Article

Subjective Affective Responses to Natural Scenes Require Understanding, Not Spatial Frequency Bands

Serena Mastria [ID], Maurizio Codispoti [ID], Virginia Tronelli [ID] and Andrea De Cesarei *[ID]

Department of Psychology, University of Bologna, Alma Mater Studiorum, Viale Berti Pichat 5, 40127 Bologna, Italy; serena.mastria4@unibo.it (S.M.); maurizio.codispoti@unibo.it (M.C.); virginia.tronelli2@unibo.it (V.T.)
* Correspondence: andrea.decesarei@unibo.it

Abstract: It is debated whether emotional processing and response depend on semantic identification or are preferentially tied to specific information in natural scenes, such as global features or local details. The present study aimed to further examine the relationship between scene understanding and affective response while manipulating visual content. To this end, we presented affective and neutral natural scenes which were progressively band-filtered to contain global features (low spatial frequencies) or local details (high spatial frequencies) and assessed both affective response and scene understanding. We observed that, if scene content was correctly reported, subjective ratings of arousal and valence were modulated by the affective content of the scene, and this modulation was similar across spatial frequency bands. On the other hand, no affective modulation of subjective ratings was observed if picture content was not correctly reported. The present results indicate that subjective affective response requires content understanding, and it is not tied to a specific spatial frequency range.

Keywords: emotion; spatial frequencies; scene understanding

1. Introduction

Our perceptual system can effectively identify relevant events in the environment that are potentially rewarding, or, otherwise, represent threats, in order to take appropriate action to assess, approach, or avoid them. When people are asked to describe the affective state elicited by pictorial cues that depict motivationally relevant content (e.g., depictions of mutilation, injuries, or nude bodies), the affective dimensions of valence (i.e., pleasant vs. unpleasant) and affective arousal (i.e., intensity; [1–6]) capture most of the variance in the reports, supporting the view that affective states are organized by a limited number of continuous dimensions [3,7–10]. When large samples of stimuli representing different contents were rated, the resulting distribution of affective states in the two-dimensional space was consistent with the motivational perspective, i.e., judgments of pleasantness indicate which motivational system is active, and judgments of arousal indicate the intensity of motivational activation [6].

In real life, objects or individuals are often viewed in crowded or cluttered conditions, from different viewpoints or distances, and our visual system is surprisingly efficient in recognizing them. Previous studies have shown that the viewing of emotionally arousing scenes (pleasant and unpleasant) modulates subjective, autonomic, and cortical changes even under perceptually challenging conditions. For instance, emotional engagement is observed when participants viewed natural scenes presented in small size [11,12], in the periphery of the visual field [13], or even for a very brief duration [14]. Similarly, some studies investigated how the compositional features of pictures modulate emotional response; in doing so, scene composition was operationalized as the result of the presence of global features that span most of a picture's area (low spatial frequencies), local details that

cover small portions of the image (high spatial frequencies), and an organizing structure that informs about where each detail is (phase of the spatial frequency spectrum) [15]. When manipulating the compositional structure of natural scenes, i.e., by filtering pictures to contain only global or local information [16,17], or by scrambling the phase of the spatial frequency spectrum [18], it was observed that even filtered or phase-scrambled scenes can elicit an affective response. At the same time, the abovementioned studies demonstrated that the various indexes of emotional response (subjective, autonomic, and central) were modulated only if participants could discriminate among emotional picture contents. If participants could not categorize or discriminate picture content, no affective modulation of behavioral, cortical, autonomic, or subjective response was observed ("semantic primacy" hypothesis; [19,20]).

Several findings indicate that scene content understanding is a necessary condition for the engagement of appetitive and defensive motivational systems in the context of natural scene perception [18,21,22]. In particular, when the phase of the spatial frequency spectrum was scrambled, accuracy in a human/animal task dropped to chance level [18]; importantly, when phase scrambling was reduced, it was not until classification accuracy increased to over 80% that emotional pictures could elicit a modulation of cortical responses (Late Positive Potential, LPP; [4,23–26]). In a task prompting participants to view pictures in progressively less degraded versions and deciding when they had got the gist of the scene, LPP modulation was first observed beginning one step before the gist, suggesting a strong link between the decisional process involved in stimulus identification and the affective modulation of the LPP [21]. Moreover, in a paradigm that compared the speed of response to several categorization tasks on the same emotional and neutral pictures, affective judgements were shown to have greater latencies compared with basic, superordinate, and subordinate categorization [27]. Consistently, chronometrically examining the speed of semantic vs. affective judgments, it was observed that semantic judgments preceded affective ones by 120 ms [20]. Altogether, these results support the view that identification is a conditio sine qua non for affective response (e.g., [14,18,20–22,27]).

On the other hand, some studies suggest that the emotional content of the pictures may influence subjective emotional judgments even when participants do not achieve semantic understanding of scene content, suggesting that the elicitation of an affective experience operates independently of identification. Results supporting this view often come from studies employing backward masking, in which the temporal visual availability of a stimulus is limited and the retinal persistence is eliminated by the presentation of a visually salient image immediately after the presented scene [28,29]. However, results from subliminal presentation often failed to be replicated [30–32]. Other studies manipulated spatial frequency content and suggested that emotional responses are consequential to the analysis of low spatial frequencies, which is carried out independently of conscious analysis and identification [33,34] However, more pronounced responses at specific spatial frequency ranges can be consequential either to a preferential processing of affective contents at a specific spatial frequency range, or to a better identification. To disentangle these two possibilities, it is necessary to balance the processing efficiency of these frequency ranges. For instance, when stimuli are high- or low-pass filtered, an identification task can be used to inform on the semantic understanding of the presented scene. However, few studies to date have adopted this strategy [21,35–37].

What is the contribution of visual content and of scene understanding to the subjective affective response to natural images? Does affective response require semantic content identification to be elicited, or are they independent? The present study aimed to further examine the relationship between scene understanding and affective response while manipulating visual content. To this end, we created pictures that only contained specific, band-passed spatial frequencies (i.e., visual features of a certain range) and that could represent emotional (pleasant or unpleasant) or neutral contents. In terms of affective response, we focused on the subjective reports of affective arousal and hedonic valence. Moreover, after each picture, we asked participants to report in an open description what

they had identified in a scene (e.g., "a child playing with a ball"). The accuracy of these open descriptions was used to assess, independently of the spatial frequency presented, the semantic identification of natural scenes. If visual content is critical for affective response but identification is not, it is expected that the scene content and spatial bandpass filter will interact in modulating affective ratings of arousal and valence, regardless of semantic identification. On the other hand, if identification is critical, one might expect that the modulation of affective response is only observed when participants correctly report the content of the scene.

As a second aim, we were interested in the possible preferential role for low compared with high spatial frequencies. It has been speculated that low spatial frequencies (coarsest features of a scene) might be sufficient to elicit emotional responses [33,34,38–40]. Therefore, building on this hypothesis, the affective reaction to scenes that are presented in their coarsest details should be more pronounced, or more resistant to lack of identification, compared with scenes that are presented in their finest details. On the other hand, if identification is a critical factor for affective response, then similar effects should be observed for high and low spatial frequencies once equated for identification.

2. Method

2.1. Participants

A total of 15 participants took part in the present study (12 females, mean age = 22.06, SD = 1.39). The sample size was determined on the basis of a previous pilot sample of participants using G*Power 3.1 [41], with the parameters alpha = 0.05, power = 0.80, partial eta squared = 0.06 (medium effect according to [42]), and correlation among repeated measures = 0.75 to detect an interaction between spatial frequency band (5 levels) and picture content (3 levels). The power analysis indicated a minimum sample size of 8 participants. All participants had self-reported normal or corrected-to-normal visual acuity, and none of them reported current or past neurological or psychopathological problems. Informed consent was obtained from all participants. The experimental protocol conforms to the Declaration of Helsinki and was approved by the Ethical Committee of the University of Bologna.

2.2. Stimuli and Equipment

We selected a total of 90 pictures from various sources, including the International Affective Picture System (IAPS; [43]) and public-domain pictures available on the Internet. The following picture categories were selected: erotic couples (arousing, pleasant, N = 15); babies (arousing, pleasant, N = 15); indoor or outdoor people in daily activities (neutral, N = 30); attacks (arousing, unpleasant, N = 15); and injuries (arousing, unpleasant, N = 15). These pictures were homogeneous with respect to brightness and contrast (pixel intensity M = 128, SD = 72.11), non-blurriness of the foreground and background, number of portrayed people (between 1 and 3 persons per picture, with visible face), and picture area covered by the face of a single individual (percent of total picture size M = 10.76%, SD = 3.41%, to discard pictures in which people were too close or too far away from the viewer).

Each picture was filtered using a band-passed filter which was centered around the following spatial frequencies, which will be defined as F_0: 4, 13.5, 45.3, 152.2, 512 cpi. The bandpass filter was implemented as the combination of a low-pass and a high-pass filter, which passed all frequencies up to F_0 and reached zero at $F_0 \cdot 3$ for the low-pass filter, and passed all frequencies from the highest frequency down to F_0 and reached zero at $F_0/3$ for the high-pass filter. An example of degraded pictures participants viewed in the experiment is displayed in Figure 1A, showing the non-degraded version on the top. Each of the 90 pictures was degraded in 5 different bands of degradation. Each participant saw each image once in one type of degradation. The images were 21 × 17 cm in size, with a visual angle of 23.72° horizontal × 19.30° vertical. Pictures were presented on a 17-inch monitor, from a distance of 50 cm, using Open Sesame software, version 3.2.7 Kafkaesque Koffka [44].
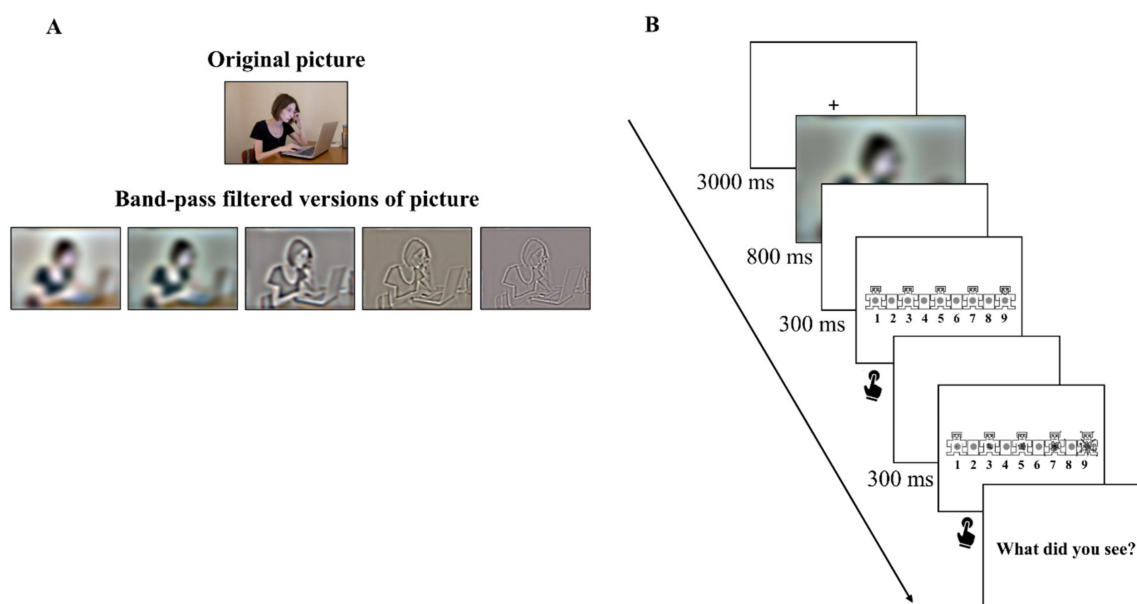
Figure 1. (**A**) Original (on the top) and band-passed (on the bottom) versions of one sample picture. The image shown in the picture is not part of the experimental database and is ©University of Bologna licensed for research use. Filtering levels were adjusted to the printed version of the picture; (**B**) Procedure of each trial. After viewing each picture, participants rated their affective state of valence and arousal on a 1–9 scale. Next, they were asked to describe the gist of the scene.

2.3. Procedure

The experiment took place inside a dimly lit laboratory room. The experimental procedure consisted of 90 trials, preceded by 4 practice trials. As depicted in Figure 1B, during each trial, a fixation cross was presented for 3000 ms. After this time, an image was displayed and remained visible for 800 ms. The visual rating scales of valence and arousal were presented using the Self-Assessment Manikin (SAM; [45,46]); each scale was preceded by a blank interval lasting 300 ms. Participants were asked to rate the affective state elicited by each picture on a 9-point scale on the dimension of affective valence and arousal. Finally, participants were asked to describe the gist of the scene by typing a short answer. The whole experimental procedure took approximately 60 min.

2.4. Data Analysis

Practice trials were excluded from the analysis. Accuracy scoring was carried out manually for each participant. The answers to the question "what did you see?" were analyzed with the following method: we classified each answer into one of the categories which were present in the study (erotic couples, babies, neutral people, attacks, and mutilation). A response was scored as accurate if the scored category matched the picture category. Two raters independently evaluated the accuracy of each response without being aware of the filtering condition that was applied. The inter-rater agreement calculated on the total number of the responses was high (Cohen's κ = 0.917). Accuracy data were submitted to repeated-measures ANOVAs with Band (five levels: 4, 13.5, 45.3, 152.2, and 512 cpi) and Content (three levels: pleasant, neutral, and unpleasant) as the within-subject factors. If a superordinate main effect or interaction was significant, we carried out separate ANOVAs on subordinate conditions or post hoc comparisons. Greenhouse–Geisser correction was applied when sphericity assumption was not met. The partial eta squared statistic (η^2_p), indicating the proportion between the variance explained by one experimental factor and the total variance, was calculated and reported.

Arousal and valence ratings were analyzed across filtering bands, content, and accuracy in the content description task. As picture degradation was minimum for intermediate levels (see Figure 1A), accuracy was unequally distributed, and missing cells in the ANOVA

design were observed for non-accurate trials in intermediate spatial frequency bands and, to a lesser extent, for accurate trials in the lowest and highest spatial frequency bands. To deal with these missing cells, we first replaced missing values using the multiple imputation method [47] using Amelia II [48] software, with Participant as cross-section variable, Degradation and Content as nominal variables, and arousal and valence as continuous variables which were constrained in the 1–9 range. A total of 5 different datasets were produced and analyzed, and F values were pooled from the analysis of these datasets. The *p* and partial eta squared values were recalculated for the pooled F values using the original degrees of freedom. Finally, a control analysis without missing data replacement was carried out which, although with smaller statistical power, confirmed the same pattern of results.

3. Results

3.1. Identification Performance

The identification performance of bandpass-filtered pictures as a function of picture content is reported in Figure 2. A significant main effect of Band was found ($F_{(4,56)} = 316.71$, $p < 0.001$, and $\eta^2_p = 0.96$), showing better identification in the intermediate bands (13.5, 45.3, and 152.2 cpi), compared with the lowest (4 cpi) and the highest (512 cpi) bands, and better identification for the lowest compared with the highest band. Significant differences were observed between all bands ($Fs_{(1,14)} > 62.93$, $ps < 0.001$, and $\eta^2_ps > 0.82$), except between the third (45.3 cpi) and the fourth (152.2 cpi) bands ($F_{(1,14)} = 3.86$, $p = 0.070$, and $\eta^2_p = 0.22$). A significant main effect of Content was also observed ($F_{(2,28)} = 61.07$, $p < 0.001$, and $\eta^2_p = 0.81$), indicating less accurate identification for unpleasant pictures compared with both pleasant ($F_{(1,14)} = 76.53$, $p < 0.001$, and $\eta^2_p = 0.85$) and neutral scenes ($F_{(1,14)} = 116.57$, $p < 0.001$, and $\eta^2_p = 0.89$), and the latter two picture contents did not significantly differ from each other ($F_{(1,14)} = 4.16$, $p = 0.061$, and $\eta^2_p = 0.23$). The Band by Content interaction was significant ($F_{(8,112)} = 5.06$, $p = 0.001$, and $\eta^2_p = 0.27$). Following this significant interaction and focusing on each spatial frequency band, a significant effect of picture content was observed in all bands ($Fs_{(2,28)} > 15.21$, $ps < 0.001$, and $\eta^2_ps > 0.52$). Specifically, a significantly more accurate identification for both neutral and pleasant pictures compared with unpleasant ones was observed in bands from the first (4 cpi) to the fourth (152.2 cpi) ($Fs_{(1,14)} > 11.92$, $ps < 0.004$, and $\eta^2_ps > 0.46$), with no differences between neutral and pleasant contents ($Fs_{(1,14)} < 4.07$, $ps > 0.063$, and $\eta^2_ps < 0.22$). In the highest band (512 cpi), neutral pictures were identified better than both pleasant and unpleasant pictures ($Fs_{(1,14)} > 12.43$, $ps < 0.01$, and $\eta^2_ps > 0.47$); the latter were significantly different from each other ($F_{(2,28)} = 8.232$, $p = 0.012$, and $\eta^2_p = 0.37$), with higher identification for pleasant pictures than unpleasant ones.
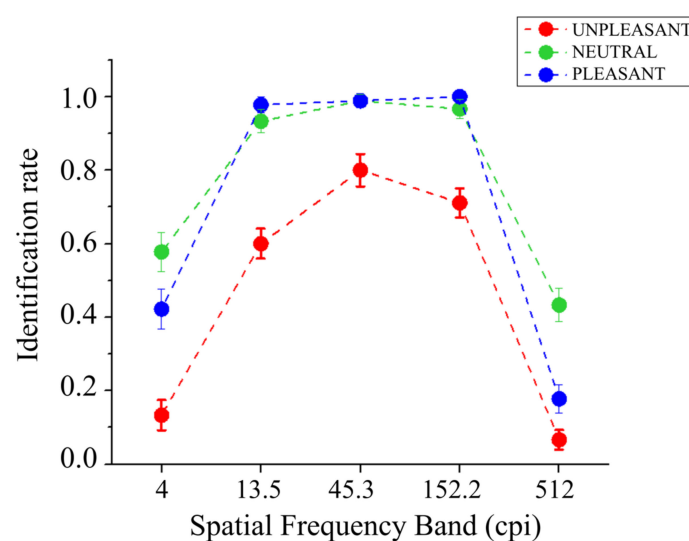
Figure 2. Identification rate (from 0 or inaccurate to 1 or accurate) of degraded natural scenes as a function of emotional content and the five spatial frequency bands.

3.2. Affective Response

Affective ratings were analyzed as a function of scene identification. As scene identification was extremely accurate in the intermediate frequency bands for most participants, it was not possible to have a fully balanced ANOVA design. While the multiple imputation method allowed us to replace missing values, the extremely high number of missing cells in non-accurate intermediate bands (up to 100%, with an average of more than 75% missing cells across participants in one content condition, with several of these participants only providing one trial per condition; see Table 1) suggested that the non-missing data in these intermediate levels might not suffice for missing data replacement, and they were therefore dropped from the analysis. For this reason, we first analyzed subjective affective reactions to scenes that were accurately described using an ANOVA design with factors Band (five levels) and Content (three levels). Then, we analyzed the effects of accuracy on the subjective ratings of arousal and valence to the most extreme spatial frequency bands using an ANOVA with factors Accuracy (accurate and inaccurate), Band (lowest and highest), and Content (pleasant, neutral, and unpleasant).

Table 1. Number of images correctly and incorrectly identified, for each emotional content and spatial frequency threshold.

		Emotional Content		
Spatial Frequency Band	**Description Accuracy**	**Unpleasant**	**Neutral**	**Pleasant**
Lowest (4 cpi)	Inaccurate	78	38	52
	Accurate	12	52	38
13.5 cpi	Inaccurate	36	6	2
	Accurate	54	84	88
54.3 cpi	Inaccurate	18	1	1
	Accurate	72	89	89
152.2 cpi	Inaccurate	26	3	0
	Accurate	64	87	90
Highest (512 cpi)	Inaccurate	84	51	74
	Accurate	6	39	16

3.2.1. Arousal Ratings

Subjective affective ratings of arousal are presented in Figure 3. In accurate trials, a significant main effect of Content was observed ($F_{(2,28)} = 39.10$, $p < 0.001$, and $\eta^2_p = 0.74$), indicating that affective responses to both unpleasant and pleasant pictures were rated as more arousing than neutral pictures ($Fs_{(1,14)} > 27.90$, $ps < 0.001$, and $\eta^2_ps > 0.67$) and that arousal ratings were higher following unpleasant compared with pleasant contents ($F_{(1,14)} = 11.70$, $p = 0.004$, and $\eta^2_p = 0.46$).

A significant interaction of Band and Content was observed ($F_{(8,112)} = 2.50$, $p = 0.016$, and $\eta^2_p = 0.15$). Following this significant interaction, the effects of Content were assessed in each band for accurate trials. In all bands, a significant effect of Content was observed ($Fs_{(2,28)} > 6.23$, $ps < 0.006$, and $\eta^2_ps > 0.31$), with higher arousal ratings for unpleasant and pleasant compared with neutral scenes ($Fs_{(1,14)} > 10.98$, $ps < 0.005$, and $\eta^2_ps > 0.44$), with the exception of pleasant vs. neutral in the highest band ($F_{(1,14)} = 3.94$, $p = 0.067$, and $\eta^2_p = 0.22$). Unpleasant scenes were rated as more arousing than pleasant scenes in the three intermediate levels ($Fs_{(1,14)} > 6.314$, $ps < 0.025$, and $\eta^2_ps > 0.31$). Finally, a significant effect of Band was observed ($F_{(4,56)} = 9.39$, $p < 0.001$, and $\eta^2_p = 0.40$), with lower ratings for scenes in the lowest and highest spatial frequency bands compared with all other spatial frequency bands ($Fs_{(1,14)} > 6.51$, $ps < 0.023$, and $\eta^2_ps > 0.32$) and lower ratings for scenes in the highest than in the lowest band ($F_{(1,14)} = 8.16$, $p = 0.013$, and $\eta^2_p = 0.37$).
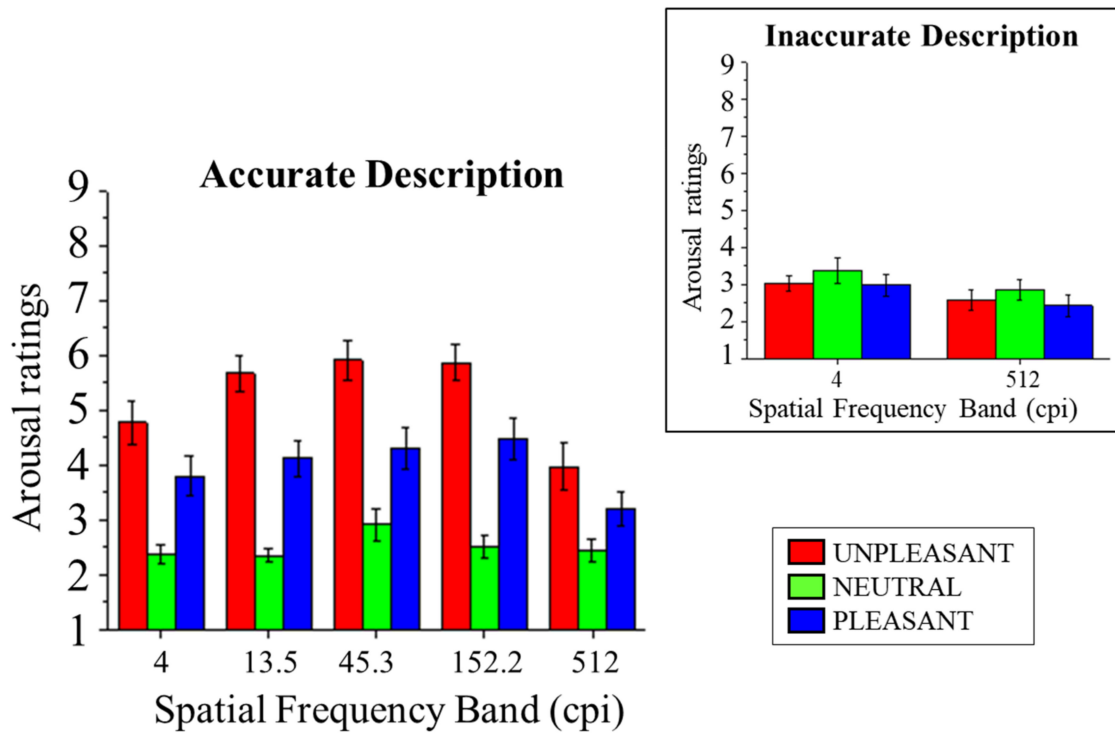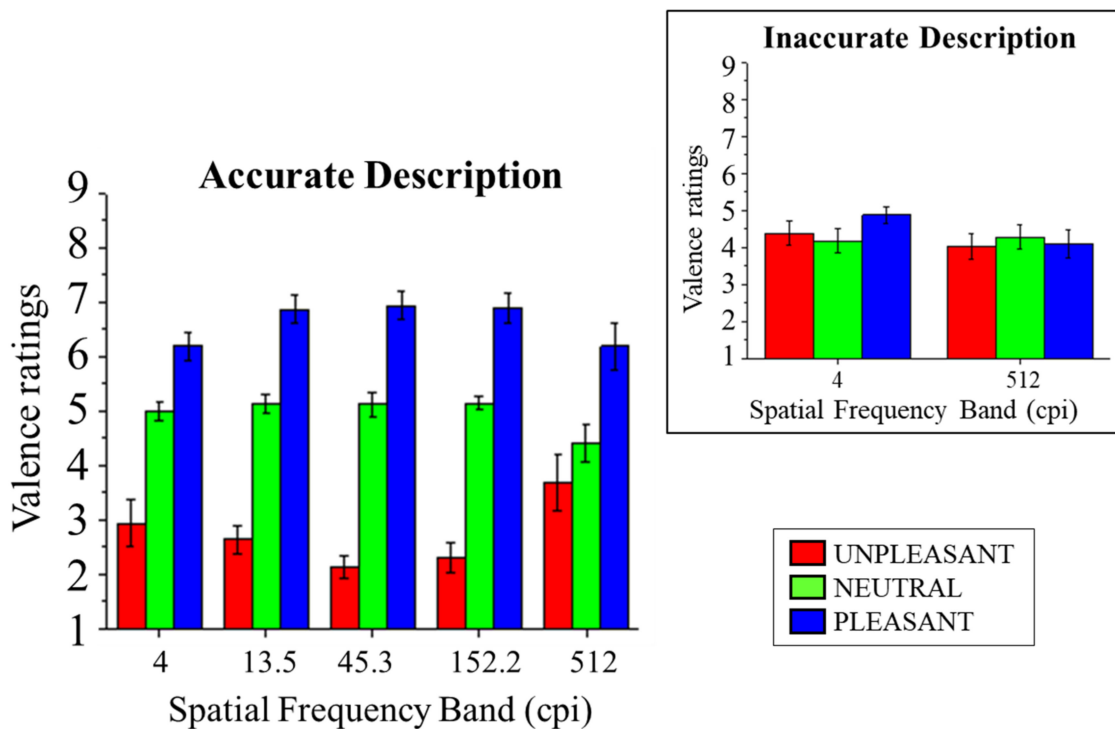
Figure 3. SAM ratings of arousal for pleasant, neutral, and unpleasant pictures as a function of the five spatial frequency bands when scene identification was achieved. The inset shows arousal scores in the lowest (4 cpi) and the highest (512 cpi) spatial frequency band as a function of emotional content when scene identification was not achieved. Error bars represent the within-participant standard error of the mean [49].

Directly focusing on the effects of identification accuracy on affective ratings in challenging conditions, i.e., the lowest and highest spatial frequency bands, we observed a significant interaction between Content and Accuracy ($F_{(2,28)}$ = 12.27, $p < 0.001$, and $\eta^2_p = 0.47$), indicating that no significant effect of Content was observed for non-accurate trials ($F_{(2,28)}$ = 2.15, $p = 0.135$, and $\eta^2_p = 0.13$), while it was observed for correctly described scenes ($F_{(2,28)}$ = 13.19, $p < 0.001$, and $\eta^2_p = 0.49$), with higher arousal ratings for pleasant and unpleasant compared with neutral scenes ($Fs_{(1,14)} > 11.46$, $ps < 0.004$, and $\eta^2_p s > 0.45$), and no significant difference between ratings for pleasant and unpleasant scenes ($F_{(1,14)}$ = 3.87, $p = 0.069$, and $\eta^2_p = 0.22$). No significant three-way interaction between Content, Accuracy, and Band was observed ($F_{(2,28)}$ = 0.99, $p = 0.386$, and $\eta^2_p = 0.07$), indicating that the spatial frequency band (lowest vs. highest) did not add to the effects of accuracy and content. Overall, a significant main effect of Content was observed ($F_{(2,28)}$ = 8.55, $p = 0.001$, and $\eta^2_p = 0.38$), indicating higher arousal ratings for unpleasant than neutral and pleasant scenes ($Fs_{(1,14)} > 5.39$, $ps < 0.036$, and $\eta^2_p s > 0.28$) but not for pleasant compared with neutral scenes ($F_{(1,14)}$ = 3.41, $p = 0.086$, and $\eta^2_p = 0.20$). Finally, a significant effect of Accuracy was observed ($F_{(1,14)}$ = 6.06, $p = 0.027$, and $\eta^2_p = 0.30$), with higher ratings for correctly compared with incorrectly described scenes, and for Band ($F_{(1,14)}$ = 8.16, $p = 0.013$, and $\eta^2_p = 0.37$), with higher ratings for the lowest compared with the highest spatial frequency band.

3.2.2. Valence Ratings

Subjective affective ratings of valence are presented in Figure 4. In accurate trials, a significant main effect of Content was observed ($F_{(2,28)}$ = 108.38, $p < 0.001$, and $\eta^2_p = 0.89$), indicating a linear effect of Content on valence ratings, with lower ratings following unpleasant scenes compared with neutral ($F_{(1,14)}$ = 110.54, $p < 0.001$, and $\eta^2_p = 0.89$) and following neutral compared with pleasant scenes ($F_{(1,14)}$ = 35.87, $p < 0.001$, and $\eta^2_p = 0.72$).

Figure 4. SAM ratings of valence for pleasant, neutral, and unpleasant pictures as a function of the five spatial frequency bands when scene identification was achieved. The inset shows valence scores in the lowest (4 cpi) and the highest (512 cpi) spatial frequency band as a function of emotional content when scene identification was not achieved. Error bars represent within-participant standard error of the mean.

A significant interaction of Band and Content was observed ($F_{(8,112)}$ = 4.01, p < 0.001, and η^2_p = 0.22). Following this significant interaction, the effects of Content were assessed at each band for accurate trials. In all bands, a significant effect of Content was observed ($Fs_{(2,28)}$ > 7.80, ps < 0.002, and η^2_ps > 0.36), with linearly increasing ratings of valence according to the pleasantness of the scenes, unpleasant vs. neutral ($Fs_{(1,14)}$ > 18.34, ps < 0.001, and η^2_ps > 0.57) and neutral vs. pleasant ($Fs_{(1,14)}$ > 7.45, ps < 0.016, and η^2_ps > 0.35), with the exception of unpleasant vs. neutral in the highest band ($F_{(1,14)}$ = 1.02, p = 0.329, and η^2_p = 0.07). No significant effect of Band was observed ($F_{(4,56)}$ = 0.42, p = 0.79, and η^2_p = 0.03).

An ANOVA with the factors Accuracy (accurate and inaccurate), Band (lowest and highest), and Content (pleasant, neutral, and unpleasant) was carried out on valence ratings. Scene Content interacted with Accuracy ($F_{(2,28)}$ = 13.47, p < 0.001, and η^2_p = 0.49), indicating that no significant effect of Content was observed for non-accurate trials ($F_{(2,28)}$ = 1.16, p = 0.33, and η^2_p = 0.08), while it was observed for correctly described scenes ($F_{(2,28)}$ = 23.15, p < 0.001, and η^2_p = 0.62), with lower valence ratings for unpleasant compared with neutral scenes ($F_{(1,14)}$ = 11.91, p = 0.004, and η^2_p = 0.46) and for neutral compared with pleasant scenes ($F_{(1,14)}$ = 15.73, p = 0.001, and η^2_p = 0.53). No significant three-way interaction between Content, Accuracy, and Band was observed ($F_{(2,28)}$ = 2.45, p = 0.105, and η^2_p = 0.15), indicating that the joint effects of Accuracy and Content did not differ between spatial frequency bands (lowest vs. highest). Finally, a significant main effect of Content was observed ($F_{(2,28)}$ = 25.54, p < 0.001, and η^2_p = 0.65), indicating linearly higher ratings of valence with increasing pleasantness of the scenes, unpleasant vs. neutral ($F_{(1,14)}$ = 10.78, p = 0.005, and η^2_p = 0.43) and neutral vs. pleasant ($F_{(1,14)}$ = 18.05, p = 0.001, and η^2_p = 0.56).

4. Discussion

Based on previous results in the picture-viewing context, in the present study, we investigated the relationship between scene identification and subjective emotional response, by manipulating the spatial frequency range of emotional and neutral scenes. The results showed that the ratings of arousal and valence were only modulated by the emotional content of the picture when scene content was correctly reported, regardless of the specific spatial frequency manipulation. No affective modulation of subjective ratings was observed when scene content was not correctly reported.

Motivationally relevant cues prompt a broad range of emotional responses, involving subjective, autonomic, facial, and cortical changes [1,5,50–52] that reflect the engagement of motivational systems, one appetitive and one defensive, that have evolved to adaptively regulate human behavior in the environment [1,6,53–55]. These responses can be elicited even under perceptually degraded conditions, such as long distance, peripheral vision, or short exposure time, providing evidence that individuals can efficiently discriminate among emotional picture contents. An example of a perceptually degraded condition is the case of images that only allow the observer to perceive their global features (low spatial frequencies) or local details (high spatial frequencies). Studies that directly examined the relation between spatial frequencies and emotional response reported that spatially filtered versions of emotional scenes may still elicit emotional responses (e.g., [16,18,21,56]). Although some studies suggested that pictures that could not be identified may still elicit emotional responses (e.g., [28,29]), the present study is consistent with previous results in indicating that emotional reactions require stimulus identification [14,18,20–22,27].

Several studies investigating the relationship between identification and emotional response assessed identification using gist understanding or closed-category categorization [18,21,27]. Here, we used a conservative measure of scene identification and asked participants to write a concise description of the scene. In terms of spatial frequency ranges, we observed that identification for all contents was higher in the intermediate frequency bands, supporting the view that no picture content relied critically on extremely low or high spatial frequency bands for identification [21,22,57,58]. In line with previous findings, we also observed that unpleasant scenes that included injuries and mutilated bodies were particularly difficult to understand and report compared with other contents [18,21]. This result suggests that discrimination difficulty is dissimilarly distributed among scene contents, possibly due to the different compositional features that are characteristic of each category. For instance, specific difficulties in understanding unpleasant pictures can come from the sometimes unusual positions of injured people or to missing body parts in some pictures. In terms of the idiosyncratic effects of each picture, future studies may use larger samples of pictures and attempt to pinpoint specific regularities (e.g., in terms of the presence of a distinctive feature or of sub-categories) that are responsible for category-specific effects. However, despite the fact that discrimination difficulty was higher for unpleasant scenes and for the highest and lower ends of the spatial frequency spectrum, affective responses to correctly described scenes were remarkably similar as a function of the spatial frequency range. In this respect, the only change in affective response was observed in the highest spatial frequency band (around 512 cpi), in which arousal ratings for pleasant and neutral scenes did not significantly differ from each other, and no significant difference between valence ratings for unpleasant and neutral pictures was observed. This result is consistent with the observation that affective categorization follows semantic categorization [27] and suggests that, although the content of these scenes was correctly described, the emotional modulation of affective states elicited by these scenes was less pronounced or absent.

Here, we observed that emotional responses, assessed using ratings of valence and arousal, did not depend on the compositional features of the natural scenes, operationalized here as the spatial frequency range. Rather, a similar affective response in terms of rated valence and arousal was observed in all visual conditions once the scene content was correctly reported. As subjective ratings represent one facet of emotional response, and responses to emotional stimuli are also expressed in several cortical and autonomic

activities and modulations, it may be asked whether other components of emotional responses are modulated by spatial frequency or by non-identified contents. In this respect, previous studies observed the affective modulation of autonomic and cortical responses only if the subjective modulation of affective ratings was achieved [14]. For instance, a study parametrically manipulated the exposure time of masked and unmasked stimuli from 25 ms to 6 s and failed to observe any affective modulation of the central, facial, and autonomic components of emotional response in the absence of the modulation of affective ratings. In the same study, participants who did not differentiate among pleasant and unpleasant contents in terms of rated pleasure showed no modulation as a function of picture content in other components of emotional response (e.g., cortical, facial, and electrodermal changes; [14]).

Along with previous studies, we observed a similar modulation of affective judgments once pictures were recognized [14,18,20–22,27]. Based on the observation that the amount of perceptual details is linearly related to the distance between an observer and an object [59], several studies investigated the pattern of emotional response to objects varying in distance, detail, or looming behavior [56,60,61]. These studies indicated that, while some components of the emotional response (e.g., modulation of electrodermal activity) are highly sensitive to the imminence of an object, other components are less modulated by picture imminence and are almost or exclusively sensitive to the content of a picture [21,56,60]. As the context in which events happen may be relevant for some components of the emotional response, future studies might manipulate picture context through perceptual manipulations as in the present study (e.g., using a perceptual manipulation as a proxy for a real-life property, i.e., spatial frequency for distance), through other contextual properties (e.g., history of the most recently viewed pictures, e.g., [62–64]), or through artificially created scenarios, i.e., in virtual-reality simulations, in which emotional responses might be elicited and measured [65,66]. Provided that emotional events are correctly detected and identified [14,18,20–22,27], it is possible that sensitivity to contextual manipulations varies between components of the emotional response (e.g., autonomic sympathetic activation vs. central activity).

As a secondary aim, we were interested in whether emotional responses rely on the semantic understanding of picture content (e.g., [18,21,22]) or are preferentially guided by specific visual information in natural scenes, such as low spatial frequencies [33,34,38–40]. In this respect, a preferential role of low spatial frequencies would have predicted that emotional responses to blurred stimuli would have been larger or less affected by the identification rate. However, neither prediction was observed, and no affective modulation was observed for either spatial frequency range when scene contents were not correctly reported. Moreover, when the combined effects of accuracy and of content were compared across spatial frequency bands, no difference between the two spatial frequency ranges was observed neither for arousal nor for valence ratings. Therefore, the current data do not support the view that a specific spatial frequency range (either low or high) contains information that is sufficient to elicit a subjective emotional response, which instead requires semantic content understanding. These results extend previous findings which showed that the emotional response, as indexed at the cortical level by the modulation of the LPP, did not depend on the compositional features of the natural scenes but instead varied with the identification of the affective content of the pictures, regardless of the spatial frequency range of natural scenes [21], and are consistent with the "semantic primacy" view of affective responses [19,20].

5. Conclusions

The present results are consistent with previous research in indicating that identification is a *conditio sine qua non* for affective response and that no isolated perceptual property or preference guides emotional response. When the emotional content of natural scenes can be recognized, it modulates evaluative judgments of pleasantness and arousal and results in the engagement of motivational systems.

References

1. Bradley, M.M. Emotion and motivation. In *Handbook of Psychophysiology*, 2nd ed.; Cacioppo, J.T., Tassinary, L.G., Berntson, G.G., Eds.; Cambridge University Press: Cambridge, UK, 2000; pp. 602–642.
2. Bradley, M.M. Natural selective attention: Orienting and emotion. *Psychophysiology* **2009**, *46*, 1–11. [CrossRef] [PubMed]
3. Bradley, M.M.; Lang, P.J. Emotion and motivation. In *Handbook of Psychophysiology*, 3rd ed.; Cacioppo, J.T., Tassinary, L.G., Berntson, G., Eds.; Cambridge University Press: New York, NY, USA, 2007; pp. 581–607.
4. Cuthbert, B.N.; Schupp, H.T.; Bradley, M.M.; Birbaumer, N.; Lang, P.J. Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biol. Psychol.* **2000**, *52*, 95–111. [CrossRef] [PubMed]
5. Lang, P.J.; Greenwald, M.K.; Bradley, M.M.; Hamm, A.O. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology* **1993**, *30*, 261–273. [CrossRef] [PubMed]
6. Lang, P.J.; Bradley, M.M.; Cuthbert, B.N. Motivated attention: Affect, activation, and action. In *Attention and Orienting*; Lang, P., Simons, R., Balaban, M., Eds.; Erlbaum: Hillsdale, NJ, USA, 1997; pp. 97–135.
7. Osgood, C.E.; Suci, G.; Tannenbaum, P. *The Measurement of Meaning*; University of Illinois: Urbana, IL, USA, 1957.
8. Russell, J.A.; Barrett, L.F. Affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *J. Personal. Soc. Psychol.* **1999**, *76*, 805–819. [CrossRef] [PubMed]
9. Schneirla, T. An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In *Nebraska Symposium on Motivation*; Jones, M., Ed.; University of Nebraska Press: Lincoln, NE, USA, 1959; pp. 27–58.
10. Wundt, W. *Grundriss der Psychologie*; Engelman: Leipzig, Germany, 1896.
11. Codispoti, M.; De Cesarei, A. Arousal and attention: Picture size and emotional reactions. *Psychophysiology* **2007**, *44*, 680–686. [CrossRef] [PubMed]
12. De Cesarei, A.; Codispoti, M. When does size not matter? Effects of stimulus size on affective modulation. *Psychophysiology* **2006**, *43*, 207–215. [CrossRef]
13. De Cesarei, A.; Codispoti, M.; Schupp, H.T. Peripheral vision and preferential emotion processing. *Neuroreport* **2009**, *20*, 1439–1443. [CrossRef] [PubMed]
14. Codispoti, M.; Mazzetti, M.; Bradley, M.M. Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology* **2009**, *46*, 731–738. [CrossRef] [PubMed]
15. Shulman, G.L.; Wilson, J. Spatial frequency and selective attention to local and global information. *Perception* **1987**, *16*, 89–101. [CrossRef]
16. Beligiannis, N.; Hermus, M.; Gootjes, L.; Van Strien, J.W. Both low and high spatial frequencies drive the early posterior negativity in response to snake stimuli. *Neuropsychologia* **2022**, *177*, 108403. [CrossRef]
17. De Cesarei, A.; Codispoti, M. Affective modulation of the LPP and alpha-ERD during picture viewing. *Psychophysiology* **2011**, *48*, 1397–1404. [CrossRef]
18. Codispoti, M.; Micucci, A.; De Cesarei, A. Time will tell: Object categorization and emotional engagement during processing of degraded natural scenes. *Psychophysiology* **2021**, *58*, e13704. [CrossRef] [PubMed]
19. Storbeck, J.; Robinson, M.D.; McCourt, M. Semantic processing precedes affect retrieval: The neurological case for cognitive primacy in visual processing. *Rev. Gen. Psychol.* **2006**, *10*, 41–55. [CrossRef]
20. Reisenzein, R.; Franikowski, P. On the latency of object recognition and affect: Evidence from temporal order and simultaneity judgments. *J. Exp. Psychol. Gen.* **2022**, *151*, 3060–3081. [CrossRef] [PubMed]
21. De Cesarei, A.; Codispoti, M. Scene identification and emotional response: Which spatial frequencies are critical? *J. Neurosci.* **2011**, *31*, 17052–17057. [CrossRef]
22. De Cesarei, A.; Codispoti, M. Spatial frequencies and emotional perception. *Rev. Neurosci.* **2013**, *24*, 89–104. [CrossRef]

23. Johnston, V.S.; Miller, D.R.; Burleson, M.H. Multiple P3s to emotional stimuli and their theoretical significance. *Psychophysiology* **1986**, *23*, 684–694. [CrossRef] [PubMed]

24. Schupp, H.T.; Cuthbert, B.N.; Bradley, M.M.; Cacioppo, J.T.; Ito, T.; Lang, P.J. Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology* **2000**, *37*, 257–261. [CrossRef] [PubMed]

25. Schupp, H.T.; Flaisch, T.; Stockburger, J.; Junghöfer, M. Emotion and attention: Event-related brain potential studies. *Prog. Brain Res.* **2006**, *156*, 31–51. [CrossRef]

26. Radilova, J. The late positive component of visual evoked response sensitive to emotional factors. *Act. Nerv. Super.* **1982**, *3*, 334–337.

27. Nummenmaa, L.; Hyönä, J.; Calvo, M.G. Semantic categorization precedes affective evaluation of visual scenes. *J. Exp. Psychol. Gen.* **2010**, *139*, 222–246. [CrossRef] [PubMed]

28. Mogg, K.; Bradley, B.P.; Williams, R.; Mathews, A. Subliminal processing of emotional information in anxiety and depression. *J. Abnorm. Psychol.* **1993**, *102*, 304–311. [CrossRef] [PubMed]

29. Prochnow, D.; Kossack, H.; Brunheim, S.; Müller, K.; Wittsack, H.J.; Markowitsch, H.J.; Seitz, R.J. Processing of subliminal facial expressions of emotion: A behavioral and fMRI study. *Soc. Neurosci.* **2013**, *8*, 448–461. [CrossRef] [PubMed]

30. Baier, D.; Kempkes, M.; Dítye, T.; Ansorge, U. Do Subliminal Fearful Facial Expressions Capture Attention? *Front. Psychol.* **2022**, *13*, 840746. [CrossRef] [PubMed]

31. Tipura, E.; Pegna, A.J. Subliminal emotional faces do not capture attention under high attentional load in a randomized trial presentation. *Vis. Cogn.* **2022**, *30*, 280–288. [CrossRef]

32. Tipura, E.; Renaud, O.; Pegna, A.J. Attention shifting and subliminal cueing under high attentional load: An EEG study using emotional faces. *Neuroreport* **2019**, *30*, 1251–1255. [CrossRef] [PubMed]

33. Vuilleumier, P.; Armony, J.L.; Driver, J.; Dolan, R.J. Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nat. Neurosci.* **2003**, *6*, 624–631. [CrossRef]

34. Mermillod, M.; Droit-Volet, S.; Devaux, D.; Schaefer, A.; Vermeulen, N. Are coarse scales sufficient for fast detection of visual threat? *Psychol. Sci.* **2010**, *21*, 1429–1437. [CrossRef]

35. Nakashima, T.; Goto, Y.; Abe, T.; Kaneko, K.; Saito, T.; Makinouchi, A.; Tobimatsu, S. Electrophysiological evidence for sequential discrimination of positive and negative facial expressions. *Clin. Neurophysiol.* **2008**, *119*, 1803–1811. [CrossRef]

36. Schyns, P.G.; Petro, L.S.; Smith, M.L. Transmission of facial expressions of emotion co-evolved with their efficient decoding in the brain: Behavioral and brain evidence. *PLoS ONE* **2009**, *4*, e5625. [CrossRef]

37. Smith, F.W.; Schyns, P.G. Smile through your fear and sadness: Transmitting and identifying facial expression signals over a range of viewing distances. *Psychol. Sci.* **2009**, *20*, 1202–1208. [CrossRef] [PubMed]

38. Alorda, C.; Serrano-Pedraza, I.; Campos-Bueno, J.J.; Sierra-Vázquez, V.; Montoya, P. Low spatial frequency filtering modulates early brain processing of affective complex pictures. *Neuropsychologia* **2007**, *45*, 3223–3233. [CrossRef] [PubMed]

39. Gomes, N.; Soares, S.C.; Silva, S.; Silva, C.F. Mind the snake: Fear detection relies on low spatial frequencies. *Emotion* **2018**, *18*, 886. [CrossRef] [PubMed]

40. Öhman, A. The role of the amygdala in human fear: Automatic detection of threat. *Psychoneuroendocrinology* **2005**, *30*, 953–958. [CrossRef] [PubMed]

41. Erdfelder, E.; Faul, F.; Buchner, A. GPOWER: A general power analysis program. *Behav. Res. Methods Instrum. Comput.* **1996**, *28*, 1–11. [CrossRef]

42. Cohen, J. *Statistical Power Analysis for the Behavioural Sciences*; Academic Press: Cambridge, MA, USA, 1969.

43. Lang, P.J.; Bradley, M.M.; Cuthbert, B.N. *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*; Technical Report A-8; University of Florida: Gainesville, FL, USA, 2008.

44. Mathôt, S.; Schreij, D.; Theeuwes, J. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* **2012**, *44*, 314–324. [CrossRef] [PubMed]

45. Lang, P.J. Behavioral treatment and bio-behavioral assessment: Computer applications. In *Technology in Mental Health Care Delivery*; Sidowski, B., Johnson, J.H., Williams, T.A., Eds.; Ablex: Norwood, NJ, USA, 1980; pp. 119–137.

46. Bradley, M.M.; Lang, P.J. Measuring emotion: The self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **1994**, *25*, 49–59. [CrossRef] [PubMed]

47. Schlomer, G.L.; Bauman, S.; Card, N.A. Best practices for missing data management in counseling psychology. *J. Couns. Psychol.* **2010**, *57*, 1–10. [CrossRef] [PubMed]

48. Honaker, J.; King, G.; Blackwell, M. Amelia II: A Program for Missing Data. *J. Stat. Softw.* **2011**, *45*, 1–47. [CrossRef]

49. O'Brien, F.; Cousineau, D. Representing error bars in within-subject designs in typical software packages. *Quant. Methods Psychol.* **2014**, *10*, 56–67. [CrossRef]

50. Cuthbert, B.N.; Schupp, H.T.; Bradley, M.; McManis, M.; Lang, P.J. Probing affective pictures: Attended startle and tone probes. *Psychophysiology* **1998**, *35*, 344–347. [CrossRef] [PubMed]

51. Greenwald, M.K.; Cook, E.W., III; Lang, P.J. Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *J. Psychophysiol.* **1989**, *3*, 51–64.

52. Lang, P.; Bradley, M. Emotion and the motivational brain. *Biol. Psychol.* **2010**, *84*, 437–450. [CrossRef]

53. Bradley, M.M.; Codispoti, M.; Cuthbert, B.N.; Lang, P.J. Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion* **2001**, *1*, 276–298. [CrossRef]

54. Dickinson, A.; Dearing, M.F. Appetitive–aversive interactions and inhibitory processes. In *Mechanisms of Learning and Motivation*; Dickinson, A., Boakes, R.A., Eds.; Erlbaum: Hillsdale, NJ, USA, 1979; pp. 203–231.

55. Konorski, J. *Integrative Activity of the Brain: An Interdisciplinary Approach*; University of Chicago Press: Chicago, IL, USA, 1967.

56. De Cesarei, A.; Codispoti, M. Effects of Picture Size Reduction and Blurring on Emotional Engagement. *PLoS ONE* **2010**, *5*, e13399. [CrossRef] [PubMed]

57. De Cesarei, A.; Loftus, G.R. Global and local vision in natural scene identification. *Psychon. Bull. Rev.* **2011**, *18*, 840–847. [CrossRef] [PubMed]

58. De Cesarei, A.; Loftus, G.R.; Mastria, S.; Codispoti, M. Understanding natural scenes: Contributions of image statistics. *Neurosci. Biobehav. Rev.* **2017**, *74*, 44–57. [CrossRef] [PubMed]

59. Loftus, G.R.; Harley, E.M. Why is it easier to identify someone close than far away? *Psychon. Bull. Rev.* **2005**, *12*, 43–65. [CrossRef] [PubMed]

60. De Cesarei, A.; Codispoti, M. Fuzzy picture processing: Effects of size reduction and blurring on emotional processing. *Emotion* **2008**, *8*, 352–363. [CrossRef]

61. Löw, A.; Lang, P.J.; Smith, J.C.; Bradley, M.M. Both Predator and Prey: Emotional Arousal in Threat and Reward. *Psychol. Sci.* **2008**, *19*, 865–873. [CrossRef]

62. Flaisch, T.; Junghöfer, M.; Bradley, M.M.; Schupp, H.T.; Lang, P.J. Rapid picture processing: Affective primes and targets. *Psychophysiology* **2008**, *45*, 1–10. [CrossRef] [PubMed]

63. Flaisch, T.; Stockburger, J.; Schupp, H.T. Affective prime and target picture processing: An ERP analysis of early and late interference effects. *Brain Topogr.* **2008**, *20*, 183–191. [CrossRef] [PubMed]

64. Schupp, H.T.; Schmälzle, R.; Flaisch, T.; Weike, A.I.; Hamm, A.O. Affective picture processing as a function of preceding picture valence: An ERP analysis. *Biol. Psychol.* **2012**, *91*, 81–87. [CrossRef] [PubMed]

65. Andreatta, M.; Winkler, M.H.; Collins, P.; Gromer, D.; Gall, D.; Pauli, P.; Gamer, M. VR for Studying the Neuroscience of Emotional Responses. In *Virtual Reality in Behavioral Neuroscience: New Insights and Methods*; Current Topics in Behavioral Neurosciences; Maymon, C., Grimshaw, G., Wu, Y.C., Eds.; Springer: Cham, Switzerland, 2023; Volume 65. [CrossRef]

66. Maymon, C.; Wu, Y.C.; Grimshaw, G. The Promises and Pitfalls of Virtual Reality. In *Virtual Reality in Behavioral Neuroscience: New Insights and Methods*; Current Topics in Behavioral Neurosciences; Maymon, C., Grimshaw, G., Wu, Y.C., Eds.; Springer: Cham, Switzerland, 2023; Volume 65. [CrossRef]