Motivation and Emotion (2019) 43:929–939
https://doi.org/10.1007/s11031-019-09780-y

ORIGINAL PAPER

EmoMadrid: An emotional pictures database for affect research

L. Carretié[1] · M. Tapia[1] · S. López-Martín[1] · J. Albert[1]

Published online: 1 July 2019

Abstract
Emotional scenes are, along with facial expressions, the most employed stimuli in Affective Sciences. However, as compared to facial expressions, available emotional scene databases are scarce and, in some cases, obsolete and overused. This paper describes EmoMadrid (http://www.psicologiauam.es/CEACO/EmoMadrid.htm), an open access database currently consisting of 813 emotional pictures. Valence and Arousal of each of these pictures were assessed by an average sample of 146 volunteers per session, who evaluated an average of 155 pictures each. Total participants up to the present is 768. EmoMadrid includes information, not provided in other databases, on low order visual parameters such as spatial frequency, luminosity, and chromatic complexity. These parameters are of crucial interest, since they have been revealed to interact with the affective content of pictures. EmoMadrid shows a robust short and long term reliability (under and over 5 years, respectively) and has already been employed in 15 Human Neuroscience and Behavior published studies, despite it has only been described in its web page.

Keywords Emotional pictures database · Valence · Arousal · Spatial frequency · Luminosity · Chromatic complexity

Introduction

Experimental research has categorized the different stimuli employed in the study of emotion according to their physical nature. Within the visual modality, which is the perceptive modality most frequently employed, stimuli may be classified into symbolic (e.g., written emotional language, signs or simple drawings such as ASCII-based emoticons) and non-symbolic (e.g., Carretié et al. 2013; Frühholz et al. 2011). Among the later, a distinction has been made between emotional facial expressions and those images in which the emotion is not produced by faces, which from now on will be referred to as "emotional scenes". Studies on emotional perception support this segregation through at least two lines of evidence. First, these two categories have each one its own affective frame. For example, emotional faces are usually evaluated and manipulated within a discrete frame (i.e., emotions are studied as a relative numerous list of discrete affective states, such as happiness, anger, fear, etc.; e.g., Ekman 1992), whereas scenes are more frequently assessed and manipulated within a dimensional frame (i.e., emotional processes are studied through a reduced number of subjacent dimensions, usually Valence—unpleasant to pleasant- and Arousal—calming to arousing-; e.g.: Russell 1979). Additionally, even when both faces and scenes are evaluated within a common dimensional frame, scenes can reach more extreme Valence and Arousal values than facial expressions: Thom et al. 2014). Second, facial expressions and scenes are processed by a differential neural circuitry, the former but not the latter involving face-specific brain areas (Haxby and Gobbini 2011), whose activity is reflected in specific temporal and spatial patterns in the main signals of neural activity (e.g., Carretié et al. 2013; Sabatinelli et al. 2011). Indeed, research-oriented databases of emotional faces and some others of emotional scenes have been created separately, as explained later.

Emotional scenes are especially interesting for studying affective processes for several reasons. First, as already explained, certain scenes may reach more extreme Valence and Arousal subjective values than facial expressions (Thom et al. 2014). Second, and in line with previous point, scenes evoke greater central and peripheral physiological indexes of emotional arousal (Mavratzakis et al. 2016; Thom et al.

L. Carretié and M. Tapia contributed equally to this work.

M. Tapia
manuel.tapia@uam.es

[1] Facultad de Psicología, Universidad Autónoma de Madrid, Madrid, Spain

2014). At the same time, this kind of emotional stimuli allow for greater "resolution" in Valence and Arousal. In other words, scenes may represent many different degrees of Arousal and Valence. Consequently, emotional scenes are recurrently employed as stimuli in studies on affective processes.

Emotional scene databases that additionally include normative ratings on the affective content of each stimulus are an advisable and widely used tool in experimental and clinical research since these ratings allow researchers to select stimuli using quantitative criteria usually based on big samples of evaluators. Surprisingly, the number of emotional scene databases is conspicuously scarce as compared to those of facial expressions. Thus, while there exist more than 40 facial expression databases (some of most known are KDEF—Lundqvist et al. 1998; NimStim—Tottenham et al. 2009; and FACES—Ebner et al. 2010; but see an exhaustive list in www.face-rec.org/databases/), to the best of our knowledge only four emotional scene databases exist up to the moment (we will refer here to "general" scene databases; note that others exist focused on specific emotions, such as SFIP—Michałowski et al. 2017 or DIRTI—Haberkamp et al. 2017, consisting of fear- and disgust-related scenes, respectively).

Half a century ago, the Center for the Study of Emotion and Attention developed the first set of affective pictures for research, the International Affective Picture (IAPS) database (Lang et al. 2005). Concretely, first reference to IAPS in a scientific report was made in 1970 (Verschoor and van Wieringen 1970), to the best of our knowledge. Initially consisting of emotional slides, IAPS has been digitized and regularly updated with new pictures and additional participants in their assessment samples. Currently, IAPS (http://csea.phhp.ufl.edu/Media.html#topmedia) provides 1195 pictures and their normative ratings, including Valence and Arousal mean and dispersion values based on the assessments of wide samples of participants (100 each picture).

IAPS is, also nowadays, the most employed affective picture database in affective research[1] and is, without any doubt, one of the most methodologically useful tools in this field. However, several reasons pull to generate some additional databases. First, the fact that IAPS was the only emotional picture database available for researchers has produced some familiarization or habituation effect, particularly when the same participant takes part in different studies regarding emotional stimulation (Dan-Glauser and Scherer 2011). In relation to this, it is often necessary to complete, adding new pictures, some motifs insufficiently covered by IAPS (e.g., spider pictures; Dan-Glauser and Scherer 2011). Two additional elements we would like to stress to complement IAPS pictures with new databases in affective research are related to cultural issues. On the one hand, a cultural-temporal element: several pictures of IAPS depict issues such as clothes, vehicles, furniture, hairstyles, etcetera, from decades ago, and this may lead to discard these stimuli in several circumstances (i.e. experimental settings where old fashioned objects can be considered particularly striking). On the other hand, there is a cultural–geographical element: despite globalization has led to an evident cross-national uniformization, mainly in western countries, there is an unequivocal "typical USA look" in several IAPS pictures that reflect idiosyncratic contextual variables that may introduce some interference that could need to be controlled in certain circumstances. Whereas the temporal context is difficult (or impossible) to avoid, pursuing a wider geographical representation may be of interest in affective research. In relation to this, a recent comparison of IAPS adaptations in eight countries shows important coincidences in Valence and Arousal ratings, but also significant differences between Asian and Western samples and among different Western countries (Soares et al. 2015).

Therefore, the necessity of complementing IAPS pictures with new databases has been increasingly evident, as affective research has grown. Three other databases, also open-access (as IAPS) to the scientific community, have been created in the twenty-first century. One of them is the Geneva affective picture database (GAPED; Dan-Glauser and Scherer 2011), which was developed by the Swiss Center for Affective Sciences and has 730 pictures. Secondly, the Nencki Affective Picture System (NAPS; Marchewka et al. 2014), was created in the Nencki Institute of Experimental Biology with 1356 photographs. The third one is the Open Affective Standardized Image Set (OASIS), designed by Kurdi, Lozano and Banaji at Harvard University (Kurdi et al. 2016) with 900 pictures. As IAPS, these databases provide both the digitized pictures and the Valence and Arousal normative ratings, among other information.

A critical handicap of emotional scenes as compared to other visual stimuli such as faces or even words is that they present a considerable variability regarding low order visual parameters such as luminosity (brightness), spatial frequency or chromatic complexity. Indeed, these parameters are practically invariant across different emotional categories of faces (mainly when they belong to the same models) or words, but may present conspicuous differences in the case of scenes. Importantly, these low level parameters are critical in Affective Sciences, since they have been revealed to modulate the behavioral and/or neural response to emotional pictures. For example, brighter images are

[1] A search carried out in February 2019 in Google Scholar introducing, enclosed in double quotes, the complete name of all emotional scene databases described here (from 2016, the year in which the most recent was created, to present) yielded 4380 results for IAPS and less than 200 for each of the rest.

Table 1 Number (total, male and female) and age (mean and SD) of participants, and number of pictures assessed in each evaluation since 2006

Nr.	Year	N (male/female)	Mean age (SD)	N pictures
1	2006	17 (3/14)	20.76 (1.30)	224
2	2006	74 (25/49)	19.74 (3.49)	150
3	2007	35 (6/29)	21.77 (4.12)	160
4	2008	44 (4/40)	21.66 (3.15)	160
5	2008	38 (17/21)	19.66 (2.62)	160
6	2008	48 (4/44)	19.48 (4.16)	160
7	2010+	86 (10/74)	19.76 (4.02)	180
8	2011	24 (3/19)	19.18 (3.70)	150
9	2012+	184 (33/148)	19.78 (2.88)	150
10	2014	37 (5/32)	19.68 (1.23)	115
11	2017+	103 (33/70)	20.97 (7.23)	200
12	2017+	85 (32/53)	19.80 (2.07)	140
Total		768 (175/593)	20.19 (3.33)	813[a]

Evaluations marked "+" were not conducted in a unique session: Evaluation Nr. 7 = 4 sessions; Evaluation Nr. 9 = 7 sessions; Evaluation Nr. 11 = 3 sessions; Evaluation Nr. 12 = 2 sessions

[a]Some pictures were assessed in more than one evaluation, in order to balance sample sizes evaluating each picture

evaluated as more positive (Lakens et al. 2013), greater complexity elicits more positive event-related potentials at occipital areas in response to emotional scenes (Bradley et al. 2007), and a typical selection of emotional scenes shows greater spectral density than neutral in low spatial frequency bands (Delplanque et al. 2007).

In other words, behavioral or neural responses may differ among stimulus categories (e.g., negative, positive and neutral) not (or not only) because their affective content, but because they differ in spatial frequency, luminosity or chromatic complexity. The lack of information on these parameters in the databases mentioned above, except in NAPS (which includes luminance and complexity details of the photographs, but not of the spatial frequencies), may hinder emotional scenes from being even more employed in experimental research and/or the clinical practice regarding affective processes.

An additional issue worth to be considered in an emotional scene database is stimulus duration. Current databases use relatively long exposure times during stimulus evaluation (IAPS 5 s, GAPED 4 and NAPS 3; in OASIS, stimuli were self-paced by evaluators). Taking into account that stimulus duration influences affective reactions (e.g., Codispoti et al. 2009), there is a lack of databases approaching to the numerous behavioral and electro/magnetophysiological studies that employ stimulus durations equal to, and shorter than, 1 s (e.g., 59.6% of event-related potential studies carried out from 1966 to 2008 using emotional scenes as stimuli employed durations ≤ 1 s, and total average duration was < 2 s: see Table 1 in the review by Olofsson et al. 2008).

Thus, the increasing relevance of the study of emotion in Neuroscience, and in other fields of Psychology, Psychiatry, Philosophy, and others, demand the introduction of additional databases for the reasons exposed above: avoiding habituation, providing enough stimuli of particular motifs, widening and adapting temporal, geographical and cultural contexts avoiding any bias towards a particular country or nationality, using shorter stimulus exposition in evaluation sessions and, importantly, providing information on low order visual parameters. The complementariness and synergy of combining these databases may only enrich the study of affective processes.

The research group "Brain, Affect and Cognition" ("Cerebro, Afecto y Cognición"—CEACO) has developed EmoMadrid during several years. EmoMadrid is conceived as an affective pictures database for scientific research, and is available at www.psicologiauam.es/CEACO/EmoMadrid.htm since 2012 (November 6). Apart from the information included in its web page, EmoMadrid has not been described yet in any scientific paper nor scientific meeting. Despite this, up to date, 184 researchers from different universities and research centers have requested permission to use EmoMadrid; in fact, it has already been employed in 15 published studies from 9 different laboratories (Andreu et al. 2018; Anil Kumar et al. 2015; Carboni et al. 2017; Carretié et al. 2011; 2013; Kosonogov et al. 2016; Jun et al. 2018; López-Martín et al. 2015; Mel'nikov et al. 2018; Román et al. 2015; Romero-Ferreiro et al. 2018; Ruiz-Padial et al. 2018; Xu et al. 2015; 2016a, b). In some studies, EmoMadrid pictures are used along with pictures from other image sets, demonstrating also the interest of combining them (e.g., Anil Kumar et al. 2015; Kosonogov et al. 2016; Ruiz-Padial et al. 2018; Xu et al. 2015).

Method

Participants

Since year 2006 until 2018, 768 undergraduate Psychology student volunteers (693 women, 175 men), with an age between 17 and 50 years (mean = 20.11; SD = 3.38), have participated in 12 evaluations. Each participant assessed an average of 155 pictures, and received an academic reward for his/her collaboration. Specific data for each evaluation are collected in Table 1.

Materials

Up to date, a total of 813 digital photographs depicting scenes with diverse affective content have been included to set up the database. A first selection of the photographs was made by the authors for each assessment session, so the images met some minimum quality criteria. This first selection also pursued to obtain an ample variety of motifs and categories. Thus, these scenes represent animals (14.62%), objects (20.15%), food (10.93%), landscapes (15.85%), and humans as protagonists (38.45%). Since humans also appear, at least partially, in some pictures of other categories in which they are not protagonists, they are finally present in 44.23% of EmoMadrid pictures. Different valence levels (negative, positive and neutral), as well as different arousal levels (arousing, calming and neutral), are present in each of these categories. Some examples of animals are dangerous animals, pets, anodyne birds; objects: hazardous objects, relax-related items, neutral furniture; food: contaminated food, appetitive desserts, vegetables; humans: erotica, wounds, or people walking on the street. All pictures are 1024×768 pixels size, and have been lend to the project by amateur and professional collaborator photographers, or have been collected from public internet sites by CEACO members.

Luminosity, chromatic complexity and spatial frequency, which, as explained above, are relevant parameters in affective reactions to emotional visual stimuli, were calculated for each picture. Regarding *luminosity*, average luminosity of each picture was computed through the Adobe Photoshop's© Histogram tool. Luminosity ranges from 0 (complete black) to 255 (complete white). Analysis on luminosity was carried out on pictures, including the vertical or the horizontal black margins some of them need as mentioned above for the Spatial frequency.

For *chromatic complexity*, JPG size in Kilobytes was used as a measure of chromatic complexity. In brief, JPG (or JPEG) format provides file sizes which are proportional to color changes from one pixel to the neighbor one. Thus, areas of the image representing the same flat color are codified using less digital information than areas including color changes. As a consequence, JPG sizes may be employed as an index of chromatic complexity of pictures (e.g., Müller et al. 2008; Marchewka et al. 2014; please note that this size–complexity correlation does not occur in other image file formats).

With respect to *spatial frequency*, spectral energies were computed for nine frequency bands (768-384 pixels/cycle or p/c, 384-192 p/c, 192-96 p/c, 96-48 p/c, 48-24 p/c, 24-12 p/c, 12-6 p/c, 6-3 p/c and residuals) within each picture following the procedure described by Delplanque et al. (2007), in which the gray 709 option was selected. We recommend the reader to visit http://www.affective-sciences.org/spatfreq for details on this procedure. Analyses on spatial frequencies were carried out on all pictures, including the black margins some of them needed in the vertical or the horizontal dimension to fit the 1024x768 pixel format.

Procedure

Participants were recruited in their own study centers yearly since 2006, via either institutional email or advertisements posted on their university or in class. Before each evaluation session, all participants were informed about the objective of the study and their rights as volunteer participants. Within each session, participants were instructed to assess pictures on both Valence and Arousal scales, attending to their first personal impression. Importantly, instructions—provided in EmoMadrid web page—were always given in person so participants could ask for further directions if needed. Additionally, the face to face format during evaluation sessions allowed for the monitoring on participants' involvement in the task, avoiding the distractibility often observed in other situations with no responsible instructor present, such as online or self-administered procedures. In all evaluation sessions, pictures were presented in environments of dim light and silence, using Power Point presentation software displayed through an RGB projector. The projection screen size and position allowed its optimal visualization from any seat in the room.

The average number of pictures presented in each evaluation session was 155 (maximum: 224; minimum: 75; standard deviation: 38). The pictures set varied partially or totally from year to year in order to balance the cumulated sample sizes evaluating each picture. Also, the picture sets were selected in such a way that the number of arousing negative, arousing positive, neutral and relaxing stimuli were balance. Additionally, with the aim of avoiding possible anchoring effects (i.e., biased responses influenced by previous responses) the order of pictures was semi-randomized so no more than three consecutive scenes of the same affective category appeared.

Pictures were presented twice so only Valence or only Arousal was assessed in each run. The order in which the Valence run and the Arousal run was assigned to each participant was counterbalanced, so it was balanced for each year's sample. The aim of this strategy -two runs to assess one dimension each time—was to avoid the "halo effect" (Saal et al. 1980; multidimensional assessments to the same item tend to be similar and to show inflated intercorrelations), which potentially affect assessments when Valence and Arousal are simultaneously evaluated. Five points Likert scales were used to rate Valence and Arousal dimensions. In the case of Valence the scale levels were *Muy negativo–Algo negativo–Ni negativo ni positivo–Algo positivo–Muy positivo* (Very negative, negative, neutral, positive, very positive). With respect to Arousal, levels were *Muy relajante–Algo relajante–Ni relajante ni activante–Algo activante–Muy activante* (Very calming, calming, neutral, arousing, very arousing). Both scales were operationalized as -2 to 2 scales, zero being the neutral value. Assessments were collected through pencil and paper questionnaires until 2008, and through electronic forms in individual computers from that date, while pictures were presented through a projection screen in the classroom.

Each picture appeared during 1 s and the inter trial interval (ITI) was 4 s, within which subjects had to assess Valence or Arousal of the picture just presented. This exposition time, shorter than the employed in other databases (three seconds minimum),[2] was chosen to approach to the numerous behavioral and electro/magnetophysiological studies that employ stimulus durations equal to, and shorter than, 1 s, as indicated in the Introduction. The four-second ITI was tested as enough to fill-in the single-dimension questionnaire and ensured reasonably short evaluation sessions (45′ maximum after both Valence and Arousal runs). Groups of 44 volunteers maximum and 17 minimum participated in each evaluation session. Assessments were recorded in the Faculty of Psychology of the Universidad Autónoma de Madrid, in the Faculty of Health Sciences of the Universidad Rey Juan Carlos, and in the Faculty of Psychology of the Universidad de Jaén by members and collaborators of CEACO group.

Results

Results of the assessments made since 2006 have been being included yearly in our database, along with the new pictures added to the sets in each new evaluation. EmoMadrid web page (www.psicologiauam.es/CEACO/EmoMadrid.htm) contains all the group data from the assessments (i.e., averaged across participants) provided to each picture, as well as the results of the analyses made by the authors on some characteristics of the images (individual assessments to each picture have been uploaded to the Open Science Framework—OSF-repository: https://osf.io/up7ch/). The *index* section of the web connects to a spreadsheet which contains each picture's name (EmoMadrid Code) and thumbnail (that allow for a rapid preview of the images), and provides the following information: (i) Valence-related information for each picture (sample size evaluating this dimension, mean score, and standard deviation), (ii) Arousal-related information (the same information as for Valence), (iii) luminosity, and (iv) complexity (JPG size), and (v) spatial frequency in nine different frequency bands. Information on these low order parameters (iii–v), whereas not included in previous databases, was considered crucial since spatial frequency, luminosity and complexity has been found to modulate affective effects of visual stimuli (Delplanque et al. 2007; Lakens et al. 2013; Kuchinke et al. 2011, respectively), as already indicated. Next, details on how these five types of information were measured is presented. Finally, a selection of negative, positive and neutral stimuli showing non-significant differences in low order visual parameters, named EMselec, is described.

Valence and Arousal

Information provided on Valence includes the sample size (762 participants, 594 women and 171 men; average sample evaluating Valence of each picture: 146, standard deviation: 47.73, minimum/maximum sample evaluating each picture: 85/256, minimum men sample size: 15), mean (−2, very negative, to 2, very positive; 0 being the neutral value) and standard deviation. The same parameters are provided for Arousal sample size (768 participants, 593 women and 175 men; average sample evaluating Arousal of each picture: 147, standard deviation: 49.06, minimum/maximum: 84/257, minimum men sample size: 15), mean (−2, very calming, to 2, very arousing; 0 is the neutral value) and standard deviation. A cell-color code helps to rapidly and visually interpret Valence and Arousal values (blue intensity increasing towards −2 and red increasing towards 2; light red/blue, or white, meaning neutrality).

Reliability of Valence and Arousal scores was analyzed via Pearson correlation for Valence and Arousal of those pictures which have been evaluated twice (i.e., were evaluated by different participants in different years), which are, at this moment, 369. We computed two types of reliability, one for intervals shorter than 5 years between the first and the second evaluation ("short term reliability" hereafter), and the other for intervals of 5 up to 11 years, the current possible maximum ("long term reliability"). Short term reliability ($n = 449$, provided that some pictures were evaluated more than twice and that, in these cases, all pairwise test–retest combinations were analyzed) is 0.969 for Valence and 0.957 for Arousal ($p < 0.001$ in both cases). Long term reliability ($n = 242$) is 0.964 for Valence and 0.922 for arousal ($p < 0.001$ in both cases). In relation to this, we also analyzed whether Valence and Arousal ratings varied across time (both in the Short and Long terms described above) through Student's t tests for independent samples. Results yielded non-significant differences (Valence Short: $t(896) = 0.531$, $p = 0.595$; Arousal Short: $t(896) = -0.227$, $p = 0.820$; Valence Long: $t(482) = 0.600$, $p = 0.549$; Arousal Long: $t(482) = 1.562$, $p = 0.119$).

Two additional analyses were carried out to characterize EmoMadrid's Valence and Arousal scores for the total set of pictures (813, currently). First, gender differences were evaluated through two one-way ANOVAs, one on average Valence and the other on the average Arousal assessments provided by men versus women. Results

[2] IAPS: 5 s. GAPED: 4 s. NAPS: 3 s. In OASIS, stimulus duration is self-paced by evaluators.

yielded non-significant differences in the case of Valence ($F(1,1624)=0.053$, $p=0.817$), but reached significance in the case of Arousal ($F(1, 1624)=3.956$, $p=0.047$), women judging EmoMadrid pictures as more arousing than men. A subsequent analysis was carried out on Arousal assessments to further characterize this effect by dividing pictures into positive (Valence over 0) and negative (Valence under 0), and showed that gender-related Arousal differences were circumscribed to negative pictures ($F(1,527)=18.141$, $p<0.001$), women assessing them as more arousing than men (results por positive pictures: $F(1,1097)=1.129$, $p=0.288$).

The second additional analyses consisted of comparing the pencil and paper method to record Valence and Arousal assessments employed during the first two years (see "Methods" section) with the electronic form method employed from 2008. ANOVAs where carried out on Valence and Arousal of those pictures evaluated under both formats ($n=369$). Differences were not significant for any of both dimensions ($F(1,737)=1.250$, $p=0.264$, and $F(1,737)=0.534$, $p=0.465$, respectively).

Valence × Arousal scatterplot

Along with the Index, EmoMadrid provides an interactive scatterplot where the user may visualize the location of each EmoMadrid picture in the Valence × Arousal space. Whereas numerical information on both dimensions is also provided in the Index, the scatterplot allows for rapid and visual selection of pictures according to their values in both dimensions. Thus, hovering in the web version over each circle (i.e., each stimulus), a legend opens showing the picture code and its Valence and Arousal average values. Figure 1 shows the distribution of stimuli in the scatterplot, but without the information on each stimulus provided in the online version (www.psicologiauam.es/CEACO/EmoMadrid.htm, Section 2, or directly through this link: goo.gl/XW73nG). This online scatterplot also shows at a glance the distribution of the main motifs (animals, objects, landscapes, humans and food, which are coded through different colors), revealing, for example, that the low arousal, high valence quadrant presents a relatively higher number of "landscapes", or that the areas of maximal arousal (both positive and negative) present a relatively high number of "humans".

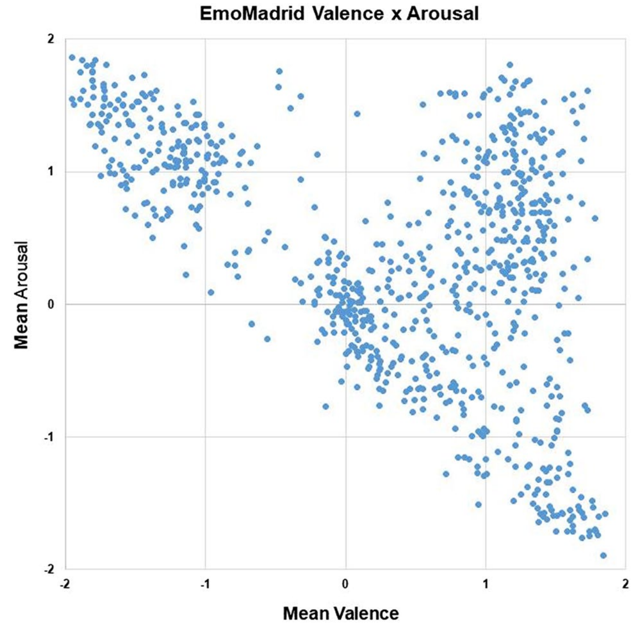

Fig. 1 Scatterplot representing average Valence and average Arousal ratings provided to each picture by the experimental samples described in the main text. In the internet version (www.psicologiauam.es/CEACO/EmoMadrid.htm, Section 2, or directly in this link: goo.gl/XW73nG), a legend opens hovering over each circle showing the picture code, Valence and Arousal average values, and motif (animal(s), object(s), landscape, human(s) or food)

Correlations between affective and visual parameters

In order to analyze whether average Valence and Arousal provided by participants to EmoMadrid pictures were associated with spatial frequency, luminosity, or complexity, bivariate Pearson correlations were carried out. Table 2 shows the main statistics yielded by these analyses. Firstly, and as previously reported with respect to other picture databases, Valence and Arosual show a significant correlation (e.g., Bradley and Lang 2000). Concretely, Arousal increases as Valence decreases (i.e., as it becomes more negative). Secondly, both Luminosity and JPG size (complexity) significantly correlate with Arousal, but not with Valence. Thus, Arousal correlates positively with JPG size, and negatively with Luminosity. Finally, spatial frequency presents a significant correlation with both Valence and Arousal in different frequency bands. In general, this correlation is positive and circumscribed to the lowest frequency bands in the case of Valence. With respect to Arousal, its correlation with spatial frequency is dual: negative with the lowest frequency bands and positive with the highest (Table 2). Figure 2 graphically shows the relationship between each affective and visual parameter through scatterplots.

EMselec: a selection of negative, positive and neutral stimuli yielding non-significant spatial frequency, luminosity, and complexity differences

In order to minimize the affective–physical interactions described in previous section, and as an additional tool, EmoMadrid provides a zip-compressed folder named EMselec which contains a possible selection (among many others) of 40 negative, 40 neutral and 40 positive pictures in which

Table 2 Main statistics resulting from bivariate Pearson correlation analyses introducing average Valence, average Arousal, spectral density of the nine frequency bands analyzed, luminosity and JPG size (color complexity): r (correlation coefficients), *p* (probability)

	Valence	Arousal
Valence		
r	1.000	−0.493
p		**0.000**
N	813.000	813.000
Arousal		
r	−0.493	1.000
p	**0.000**	
N	813.000	813.000
SF 768-384		
r	0.076	−0.141
p	**0.031**	**0.000**
N	810.000	810.000
SF 384-192		
r	0.091	−0.094
p	**0.009**	**0.007**
N	813.000	813.000
SF 192-96		
r	0.067	−0.076
p	0.055	**0.031**
N	813.000	813.000
SF 96-48		
r	0.073	−0.025
p	**0.036**	0.472
N	813.000	813.000
SF 48-24		
r	0.083	0.056
p	**0.017**	0.112
N	813.000	813.000
SF 24-12		
r	0.067	0.098
p	0.056	**0.005**
N	813.000	813.000
SF 12-6		
r	0.067	0.112
p	0.055	**0.001**
N	813.000	813.000
SF 6-3		
r	0.049	0.120
p	0.166	**0.001**
N	813.000	813.000
SF residual		
r	0.018	0.076
p	0.605	**0.030**
N	813.000	813.000
Size		
r	−0.049	0.057
p	0.164	**0.103**
N	813.000	813.000
Luminance		
r	0.063	−0.131
p	0.071	**0.000**
N	813.000	813.000

Note that the 768-384 p/c (pixels/cycle) and the 8-3 p/c bands correspond to the lowest and highest SFs, respectively

Probability (*p*) values are highlighted in bold when significant

SF spatial frequency

visual parameters do not yield significant differences. Thus, these pictures have been selected attending to the following criteria, which were tested through one-way ANOVAs using Stimulus Category as factor (3 levels, Negative, Neutral and Positive) and post hoc pairwise comparison tests (Tukey's HSD). First, negative and positive images do not differ in Arousal but both groups significantly differ from neutral (negative-neutral-positive averages in Arousal form a "V" shape: Fig. 2): $F(2,117) = 171.767$, $p < 0.001$. Second, Valence significantly differs among the three groups of stimuli (negative-neutral-positive averages in Valence form a linear ascending slope: Fig. 3): $F(2,117) = 1667.407$, $p < 0.001$. Third, differences in spatial frequencies among negative, neutral and positive categories are non-significant with respect to the nine frequency bands ($F(2,117) < 1.2$ in all cases, $p > 0.3$ in all cases; see Table 3). Fourth, differences in luminosity among negative, neutral and positive categories are non-significant ($F(2,117) = 0.156$, $p = 0.856$). And fifth, differences in chromatic complexity (JPG size) among negative, neutral and positive categories are non-significant ($F(2,117) = 0.454$, $p = 0.636$).

Details on statistical contrasts on Valence, Arousal, Spatial frequency, Luminosity and Complexity regarding this selection of 40 negative, 40 neutral and 40 positive EmoMadrid pictures, are available here: http://www.psicologiauam.es/gruposinv/ceaco/EmoMadrid/EMestadistica.htm.

Discussion

EmoMadrid currently provides information beyond other databases, such as spatial frequency, luminosity, color complexity, source of each picture, main motif depicted and the presence or not of humans (along with "traditional" Valence and Arousal measures). However, four lines of future development are scheduled in the short and middle term to further reinforce this picture database. First, it is under scopes to gradually increase the number of pictures in order to complete some areas of the Valence × Arousal plane less covered

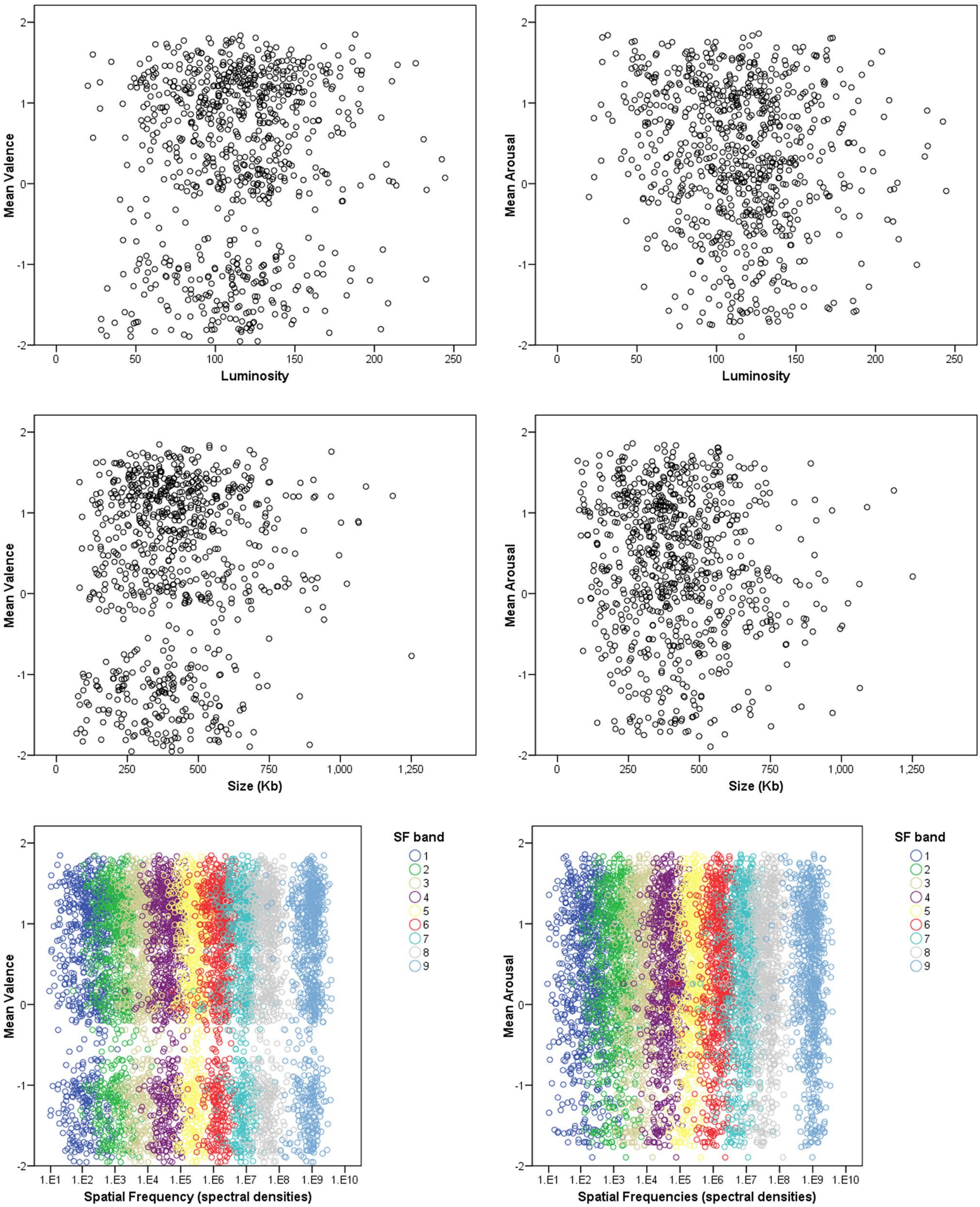

Fig. 2 Scatterplots depicting the relationship between each affective parameter (Valence left, Arousal right) and each visual parameter. *SF* spatial frequency. SF bands from the lowest to the highest: (1) 768-384 pixels/cycle or p/c, (2) 384-192 p/c, (3) 192-96 p/c, (4) 96-48 p/c, (5) 48-24 p/c, (6) 24-12 p/c, (7) 12-6 p/c, (8) 6-3 p/c and (9) residuals. In spatial frequency scatterplots, abscissa is in logarithmic scale

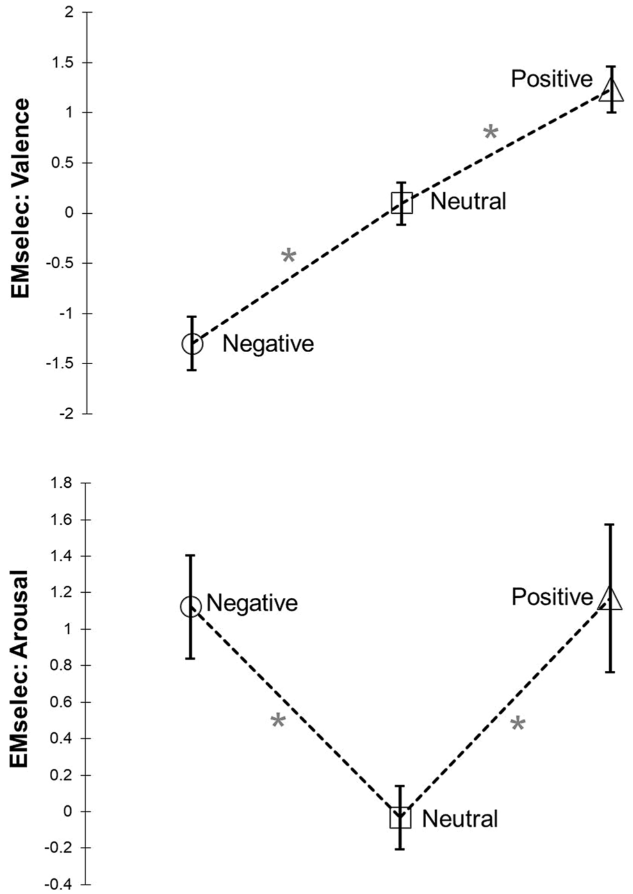

Fig. 3 Means and standard deviations (error bars) of Valence and Arousal assessments provided to stimuli included in EMselec, a selection of 38 negative, 38 neutral and 38 positive EmoMadrid pictures (*p < 0.001 in pairwise Tukey's HSD post hoc comparisons)

Table 3 EMselec: main statistics resulting from one-way ANOVAs for each spatial frequency band introducing Stimulus Category (three levels: negative, neutral, and positive, 40 pictures each) as factor (p: probability). *SF* spatial frequency

	F(2,117)	p
SF 768-384	0.867	0.423
SF 384-192	0.155	0.857
SF 192-96	0.310	0.734
SF 96-48	0.627	0.536
SF 48-24	0.622	0.539
SF 24-12	0.629	0.535
SF 12-6	1.146	0.321
SF 6-3	0.505	0.605
SF residual	0.102	0.903

and to balance, as far as possible, the distribution of motifs (e.g., increasing the number of negatively valenced foods, which are unbalanced with respect to positively valenced foods in current EmoMadrid version). However, it is important to stress at this respect that the four quadrants of the Valence × Arousal plane are not equally filled by real-life stimuli; for example, negative-calming stimulation is rare (and theoretically improbable). Second, it is also our intention trying to increase evaluation samples to a minimum of 100 hundred people per picture, maintaining the face-to-face (non-online) scheme, in order to increase the reliability of assessments. This data collection procedure may complement other non-presential methodologies employed in currently available emotional scene databases, which may help researchers to decide which one is more convenient to their purposes. A related scope is trying to gradually balance the male–female proportion of participants, currently biased towards the latter. At this respect, our analyses reveal that significant gender differences exist regarding Arousal (but not Valence) assessments to negative (but not positive) EmoMadrid pictures. These gender differences are narrower than those found in the assessments of pictures from other databases (which extend to Valence and to Arousal assessments to positive pictures: e.g., Bradley et al. 2001), but may vary in the future as sample sizes are gradually increased.

Third, EmoMadrid should take into account other cultural and age groups. In the former respect, EmoMadrid scores have only been obtained from volunteers living in Spain. Whereas Spanish samples show, essentially, similar assessments of emotional pictures than in other Western countries (Moltó et al. 1999; Vila et al. 2001), cultural differences in the evaluation of the same scenes are undeniable, as explained in the Introduction (see a review in Soares et al. 2015). EmoMadrid is now open to include assessments from other international laboratories in order to increase its cultural coverage (www.psicologiauam.es/CEACO/EmoMadrid.htm, Section 1). In the latter respect (age), our assessing samples were relatively young, but age-related changes in emotional processing have been reported (e.g., Charles et al. 2001; Leigland et al. 2004; Perlman and Pelphrey 2001; Urry and Gross 2010), so our normative ratings should also include older adults in future samples. And fourth, future evaluation sessions will also ask on the discrete emotion(s) considered by participants to be involved in pictures, which may be of interest for researchers interested in specific affective categories.

The introduction of visual parameters on EmoMadrid and the analyses on how they correlate with affective parameters are of especial interest, since they reveal significant correlations indexing mutual physical–affective relationship. Thus, Arousal has shown to correlate with color complexity and with luminosity, and both Arousal and Valence correlate with different spatial frequency bands. These correlations between physical and affective parameters, which are also found in other databases (e.g., Delplanque et al. 2007), are methodologically crucial for researchers, who may neutralize them using the information provided in EmoMadrid to manipulate and control these parameters during picture selection. They also stress the relevance of designing sets of pictures varying in their affective content but not in their physical parameters, such as EMselec (which is only one of the possible sets accomplishing these criteria).

Finally, we summarize some recommendations to increase the usefulness of EmoMadrid (and other databases) in research on affective processes. First, Valence and Arousal normative ratings provided in the Index section should be considered as an orientative a priori criterion to select pictures. Indeed, these ratings cannot be used to replace those that should be collected in each study from their actual experimental samples. It is strongly recommended that each experimental participant evaluates the selected pictures at the end of the experiment trying to mimic the actual experimental parameters (stimulus duration, stimulus distance, etc.) since, as explained, may modulate affective reactions. Finally, and in order to avoid the ''halo effect'', previously defined (Saal et al. 1980), the possibility of asking for the assessment of Valence and Arousal in different runs (and counterbalance their order) should be considered.

Acknowledgements We wish to thank Elisabeth Ruiz Padial and Francisco Mercado for their help in the assessment of EmoMadrid pictures at the Universidad de Jaén and Universidad Rey Juan Carlos, respectively. We also acknowledge the disinterested contribution made by the photographers Galder Izaguirre, Víctor Nogales, Óscar Rodrigo and Jordi García-Pons.

Funding This work was supported by the FEDER/Ministerio de Ciencia, Innovación y Universidades (PGC2018-093570-B-I00) and by the Comunidad de Madrid (S2015/HUM-3327).

Compliance with ethical Standards

Conflict of interest The authors declared that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

Andreu, C. I., Cosmelli, D., Slagter, H. A., & Franken, I. H. (2018). Effects of a brief mindfulness-meditation intervention on neural measures of response inhibition in cigarette smokers. *PLoS ONE, 13*(1), e0191661.

Anil Kumar, K. M., Kiran, B. R., Shreyas, B. R., & Victor, S. J. (2015). A multimodal approach to detect user's emotion. *Procedia Computer Science, 70,* 296–303.

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion, 3,* 300–319.

Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology, 44,* 364–373.

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 49–59). Oxford: Oxford University Press.

Carboni, A., Kessel, D., Capilla, A., & Carretié, L. (2017). The influence of affective state on exogenous attention to emotional distractors: Behavioral and electrophysiological correlates. *Scientific Reports, 7,* 8068.

Carretié, L., Kessel, D., Carboni, A., López-Martín, S., Albert, J., Tapia, M., et al. (2013). Exogenous attention to facial vs nonfacial emotional visual stimuli. *Social Cognitive and Affective Neuroscience, 8,* 764–773.

Carretié, L., Ruiz-Padial, E., López-Martín, S., & Albert, J. (2011). Decomposing unpleasantness: Differential exogenous attention to disgusting and fearful stimuli. *Biological Psychology, 86,* 247–253.

Charles, S. T., Reynolds, C. A., & Gatz, M. (2001). Age-related differences and change in positive and negative affect over 23 years. *Journal of Personality and Social Psychology, 80*(1), 136.

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology, 46,* 731–738.

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance. *Behavior Research Methods, 43,* 468–477.

Delplanque, S., N'diaye, K., Scherer, K., & Grandjean, D. (2007). Spatial frequencies or emotional effects? A systematic measure of spatial frequencies for IAPS pictures by a discrete wavelet analysis. *Journal of Neuroscience Methods, 165,* 144–150.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42,* 351–362.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6,* 169–200.

Frühholz, S., Jellinghaus, A., & Herrmann, M. (2011). Time course of implicit processing and explicit processing of emotional faces and emotional words. *Biological Psychology, 87,* 265–274.

Haberkamp, A., Glombiewski, J. A., Schmidt, F., & Barke, A. (2017). The DIsgust-RelaTed-images (DIRTI) database: Validation of a novel standardized set of disgust pictures. *Behaviour Research and Therapy, 89,* 86–94.

Haxby, J. V., & Gobbini, M. I. (2011). Distributed neural systems for face perception. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 93–107). Oxford: Oxford University Press.

Jun, S., Lee, S. K., & Han, S. (2018). Differences in large-scale and sliding-window-based functional networks of reappraisal and suppression. *Science of Emotion and Sensibility, 21,* 83–102.

Kosonogov, V., Sánchez-Navarro, J. P., Martínez-Selva, J. M., Torrente, G., & Carrillo-Verdejo, E. (2016). Social stimuli increase physiological reactivity but not defensive responses. *Scandinavian Journal of Psychology, 57,* 393–398.

Kuchinke, L., Schlochtermeier, L., & Jacobs, A. M. (2011). Differences in the neural processing of emotional pictures and words are modulated by stimulus complexity. *Psychophysiology, 48,* S3.

Kurdi, B., Lozano, S., & Banaji, M. R. (2016). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods, 37,* 626–630.

Lakens, D., Fockenberg, D. A., Lemmens, K. P., Ham, J., & Midden, C. J. (2013). Brightness differences influence the evaluation of affective pictures. *Cognition and Emotion, 27,* 1225–1246.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Gainesville, FL: University of Florida.

Leigland, L. A., Schulz, L. E., & Janowsky, J. S. (2004). Age related changes in emotional memory. *Neurobiology of Aging, 25,* 1117–1124.

López-Martín, S., Albert, J., Fernández-Jaén, A., & Carretié, L. (2015). Emotional response inhibition in children with attention-deficit/hyperactivity disorder: Neural and behavioural data. *Psychological Medicine, 45,* 2057–2071.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces (KDEF)*. CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet.

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods, 46*(2), 596–610.

Mavratzakis, A., Herbert, C., & Walla, P. (2016). Emotional facial expressions evoke faster orienting responses, but weaker emotional responses at neural and behavioural levels compared to scenes: A simultaneous EEG and facial EMG study. *Neuroimage, 124,* 931–946.

Mel'nikov, M., Petrovskii, E., Bezmaternykh, D., Kozlova, L., Shtark, M., Savelov, A., et al. (2018). fMRI responses in healthy individuals and in patients with mild depression to presentation of pleasant and unpleasant images. *Bulletin of Experimental Biology and Medicine, 164,* 601–604.

Michałowski, J. M., Droździel, D., Matuszewski, J., Koziejowski, W., Jednoróg, K., & Marchewka, A. (2017). The set of fear inducing pictures (SFIP): Development and validation in fearful and nonfearful individuals. *Behavior Research Methods, 49,* 1407–1419.

Moltó, J., Montañés, S., Poy, R., Segarra, P., Pastor, M. C., Tormo, M. P., et al. (1999). Un nuevo método para el estudio experimental de las emociones: El International Affective Picture System (IAPS). Adaptación española [A new method for the experimental study of emotions: The International Affective Picture System (IAPS). Spanish adaptation]. *Revista de Psicología General y Aplicada, 52,* 55–87.

Müller, M. M., Andersen, S. K., & Keil, A. (2008). Time course of competition for visual processing resources between emotional pictures and foreground task. *Cerebral Cortex, 18,* 1892–1899.

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology, 77,* 247–265.

Perlman, S. B., & Pelphrey, K. A. (2001). Developing connections for affective regulation: Age-related changes in emotional brain connectivity. *Journal of Experimental Child Psychology, 108*(3), 607–620.

Román, F. J., García-Rubio, M. J., Privado, J., Kessel, D., López-Martín, S., Martínez, K., et al. (2015). Adaptive working memory training reveals a negligible effect of emotional stimuli over cognitive processing. *Personality and Individual Differences, 74,* 165–170.

Romero-Ferreiro, V., Aguado, L., Torío, I., Sánchez-Morla, E. M., Caballero-González, M., & Rodriguez-Jimenez, R. (2018). Influence of emotional contexts on facial emotion attribution in schizophrenia. *Psychiatry Research, 270,* 554–559.

Ruiz-Padial, E., Medialdea, M. M., del Paso, G. R., & Thayer, J. F. (2018). Individual differences in attentional capture by pictures of fear and disgust as indexed by cardiac responses. *Journal of Psychophysiology, 32*(4), 191–201.

Russell, J. A. (1979). Affective space is bipolar. *Journal of Personality and Social Psychology, 37,* 345–356.

Saal, F. E., Downey, R. G., & Lahey, M. A. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin, 88,* 413–428.

Sabatinelli, D., Fortune, E. E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W. T.,… & Jeffries, J. (2011). Emotional perception: meta-analyses of face and natural scene processing. *Neuroimage, 54*, 2524–2533.

Soares, A., Pinheiro, A., Costa, A., Frade, C., Comesaña, M., & Pureza, R. (2015). Adaptation of the international affective picture system (IAPS) for european portuguese. *Behavior Research Methods, 47,* 1159–1177.

Thom, N., Knight, J., Dishman, R., Sabatinelli, D., Johnson, D. C., & Clementz, B. (2014). Emotional scenes elicit more pronounced self-reported emotional experience and greater EPN and LPP modulation when compared to emotional faces. *Cognitive, Affective, & Behavioral Neuroscience, 14,* 849–860.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168,* 242–249.

Urry, H. L., & Gross, J. J. (2010). Emotion regulation in older age. *Current Directions in Psychological Science, 19,* 352–357.

Verschoor, A. M., & van Wieringen, P. C. W. (1970). Vigilance performance and skin conductance. *Acta Psychologica, 33,* 394–401.

Vila, J., Sánchez, M., Ramírez, I., Fernández, M. C., Cobos, P., Rodríguez, S.,… Moltó, J. (2001). El Sistema Internacional de Imágenes Afectivas (IAPS): Adaptación española. Segunda parte [The International Affective Picture System (IAPS): Spanish adaptation. Second part]. *Revista de Psicología General y Aplicada, 54*, 635–657.

Xu, M., Ding, C., Li, Z., Zhang, J., Zeng, Q., Diao, L., et al. (2016a). The divergent effects of fear and disgust on unconscious inhibitory control. *Cognition and Emotion, 30,* 731–744.

Xu, M., Li, Z., Ding, C., Zhang, J., Fan, L., Diao, L., et al. (2015). The divergent effects of fear and disgust on inhibitory control: An ERP study. *PLoS ONE, 10*(6), e0128932.

Xu, M., Li, Z., Fan, L., Sun, L., Ding, C., Li, L., et al. (2016b). Dissociable effects of fear and disgust in proactive and reactive inhibition. *Motivation and Emotion, 40,* 334–342.