



SCIREA Journal of Education

ISSN: 2995-3901

<http://www.scirea.org/journal/Education>

April 21, 2026

Volume 11, Issue 1, February 2026

<https://doi.org/10.54647/education880658>

Development and Initial Normative Validation of the Affective Valence Image Task (AVIT)

Alfonso García-Monge ¹, Ana Escudero Montero ²

¹Universidad de Valladolid, Spain

²Universidad Complutense de Madrid, Spain

Abstract

The aim of the present study was to develop and provide an initial validation of the Affective Valence Image Task (AVIT), a visual instrument designed to assess basic hedonic polarity in terms of attraction and rejection. The image bank consisted of 60 color digital images, divided into two sets of 30 stimuli: one attraction set and one rejection set. The images were selected to represent affectively clear contents and to serve as functionally useful reference stimuli for later studies on affective responses, behavioral classification, and neurophysiological recording.

The initial normative validation was conducted with 52 young adult participants aged 19 to 23 years, who rated all images using 1-to-5 Likert scales. The results showed high and consistent ratings in the expected direction for both sets. The attraction images obtained a mean score of 4.61 (SD = 0.37), and the rejection images obtained a mean score of 4.73 (SD = 0.30). Internal consistency was excellent in both categories ($\alpha = .960$ for attraction and $\alpha = .953$ for rejection), and inter-rater agreement was high, both in terms of the intraclass correlation coefficient (ICC = .845 and .907, respectively) and Kendall's W, which was significant in both cases. The mean ratings for both sets were significantly above the midpoint of the scale, with no significant differences between the overall mean intensity of attraction and rejection.

Overall, the results support the initial psychometric adequacy of the AVIT as a brief, clear, and functionally focused tool for the assessment of attraction–rejection, and support its use as a reference system in future research involving both generic affective stimuli and domain-specific stimuli.

Keywords

Hedonic valence; affective images; emotional assessment; visual instrument; approach–avoidance

1. Introduction

The use of standardized affective image sets has played a central role in the experimental study of emotion because it makes it possible to induce affective states under controlled conditions and to compare findings across studies and laboratories (Branco et al., 2023). Within this field, the International Affective Picture System (IAPS) has become one of the most widely used resources, and its normative tradition has been closely linked to the Self-Assessment Manikin (SAM), a pictorial and nonverbal measure that assesses pleasure, arousal, and dominance through brief ratings (Bradley & Lang, 1994; Branco et al., 2023).

On this methodological basis, new stimulus sets and response tools have subsequently been developed, showing that affect induction systems can be adapted to specific experimental purposes without abandoning the classical dimensional framework (Betella & Verschure, 2016; Kurdi et al., 2017; Marchewka et al., 2014). The Nencki Affective Picture System (NAPS), for example, included ratings of valence, arousal, and approach–avoidance using computerized bipolar sliding scales, and also reported physical properties of the images such as luminance, contrast, and entropy (Marchewka et al., 2014). In turn, OASIS provided an open-access image set with valence and arousal norms (Kurdi et al., 2017), and the Affective Slider showed that continuous digital assessment of pleasure and arousal can be calibrated against the SAM, without requiring written instructions and with good compatibility across digital devices (Betella & Verschure, 2016). Taken together, these developments show that the construction of affective instruments can be oriented toward specific experimental needs, provided that both stimulus selection and the response system are explicitly described and validated (Betella & Verschure, 2016; Kurdi et al., 2017; Marchewka et al., 2014).

The literature also shows that the validity of an affective system should not be taken for granted outside the population in which it was originally normed (Branco et al., 2023). The systematic review by Branco et al. (2023) identified IAPS validation studies conducted in numerous countries and discussed cross-cultural, age-related, and sex-related differences, thereby reinforcing the need to test stimuli in the specific sample in which they are intended to be used. In Spain, the third part of the IAPS adaptation found high correlations with the North American norms, but also small cross-cultural differences in arousal and dominance, such that Spanish participants tended to assign higher arousal ratings and lower dominance ratings than American participants (Moltó et al., 2013). These findings support the idea that cultural adaptation is not a secondary detail, but rather a methodologically relevant condition when affective images are intended to be used as experimental reference material (Branco et al., 2023; Moltó et al., 2013).

This need for adaptation becomes even more important when the target population is composed of children and young people. In studies with minors, several authors have deliberately selected developmentally appropriate images and excluded scenes involving erotica, mutilation, and other highly arousing or violent content (Cohen-Gilbert & Thomas, 2013; Hajcak & Dennis, 2009). Hajcak and Dennis (2009), for example, worked with images that children might see on television or in the news and explicitly excluded erotica and mutilation, whereas Cohen-Gilbert and Thomas (2013) removed the most arousing images from their set, including erotica, mutilation, and some violent scenes, in order to adapt it for younger adolescents. In the same direction, Zamora et al. (2020) argued that emotion research in childhood should ideally rely on databases rated by children themselves, because much previous research has used sets rated by adults or images taken from the Internet without sufficient procedures for standardization and content control. Convergingly, Cordon et al. (2013) developed the Developmental Affective Photo System (DAPS) and obtained ratings of valence, arousal, and complexity from both children and adults precisely in order to provide material tailored to developmental research.

Another relevant methodological issue is that the affective response to an image does not depend solely on its semantic content, but also on certain low- and mid-level perceptual properties (Marchewka et al., 2014; Redies et al., 2020). NAPS specifically incorporated information on luminance, contrast, and entropy, and justified this decision by noting that the physical properties of images can influence the affective processing of visual stimuli (Marchewka et al., 2014). More recently, Redies et al. (2020) showed that global image

properties related to structure and composition can predict part of the subjective ratings of valence and arousal. Therefore, the construction of an affective image system oriented toward specific experimental purposes requires not only classifying content as positive or negative, but also documenting and, as far as possible, controlling visual variables that may act as confounding factors (Marchewka et al., 2014; Redies et al., 2020).

In this context, the objective of an instrument focused on attraction–rejection can be formulated more precisely than that of a general affective image bank. The aim is not necessarily to reproduce the entire classical affective space exhaustively, but rather to construct a controlled reference of hedonic polarity and motivational direction that makes it possible to classify stimuli as positive or negative in a specific population and, subsequently, to compare that reference with other contents of interest. This rationale is consistent with recent findings showing that negative affective experience combines common representations and stimulus-type-specific representations, such that both general components and content-dependent components jointly contribute to subjective experience (Čeko et al., 2022). Moreover, this strategy appears particularly relevant in educational contexts, where certain contents may acquire an affective meaning of their own. Linares et al. (2025) found that, in individuals with high math anxiety, math-related images were rated as lower in valence and dominance and higher in arousal than neutral images, and were processed similarly to low-arousal unpleasant images. From this perspective, a validated attraction–rejection system may function as a basic affective reference from which to examine how much of the response to later stimuli—such as mathematical content or specific pedagogical situations—is generic and how much is stimulus-specific (Čeko et al., 2022; Linares et al., 2025).

Within this framework, the present study aims to develop and validate an image-based system specifically oriented toward the rapid assessment of hedonic polarity in terms of attraction–rejection in Spanish children and young people. The purpose is not to replace general databases such as IAPS, NAPS, or OASIS, but rather to provide a methodologically focused, culturally proximal, and developmentally appropriate tool that can be used as an experimental reference in later studies and, more specifically, in paradigms designed to compare generic attraction–rejection with affective responses to domain-specific stimuli.

2. Method

2.1. General Design

The present study was conceived as an instrumental study aimed at the development and initial validation of a visual system for the assessment of attraction–rejection. The objective was to construct a set of affective stimuli organized into two opposite hedonic poles and to examine, in an initial normative phase, whether these stimuli were rated clearly, consistently, and concordantly in the expected direction. The applied purpose of the system is its subsequent use as a basic affective reference in experimental studies involving behavioral and neurophysiological classification.

2.2. Conceptual Design of the Instrument

The instrument was designed to operationalize a basic hedonic polarity in terms of attraction and rejection. The attraction pole was defined as a positive orientation toward the stimulus, associated with pleasantness, liking, closeness, or an approach-oriented disposition. The rejection pole was defined as a negative orientation, associated with unpleasantness, aversion, repulsion, or an avoidance-oriented disposition. The aim was not to represent a full spectrum of discrete emotions, but rather to provide a system functionally centered on the hedonic direction of the response.

From this perspective, the image bank was not constructed around exceptionally extreme contents or materials that might be inappropriate for younger populations, but rather around scenes with clear affective meaning, easily recognizable, and suitable for subsequent experimental use without resorting to excessively intrusive materials. This approach responds to the goal of generating a basic affective reference that may later allow comparison between responses to these generic stimuli and responses to more specific contents, such as mathematical materials, pedagogical scenes, or teacher explanations.

2.3. Development and Selection of the Stimuli

The final bank consisted of 60 color digital images, distributed into two sets of 30 stimuli each. The attraction images included scenes associated with positively valenced content, such as babies, pets, and natural landscapes. The rejection images included scenes designed to elicit a negative affective response, such as garbage, spoiled food, and intestinal surgery scenes. The stimuli were presented in standardized color digital format.

The images were generated using artificial intelligence tools. This procedure made it possible to work with scenes adjusted to previously defined categories and to maintain relative formal homogeneity across stimuli in global aspects such as brightness, color, and visual style. In the final selection, priority was given to images with clear semantic content, without embedded text or accessory elements that might hinder the immediate interpretation of the stimulus.

For internal organizational purposes, the images were grouped into two series, one for attraction and one for rejection, so that each stimulus constituted an independent unit of evaluation. This structure allows the full bank to be used or, alternatively, specific subsets to be selected depending on the objectives of later studies.

2.4. Presentation System and Response Logic

The system was conceived for integration into a computerized paradigm for the rapid assessment of hedonic valence. In its intended experimental configuration, each image is presented for 1500 ms, followed by a 50 ms blank screen and then by a vertical response bar. When the bar appears, the participant is required to use the mouse wheel to indicate the direction of the evaluation: movement in one direction is interpreted as a positive or approach-oriented response, whereas movement in the opposite direction is interpreted as a negative or avoidance-oriented response.

However, the validation phase presented in this study was carried out using an explicit rating procedure based on Likert scales, with the aim of first refining the image bank and determining whether each set functioned clearly and consistently before its full integration into the final computerized paradigm.

2.5. Participants

As summarized in Table 1, 52 individuals participated in the study, including 25 women and 27 men, aged between 19 and 23 years ($M = 21.04$, $SD = 1.44$). All participants provided informed consent before beginning the task.

Table 1 Sample Characteristics and Evaluation Procedure (Normative Ratings)

Characteristic	Value
Participants (N)	52
Age	19–23 years

Characteristic	Value
Women, <i>n</i> (%)	25 (48.1%)
Men, <i>n</i> (%)	27 (51.9%)
Attraction images (<i>n</i>)	30
Rejection images (<i>n</i>)	30
Response scale	1 (minimum) – 5 (maximum)
Responses per set (52 × 30)	1,560

Because the aim of this phase was to obtain an initial normative validation of the material, a homogeneous sample of young adults was used. Accordingly, the results should be interpreted as a preliminary validation of the image bank rather than as a definitive normative standard for all potential target populations.

2.6. Normative Validation Procedure

The rating task was administered individually in a computerized format. Before the session began, participants received instructions regarding the procedure and the meaning of the response scales. They were asked to rate each image according to the affective impression it produced and to respond directly, without spending excessive time on each judgment.

Each participant evaluated all 60 images from the full bank. For the attraction images, a 1-to-5 Likert scale was used, where 1 indicated no attraction and 5 indicated maximum attraction. For the rejection images, a 1-to-5 Likert scale was also used, where 1 indicated no rejection and 5 indicated maximum rejection. The images were presented in randomized order for each participant.

The procedure generated a total of 1,560 ratings for the attraction set and 1,560 ratings for the rejection set, corresponding to 52 participants who evaluated 30 images in each category.

2.7. Validation Strategy

The validation of the image bank was approached at several levels. First, the descriptive adequacy of the stimuli was examined, taking into account the expected direction of the ratings and the mean intensity reached by each set. Second, the internal consistency of the attraction and rejection images was assessed in order to determine whether each group of

stimuli functioned coherently as a differentiated dimension. Third, the degree of agreement among raters was analyzed as an indicator of the intersubjective stability of the ratings. Finally, inferential tests were conducted to determine whether the mean scores were significantly above the theoretical midpoint of the scale and to explore possible differences between categories and as a function of sex.

As a practical adequacy criterion, it was considered desirable for the images to obtain high ratings in the expected direction. In the present study, 100% of the attraction images and 100% of the rejection images reached mean scores equal to or above 4.0, which supported the retention of the full bank for subsequent uses.

2.8. Data analysis

Descriptive statistics were calculated to summarize the behavior of the ratings in both stimulus sets, including means, standard deviations, medians, and frequency distributions. The attraction images obtained a mean score of 4.61 (SD = 0.37), whereas the rejection images reached a mean of 4.73 (SD = 0.30). No ratings of 1 or 2 were recorded in either category.

The internal consistency of each set was assessed using Cronbach's alpha coefficient and corrected item-total correlations. Alpha was .960 for the attraction images and .953 for the rejection images. Corrected item-total correlations ranged from .415 to .814 for attraction and from .422 to .795 for rejection. The removal of any individual item did not substantially improve the alpha coefficients.

The degree of agreement among participants was examined using the intraclass correlation coefficient (ICC, model 2,k) and Kendall's W coefficient. The ICC was .845 for attraction and .907 for rejection. Kendall's W was significant in both cases, with $W = .174$, $\chi^2(29) = 261.88$, $p < .001$ for attraction, and $W = .252$, $\chi^2(29) = 380.25$, $p < .001$ for rejection.

In addition, one-sample t tests were conducted against the theoretical midpoint of the scale (value = 3) to determine whether the mean ratings were significantly above the level of indifference. Attraction ratings were significantly above the midpoint, $t(51) = 31.70$, $p < .001$, $d = 4.40$, and the same pattern was found for rejection ratings, $t(51) = 41.09$, $p < .001$, $d = 5.70$. A paired-samples t test was also conducted to compare the mean intensity of the two scales, and no statistically significant differences were found, $t(51) = -1.84$, $p = .072$, $d = -0.36$. Independent-samples t tests were also performed to explore possible differences as a function of sex. In all analyses, the significance level was set at $\alpha = .05$.

2.9. Ethical considerations

Participation was voluntary and was preceded by the provision of informed consent and approval by the University Ethics Committee. The data were recorded confidentially and analyzed in aggregate form. Before beginning the task, participants were informed that some images might be unpleasant and that they could discontinue their participation at any time without any consequence.

3. Results

3.1. Descriptive statistics

Table 2 presents the descriptive statistics for attraction and rejection ratings for the total sample and by sex. In the full sample, the attraction images obtained a mean score of 4.61 (SD = 0.37), whereas the rejection images reached a mean of 4.73 (SD = 0.30). The means observed for women and men followed a similar pattern, with high values in both categories. For attraction, women obtained a mean of 4.50 (SD = 0.40) and men a mean of 4.71 (SD = 0.31). For rejection, women obtained a mean of 4.75 (SD = 0.32) and men a mean of 4.71 (SD = 0.29). In addition, all images in the bank obtained mean scores equal to or above 4.0 on the corresponding dimension.

Table 2. Descriptive statistics of attraction and rejection ratings by sex

	n	Attraction		Rejection	
		M	SD	M	SD
Total	52	4.61	0.37	4.73	0.30
Women	25	4.50	0.40	4.75	0.32
Men	27	4.71	0.31	4.71	0.29

Note. Scores are expressed on a 1-to-5 scale.

The frequency distribution, shown in Table 3, confirmed that responses were concentrated in the upper range of the scale. For the attraction images, 65.1% of the ratings corresponded to 5 points, 31.0% to 4 points, and 4.0% to 3 points. For the rejection images, 76.1% of the responses corresponded to 5 points, 21.0% to 4 points, and 2.9% to 3 points. No ratings of 1 or 2 were recorded in either category. This pattern indicates that the stimuli were rated clearly and consistently in the expected direction.

Table 3. Distribution of Ratings by Stimulus Type and Approximate Mean (1–5)

Stimulus type	3	4	5	M
Attraction (30 images)	4%	31%	65%	4.61
Rejection (30 images)	3%	21%	76%	4.73

Note. Percentages reflect the average distribution of responses across the 30 images in each set. No ratings of 1 or 2 were recorded (0%). The mean (M) is a weighted estimate based on the observed proportions (3–5).

3.2. Internal consistency

Table 4 presents the internal reliability indicators for both stimulus sets. Cronbach’s alpha coefficient was .960 for the attraction scale and .953 for the rejection scale. Corrected item–total correlations ranged from .415 to .814 for attraction ($M = .657$) and from .422 to .795 for rejection ($M = .632$). In neither set did the removal of any item substantially improve the alpha value. Overall, these results indicate excellent internal consistency for both scales.

Table 4. Internal reliability indicators for both stimulus sets

Category	Cronbach’s α	McDonald’s ω	Item–total (range)	r Item–total (M)	r α if item deleted (range)
Attraction	.960	.962	.415 – .814	.657	.957 – .960
Rejection	.953	.961	.422 – .795	.632	.951 – .953

Additionally, McDonald’s omega coefficient (ω) was calculated from a unifactorial model estimated on the polychoric correlation matrix, given the ordinal level of measurement of the responses. The values obtained were $\omega = .962$ for attraction and $\omega = .961$ for rejection. The convergence between alpha and omega indicates that the factor loadings of the items are sufficiently homogeneous and confirms that both sets of stimuli function as essentially unidimensional scales. In the rejection scale, Image 27 was excluded from the calculation of omega because it showed zero variance, as all participants assigned the maximum score. This unidimensional character was also supported by the factorial structure: the first factor explained 46.0% of the variance in attraction and 46.6% in rejection, with ratios between the first and second eigenvalues of 6.62 and 6.39, respectively. These values clearly exceed the

conventional criterion of a ratio greater than 3 as an indicator of predominant unidimensionality.

3.3. Inter-rater agreement

The indices of agreement among participants are presented in Table 5. The intraclass correlation coefficient (ICC, model 2,k) was .845 for the attraction images and .907 for the rejection images, indicating a high degree of agreement among raters. Complementarily, Kendall’s W coefficient was significant in both cases: $W = .174$, $\chi^2(29) = 261.88$, $p < .001$ for attraction, and $W = .252$, $\chi^2(29) = 380.25$, $p < .001$ for rejection. These results show that the ratings were not only intense, but also notably convergent across participants.

Table 5. Inter-rater agreement indicators by stimulus category

Category	ICC (2,k)	Interpretation	Kendall’s W	$\chi^2(29)$	p
Attraction	.845	Excellent	.174	261.88	< .001
Rejection	.907	Excellent	.252	380.25	< .001

Note. ICC = intraclass correlation coefficient (two-way random-effects model, 2,k). Cicchetti’s (1994) criteria: < .40 poor, .40–.59 fair, .60–.74 good, $\geq .75$ excellent.

It should be noted that the absolute values of Kendall’s W (.174 for attraction and .252 for rejection) were modest, which contrasts with the high ICC values. This discrepancy is, however, expected and consistent with the characteristics of the image bank. Kendall’s W evaluates agreement in the relative ordering of the stimuli, and its magnitude depends directly on the variability across items in the ratings assigned by the evaluators. Because the bank was designed to maximize the hedonic clarity of each image, the ratings were concentrated in the upper range of the scale (all means ≥ 4.0), which substantially reduced dispersion across stimuli and, consequently, the available range for evaluators to establish matching rank orders. In other words, W is low not because participants disagreed about the direction or intensity of their ratings—which is precisely what the ICC reflects—but because the restriction of range inherent in the design of the instrument limits ordinal variability across items. In this context, the statistical significance of W in both sets ($p < .001$) confirms that, even within this compressed range, there was a degree of ordinal agreement greater than would be expected by chance.

3.4. Comparison with the midpoint of the scale

Table 6 presents the one-sample t tests against the theoretical midpoint of the scale (value = 3). Attraction ratings were significantly higher than this value, $t(51) = 31.70$, $p < .001$, $d = 4.40$, 95% CI [1.50, 1.72]. Likewise, rejection ratings also exceeded the midpoint significantly, $t(51) = 41.09$, $p < .001$, $d = 5.70$, 95% CI [1.65, 1.82]. Effect sizes were very large in both cases, indicating that the ratings were far above the level of indifference.

Table 6. One-sample t tests against the midpoint of the scale (value = 3)

Category	M	SD	t(51)	p	Cohen's d	95% CI
Attraction	4.61	0.37	31.70	< .001	4.40	[1.50, 1.72]
Rejection	4.73	0.30	41.09	< .001	5.70	[1.65, 1.82]

Note. 95% CI = 95% confidence interval of the difference from the midpoint (3).

3.5. Comparison between attraction and rejection

The direct comparison between the overall mean scores for attraction and rejection did not show statistically significant differences, $t(51) = -1.84$, $p = .072$, $d = -0.36$. This result suggests that both stimulus sets elicited comparable levels of intensity in their respective dimensions, which constitutes a methodologically favorable property for future comparative designs.

3.6. Sex differences

Table 7 presents the comparisons by sex. Levene's test confirmed the homogeneity of variances for both attraction ($p = .253$) and rejection ($p = .945$). Significant differences were observed in attraction ratings, $t(50) = -2.12$, $p = .039$, $d = -0.59$, with slightly higher scores in men than in women. No significant differences were found in rejection ratings, $t(50) = 0.44$, $p = .665$, $d = 0.12$.

Table 7. Sex differences in attraction and rejection ratings

Scale	Women (n = 25)		Men (n = 27)		t(50)	d
	M	SD	M	SD		
Attraction	4.50	0.40	4.71	0.31	-2.12*	-0.59
Rejection	4.75	0.32	4.71	0.29	0.44	0.12

Note. * $p < .05$.

3.7. Item-level analysis

The item-level analyses, presented in Tables 8 and 9, showed that all images in the bank reached high mean scores in the expected direction. In the attraction set, mean ratings ranged from 4.31 to 4.96. In the rejection set, means ranged from 4.02 to 5.00. In the latter case, one image obtained a mean score of 5.00 with zero variance, indicating complete unanimity in participants' responses. In both sets, corrected item–total correlation values remained within adequate ranges, further supporting the stability of the bank at the individual stimulus level.

Table 8. Descriptive statistics and reliability of the attraction images

Image	<i>M</i>	<i>DT</i>	<i>Mdn</i>	% 5	% 4	% 3	<i>Item–total r</i>	<i>α if item deleted</i>
Img1	4.44	0.64	5.0	51.9	40.4	7.7	0.698	0.958
Img2	4.31	0.67	4.0	42.3	46.2	11.5	0.814	0.957
Img3	4.67	0.51	5.0	69.2	28.8	1.9	0.481	0.960
Img4	4.52	0.50	5.0	51.9	48.1	0.0	0.650	0.959
Img5	4.56	0.57	5.0	59.6	36.5	3.8	0.699	0.958
Img6	4.60	0.63	5.0	67.3	25.0	7.7	0.791	0.957
Img7	4.67	0.51	5.0	69.2	28.8	1.9	0.701	0.958
Img8	4.56	0.54	5.0	57.7	40.4	1.9	0.645	0.959
Img9	4.42	0.54	4.0	44.2	53.8	1.9	0.635	0.959
Img10	4.62	0.57	5.0	65.4	30.8	3.8	0.695	0.958
Img11	4.71	0.46	5.0	71.2	28.8	0.0	0.581	0.959
Img12	4.69	0.54	5.0	73.1	23.1	3.8	0.697	0.958
Img13	4.75	0.56	5.0	80.8	13.5	5.8	0.629	0.959
Img14	4.35	0.68	4.0	46.2	42.3	11.5	0.721	0.958
Img15	4.35	0.65	4.0	44.2	46.2	9.6	0.664	0.959

Img16	4.81	0.40	5.0	80.8	19.2	0.0	0.620	0.959
Img17	4.38	0.63	4.0	46.2	46.2	7.7	0.674	0.959
Img18	4.85	0.36	5.0	84.6	15.4	0.0	0.447	0.960
Img19	4.81	0.40	5.0	80.8	19.2	0.0	0.658	0.959
Img20	4.90	0.30	5.0	90.4	9.6	0.0	0.567	0.959
Img21	4.38	0.60	4.0	44.2	50.0	5.8	0.601	0.959
Img22	4.58	0.61	5.0	63.5	30.8	5.8	0.797	0.957
Img23	4.58	0.61	5.0	63.5	30.8	5.8	0.765	0.958
Img24	4.71	0.50	5.0	73.1	25.0	1.9	0.669	0.959
Img25	4.96	0.19	5.0	96.2	3.8	0.0	0.582	0.960
Img26	4.85	0.41	5.0	86.5	11.5	1.9	0.415	0.960
Img27	4.63	0.53	5.0	65.4	32.7	1.9	0.737	0.958
Img28	4.63	0.53	5.0	65.4	32.7	1.9	0.755	0.958
Img29	4.73	0.49	5.0	75.0	23.1	1.9	0.647	0.959
Img30	4.31	0.67	4.0	42.3	46.2	11.5	0.683	0.959

Nota. N = 52. α total = .960.

Table 9. Descriptive statistics and reliability of the rejection images

Image	<i>M</i>	<i>DT</i>	<i>Mdn</i>	% 5	% 4	% 3	<i>Item-total r</i>	<i>α if item deleted</i>
Img1	4.83	0.43	5.0	84.6	13.5	1.9	0.757	0.951
Img2	4.90	0.30	5.0	90.4	9.6	0.0	0.546	0.952
Img3	4.73	0.45	5.0	73.1	26.9	0.0	0.691	0.951

Img4	4.90	0.30	5.0	90.4	9.6	0.0	0.561	0.952
Img5	4.79	0.41	5.0	78.8	21.2	0.0	0.610	0.952
Img6	4.92	0.33	5.0	94.2	3.8	1.9	0.620	0.952
Img7	4.35	0.74	4.5	50.0	34.6	15.4	0.677	0.952
Img8	4.52	0.70	5.0	63.5	25.0	11.5	0.739	0.951
Img9	4.92	0.27	5.0	92.3	7.7	0.0	0.703	0.952
Img10	4.58	0.57	5.0	61.5	34.6	3.8	0.631	0.952
Img11	4.79	0.41	5.0	78.8	21.2	0.0	0.777	0.951
Img12	4.77	0.47	5.0	78.8	19.2	1.9	0.524	0.953
Img13	4.96	0.19	5.0	96.2	3.8	0.0	0.560	0.953
Img14	4.52	0.58	5.0	55.8	40.4	3.8	0.701	0.951
Img15	4.92	0.27	5.0	92.3	7.7	0.0	0.422	0.953
Img16	4.87	0.34	5.0	86.5	13.5	0.0	0.795	0.951
Img17	4.50	0.58	5.0	53.8	42.3	3.8	0.656	0.951
Img18	4.87	0.34	5.0	86.5	13.5	0.0	0.513	0.953
Img19	4.63	0.53	5.0	65.4	32.7	1.9	0.664	0.951
Img20	4.63	0.53	5.0	65.4	32.7	1.9	0.616	0.952
Img21	4.62	0.63	5.0	69.2	23.1	7.7	0.716	0.951
Img22	4.02	0.64	4.0	21.2	59.6	19.2	0.640	0.952
Img23	4.69	0.47	5.0	69.2	30.8	0.0	0.615	0.952
Img24	4.77	0.47	5.0	78.8	19.2	1.9	0.680	0.951
Img25	4.92	0.27	5.0	92.3	7.7	0.0	0.753	0.952

Img26	4.81	0.44	5.0	82.7	15.4	1.9	0.704	0.951
Img27	5.00	0.00	5.0	100.0	0.0	0.0	—	0.954
Img28	4.83	0.43	5.0	84.6	13.5	1.9	0.603	0.952
Img29	4.69	0.54	5.0	73.1	23.1	3.8	0.703	0.951
Img30	4.69	0.54	5.0	73.1	23.1	3.8	0.776	0.950

Note. N = 52. Total α = .953. The dash (-) indicates zero variance in the item (all participants assigned the same score).

4. Discussion

The results of the present study indicate that the stimulus bank produced a clear and consistent hedonic polarization: the attraction and rejection images received high ratings in the expected direction, and all of them exceeded the criterion of 4 points on their corresponding dimension. This pattern is consistent with the general logic of standardized affective image sets, whose purpose is to induce affective states in a controlled and replicable manner in experimental contexts (Lang et al., 2008). The literature further shows that the International Affective Picture System (IAPS) has served as a reference model for the development of many later resources and that its validation has been extended to multiple cultural and population contexts (Branco et al., 2023). Within this same tradition, NAPS was designed as a broad database of realistic photographs with ratings of valence, arousal, and approach–avoidance, whereas OASIS offered an open image set with affective norms for experimental research (Kurdi et al., 2017; Marchewka et al., 2014). In contrast to these generalist systems, the present bank adopts a more focused strategy: it does not aim to cover the entire affective space, but rather to construct two clearly differentiated hedonic poles that may function as an instrumental reference in subsequent studies.

From this perspective, the concentration of responses in the upper range of the scale should not be interpreted as a limitation of the instrument, but rather as a consequence of the construction criterion that was followed. Whereas IAPS, NAPS, and OASIS were designed to cover a wide range of contents and affective gradations, the present system was deliberately developed to maximize the clarity of the attraction and rejection poles (Kurdi et al., 2017; Lang et al., 2008; Marchewka et al., 2014). For this reason, the absence of significant

differences between the overall mean intensity of the two sets constitutes a methodological strength, as it suggests that the positive and negative categories are reasonably balanced in hedonic strength, thereby facilitating their later use in experimental comparisons across conditions.

The excellent internal consistency and the high level of agreement among raters reinforce the idea that the bank not only differentiates the expected affective direction, but does so in a stable manner across participants. This stability is consistent with the logic of normative databases of emotional stimuli, which are based precisely on the aggregation of intersubjective judgments in order to provide materials that can be reused in other studies (Lang et al., 2008; Kurdi et al., 2017; Marchewka et al., 2014). It is also consistent with the emergence of resources specifically developed for certain age ranges, such as the Developmental Affective Photo System (DAPS) and the child-rated subset of NAPS, both of which were constructed from normative ratings obtained to ensure that the stimuli were appropriate and useful for developmental research (Cordon et al., 2013; Zamora et al., 2020). In this sense, although direct comparisons of psychometric coefficients across image banks are not always appropriate—because not all of them were designed as brief and unidimensional scales—the results of the present study do indicate that the set functions coherently as a focused instrument.

Another relevant aspect is that the present validation was carried out using explicit Likert scales, whereas the instrument is intended to be subsequently integrated into a computerized procedure involving rapid responding and low verbal load. This methodological evolution is consistent with trends already present in the literature. In NAPS, for example, ratings were collected through computerized bipolar semantic scales, and the Affective Slider was developed precisely to provide a digital, continuous, and nonverbal measure of affective experience equivalent to the SAM for the assessment of pleasure and arousal (Betella & Verschure, 2016; Marchewka et al., 2014). Therefore, the normative results obtained here not only support the selection of the images, but also provide a solid basis for transferring the bank to a more dynamic response format that is closer to its intended experimental use.

Sex differences were limited and restricted to attraction ratings. This finding should be interpreted with caution, but it is not incompatible with the previous literature. Bradley et al. (2001) showed that men and women do not react identically to affective images and that the most marked differences tend to emerge with highly motivational contents, especially erotica and threat/mutilation. In the same vein, the Spanish adaptation of the IAPS confirmed that

affective norms may vary as a function of cultural context and showed particularly visible differences in arousal and dominance compared with North American norms (Moltó et al., 2013). In addition, ERP studies have shown that contents such as erotica and mutilation tend to elicit especially intense responses compared with other affective categories (Weinberg & Hajcak, 2010). Within this framework, the absence of sex differences in rejection ratings in the present study may be related to the methodological decision to avoid extreme categories and to focus the instrument on a more controlled rather than maximalist hedonic polarity.

The main limitation of the study is that this validation was conducted with young adults, whereas the image bank is intended as a potentially useful tool for research with children, adolescents, and young people. The literature clearly indicates that adult affective norms cannot simply be extrapolated to other developmental stages. DAPS was created precisely to provide images suitable for both children and adults, with specific ratings of valence, arousal, and complexity, and the child-rated subset of NAPS was developed to offer norms obtained directly from child populations (Cordon et al., 2013; Zamora et al., 2020). The systematic review by Branco et al. (2023) also emphasizes that differences in age, sex, and culture are part of the normal process of adaptation and validation of this type of material. Therefore, the present data should be understood as an initial normative validation of the bank rather than as its definitive validation for school-age or adolescent populations.

Despite this limitation, the study provides a useful basis for the broader aim of the project: to use a set of generic attraction–rejection stimuli as a reference for comparing responses to other, more specific contents. This idea is consistent with recent findings showing that negative affect combines common components and stimulus-type-specific components, rather than being reducible to a single nonspecific dimension (Čeko et al., 2022). Moreover, evidence already exists in the academic domain that certain school-related contents may acquire an affective meaning of their own. Linares et al. (2025), for example, found that math-related images were rated by individuals with high math anxiety as less positive, less controllable, and more arousing than neutral images. From this perspective, having a generic, brief, and clearly polarized bank may be methodologically very useful for studying how much such domain-specific stimuli share with a more general attraction–rejection response and how much is specific to them.

Taken together, the results suggest that a brief, culturally proximal, and deliberately attraction–rejection-focused bank can achieve adequate levels of hedonic clarity, internal coherence, and intersubjective stability. Its main contribution does not lie in replacing

generalist systems such as IAPS, NAPS, or OASIS, but in providing a reference instrument functionally suited to research in which the aim is to isolate basic hedonic polarity and, subsequently, examine its relationship with other, more complex or more context-specific stimuli in educational settings (Branco et al., 2023; Kurdi et al., 2017; Lang et al., 2008; Marchewka et al., 2014).

References

- [1] Betella, A., & Verschure, P. F. M. J. (2016). The Affective Slider: A digital self-assessment scale for the measurement of human emotions. *PLOS ONE*, *11*(2), e0148037. <https://doi.org/10.1371/journal.pone.0148037>
- [2] Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion*, *1*(3), 300–319. <https://doi.org/10.1037/1528-3542.1.3.300>
- [3] Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- [4] Branco, D., Gonçalves, Ó. F., & Bermúdez i Badia, S. (2023). A systematic review of International Affective Picture System (IAPS) around the world. *Sensors*, *23*(8), 3866. <https://doi.org/10.3390/s23083866>
- [5] Čeko, M., Kragel, P. A., Woo, C.-W., López-Solà, M., & Wager, T. D. (2022). Common and stimulus-type-specific brain representations of negative affect. *Nature Neuroscience*, *25*(6), 760–770. <https://doi.org/10.1038/s41593-022-01082-w>
- [6] Cohen-Gilbert, J. E., & Thomas, K. M. (2013). Inhibitory control during emotional distraction across adolescence and early adulthood. *Child development*, *84*(6), 1954–1966. <https://doi.org/10.1111/cdev.12085>
- [7] Cordon, I. M., Melinder, A. M. D., Goodman, G. S., & Edelman, R. S. (2013). Children's and adults' memory for emotional pictures: Examining age-related patterns using the Developmental Affective Photo System. *Journal of Experimental Child Psychology*, *114*(2), 339–356. <https://doi.org/10.1016/j.jecp.2012.08.004>
- [8] Hajcak, G., & Dennis, T. A. (2009). Brain potentials during affective picture processing in children. *Biological psychology*, *80*(3), 333–338. <https://doi.org/10.1016/j.biopsycho.2008.11.006>

- [9] Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, 49(2), 457–470. <https://doi.org/10.3758/s13428-016-0715-3>
- [10] Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual* (Technical Report A-8). University of Florida.
- [11] Linares, R., Pelegrina, S., & Delgado-Rodríguez, R. (2025). Mathematics on the blackboard! Emotional processing of math-related pictures in individuals with math anxiety. *Scientific Reports*, 15, 26888. <https://doi.org/10.1038/s41598-025-12347-2>
- [12] 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. <https://doi.org/10.3758/s13428-013-0379-1>
- [13] Moltó, J., Segarra, P., López, R., Esteller, À., Fonfría, A., Pastor, M. C., & Poy, R. (2013). Adaptación española del “International Affective Picture System” (IAPS): Tercera parte. *Anales de Psicología*, 29(3), 965–984. <https://doi.org/10.6018/analesps.29.3.153591>
- [14] Redies, C., Grebenkina, M., Mohseni, M., Kaduhm, A., & Dobel, C. (2020). Global Image Properties Predict Ratings of Affective Pictures. *Frontiers in psychology*, 11, 953. <https://doi.org/10.3389/fpsyg.2020.00953>
- [15] Solomon, B., DeCicco, J. M., & Dennis, T. A. (2012). Emotional picture processing in children: an ERP study. *Developmental cognitive neuroscience*, 2(1), 110–119. <https://doi.org/10.1016/j.dcn.2011.04.002>
- [16] Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: The time-course of neural activity elicited by specific picture content. *Emotion*, 10(6), 767–782. <https://doi.org/10.1037/a0020242>
- [17] Zamora, E. V., Richard's, M. M., Introzzi, I., Aydmune, Y., Urquijo, S., Guàrdia Olmos, J., & Marchewka, A. (2020). The Nencki Affective Picture System (NAPS): A children-rated subset. *Trends in Psychology*, 28, 477–493. <https://doi.org/10.1007/s43076-020-00029-z>