. 5).

Since there was a limited set of studies approaching the variation of finger temperature as a function of distinct emotional categories, studies were aggregated according to their emotional valence. For the set of studies with negative emotional content, there was a trend for a reduction on finger temperature (*d*=-.12, *p*=.335*; k*=13*)*. For the set of studies with positive valence, considering the very low number of studies (*k*=5) there was an overall reduction of the signal of this metric, although not statistically significant (*d*=-.57, *p*=.330) (Fig. 6).

No meta-regression aggregation was implemented for the set of cardiovascular measures, due to the reduce number of studies reporting self-reported arousal along with enough statistical information for the computation of effect-sizes. The same was also observed for the correlates of respiratory and facial systems.

3.2.3. Respiratory system

There were small-to-moderate increases in RR for happiness (*d*=.49, *p*=.090), fear (*d*=.19, *p*=.110), anger (*d*=.22, *p*=.18), sadness (*d*=.31, *p*=.35) and disgust (*d*=.92, *p*=.36) (Fig. 7). Nevertheless, none of these emotional categories was found to produce significant overall estimates. Similarly, the combined effects of the categories with emotional valence did not achieve statistical significance (*d*=.65, *p*=.139).

3.2.4. Facial system

Two EMG facial measures were considered for this meta-analytic aggregation: The Corrugator Supercilii and the Zygomaticus Major. For both measures, there was a reduced number of studies reporting essential measures for the computation of effect sizes, considering individual emotional categories. As such, we grouped the different contents according to the emotional valence of the stimuli, *i.e.,* sets of positive and negative emotions. For the set of positive emotions, non-significant changes from neutral stimulation were found for the Corrugator Supercilii (*d*=.19, *p*=.510*; k*=15); on the other hand, the set of negative emotions was associated with a significant, large, increase of the activity of the Corrugator Supercilii (*d*=1.00, *p*=.010*; k*=12). Considering the Zygomaticus Major, small, non-significant differences were obtained for the set of negative stimuli, in comparison with baseline scores (*d*=-.24, *p*=.378*; k*=9); the set of positive stimuli was characterized by significant, large increases of the Zygomaticus Major activity (*d*=2.31, *p*=.020*; k*=7) (Fig. 8).

3.3. Association between ANS effect-sizes and self-reported measures of arousal

Across all the individual emotional categories, the effect size estimates of SCL were significantly associated with the self-reported measures of arousal (r=.59, p<.001). This trend was also observed for the studies examining SCR and HR, although with considerably lower magnitudes (r=0.28 and r=0.26, respectively). For studies examining RR, there was an inverse relationship between effect-size estimates and self-reported arousal (r=-.34). For the remaining meta-analyzed ANS measures, we did not compute correlation coefficients, considering the reduced number of studies reporting self-reported measures together with these psychophysiological indices. The patterns association between ANS measures and self-reported arousal is visually represented on Fig. 9.

4. Discussion

In this study, we conducted a systematic review and a meta-analytic pipeline to investigate a set of ANS correlates to emotional induction. We specifically focused on audio-visual emotional induction (*i.e.,* video) strategies, with the goal of (1) yielding a more naturalistic approach of real-life emotional processing and (2) focusing on a single modality of emotional induction, to reduce inter-modalities variability. We focused on three categories of ANS correlates, including electrodermal, cardiovascular and respiratory measures. We found little evidence for the discrimination of individual categories, in comparison to stimuli of neutral valence, or baseline conditions. Nevertheless, specific ANS measures were associated with subjectively-rated arousal levels.

Dimensional approaches, which conceptualizes the affective experience as a continuum of ambiguous states of emotional processing (Posner et al., 2005), have gained cumulative acceptance among the scientific community. The most popular view within this doctrine is the circumplex model of affect (Russell, 1980), which states that any affective state results from the combination of two basic neurophysiological systems: hedonic valence (a continuum that varies from pleasure to displeasure) and arousal (a continuum that varies from calm to excited) (Russell, 2003). This meta-analytic aggregation confirms previous hypotheses suggesting that physiological responses vary incrementally with subjective ratings of valence and arousal (Posner et al., 2005). Specifically, (1) the level of subjective arousal has been demonstrated to be associated with increases of heart-rate and skin conductance (Lang, Greenwald, Bradley, & Hamm, 1993), (2) augmentation of the blood-oxygen-level-dependent (BOLD) contrast of the occipital cortex was linked to increased subjective arousal during the visualization of emotional static pictures (Bradley et al., 2003), (3) high arousal was associated with larger late positive event-related potentials (ERPs) (Rozenkrants, Olofsson, & Polich, 2008) and N170 (Almeida, Ferreira-Santos et al., 2016), while valence was associated with the amplitude of early to middle-range components (Olofsson, Nordin, Sequeira, & Polich, 2008). However, the magnitude of association between self-reported arousal and ANS variation varied from small to large, which raises cautious when interpreting these findings.

An important aspect that needs to be highlighted pertains to the high levels of between-studies' heterogeneity, which could not be accounted by any of the pre-defined variables of interest. This raises the possibility that some methodological aspects, including characteristics of the acquisition and processing of the psychophysiological signals may yield between-studies' differences. Having this in mind, the inclusion of structured forms for the reporting of psychophysiological signals may be an important step towards a clearer between-studies' comparability and reproducibility. In our meta-analytic investigation, a considerable portion of the included studies did not report individual measures of central tendency and dispersion (*i.e.,* means and standard deviations/standard errors/confidence-intervals) in text or tabular forms; in contrast, these measures were frequently either represented in plots or omitted, in detriment of omnibus F-statistics. When reporting between-groups' or between-conditions' comparisons, researchers tend to place an excessive focus on statistical significance, despite the limitations of this approach on the representation of the magnitude of differences (Quintana, 2017). Furthermore, as with most of psychology research, psychophysiology relies on p-values as the main source of statistical evidence(Baldwin, 2017). These practices are likely to provide overestimates of the effects' magnitude, particularly in low-powered studies (Groppe, 2017).

Following the current movements to face the reproducibility crisis, which has already been highlighted as a priority for the field of psychophysiological research (Kappenman & Keil, 2017), psychophysiology experiments would benefit from research practices that promote a comprehensive methodological description, as well as data-sharing practices that allow others to integrate the findings from multiple studies, such as meta-analytic investigations. It is our perspective that the research on this field will benefit from (1) the promotion of open-science research, namely the publication of datasets and code for processing and statistical analysis in public repositories, such as the Open Science Framework (OSF; https://osf.io/); (2) the incentive for the publication of pre-registered reports, in which the methodological protocol (including sample size determination, signal processing pipeline, analytical plan, etc.) is submitted to peer-review before the beginning of data collection; (3) the creation of online repositories/databases for the aggregation of peripheral psychophysiology investigations (following the examples of neuroimaging field, in which web-based platform allow the aggregation of multiple studies – *e.g.,* Neurosynth); (4) the development of structured checklists to enable proper and comparable reporting of psychophysiology experiments, such as the Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH) (Quintana, Alvares, & Heathers, 2016).

Even though our work shares similarities with the work from Siegel and colleagues (2008), there are some methodological and conceptual differences that deserve to be outlined. First, the abovementioned investigation provided a deep, comprehensive approach, of a set of modalities of emotion induction: with this strategy, the authors could cover static and dynamic, visual and auditory, methods for inducing emotion; in contrast, we focused on the analysis of a specific modality of emotional induction – which allowed us to approximate to a more ecological fashion for emotional induction. Second, we decided to compare different emotional categories with neutral contents, as well as with baseline levels. The engagement of participants' attention to a specific content affects by itself the ANS behavior – for this reason, we considered that the comparison with neutral stimuli would provide valuable knowledge. Third, the fact that our literature search was concluded on 2017 allowed us to include a considerable wider extension of studies using the same emotion induction modality. Therefore, we could perform a set of meta-regression analyses, which allowed us to properly assess the fit of a dimensional approach to emotion induction, by considering the impact of emotional valence and arousal. We could also include a more comprehensive characterization of ANS measures, including electromyography.

Strengths and Limitations

Despite the comprehensive approach here implemented, there are some drawbacks associated with this work. One issue pertains to the inclusion of emotional induction strategies exclusively based on films. Other modalities, such as the visualization of static (*i.e.,* pictures) or purely acoustic (*i.e.,* music or sounds) stimuli, could also be considered. It has been previously suggested that different emotional induction modalities are associated with dissociable ANS correlates. However, Siegel and colleagues (2018) provided meta-analytic evidence for a lack of significant differences between distinct emotion induction modalities. As such, it would be interesting to compare the magnitude of pooled estimates of effect sizes across different modalities. However, with this work, we favored a deeper exploration of multiple ANS correlates of different emotional categories, while focusing on a single emotional induction modality. Being characterized by a dynamic audiovisual nature, this modality may provide an increased ecological value as it may better represent a more natural induction of affective states. While the dynamic structure of these stimuli has great advantages, it also raises additional challenges. One major question is related to the fact that while a single film can evoke a multiplicity of emotions, it is typically used as representing a given valence/emotional category. Recent reports have questioned the notion that positive and negative feelings are mutually exclusive (Kreibig, Samson, & Gross, 2013) and that the implementation of paradigms that accommodate the co-occurrence of mixed emotional states are a most adequate representation of the multifactorial nature of emotion (Kreibig & Gross, 2017). In fact, whereas the use of films is likely to potentiate an increased ecological validity with respect to emotional induction, the fact that these stimuli are typically analyzed from a static perspective (*e.g.,* assuming global self-reported valence and arousal scores to the whole film) diminishes the richness that may be captured from the dynamic variation occurring during their full length. It is relevant to emphasize that the maximum shared variance between arousal and ANS variation (namely, SCL) was, approximately, 35% (R^2 value). This means that 65% of the variation of SCL is not accounted by self-reported arousal. As referred above, the dynamic variations of the arousal throughout the duration of stimuli may not be accurately represented by one single self-reported measure. The humans' ability for self-reporting their own experience is limited and often “determined” by heuristics, or shortcuts. Previous research has long demonstrated that human individuals are particularly unreliable when retrieving information from the memory of past events. People typically make their judgements based on prototypical moments, or snapshots (the most intense affective experiences) and have a difficulty in correcting for the atypicality of these moments across the overall experience – a “failure” that is known the representativeness heuristic (Kahneman & Tversky, 1972). According to this perspective, the duration of an event is typically neglected by the individual’s retrieval, in detriment of a combined over-focus on the most intense timepoints (peaks) and the end of the experience, also known as the peak-the-end heuristic (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). These constitute plausible, but theoretical explanations for the findings here reported. Complementary analyses focusing on the magnitude of the peak of skin conductance responses could better test this notion. The amount of studies reporting this measure was, however, very limited, which precluded us to implement any type of quantitative analysis.

Altogether, we consider that future studies are required to further explore the advantages of these dynamic stimuli, by continuously assessing self-reported measures of valence/arousal or equivalent. With this strategy, researchers might be able to obtain fine-grained characterizations of different intervals of emotional films, which may contribute to a better understanding of the psychophysiological correlates of emotional induction. In addition to this, future investigations may also benefit from the combined acquisition of complementary ANS signals. This multivariate strategy may be of relevance to predict, or classify, distinct emotional states, based on the psychophysiological signatures being continuously collected.

One strength of this work concerns the fact that it did not depart from a pre-conceived perspective towards the corroboration of a specific theory of emotion. Retrieving both the categorical classification of the different stimuli and, when available, the self-reported arousal and valence ratings, allowed us to pool the effects from categorical and dimensional perspectives.

In sum, this investigation provides evidence to the perspective that individual categories of emotion are not fingerprinted in individual ANS correlates. Instead, while some ANS measures were sensitive to the valence of the stimuli (*e.g.,* EMG), measures from the electrodermal and cardiovascular system were generally sensitive to the intensity (*i.e.,* arousal) of the content being perceived. Nevertheless, the increased between-studies’ heterogeneity (which had also been described in previous aggregations of psychophysiological measures) raises cautious on the interpretation of the pooled estimates.

Acknowledgements

Pedro Silva Moreira is supported by an FCT fellowship grant (PhD-iHES program) with the reference PDE/BDE/113601/2015.

Financial Disclosures

Pedro Chaves, Nuno Dias and Pedro R. Almeida are part of MindProber Labs, a technological provider for market research – which partially supported this research.

The paper has not been published previously, or is under consideration for publication elsewhere, in English or in any other language.

References

Almeida, P.R., Ferreira-Santos, F., Chaves, P.L, Piva, T.O., Barbosa, F, & Marques-Teixeira, J. (2016). Perceived arousal of facial expressions of emotion modulates the N170, regardless of emotional category: Time domain and time-frequency dynamics. *International Journal of Psychophysiology, 99*, 48-56.

Baldwin, S. A. (2017). Improving the rigor of psychophysiology research. *International Journal of Psychophysiology, 111*, 5-16.

Barrett, L. F. (2017a). *How emotions are made: The secret life of the brain*: Houghton Mifflin Harcourt.

Barrett, L. F. (2017b). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience, 12*(1), 1-23.

Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology, 60*(1), 34-43.

Bos, M. G., Jentgens, P., Beckers, T., & Kindt, M. (2013). Psychophysiological response patterns to affective film stimuli. *PLoS One, 8*(4), e62661.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry, 25*(1), 49-59.

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral neuroscience, 117*(2), 369.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. *Handbook of emotions, 2*, 173-191.

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American journal of psychology, 39*(1/4), 106-124.

Carvalho, S., Leite, J., Galdo-Álvarez, S., & Gonçalves, Ó. F. (2012). The emotional movie database (EMDB): A self-report and psychophysiological study. *Applied psychophysiology and biofeedback, 37*(4), 279-294.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd. In: Hillsdale, NJ: erlbaum.

Fedorov, S. (2008). GetData graph digitizer. *available at www. getdata-graph-digitizer. com*.

Gilman, T. L., Shaheen, R., Nylocks, K. M., Halachoff, D., Chapman, J., Flynn, J. J., . . . Coifman, K. G. (2017). A film set for the elicitation of emotion in research: A comprehensive catalog derived from four decades of investigation. *Behavior research methods, 49*(6), 2061-2082.

Groppe, D. M. (2017). Combating the scientific decline effect with confidence (intervals). *Psychophysiology, 54*(1), 139-145.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & emotion, 9*(1), 87-108.

Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., & Bartussek, D. (2005). A revised film set for the induction of basic emotions. *Cognition and emotion, 19*(7), 1095-1109. doi:10.1080/02699930541000084

James, W. (1884). What is an emotion? *Mind, 9*(34), 188-205.

Jenkins, L. M., & Andrewes, D. G. (2012). A new set of standardised verbal and non-verbal contemporary film stimuli for the elicitation of emotions. *Brain Impairment, 13*(2), 212-227.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological science, 4*(6), 401-405.
Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. In *The concept of probability in psychological experiments* (pp. 25-48): Springer.
Kalluri, H. V., Zhang, H., Caritis, S. N., & Venkataramanan, R. (2017). A physiologically based pharmacokinetic modelling approach to predict buprenorphine pharmacokinetics following intravenous and sublingual administration. *British journal of clinical pharmacology, 83*(11), 2458-2473.
Kappenman, E. S., & Keil, A. (2017). Introduction to the special issue on recentering science: replication, robustness, and reproducibility in psychophysiology. *Psychophysiology, 54*(1), 3-5.
Kaviani, H., Gray, J. A., Checkley, S. A., Kumari, V., & Wilson, G. D. (1999). Modulation of the acoustic startle reflex by emotionally-toned film-clips. *International Journal of Psychophysiology, 32*(1), 47-54.
Kreibig, S. D., & Gross, J. J. (2017). Understanding Mixed Emotions: Paradigms and Measures. *Curr Opin Behav Sci, 15*, 62-71. doi:10.1016/j.cobeha.2017.05.016
Kreibig, S. D., Samson, A. C., & Gross, J. J. (2013). The psychophysiology of mixed emotional states. *Psychophysiology, 50*(8), 799-811.
Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261-273.
Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The psychophysiology of emotion. *Handbook of emotions, 3*, 180-195.
Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychological bulletin, 137*(5), 834.
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.
Maffei, C., Roder, E., Cortesan, C., Passera, F., Rossi, M., Segrini, M., . . . Fossati, A. (2014). Kinematic elicitation of basic emotions: A validation study in an Italian sample. *Psychology, 5*(09), 1065.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine, 6*(7), e1000097.
Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*(5827), 1002-1005.
Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological psychology, 77*(3), 247-265. doi:10.1016/j.biopsycho.2007.11.006
Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and emotion, 7*(2), 171-193.
Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology, 17*(3), 715-734.
Pustejovsky, J. E., & Tipton, E. (2014). Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models. *Journal of Business & Economic Statistics*, 1-12.
Quintana, D. S. (2017). Statistical considerations for reporting and planning heart rate variability case-control studies. *Psychophysiology, 54*(3), 344-349.

Quintana, D. S., Alvares, G. A., & Heathers, J. A. J. (2016). Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): recommendations to advance research communication. *Translational psychiatry, 6*(5), e803.
Ray, R. D. (2007). Emotion elicitation using films. *Handbook of emotion elicitation and assessment, 9*.
Rozenkrants, B., Olofsson, J. K., & Polich, J. (2008). Affective visual event-related potentials: Arousal, valence, and repetition effects for normal and distorted pictures. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology, 67*(2), 114-123. doi:10.1016/j.ijpsycho.2007.10.010
Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology, 39*(6), 1161.
Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review, 110*(1), 145.
Samson, A. C., Kreibig, S. D., Soderstrom, B., Wade, A. A., & Gross, J. J. (2016). Eliciting positive, negative and mixed emotional states: A film library for affective scientists. *Cognition and emotion, 30*(5), 827-856.
Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological review, 69*(5), 379.
Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and emotion, 24*(7), 1153-1172.
Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., . . . Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological bulletin, 144*(4), 343.
Solomon, R. C. (1993). The philosophy of emotions. *Handbook of emotions, 2*, 5-13.
Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *2010, 36*(3), 48. doi:10.18637/jss.v036.i03
Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology, 54*(6), 1063.

Supplementary Material

Supplementary File 1. PRISMA checklist

Table 1. Description of ANS Measures

Group		Measure	Definition
Electrodermal system			
	Tonic measures		
		Skin conductance level (SCL)	Tonic level of electrical conductivity of skin
	Phasic measures		
		Change in SCL	Gradual changes in SCL measured at two or more points in time
		Frequency of NS-SCRs	Number of SCRs in absence of identifiable eliciting stimulus
		SCR amplitude	Phasic increase in conductance shortly following stimulus onset
		SCR latency	Temporal interval between stimulus onset and SCR initiation
		SCR rise time	Temporal interval between SCR initiation and SCR peak
		SCR half recovery time	Temporal interval between SCR peak and point of 50% recovery of SCR amplitude
		SCR habitation (trials to habituation)	Number of stimulus presentations before two or three trials with no response
		SCR habituation (slope)	Rate of change of ER-SCR amplitude
Cardiovascular system			
	Electrocardiogram		
		Heart rate (HR)	Temporal interval between successive R spikes
		SDNN	Standard deviation of the normal beat to normal beat intervals (normal-to-normal or NN)
		RMSSD	Root Mean Square Successive Difference) statistic
		Low Frequency (LF)	Power in low-frequency range. Mixture of sympathetic and parasympathetic rhythms
		High Frequency (HF)	Heart rate fluctuations occurring within the respiratory frequency band - Power in high-frequency range
		Respiratory sinus arrhythmia (RSA)	Respiratory gating of autonomic control by afferent input from lung stretch receptors
	Blood Pressure		
		Systolic	Systolic Blood Pressure
		Diastolic	Diastolic Blood Pressure
		Finger Temperature	
Respiratory system			
		Respiratory rate	Number of breaths
Facial system			
	EMG		
		Corrugator Supercilii	Group of facial muscles associated with frowning
		Zygomaticus Major	Group of facial muscles associated with smiling

Table 2. Characteristics of the included studies

Author	Year	N	Country	Provenience	Mean Age	% Males
Aguado	2016	38	Spain	University	22.3	50%
Aldao	2013	17	US	University/Community	30.5	--
Anastassiou-Hadjicharalambous	2008	44	UK	Community	8.7	50%
Arnaudova	2017	78	Netherlands	University	19.6	6%
Austin	2007	11	US	Community	--	0%
Ayala	2010	20	US	Community	28.2	15%
Baldaro	2001	45	Italy	University	--	0%
Baldaro	1996	30	Italy	University	--	47%
Baldaro	1990	24	Italy	University	--	33%
Beevers	2011	67	US	University	18.4	0%
Bensafi	2004	72	US	University	12.8	50%
Berna	2014	63	France	University	20.9	0%
Blau	2009	86	Israel	kindergarten children	4.7	37%
Bogdanov	2013	21	Ucraine	University	--	48%
Bos	2013	35	Netherlands	University	20.6	34%
Bos	2013	35	Netherlands	University	20.6	34%
Bradley	2009	96	UK	Community	35.6	41%
Bradley	2007	50	US	University	19.7	44%
Bride	2014	408	US	University/Community	24.4	37%
Britton	2006	40	US	Community	19.3	48%
Brumbaugh	2013	169	US	Community	27.0	40%
Brzozowski	2017	49	UK	University	18.9	43%
Busscher	2010	36	Netherlands	Community	43.4	47%
Butler	2006	36	US	University	20.0	0%
Carboni	2017	30	Spain	University	24.8	37%
Carvalho	2012	32	Portugal/Spain	University	23.3	50%
Chentsova-Dutton	2010	34	US	Community	30.2	0%
Chentsova-Dutton	2010	60	US	University	19.4	49%
Chentsova-Dutton	2010	18	US	Community	32.1	0%
Chentsova-Dutton	2010	16	US	Community	28.4	0%
Chentsova-Dutton	2014	114	US	University/Community	21.3	35%
Clapp	2015	192	US	University	19.9	43%
Codispoti	2008	60	Italy	University	23.1	45%
Costa	2009	60	Italy	--	27.6	50%
Coyne	2011	50	UK	University	24.7	40%
Crowell	2017	116	US	Community	35.0	0%
Davis	2016	101	US	Community	5.8	54%
Davydov	2011	26	Belgium	University	20.0	0%
Davydov	2011	26	Belgium	University	20.0	0%
Davydov	2013	26	Belgium	University	20.0	0%
de Groot	2014	52	Netherlands	--	22.4	50%
de Jong	2011	60	Netherlands	University	21.6	13%
de Sousa	2012	25	Australia	Community	29.0	56%
de Wied	2012	32	Netherlands	Community	13.8	100%
Demaree	2004	52	US	University	18.5	48%
Demaree	2005	69	US	University	19.3	48%
Deng	2017	110	China	University	21.2	28%
Deng	2016	79	China	Community	20.9	39%
Eberhardt	2016	17	Germany	University	21.5	0%
Eisenberg	1992	117	US	School	7.3	56%
Elices	2012	30	Spain	University/Community	26.9	0%
Erisman	2010	15	US	University	24.1	50%
Evans	2013	87	UK	Community	33.1	33%
Fang	2001	62	US	University	19.7	100%
Fanti	2016	56	Cyprus	University/Community	20.5	46%
Fanti	2017	82	Cyprus	University/Community	21.0	50%
Fernández	2012	123	Spain	University/Community	29.2	26%

Fortunato	2013	273	Germany	Kindergarten	6.3	36%
Fowles	2000	92	US	Community	4.4	51%
Francis	2016	58	Australia	University/Community	23.8	26%
Fredrickson	1998	72	US	Community	--	50%
Gatzke-Kopp	2014	209	US	kindergarten children	6.0	63%
Gentzler	2009	65	US	Community	7.9	54%
Gilbert	2016	83	US	Community	19.7	0%
Gilchrist	2016	60	UK	University/Community	22.0	75%
Gilchrist	2016	42	UK	University/Community	22.0	26%
Giuliani	2008	16	US	University	18.8	0%
Glissen	2008	78	Netherlands	Community	3.8	49%
Glissen	2008	92	Netherlands	Community	7.4	47%
Glissen	2007	78	Netherlands	Community	3.9	49%
Golland	2015	78	Israel	University	--	0%
Golland	2014	27	Israel	University	20.0	33%
Gomez	2005	73	Germany	University	24.0	51%
Gomez	2009	76	Germany	University/Community	24.0	51%
Gračanin	2007	65	--	University	21.5	22%
Gross	1998	40	US	University	21.0	50%
Gross	1993	43	US	University	19.3	100%
Gross	1994	150	US	University	19.1	0%
Gruber	2011	24	US	Community	35.5	50%
Gruber	2011	31	US	University	20.4	35%
Hagenaars	2014	50	Netherlands	University	21.0	0%
Hamilton	2011	19	US	Community	24.4	0%
Harrison	2000	30	UK	University	21.0	50%
Hastings	2009	215	US	Community	13.3	51%
Hastings	2014	220	US	Community	13.7	50%
Hendriks	2007	60	Netherlands	University	20.4	0%
Herring	2011	39	US	University	21.5	31%
Herring	2011	39	US	University	21.5	31%
Hsieh	2016	34	Taiwan	University	22.0	18%
Ivonin	2015	23	Spain	University	27.8	57%
Ivonin	2015	25	Netherlands	University	24.0	52%
Jang	2015	20	Korea	University	21.0	50%
Jin	2015	25	US	Community	31.0	40%
Jones	2014	20	US	Community	3.1	75%
Jones	2014	21	US	Community	5.9	52%
Jönsson	2008	30	Sweden	University	23.3	50%
Kalvin	2016	169	US	kindergarten children	5.6	66%
Kaviani	2010	16	UK	Community	28.3	50%
Kaviani	2005	16	Iran	University	27.4	50%
Kaviani	2006	20	Iran	Community	--	100%
Kindt	2005	50	Netherlands	University	20.7	30%
Kornreich	1998	14	Belgium	Community	--	--
Krahé	2011	303	Germany	University	23.8	71%
Kreibig	2007	34	US	University	21.0	44%
Kreibig	2011	32	US	University	20.9	47%
Kreibig	2013	43	US	University	20.8	0%
Kreibig	2015	48	US	University	20.7	0%
Kreibig	2007	6			0.0	44%
Kuijsters	2016	15	Netherlands	University	22.4	53%
Kuijsters	2016	15	Netherlands	University	22.4	53%
Kuijsters	2015	38	Netherlands	Community	78.8	50%
Kumari	2001	10	UK	Community	--	30%
Kunzmann	2005	48	US	Community	21.0	50%
Kunzmann	2005	47	US	Community	71.0	51%
Kunzmann	2005	48	Germany	Community	23.9	50%
Kuo	2009	20	US	Community	23.3	0%
Kuo	2013	20	Canada	Community	23.3	0%

Kuypers	2017	80	Netherlands	University	22.5	50%
Kyranides	2016	82	Cyprus	Community	20.0	51%
Laan	1995	49	Netherlands	University	22.3	0%
Lackner	2014	48	Austria	University	21.0	0%
Lane	2009	12	US	Community	23.3	0%
LeBlanc	2016	26	US	Community	38.7	27%
Lee	2009	80	Korea	University	20.8	46%
Lin	2017	50	Israel	Soldiers	18.9	100%
Llera	2014	95	US	University	19.0	28%
Lobbestael	2006	64	Netherlands	University (7 were not)	23.4	50%
López-Benítez	2017	31	Spain	University	21.1	19%
López-Benítez	2017	33	Spain	University	21.1	39%
Maras	2012	19	UK	--	37.1	83%
Marsh	2007	23	US	Community	10.5	100%
Matsunaga	2009	12	Japan	--	--	100%
Matsunaga	2008	11	Japan	University	--	50%
Merrifield	2014	72	Canada	University	18.9	39%
Mohino-Herranz	2015	40	Spain	University	--	70%
Montoya	2005	32	Spain	--	26.0	63%
Morawetz	2016	23	Germany	--	23.0	35%
Musser	2013	75	US	Community	7.6	49%
Olatunji	2015	95	US	University	19.0	24%
Pallavicini	2013	34	Italy	University	21.2	--
Palomba	2000	46	Italy	University	23.8	33%
Pang	2013	207	US	Community	9.9	--
Park	2013	12	Korea	University	20.0	50%
Park	2011	20	Korea	University/Community	29.3	50%
Pfabigan	2015	15	Austria	Community	35.6	100%
Pichon	2014	25	--	--	23.2	48%
Pu	2010	136	US	University	18.8	49%
Quas	2007	109	US	Community	6.1	51%
Radstaak	2011	110	Netherlands	University	21.1	13%
Ramos	2015	70	Spain	University	31.7	--
Renshon	2015	138	US	University	22.8	100%
Reynaud	2012	33	France	Community	27.5	12%
Rickard	2004	21	Australia	University	25.5	57%
Rimes	2016	80	UK	Community	33.0	29%
Rimm-Kaufman	1996	32	US	University	20.0	0%
Rimm-Kaufman	1996	32	US	University	--	0%
Ripley	2017	184	US	University	19.9	45%
Ritz	2010	25	Germany	Community	28.0	36%
Ritz	2013	20	US	University/Community	27.7	18%
Ritz	2012	14	US	Community	36.4	72%
Ritz	2005	14	Germany	Community	36.4	29%
Ritz	2011	14	America	Community	36.4	29%
Ritz	2010	25	Germany	University/Community	28.0	36%
Ritz	2011	14	US	Community	36.4	29%
Roberts	2008	160	US	University	20.8	40%
Robinson	2007	55	US	University	19.1	47%
Rohrmann	2008	89	Germany	University	27.9	53%
Rohrmann	2008	89	Germany	University	27.9	53%
Rohrmann	2009	120	Germany	University	25.5	100%
Rommel	2015	24	France	University	19.0	0%
Rosselló	2015	30	Spain	Community	48.1	0%
Roth	2014	33	Israel	University	24.9	40%
Rottenberg	2003	31	US	--	33.5	0%
Rottenberg	2002	33	US	Community	32.3	30%
Rushby	2013	25	Australia	Community	31.0	56%
Salters-Pedneault	2007	37	US	University/Community	26.7	0%
Sarlo	2008	17	Italy	University	22.7	0%

Schaich	2013	66	Netherlands	University	20.1	0%
Schallcross	2017	142	US	University/Community	22.1	30%
Schmeichel	2006	50	US	University	18.9	46%
Schneider	2012	28	Germany	Community	34.7	54%
Schneiderman	2011	112	US	University/Community	23.4	51%
Schröder	2015	16	Germany	Community	30.2	81%
Seeley	2016	76	US	University/Community	26.6	41%
Seider	2011	76	US	Community	25.4	49%
Seider	2011	73	US	Community	43.7	49%
Seider	2011	73	US	Community	64.6	47%
Shenhav	2014	80	US	University/Community	27.1	51%
Shenhav	2014	80	US	University/Community	27.1	51%
Sheppes	2009	45	Israel	University	22.9	0%
Shi	2017	48	China	University	23.5	48%
Silvestrini	2007	43	Switzerland	University	24.0	84%
Simon	2017	20	US	Community	27.9	19%
Simon	2017	20	US	University/Community	27.2	15%
Šolcová	2017	124	Czech Republic	University	22.5	41%
Soto	2016	59	US	University/Community	19.5	46%
Stange	2017	134	US	University	21.9	42%
Stephens	2010	49	US	University	19.3	45%
Stoléru	1999	8	France	University	23.0	100%
Svaldi	2012	17	Germany	University	22.8	0%
Svaldi	2012	17	Germany	University	22.7	0%
Tramoni	2008	13	France	Community	24.6	38%
Tsai	2000	24	US	Community	27.9	50%
Tsai	2000	24	US	Community	75.7	50%
Tsai	2000	24	US	Community	26.7	50%
Tsai	2000	24	US	Community	73.6	50%
Tuck	2017	117	New Zealand	Community	41.8	40%
Tull	2007	17	US	University	22.0	12%
Tull	2010	34	US	University	25.9	100%
Uy	2013	7	US	Community	29.8	43%
Valiente	2004	157	US	School	7.7	53%
van den Broek	2009	24	Netherlands	Community	43.0	17%
Vasilev	2009	69	US	Community	9.8	--
Vianna	2006	16	US	15.88 (0.33)	26.7	44%
Wang	2013	98	China	University	20.0	21%
Wegerer	2013	66	Austria	University	23.4	0%
Wegerer	2013	66	Austria	University	23.4	0%
Wegerer	2014	37	Austria	University	23.9	0%
Wen	2014	27	China	University	20.0	33%
Werner	2007	16	California	Community	67.0	84%
Werner	2015	29	Austria	University	23.6	0%
Werner	2015	29	Austria	University	23.6	0%
Wittling	1998	45	--	--	--	24%
Wolgast	2011	94	Sweden	University	27.4	49%
Wu	2014	8	Belgium	University	--	63%
Wu	2014	8	Belgium	Community	--	53%
Yaroslavsky	2013	75	US	Community	29.6	0%
Yaroslavsky	2013	94	US	Community	29.0	26%
Yaroslavsky	2016	161	US	Community	16.5	64%
Yaroslavsky	2014	170	US	Community	30.9	--
Zantinge	2017	45	Netherlands	School	4.6	82%

Table 3. Summary of the pooled estimates

		k	Pooled ES	Lower CI	Upper CI	I2	Cluster-Robust ES	p-value
SCL								
	Sadness	31	0.27	0.09	0.46	91.709	0.23	0.08
	Disgust	20	0.942	0.58	1.31	97	0.76	<.001
	Fear	16	1	0.57	1.3	95.7	0.775	<.001
	Anger	4	0.53	0.076	0.98	92.95	0.528	0.11
	Happiness	10	0.66	0	1.04	91.79	0.7	0.027
	Negative	87	0.48	0.36	0.59	92.64	0.452	<.001
SCR								
	Sadness	7	0.42	0.1	0.74	84.51	0.42	0.04
	Disgust	4	3	-2.275	8.7	99.85	3.21	0.33
	Fear	6	3.11	-1	7.089	99.8	3.11	0.18
	Anger	4	0.24	-0.17	1	82.044	0.24	0.34
	Happiness	5	0.75	0.08	1.42	94	0.751	0.09
	Negative	25	1.36	0.25	2.48	99.59	1	0.144
HR								
	Sadness	32	-0.042	-0.23	0.15	94.26	-0.05	0.68
	Disgust	26	0	-0.752	0.02	98.19	-0.39	0.09
	Fear	17	0.14	0	0.547	97.96	0.14	0.52
	Anger	16	0.46	-0.06	1	98.568	0.46	0.1
	Happiness	28	-0.25	-0.51	0.01	95	-0.245	0.09
	Negative	96	-0.13	-0.68	0.42	99.74	0	0.773
HF-HRV								
	Sadness	9	1.483	-2.12	5.09	99.93	1.48	0.44
	Disgust	6	0	-0.427	0.27	86.22	-0.08	0.66
	Fear	8	-0.36	-1	0.148	97.12	-0.36	0.21
	Anger	2	-0.59	-0.96	0	29.752	-0.59	0.2
	Happiness	6	-0.16	-0.53	0.21	87	-0.163	0.42
	Negative	26	0.19	-0.85	1.23	99.76	0	0.116
RR								
	Sadness	6	0.312	-0.28	0.9	96.17	0.31	0.35
	Disgust	6	1	-0.876	2.72	99.52	0.92	0.36
	Fear	5	0.19	0	0.385	61.36	0.19	0.11
	Anger	3	0.22	0	0	57.89	0.22	0.18
	Happiness	7	0.47	0.08	0.86	90	0.485	0.09
	Negative	20	0.44	-0.08	0.97	98.74	1	0.139
EMG-Cor								
	Positive	15	0.27	-0.06	0.6	92.05	0.189	0.51
	Negative	12	0.88	0.5	1.263	91.46	1	0.01
EMG-Zyg								
	Positive	7	2.035	0.74	3	98.35	2.31	0.02
	Negative	9	0	-0.6	0.15	91.15	-0.24	0.378
FT								
	Positive	5	-0.42	-1.29	0.439	90.27	-0.57	0.33
	Negative	13	-0.12	-0.28	0	77.44	-0.12	0.335

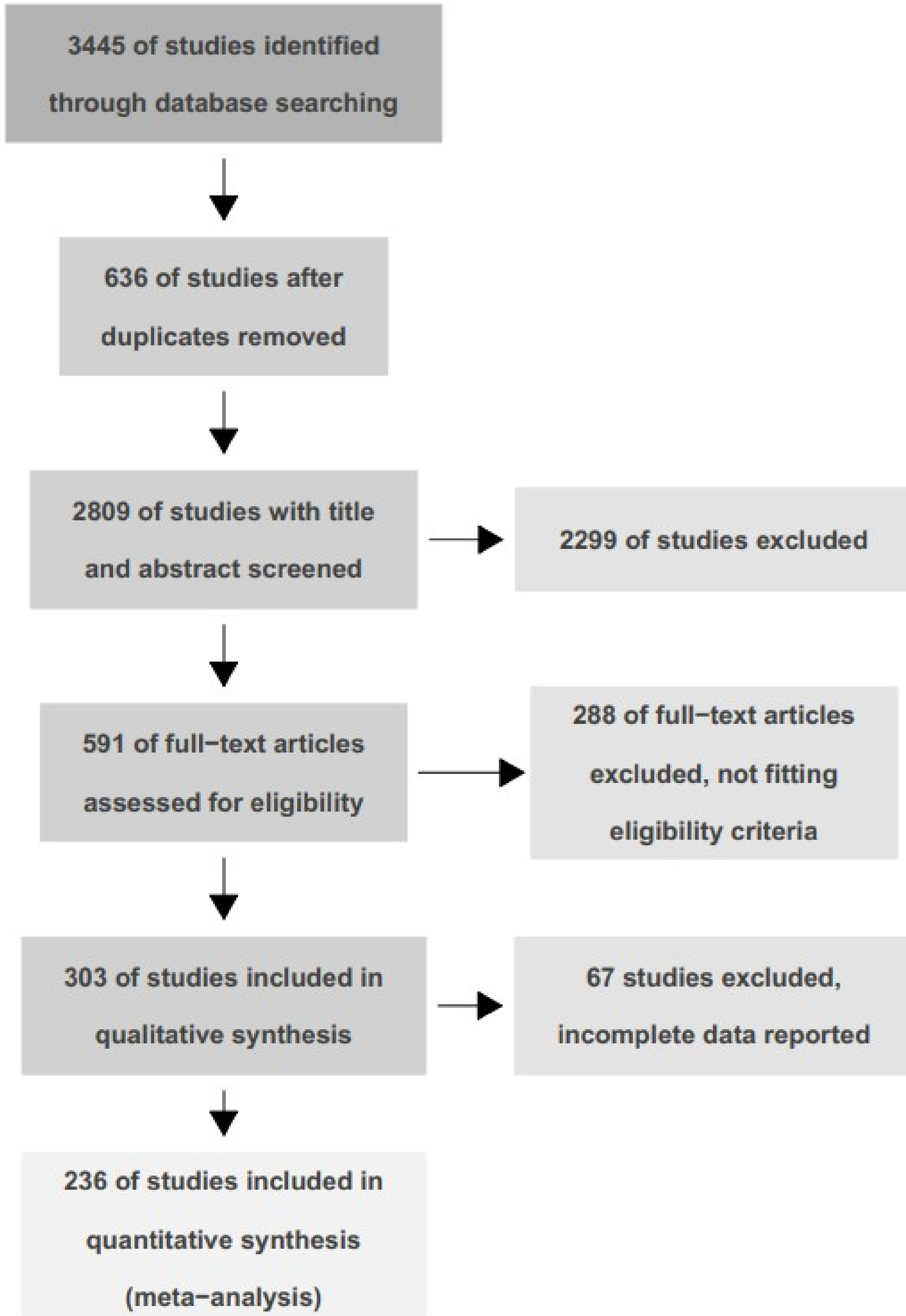

Figure 1. PRISMA flowchart

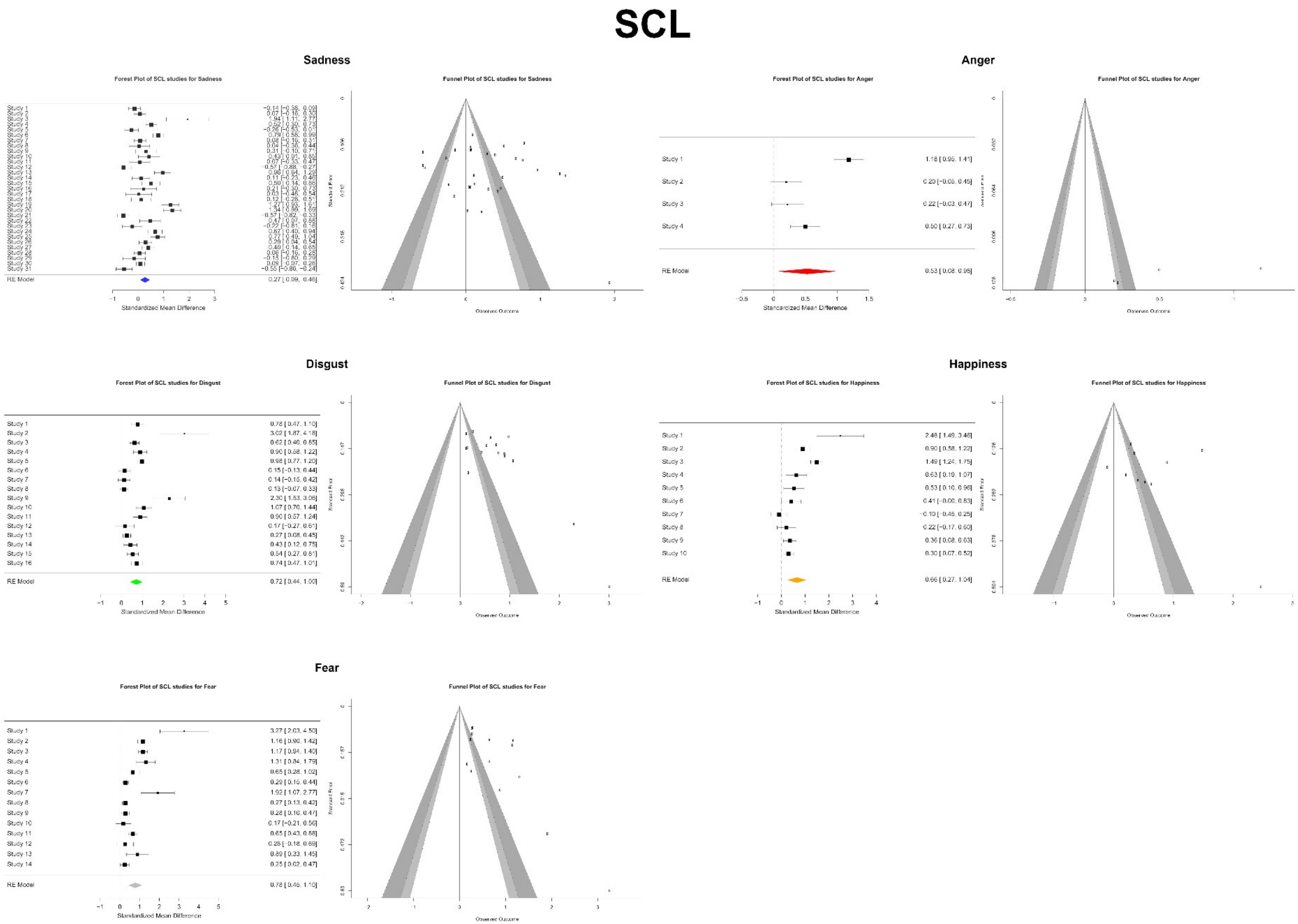

Figure 2. Forest and funnel plots for SCL correlates of the different emotional categories. The summary statistics represented in forest plots differ from the conservative cluster-robust estimates, which considers the statistical dependency between effect-sizes obtained reported on the same manuscript.

Figure 3. Forest and funnel plots for SCR correlates of the different emotional categories. The summary statistics represented in forest plots differ from the conservative cluster-robust estimates, which considers the statistical dependency between effect-sizes obtained reported on the same manuscript.

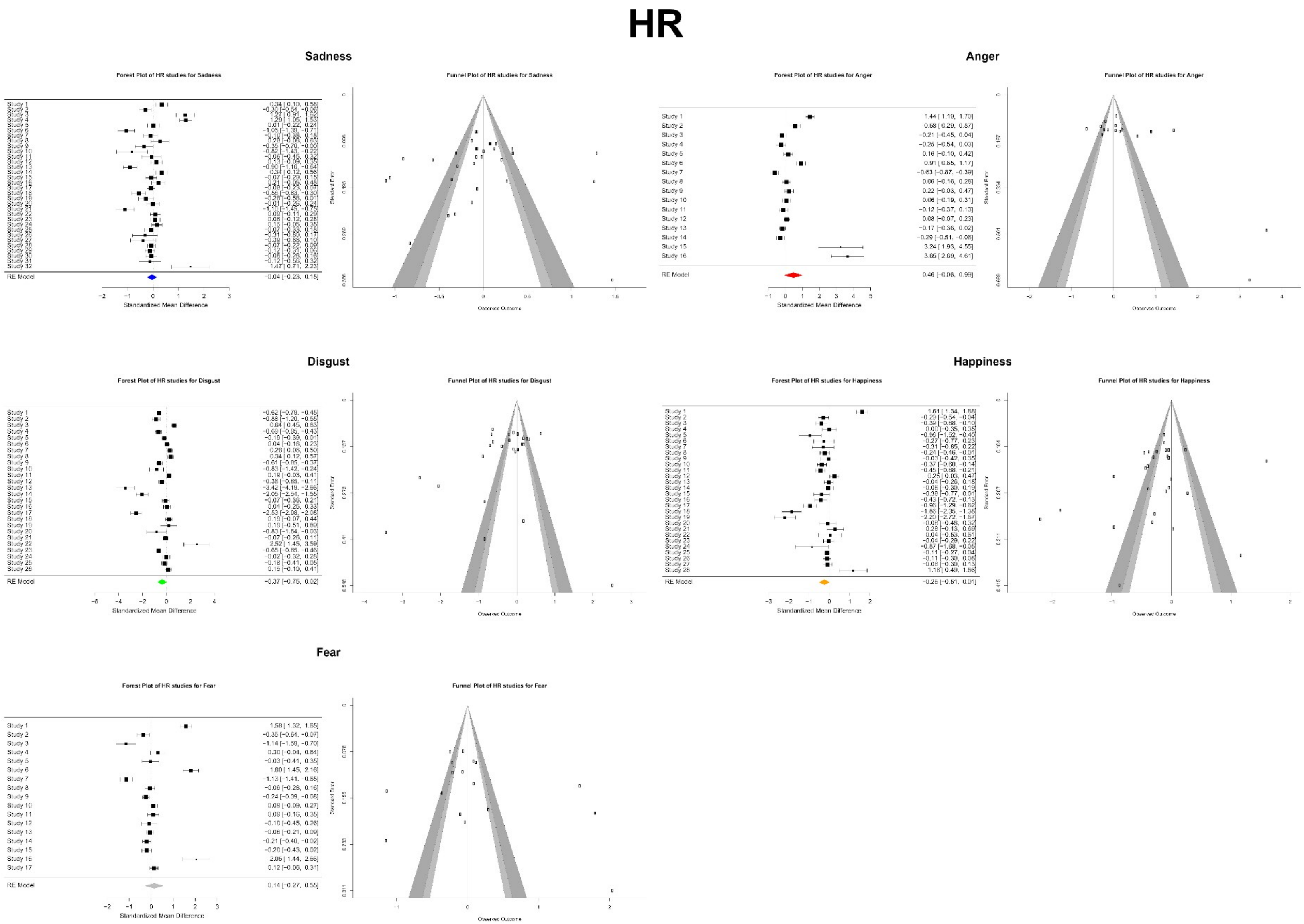

Figure 4. Forest and funnel plots for HR correlates of the different emotional categories. The summary statistics represented in forest plots differ from the conservative cluster-robust estimates, which considers the statistical dependency between effect-sizes obtained reported on the same manuscript.

HF-HRV

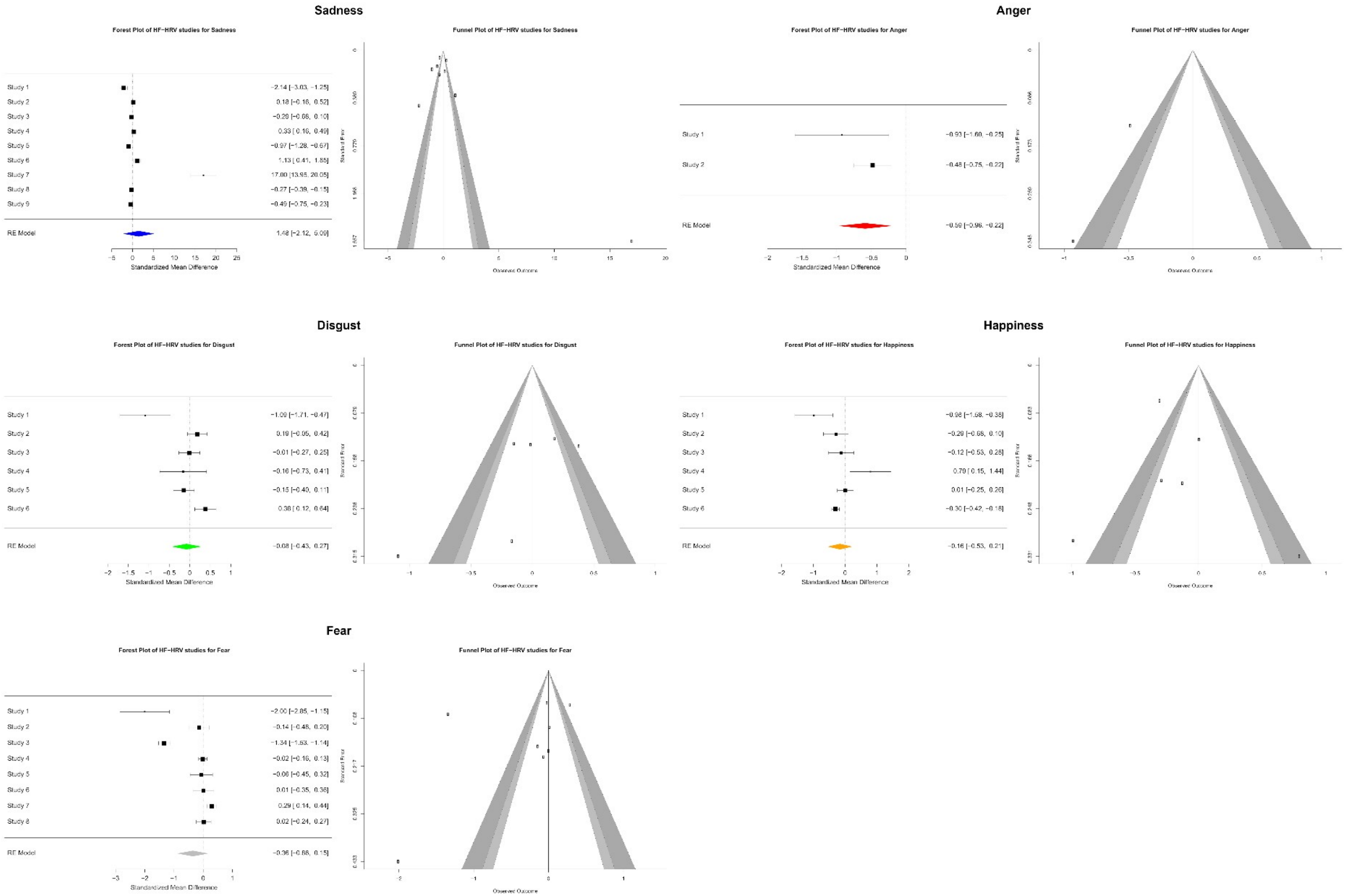

Figure 5. Forest and funnel plots for HF-HRV correlates of the different emotional categories. The summary statistics represented in forest plots differ from the conservative cluster-robust estimates, which considers the statistical dependency between effect-sizes obtained reported on the same manuscript.

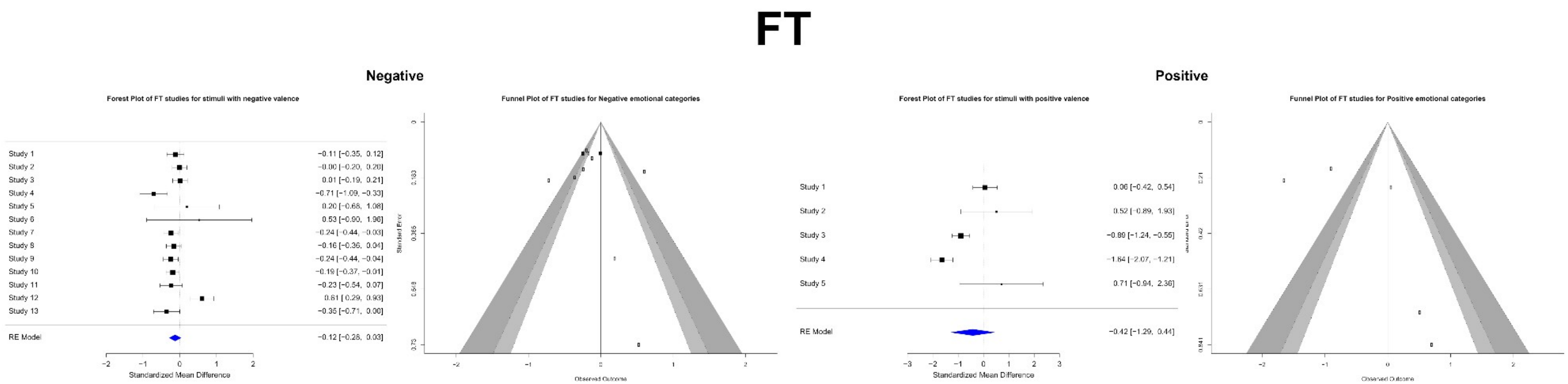

Figure 6. Forest and funnel plots for FT correlates of positive and negative valence. The summary statistics represented in forest plots differ from the conservative cluster-robust estimates, which considers the statistical dependency between effect-sizes obtained reported on the same manuscript.

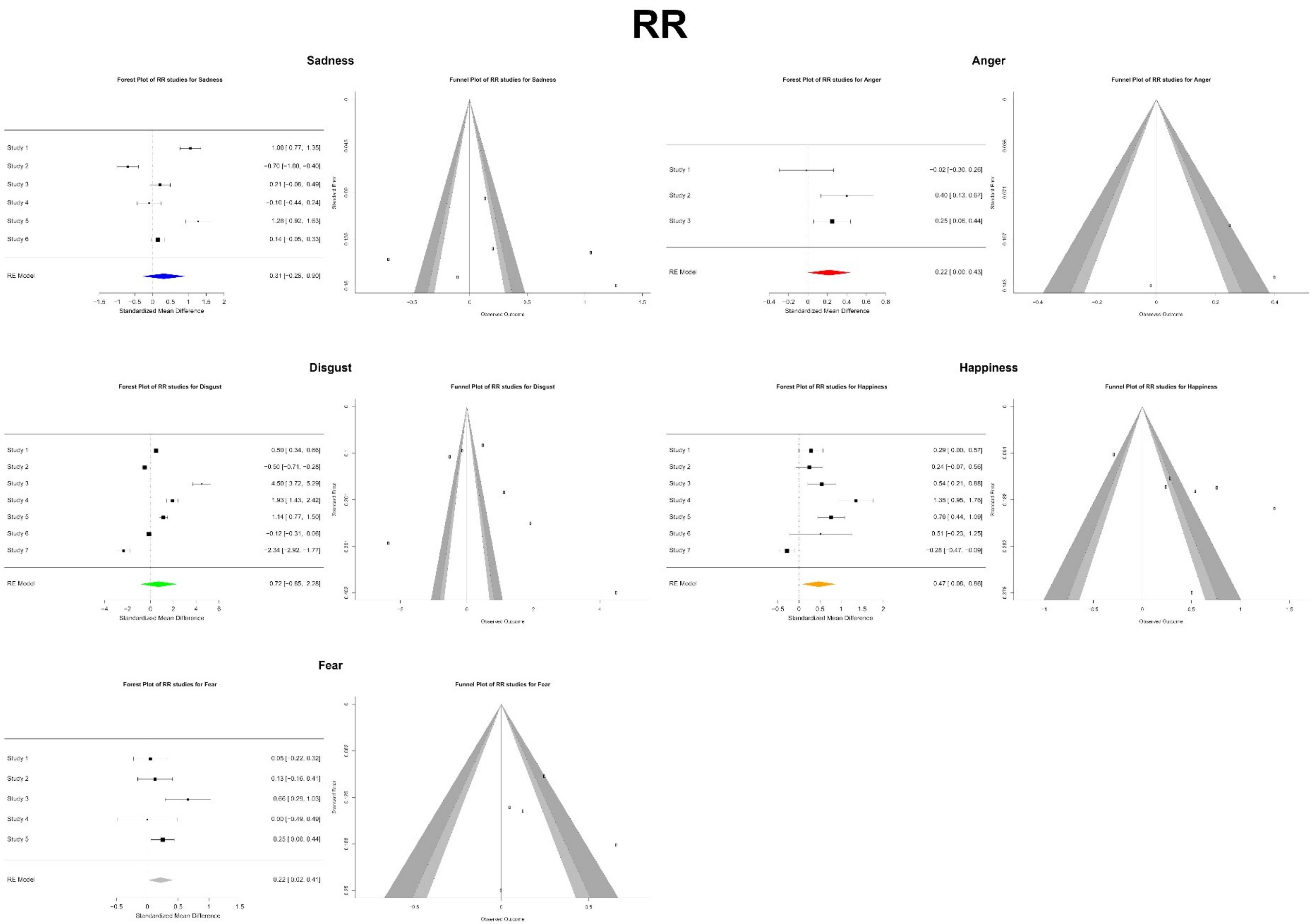

Figure 7. Forest and funnel plots for RR correlates of the different emotional categories. The summary statistics represented in forest plots differ from the conservative cluster-robust estimates, which considers the statistical dependency between effect-sizes obtained reported on the same manuscript.

EMG

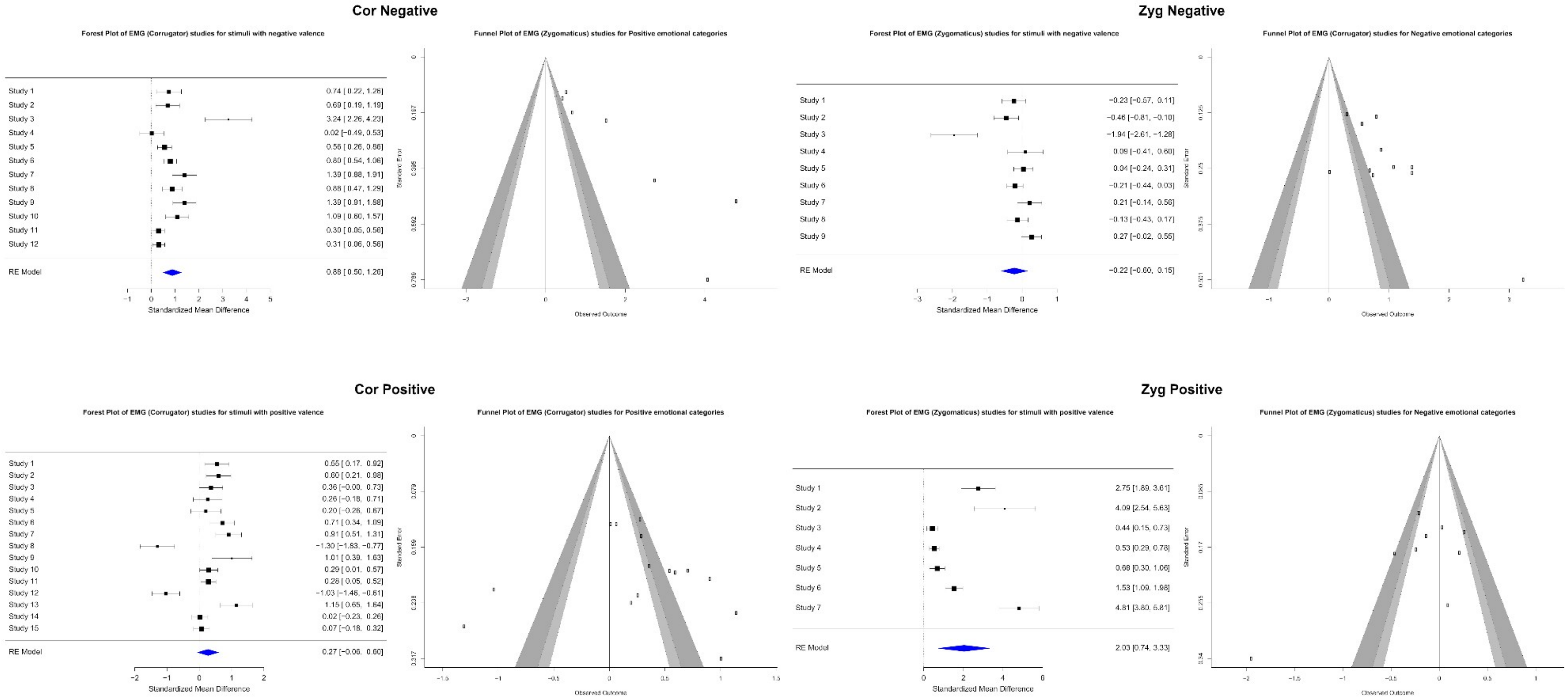

Figure 8. Forest and funnel plots for EMG correlates of positive and negative valence. The summary statistics represented in forest plots differ from the conservative cluster-robust estimates, which considers the statistical dependency between effect-sizes obtained reported on the same manuscript.

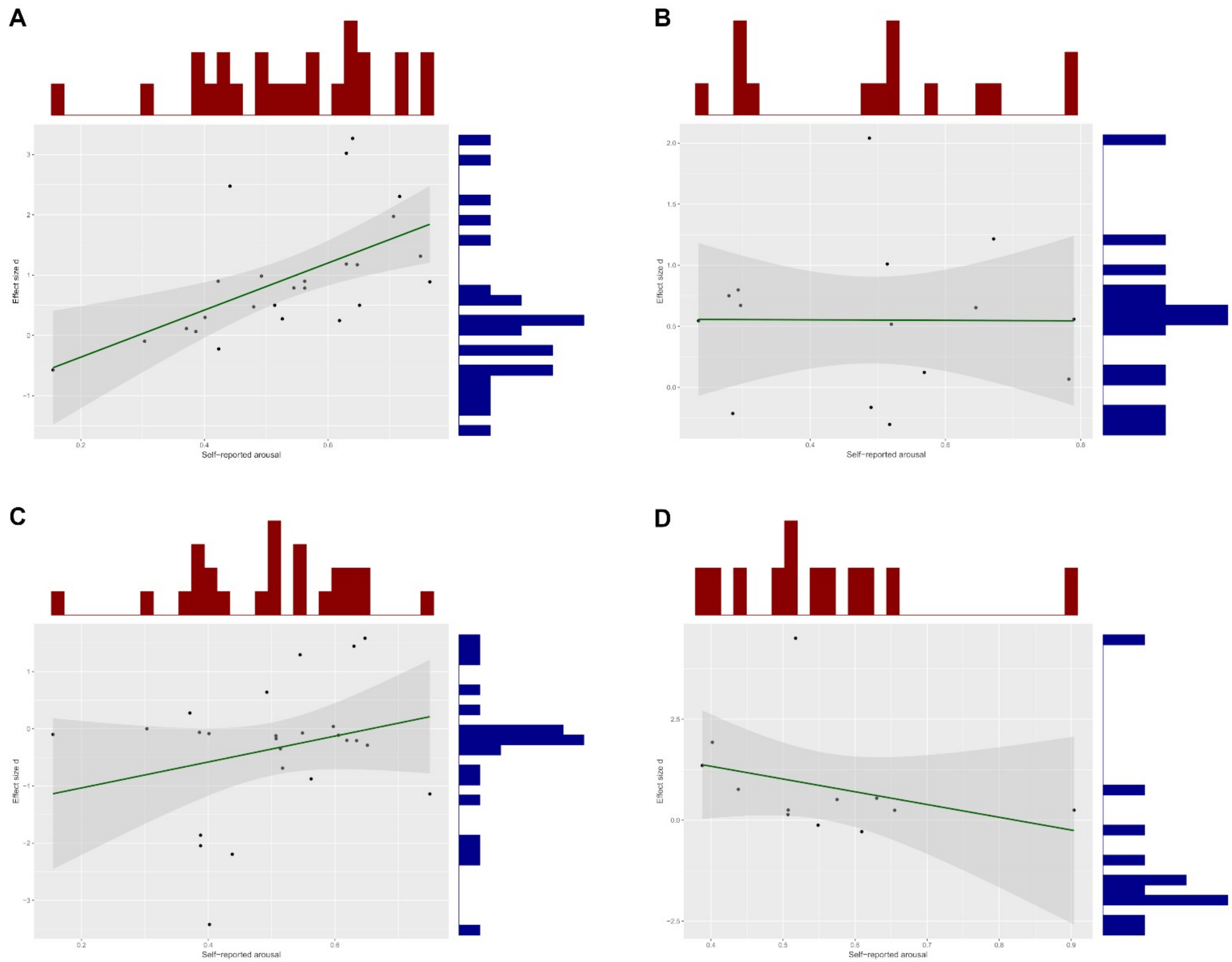

Figure 9. Scatter plots representing the association between self-reported arousal and ANS correlates for all the emotional categories, including (A) SCL, (B) SCR, (C) HR and (D) RR.