



## OPEN ACCESS

## EDITED BY

Francesca Conca,  
University Institute of Higher Studies in Pavia,  
Italy

## REVIEWED BY

Anna M. Borghi,  
Sapienza University of Rome, Italy  
Anna Gilioli,  
University of Modena and Reggio Emilia, Italy

## \*CORRESPONDENCE

Marina M. Gerges  
✉ gerges.m.m@hse.ru  
Vladimir Kosonogov  
✉ vkosonogov@hse.ru

RECEIVED 07 March 2025

ACCEPTED 02 September 2025

PUBLISHED 18 September 2025

## CITATION

Gerges MM, Shelepenkov D and  
Kosonogov V (2025) Advancing affective  
stimuli databases: challenges and solutions.  
*Front. Psychol.* 16:1589612.  
doi: 10.3389/fpsyg.2025.1589612

## COPYRIGHT

© 2025 Gerges, Shelepenkov and Kosonogov.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Advancing affective stimuli databases: challenges and solutions

Marina M. Gerges<sup>1\*</sup>, Danila Shelepenkov<sup>1</sup> and  
Vladimir Kosonogov<sup>2\*</sup>

<sup>1</sup>Institute for Cognitive Neuroscience, HSE University, Moscow, Russia, <sup>2</sup>Affective Psychophysiology Laboratory, Institute of Health Psychology, HSE University, Saint Petersburg, Russia

Affective stimulus databases are integral elements in psychological and neuroscientific research, enabling the controlled induction of emotional states. However, despite significant progress, existing databases face methodological limitations that interfere with cross-study comparability and reproducibility. This review thoroughly examines modern affective stimulus databases across visual, auditory, textual, and multimodal domains, presenting their positive attributes and deficiencies. Key challenges include variability in stimulus standardization, inconsistencies in validation procedures, cultural specificity, and reliance on either categorical or dimensional emotion assessment methods. Additionally, issues related to stimulus diversity, duration control, and ecological validity further complicate the interpretation of results in psychophysiological studies. To address these challenges, we propose strategies for improving future databases, including the integration of standardized evaluation methodologies, the expansion of multimodal and culturally diverse stimuli, and the implementation of advanced technological solutions such as virtual reality and machine learning. Improving the structure of databases and maintaining consistent methodologies will increase the reliability and applicability of emotion research, ultimately contributing to a more comprehensive understanding of affective processes across different fields.

## KEYWORDS

affective science, databases, stimuli, psychophysiological research, affective stimulus databases

## 1 Introduction

To date, numerous researchers have sought to understand emotions, their origins, physiological correlations, and how they shape human behavior. Emotions are complex phenomena, encompassing both conscious and unconscious components, modulated by a variety of social, cultural, cognitive, and physiological factors (Izard, 2009; Barrett, 2017). The definition of emotion remains a subject of debate, with some theories emphasizing discrete categories (Ekman, 1992), such as anger or fear, and others advocating for dimensional models (Russell, 1980), which conceptualize emotion along continuous axes like valence and arousal. This distinction between models also shapes how emotional stimuli are developed and validated in experimental databases. Affective neuroscience continues to investigate the neural substrates underlying emotions (Lindquist et al., 2012; Kober et al., 2008). However, the emotion study relies on diverse protocols of induction, from visual and auditory stimuli to complex social scenarios. This heterogeneity often obscures findings and impedes progress in the field (Lang and Bradley, 2010; Schaefer et al., 2010). This review critically evaluates how this imbalance affects methodological consistency and cross-study comparability. We provide an overview of stimulus databases used to induce emotions. Our aim is not to outline existing

databases but rather to critically evaluate them, highlighting their strengths and limitations. Hence, we propose several ideas for improving the organization and standardization of future databases to enhance the quality of emotion research. By addressing the current limitations of affective stimuli databases and advocating for standardized and multimodal approaches, this review aims to contribute to more replicable, and generalizable findings in the study of emotions (Bradley and Lang, 2007; Kurdi et al., 2017).

## 2 Overview of emotional stimulus databases

Affective stimulus databases were developed to standardize emotion induction protocols. Some early efforts focused on visual (IAPS) (Lang et al., 1997) and auditory stimuli (IADS) (Bradley and Lang, 1999). Over subsequent decades, databases diversified to include facial expressions (e.g., NimStim) (Tottenham et al., 2009), video clips (DEVO) (Ack Baraly et al., 2020), and culturally specific stimuli (e.g., EmoMadrid) (Carretié et al., 2019). By 2020, the field had expanded to encompass multimodal and virtual reality (VR) stimuli, reflecting technological advancements and a growing emphasis on ecological validity.

A landmark effort to consolidate these resources is the KAPODI database (Diconne et al., 2022), which cataloged 35 auditory, 117 facial, 35 pictorial, 43 video, 89 textual, and 45 mixed-modality affective databases published between 1963 and 2020. KAPODI provides metadata for the year, sample sizes, descriptions, and validation methodologies, serving as a critical resource for researchers navigating the fragmented landscape of affective stimuli.

Post-2020 advancements have focused on addressing gaps in cultural specificity, multimodal integration, and ecological validity. Among these developments, The Empathy for Pain Stimuli System (Meng et al., 2023) uses visual and contextual cues to depict pain-inducing scenarios, validated for eliciting empathic responses. The ESISCA (Zhang et al., 2023) provides 274 images of social interactions conveying happiness, anger, sadness, fear, disgust, and neutral emotions for studying emotion recognition and empathy in social contexts. The SocialVidStim (Tully et al., 2024) offers video stimuli of social evaluations (e.g., praise, criticism), validated for ecological relevance. Additionally, the luVRe (Schöne et al., 2023) provides immersive 3D/360° stimuli for VR, while An Open-Access Database of Video Stimuli for Action Observation Research supports neuroimaging studies of motor cognition (Georgiev et al., 2024). These databases emphasize open-access principles, cultural sensitivity, and technological innovation.

## 3 Classification of emotional stimuli

There are several approaches for classifying emotional stimuli; however, we argue that the most fundamental distinction is based on modality. Different modalities evoke distinct physiological and psychological responses, making this classification particularly relevant for affective neuroscience and psychology (Schirmer and Adolphs, 2017).

Visual stimuli, particularly static images, are widely used in affective research due to their ease of control, reproducibility, and

standardization (Lang et al., 1997). Prominent databases such as the International Affective Picture System (IAPS) and OASIS provide a broad range of affective ratings (Kurdi et al., 2017), though they primarily used the dimensional model of emotion. Among image-based databases, several specialized resources stand out. For example, GAPED focuses on fear-related stimuli (Dan-Glauser and Scherer, 2011), NAPS provides culturally neutral content (Marchewka et al., 2014), and ESISCA captures culturally informed social interactions (Zhang et al., 2023), thereby broadening the scope of affective visual material.

A notable subcategory within visual stimuli includes databases of emotional facial expressions, such as KDEF (Lundqvist et al., 2015) and Ekman's series (Ekman and Friesen, 1976). These databases typically follow the discrete model of emotion, focusing on basic emotions and their corresponding facial expressions.

Dynamic visual stimuli—such as video clips and virtual reality (VR) content—offer enhanced ecological validity compared to static images, as they better approximate real-life emotional experiences. However, they introduce challenges in standardization, including variability in video duration, resolution, and audiovisual quality. Many of these databases utilize videoclips from unfamiliar sources (Ack Baraly et al., 2020) or well-known films. Notable examples include the Emotional Film Clips database (Gross and Levenson, 1995) and FilmStim (Schaefer et al., 2010), both of which offer carefully curated and validated video excerpts for emotion elicitation.

Recent advances have led to the development of more socially and contextually rich stimuli. For instance, SocialVidStim (Tully et al., 2024) could be fruitful for studying cognitive processes related to social interaction, while luVRe (Schöne et al., 2023) offers immersive 3D/360° VR environments. VR-based affective databases such as PanoEmo (Kosonogov et al., 2024), AVDOS-VR (Gnacek et al., 2024), and IAVRS (Mancuso et al., 2024) provide standardized emotional content within immersive virtual settings. These tools aim to increase emotional realism and participant engagement.

Not Only Visual Stimuli but also auditory ones, including music, speech, and environmental sounds, effectively induce emotions due to their strong link to autonomic responses (Bradley and Lang, 1999). These stimuli can be either nonverbal or verbal. Nonverbal auditory stimuli include emotionally expressive sounds, such as those found in the International Affective Digitized Sounds (IADS), emotionally annotated music tracks provided by EmoMusic (Soleymani et al., 2013), and domain-specific classifications of musical emotions featured in the Geneva Emotional Music Scale (GEMS) (Zentner et al., 2008). Verbal auditory stimuli encompass speech, pseudospeech, and vocal expressions. These types of stimuli often require linguistic and cultural adaptation to ensure emotional relevance and accuracy (Pell et al., 2015). Currently, most verbal auditory databases are available in a limited number of languages. However, there are notable exceptions, such as the VENEC and Vocally Expressed Emotions datasets, which aim for broader applicability. Notable databases include Brussels Mood Inductive Audio Stories (Bertels et al., 2014) for mood elicitation, as well as culturally specific sets like Chinese Vocal Emotional Stimuli (Liu and Pell, 2012) and the Italian Emotional Speech Database (Costantini et al., 2014).

Text-based stimuli provide an alternative method for emotion induction, particularly valuable in semantic and cognitive research. However, their effectiveness largely depends on participants' linguistic proficiency and cultural background (Kanske and Kotz, 2010). Most

text-based databases rely on dimensional models of emotion, primarily focusing on affective dimensions such as valence and arousal, occasionally incorporating additional dimensions. A widely used example is the Affective Norms for English Words (ANEW) database (Bradley and Lang, 1999), along with its various adaptations, such as the Spanish version (Redondo et al., 2007). While textual stimuli offer high experimental control and ease of implementation, they lack the sensory and contextual richness inherent in visual and auditory modalities, which can limit the depth and immediacy of emotional responses they elicit.

Multimodal stimuli combine multiple sensory inputs, such as images and sounds, to enhance ecological validity (Baumgartner et al., 2006). Although relatively few studies directly compared the effectiveness of emotional elicitation across different modalities, multimodal stimuli hold significant promise. They evoke stronger autonomic responses—such as skin conductance, pupil dilation, and heart rate changes—which are widely used as physiological indicators of emotional arousal and engagement (Bradley et al., 2001; Bradley et al., 2008). These responses correlate with both valence and arousal dimensions and contribute to validating the emotional impact of stimuli (Baumgartner et al., 2006; Fan et al., 2020). Additionally, multimodal input activates broader neural networks. For instance, increased alpha power during audiovisual conditions suggests deeper emotional processing (Fan et al., 2020). Baumgartner et al. (2006) found that combining music with emotional images enhanced emotional responses across self-report and physiological measures.

Moreover, multimodal stimuli often possess greater ecological validity, simulating real-world scenarios more effectively. This may improve the generalizability of findings to everyday experiences (Brück et al., 2011). However, multimodal results can sometimes be less stable than those elicited by simpler stimuli. Static images may offer superior control and reproducibility, making them ideal for eliciting certain emotions under tightly controlled conditions (Schmidt and Trainor, 2001), controlling over participants' focus, and reducing attentional drift (Lang et al., 1997). Conversely, prolonged exposure to dynamic or multimodal stimuli may be better suited to study sustained emotional responses, empathy, or narrative engagement (Schaefer et al., 2010). While multimodal stimuli offer greater realism, they introduce methodological complexities, including the need to control sensory integration effects and participant variability in immersive experience.

About half of all databases and most of the image databases (that do not include facial expressions) rely on the dimensional model of emotion, which conceptualizes affective experience along two continuous axes: valence describes the extent to which an emotion is positive or negative (ranging from unpleasant to pleasant), whereas arousal refers to its intensity (ranging from calm to excited) (Russell, 1980). This framework, known as the dimensional model of affect, is widely adopted but often excludes other models such as discrete emotion categories (e.g., anger, fear, sadness), potentially limiting cross-paradigm comparisons. In some cases, a third dimension—dominance (control vs. submission)—is also included, as in the ANEW word database. Some other dimensions can be added depending on the goal of a study, like social load (Kosonogov et al., 2020) and craving (Suissa-Rochelleau et al., 2019) or semantic content (Itkes et al., 2019).

However, few databases encompass both dimensional and categorical models; that is, they represent the model that supposes that

each discrete affective state should lie along valence and arousal dimensions (Russell, 1980). Notable exceptions include DEAP (Skaramagkas et al., 2023), which combines valence-arousal ratings with discrete emotion categories of musical videos with physiological data, and luVRe (Schöne et al., 2023), which offers immersive VR stimuli rated along both dimensions and discrete labels.

Other relevant variables such as ecological validity and visual complexity can also influence emotional responses and are increasingly considered in newer databases. These factors, while not part of traditional models, contribute to how stimuli can be perceived and processed. Terminology within this domain also varies across studies—for instance, “emotion elicitation,” “emotion recognition,” and “emotion perception” are often used interchangeably, though they refer to distinct processes (Lindquist et al., 2012; Schirmer and Adolphs, 2017). Greater precision and consistency in these terms would help reduce confusion and support better methodological alignment across studies.

## 4 Methodological considerations

Meta-analyses face challenges due to the variability in emotion induction methods across studies. Kirby and Robinson (2017) noted diverse protocols, including music, memory recall, facial expressions, words, and imagery. Similarly, Lindquist et al. (2012) combined studies using film clips, facial stimuli, and imagery, while other meta-analyses (Murphy et al., 2003; Kober et al., 2008; Vytal and Hamann, 2010) integrated data from mixed protocols such as visual, auditory, and memory-based methods. Despite their value, meta-analyses assume comparability among methods, though emotional perception can vary due to the stimulus modality (Schirmer and Adolphs, 2017; Fan et al., 2020; Murphy et al., 2003). This highlights the need for a nuanced approach that accounts for modality effects.

Methodological inconsistencies across affective databases often stem from the emotion model adopted. As we already mentioned, dimensional approaches, which dominate current databases, typically assess valence and arousal using continuous scales such as the Self-Assessment Manikin (Polo et al., 2024). Categorical models, by contrast, classify emotions into discrete categories, such as basic emotions or context-dependent states like distress or comfort. Each model shapes how emotions are measured, validated, and interpreted: dimensional scales allow for subtle gradations, whereas categorical labels may offer more intuitive clarity but risk oversimplifying affective complexity (Salikova and Kosonogov, 2025). Most databases rely exclusively on one framework, which may limit their adaptability across study designs and research questions.

The overall quality of affective databases also depends on their normative data and validation processes. While early databases often used large and diverse samples (Branco et al., 2023), others rely on smaller, culturally homogenous groups, reducing generalizability. In addition, the choice of rating instrument—whether Likert scales (and the number of points), visual analogue scales, or SAM—affects measurement sensitivity and interpretability. Researchers should also consider confounds such as order effects (Kosonogov, 2020), participant fatigue (Singh et al., 2025), and cultural biases (Morris, 1995), all of which can impact the reliability and replicability of emotional data. In addition, ethical concerns and cultural specificity present further challenges in emotion research. The availability of



affective stimuli that are both standardized and ethically approved is crucial. In the early stages of database development, professional photographers and videographers were often commissioned to produce stimuli, especially for content that was difficult to access (like violent or erotic material). Today, many databases rely on open-access images and films or provide web links to stimuli; however, these links may become obsolete over time as noted by Li et al. (2017). Some affective stimuli, particularly those from older databases, may no longer align with the experiences or cultural contexts of newer generations, potentially affecting the emotional responses they elicit.

Furthermore, because emotional responses can vary significantly across cultural contexts, a stimulus that is validated in one cultural setting may not evoke the same response in another. Cross-cultural research reveals that emotional responses are strongly influenced by physical and social environments. Jonauskaite et al. (2019) showed that participants in less sunny regions across 55 countries were more likely to associate yellow with joy, while Huwaë and Schaafsma (2018) found that individuals from collectivistic cultures suppress both positive and negative emotions more. Researchers should therefore validate foreign databases in their local context or develop culture-specific stimuli for the target population.

The application of emotional stimulus databases in psychophysiological research requires careful consideration. Static stimuli (e.g., pictures, words) allow precise control, whereas dynamic stimuli (e.g., videos, spoken or written texts, VR) evoke more complex nervous processes, complicating physiological measurements. Many databases have been successfully used in studies employing fMRI (Caria, 2020), EEG (Hajcak and Dennis, 2009), MEG (Styliadis et al., 2014), skin conductance (D'Hondt et al., 2010), heart rate (Bradley et al., 2001), and EMG (Baglioni et al., 2010). Additionally, abrupt sounds have been used to elicit the startle blink reflex, reflecting valence (Vrana et al., 1988; Kosonogov et al., 2016), while the eSEE-d dataset captures eye movements during emotion-evoking videos (Kreibig, 2010).

## 5 Limitations of current databases

Despite significant advances in developing affective stimulus databases, several limitations persist, affecting their applicability in psychophysiological research and affective science. One of the primary challenges is the *lack of standardization* across databases, making it difficult to compare findings across studies (Bradley and Lang, 2007). Differences in stimulus quantity, rating methodologies, and validation procedures can introduce inconsistencies that affect the reliability and reproducibility of research (Mollahosseini et al., 2019). Additionally, while some databases have been validated using large and diverse participant samples, others rely on smaller, culturally homogenous groups, limiting the generalizability of findings (Dan-Glauser and Scherer, 2011). At the same time, culturally specific databases are essential for capturing population-relevant emotional responses. We acknowledge this tension and suggest that future research should balance both goals by either validating existing databases across multiple cultural contexts or by developing parallel culturally grounded versions using shared protocols. To improve standardization more broadly, we recommend the use of common rating instruments (e.g., Self-Assessment Manikin or 9-point Likert scales), clear reporting of physical stimulus properties (e.g., duration, brightness,

resolution), and inclusion of participant demographic metadata. Standardized documentation formats and open-access repositories would further enhance reproducibility and cross-study comparability. Addressing these limitations is essential for improving the robustness and comparability of studies employing affective stimuli. There is considerable variability in the *number of stimuli* provided across studies. Although many researchers claim to use large databases, the actual stimulus count often remains inconsistent and insufficiently standardized. For example, the IAPS (Lang et al., 1997) offers a well-balanced set of stimuli across emotional categories, whereas smaller databases frequently lack sufficient stimuli per category, limiting their utility for robust experimental designs. This variability can lead to difficulties when researchers attempt to compile comprehensive stimulus sets that adequately represent a broad range of emotional states. As a positive example, the OASIS (Kurdi et al., 2017) provides 900 public domain images with normative ratings on valence and arousal, effectively addressing some of the limitations observed in earlier databases.

The *number of raters* for validating affective stimuli is often limited. For instance, while large-scale datasets like AffectNet (Mollahosseini et al., 2019) use online crowdsourcing to gather normative data, each image is typically annotated by only one rater (with a small subset receiving a second annotation for reliability). In contrast, some dynamic stimuli databases rely on samples of fewer than 50 participants, which may limit the generalizability of the ratings. Moreover, inconsistent stimulus durations can introduce variability in physiological responses (e.g., heart rate and skin conductance), further complicating result interpretation (Bradley et al., 2001).

Most affective databases rely exclusively on a single *measurement approach*—either dimensional (e.g., valence and arousal) or categorical (e.g., basic emotions)—which restricts the scope of emotional assessment. While the DEAP database (Koelstra et al., 2012) effectively combines physiological data with dimensional ratings, many databases depend solely on one type of assessment. For example, the NimStim facial expression set (Tottenham et al., 2009) relies exclusively on categorical labels obtained from a relatively small group of raters, limiting its ability to capture the full complexity of affective experiences. Because emotional responses are inherently multidimensional, the absence of a comprehensive, integrated approach can hinder cross-study comparisons and reduce the interpretability of findings (De Cesarei and Codispoti, 2010).

*Physical characteristics* such as brightness, loudness, complexity, and representational format (symbolic vs. photographic) vary significantly across affective databases, potentially introducing unintended confounds. While some databases, like EmoMadrid (Carretié et al., 2019), control for brightness and contrast, others contain substantial variability that may influence attentional engagement and neural processing. De Cesarei and Codispoti (2010) found that even factors like image size can affect psychophysiological responses, while Nejati (2021) demonstrated that 3D stimuli impose a higher cognitive load than 2D images, increasing pupil dilation and saccadic eye movements. These findings underscore the importance of standardizing physical properties within affective databases to minimize unintended influences on emotional and physiological measurements.

From a psychophysiological perspective, it is ideal for stimuli to be of *uniform duration* to ensure that evoked responses such as heart

rate, skin conductance, and neural activity are directly comparable across conditions. However, many affective video databases either include clips of varying durations or, when they standardize duration, provide too few examples to support reliable comparisons across different emotions. For example, the DEVO includes video clips ranging in length from 3 to 15 s (Ack Baraly et al., 2020), complicating the interpretation of physiological data due to inconsistencies in exposure times. In contrast, the Database of Visual Affective Stimuli (Li et al., 2022) standardizes all clips to a single 3 s duration, providing greater control over stimulus presentation. However, while such databases improve methodological consistency, they may not include sufficiently diverse clips for studies requiring detailed comparisons of specific emotions, such as distinguishing psychophysiological responses to sadness and anger. Expanding standardized databases to include a wider range of emotional content while maintaining strict temporal control remains a key challenge in affective research. A new study by Kosonogov et al. (2025) introduces a database of one-minute silent video clips with valence and arousal ratings, providing 160 affective videos. This development represents a step toward standardizing and ensuring the diversity of clips needed for more nuanced emotional research. Key methodological considerations for improving database development are summarized in Table 1.

6 Future directions

To address these limitations and move toward an ideal affective database, several improvements are proposed.

First, each subset of stimuli should consist of at least 10 items for each affective category (e.g., negative, neutral, positive) to ensure balanced representation. Using 10 stimuli per emotion ensures a sufficient variety of facial expressions to represent each emotion across different intensities and individual differences. This number strikes a balance between maintaining statistical power and minimizing participant fatigue. It aligns with the standard practice in emotion

perception research, where studies often utilize between 5 and 10 stimuli per emotion to achieve reliable and generalizable results without overburdening participants (Mikels et al., 2005; Pham and Sanchez, 2019). This can help psychophysiolgists to find a sufficient number of stimuli to provoke reliable reactions, since they depend on personal experience of subjects. In psychological studies, more stimuli may be presented, if they do not suppose baseline intervals, resting states, etc. However, in other EEG paradigms, significantly larger sets of stimuli or trials are typically required to obtain reliable results. For example, emotion recognition datasets used in machine learning, such as SEED-VII (Jiang et al., 2024), include only 12 video clips per emotion—a relatively limited amount when compared to methodological recommendations for event-related potentials (ERPs), another widely used EEG-based approach. Depending on the specific ERP component, the recommended number of trials ranges from a minimum of 8 for detecting error-related negativity (ERN) (Boudewyn et al., 2018) to approximately 50 trials for reliable motor-related responses (Borràs et al., 2022). Second, an adequate number of raters—ideally in the hundreds—should assess each stimulus to ensure robust and reliable normative data (Bradley and Lang, 2007). Increasing the number of raters improves the reliability of ratings by minimizing the influence of outliers or biases and allows for statistical power to analyze subgroup differences (Kurdi et al., 2017). For large databases, this is achieved by having different subsample rates for distinct stimulus subsets. In IAPS, thousands of images are divided into smaller sets (e.g., 100 images per subset), each rated by a unique group (Bradley et al., 2008). Similarly, ANEW, which includes hundreds of words, assigns smaller groups (e.g., 50 words per group) to ensure adequate ratings while distributing the workload efficiently (Bradley and Lang, 1999). Such strategies enhance representativeness, variability, and the reliability of normative data.

Third, future databases would benefit from the integration of both dimensional and categorical rating approaches. We suppose that such integration could contribute to the long-standing debate between discrete and dimensional models of emotion. A broad data collection

TABLE 1 Key methodological considerations for affective databases.

Methodological consideration	Recommendation	Supporting references	Relevant databases
Number of stimuli per category	Affective categories (negative, neutral, positive) should each include at least 10 stimuli to provide balanced representation and capture emotional variability across participants.	Schindler et al. (2017), Mikels et al. (2005), Bänziger et al. (2012), and Pant et al. (2023)	IAPS, OASIS, ANEW, GEMEP, PhyMER (Bänziger et al., 2012; Pant et al., 2023)
Number of raters per stimulus	Each stimulus should be rated by at least 100 participants to ensure robust and reliable normative data. Subsampling can be used for large databases, ensuring diversity in ratings.	Bradley and Lang (2007), Kurdi et al. (2017), Livingstone and Russo (2018), and Ma et al. (2015)	AffectNet, RAVDESS (Livingstone and Russo, 2018), Chicago Face Database (CFD) (Ma et al., 2015)
Dimensional vs. categorical models	Databases should integrate both dimensional (e.g., valence, arousal) and categorical (e.g., basic emotions) rating systems for more comprehensive emotion measurement.	Libkuman et al. (2007) and Morgan (2019)	IAPS (Libkuman et al., 2007), ANEW, Morgan Emotional Speech Set (Morgan, 2019)
Standardization of stimuli	Standardize non-affective properties, including visual (size, resolution), auditory (volume, pitch), and text (length, segmentation), to minimize confounds and enhance stimulus control.	Lakens et al. (2013), Jaquet et al. (2014), and de Gelder and Van den Stock (2011)	IAPS, OASIS, DEAP, BEAST (de Gelder and Van den Stock, 2011)
Cultural specificity	Future databases should either include culturally diverse samples, or be validated across multiple cultural contexts to ensure generalizability and ecological validity of emotional responses.	Pham and Sanchez (2019) and Bradley et al. (2008)	EmoMadrid, GAPED, OASIS, KDEP

of both discrete and dimensional ratings could help reveal some relationships between different variables used in both approaches and possibly reconcile them. In any case, tagging affective stimuli in both approaches may help researchers to test some hypotheses from different angles. One technical solution may involve assigning different groups of raters to distinct parts of the evaluation (for example, one group using dimensional scales and another using categorical labels) and indicating the method used in the database description. Another possibility would be to ask subjects to rate half of stimuli using one approach and another half with another approach.

Finally, it is critical to control and equalize all non-affective properties of stimuli. For visual content, this includes standardizing size, resolution, brightness, edge density, and duration; for auditory stimuli, maintaining consistent volume and pitch; and for text, ensuring uniform length and segmentation. While full standardization may not be feasible, minimizing confounds ensures responses are primarily driven by affective content. Brighter images tend to be rated as more positive (Lakens et al., 2013), higher visual complexity enhances occipital event-related potentials (Schmidt and Trainor, 2001), and emotional scenes show greater spectral density in low spatial frequency bands (Delplanque et al., 2007). In audio, pitch variations significantly affect perceived pleasantness and arousal (Jaquet et al., 2014), while in text, shorter segments elicit stronger emotional reactions (Schindler et al., 2017), emphasizing the need for careful stimulus formatting.

## 7 Conclusion

Affective stimulus databases have advanced psychophysiological research by linking emotional experiences with physiological markers. However, challenges such as variability in methods, validation, and cultural specificity continue to impede progress. Addressing these issues through standardized protocols, technological innovation, and cross-cultural validation is critical for enhancing the robustness and replicability of future studies. As the field embraces emerging methods such as virtual reality and machine learning, concerted efforts to refine affective databases will pave the way for deeper insights into the complex interplay between emotion, cognition, and physiology.

## References

- Ack Baraly, K. T., Muyingo, L., Beaudoin, C., Karami, S., Langevin, M., and Davidson, P. S. R. (2020). Database of emotional videos from Ottawa (DEVO). *Collabra Psychology* 6:10. doi: 10.1525/collabra.180
- Baglioni, C., Spiegelhalter, K., Lombardo, C., and Riemann, D. (2010). Sleep and emotions: a focus on insomnia. *Sleep Med. Rev.* 14, 227–238. doi: 10.1016/j.smrv.2009.10.007
- Bänziger, T., Mortillaro, M., and Scherer, K. R. (2012). Introducing the Geneva multimodal expression Corpus for experimental research on emotion perception. *Emotion* 12, 1161–1179. doi: 10.1037/a0025827
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Boston: Houghton Mifflin Harcourt: Houghton Mifflin Harcourt.
- Baumgartner, T., Esslen, M., and Jäncke, L. (2006). From emotion perception to emotion experience: emotions evoked by pictures and classical music. *Int. J. Psychophysiol.* 60, 34–43. doi: 10.1016/j.ijpsycho.2005.04.007
- Bertels, J., Deliëns, G., Peigneux, P., and Destrebecqz, A. (2014). The Brussels mood inductive audio stories (MIAS) database. *Behav. Res. Methods* 46, 1098–1107. doi: 10.3758/s13428-014-0445-3
- Borrás, M., Romero, S., Alonso, J. F., Bachiller, A., Serna, L. Y., Migliorelli, C., et al. (2022). Influence of the number of trials on evoked motor cortical activity in EEG recordings. *J. Neural Eng.* 19:046050. doi: 10.1088/1741-2552/ac86f5
- Boudewyn, M. A., Luck, S. J., Farrens, J. L., and Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology* 55:e13049. doi: 10.1111/psyp.13049
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion* 1, 276–298. doi: 10.1037/1528-3542.1.3.276
- Bradley, M. M., and Lang, P. J. (1999). *International Affective Digitized Sounds (IADS): Stimuli, instruction manual and affective ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
- Bradley, M. M., and Lang, P. J. (2007). “The international affective picture system (IAPS) in the study of emotion and attention” in *Handbook of Emotion Elicitation and Assessment* (New York, NY: Oxford University Press), 29–46.
- Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45, 602–607. doi: 10.1111/j.1469-8986.2008.00654.x
- Branco, D., Gonçalves, Ó. F., and Bermúdez i Badia, S. (2023). A systematic review of international affective picture system (IAPS) around the world. *Sensors* 23:3866. doi: 10.3390/s23083866
- Brück, C., Kreifelts, B., and Wildgruber, D. (2011). Emotional voices in context: a neurobiological model of multimodal affective information processing. *Phys Life Rev* 8, 383–403. doi: 10.1016/j.plrev.2011.10.002

## Author contributions

MG: Writing – original draft, Writing – review & editing. DS: Writing – original draft. VK: Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The article was prepared within the framework of the project “Mirror Laboratories” HSE University.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. ChatGPT was used.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



- Caria, A. (2020). Mesocorticolimbic interactions mediate fMRI-guided regulation of self-generated affective states. *Brain Sci.* 10:223. doi: 10.3390/brainsci10040223
- Carretié, L., Tapia, M., López-Martín, S., and Albert, J. (2019). EmoMadrid: an emotional pictures database for affect research. *Motiv. Emot.* 43, 929–939. doi: 10.1007/s11031-019-09780-y
- Costantini, G., Iaderola, I., Paoloni, A., and Todisco, M. (2014). EMOVO Corpus: An Italian Emotional Speech Database. In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). European Language Resources Association (ELRA); p. 3501–4.
- D'Hondt, F., Lassonde, M., Collignon, O., Dubarry, A. S., Robert, M., Rigoulot, S., et al. (2010). Early brain-body impact of emotional arousal. *Front. Hum. Neurosci.* 4:33. doi: 10.3389/fnhum.2010.00033
- Dan-Glauser, E. S., and Scherer, K. R. (2011). The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behav. Res. Methods* 43, 468–477. doi: 10.3758/s13428-011-0064-1
- De Cesarei, A., and Codispoti, M. (2010). Effects of picture size reduction and blurring on emotional engagement. *PLoS One* 5:e13399. doi: 10.1371/journal.pone.0013399. PMID: 20976237
- de Gelder, B., and Van den Stock, J. (2011). The Bodily Expressive Action Stimulus Test (BEAST). Construction and Validation of a Stimulus Basis for Measuring Perception of Whole Body Expression of Emotions *Front Psychol.* 2, 181. doi: 10.3389/fpsyg.2011.00181
- Delplanque, S., N'diaye, K., Scherer, K., and Grandjean, D. (2007). Spatial frequencies or emotional effects? *J. Neurosci. Methods* 165, 144–150. doi: 10.1016/j.jneumeth.2007.05.030
- Diconne, K., Kountouriotis, G. K., Paltoglou, A. E., Parker, A., and Hostler, T. J. (2022). Presenting KAPODI – the searchable database of emotional stimuli sets. *Emotion Rev.* 14, 84–95. doi: 10.1177/17540739211072803
- Ekman, P. (1992). An argument for basic emotions. *Cogn. Emot.* 6, 169–200. doi: 10.1080/02699939208411068
- Ekman, P., and Friesen, W. V. (1976). Measuring facial movement. *Environ. Psychol. Nonverb. Behav.* 1, 56–75. doi: 10.1007/BF01115465
- Fan, X., Deng, Z., Wang, K., Peng, X., and Qiao, Y. (2020). Learning Discriminative Representation For Facial Expression Recognition From Uncertainties. In: 2020 IEEE International Conference on Image Processing (ICIP). IEEE; 2020. p. 903–7. doi:10.1109/ICIP40778.2020.9190643
- Georgiev, C., Legrand, T., Mongold, S. J., Fiedler-Valenta, M., Guittard, F., and Bourguignon, M. (2024). An open-access database of video stimuli for action observation research in neuroimaging settings: psychometric evaluation and motion characterization. *Front. Psychol.* 15:15. doi: 10.3389/fpsyg.2024.1407458
- Gnacek, M., Quintero, L., Mavridou, I., Balaguer-Ballester, E., Kostoulas, T., Nduka, C., et al. (2024). AVDOS-VR: affective video database with physiological signals and continuous ratings collected remotely in VR. *Sci Data* 11:132. doi: 10.1038/s41597-024-02953-6
- Gross, J. J., and Levenson, R. W. (1995). Emotion elicitation using films. *Cogn. Emot.* 9, 87–108. doi: 10.1080/02699939508408966
- Hajcak, G., and Dennis, T. A. (2009). Brain potentials during affective picture processing in children. *Biol. Psychol.* 80, 333–338. doi: 10.1016/j.biopsycho.2008.11.006
- Huwaë, S., and Schaafsma, J. (2018). Cross-cultural differences in emotion suppression in everyday interactions. *Int. J. Psychol.* 53, 176–183. doi: 10.1002/ijop.12283
- Itkes, O., Eviatar, Z., and Kron, A. (2019). Semantic and affective manifestations of ambivalence (valence). *Cogn. Emot.* 33, 1356–1369. doi: 10.1080/02699931.2018.1564249
- Izard, C. E. (2009). Emotion theory and research: highlights, unanswered questions, and emerging issues. *Annu. Rev. Psychol.* 60, 1–25. doi: 10.1146/annurev.psych.60.110707.163539
- Jaquet, L., Danuser, B., and Gomez, P. (2014). Music and felt emotions: how systematic pitch level variations affect the experience of pleasantness and arousal. *Psychol. Music* 42, 51–70. doi: 10.1177/0305735612456583
- Jiang, W. B., Liu, X. H., Zheng, W. L., and Lu, B. L. (2024). SEED-VII: a multimodal dataset of six basic emotions with continuous labels for emotion recognition. *IEEE Trans. Affect. Comput.* doi: 10.1109/TAFFC.2024.3485057
- Jonauskaitė, D., Abdel-Khalek, A. M., Abu-Akel, A., Al-Rasheed, A. S., Antonietti, J. P., Åsgerisson, Å. G., et al. (2019). The sun is no fun without rain: physical environments affect how we feel about yellow across 55 countries. *J. Environ. Psychol.* 66:101350. doi: 10.1016/j.jenvp.2019.101350
- Kanske, P., and Kotz, S. A. (2010). Leipzig affective norms for German: a reliability study. *Behav. Res. Methods* 42, 987–991. doi: 10.3758/BRM.42.4.987
- Kirby, L. A. J., and Robinson, J. L. (2017). Affective mapping: an activation likelihood estimation (ALE) meta-analysis. *Brain Cogn.* 118, 137–148. doi: 10.1016/j.bandc.2015.04.006
- Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., and Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *NeuroImage* 42, 998–1031. doi: 10.1016/j.neuroimage.2008.03.059
- Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., et al. (2012). DEAP: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15
- Kosonogov, V. (2020). The effects of the order of picture presentation on the subjective emotional evaluation of pictures. *Psicologia.* 34, 171–178. doi: 10.17575/psicologia.v34i2.1608
- Kosonogov, V., Efimov, K., Kuskova, O., and Blank, I. (2025). One-minute silent video clips: a database of valence and arousal. *Eur. J. Psychol.* doi: 10.5964/ejop.14685 [In press].
- Kosonogov, V., Hajiyeva, G., and Zybrev, I. (2024). Panoemo, a set of affective 360-degree panoramas: a psychophysiological study. *Virtual Reality* 28:3. doi: 10.1007/s10055-023-00900-1
- Kosonogov, V., Martínez-Selva, J. M., Torrente, G., Carrillo-Verdejo, E., and Sánchez-Navarro, J. (2020). Does social content influence the subjective evaluation of affective pictures? *Span. J. Psychol.* 23:e25. doi: 10.1017/S1138741620000270
- Kosonogov, V., Sanchez-Navarro, J. P., Martínez-Selva, J. M., Torrente, G., and Carrillo-Verdejo, E. (2016). Social stimuli increase physiological reactivity but not defensive responses. *Scand. J. Psychol.* 57, 393–398. doi: 10.1111/sjop.12311
- Kreibitz, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010
- Kurdi, B., Lozano, S., and Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behav. Res. Methods* 49, 457–470. doi: 10.3758/s13428-016-0715-3
- Lakens, D., Fockenberg, D. A., Lemmens, K. P. H., Ham, J., and Midden, C. J. H. (2013). Brightness differences influence the evaluation of affective pictures. *Cogn. Emot.* 27, 1225–1246. doi: 10.1080/02699931.2013.781501
- Lang, P. J., and Bradley, M. M. (2010). Emotion and the motivational brain. *Biol. Psychol.* 84, 437–450. doi: 10.1016/j.biopsycho.2009.10.007
- Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1997). International affective picture system (IAPS): Instruction manual and affective ratings. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
- Li, B. J., Bailenson, J. N., Pines, A., Greenleaf, W. J., and Williams, L. M. (2017). A public database of immersive VR videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures. *Front. Psychol.* 8:8. doi: 10.3389/fpsyg.2017.02116
- Li, Q., Zhao, Y., Gong, B., Li, R., Wang, Y., Yan, X., et al. (2022). Visual affective stimulus database: a validated set of short videos. *Behavioral Sciences* 12:137. doi: 10.3390/bs12050137
- Libkuman, T. M., Otani, H., Kern, R., Viger, S. G., and Novak, N. (2007). Multidimensional normative ratings for the international affective picture system. *Behav. Res. Methods* 39, 326–334. doi: 10.3758/BRF03193164
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., and Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35, 121–143. doi: 10.1017/S0140525X11000446
- Liu, P., and Pell, M. D. (2012). Recognizing vocal emotions in mandarin Chinese: a validated database of Chinese vocal emotional stimuli. *Behav. Res. Methods* 44, 1042–1051. doi: 10.3758/s13428-012-0203-3
- Livingstone, S. R., and Russo, F. A. (2018). The Ryerson audio-visual database of emotional speech and song (RAVDESS): a dynamic, multimodal set of facial and vocal expressions in north American English. *PLoS One* 13:e0196391. doi: 10.1371/journal.pone.0196391
- Lundqvist, D., Flykt, A., and Öhman, A. (2015). Karolinska directed emotional faces. *PsychTESTS Dataset*. doi: 10.1080/02699930701626582
- Ma, D. S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: a free stimulus set of faces and norming data. *Behav. Res. Methods* 47, 1122–1135. doi: 10.3758/s13428-014-0532-5
- Mancuso, V., Borghesi, F., Chirico, A., Bruni, F., Sarcinella, E. D., Pedroli, E., et al. (2024). IAVRS – international affective virtual reality system: psychometric assessment of 360° images by using psychophysiological data. *Sensors* 24:4204. doi: 10.3390/s24134204
- Marchewka, A., Żurawski, Ł., Jednoróg, K., and Grabowska, A. (2014). The Nencki affective picture system (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav. Res. Methods* 46, 596–610. doi: 10.3758/s13428-013-0379-1
- Meng, J., Li, Y., Luo, L., Li, L., Jiang, J., Liu, X., et al. (2023). The empathy for pain stimuli system (EPSS): development and preliminary validation. *Behav. Res. Methods* 56, 784–803. doi: 10.3758/s13428-023-02087-4
- Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., Reuter-Lorenz, P. A., et al. (2005). Emotional category data on images from the International Affective Picture System. *Behav Res Methods* 37, 626–630.
- Mollahosseini, A., Hasani, B., and Mahoor, M. H. (2019). Affectnet: a database for facial expression, valence, and arousal computing in the wild. *IEEE Trans. Affect. Comput.* 10, 18–31. doi: 10.1109/TAFFC.2017.2740923
- Morgan, S. D. (2019). Categorical and dimensional ratings of emotional speech: behavioral findings from the Morgan emotional speech set. *J. Speech Lang. Hear. Res.* 62, 4015–4029. doi: 10.1044/2019\_JSLHR-S-19-0144
- Morris, J. D. (1995). Observations: SAM: the self-assessment manikin: an efficient cross-cultural measurement of emotional response. *J. Advert. Res.* 35, 63–68. doi: 10.1080/00218499.1995.12466497

- Murphy, F. C., Nimmo-Smith, I., and Lawrence, A. D. (2003). Functional neuroanatomy of emotions: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 207–233. doi: 10.3758/cabn.3.3.207
- Nejati, V. (2021). Effect of stimulus dimension on perception and cognition. *Acta Psychol.* 212:103208. doi: 10.1016/j.actpsy.2020.103208
- Pant, S., Yang, H. J., Lim, E., Kim, S. H., and Yoo, S. B. (2023). PhyMER: physiological dataset for multimodal emotion recognition with personality as a context. *IEEE Access* 11, 107638–107656. doi: 10.1109/ACCESS.2023.3320053
- Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., and Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biol. Psychol.* 111, 14–25. doi: 10.1016/j.biopsycho.2015.08.008
- Pham, H., and Sanchez, C. A. (2019). Text segment length can impact emotional reactions to narrative storytelling. *Discourse Process.* 56, 210–228. doi: 10.1080/0163853X.2018.1426351
- Polo, E. M., Farabbi, A., Mollura, M., Mainardi, L., and Barbieri, R. (2024). Understanding the role of emotion in decision making process: using machine learning to analyze physiological responses to visual, auditory, and combined stimulation. *Front. Hum. Neurosci.* 17:17. doi: 10.3389/fnhum.2023.1286621
- Redondo, J., Fraga, I., Padrón, I., and Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behav. Res. Methods* 39, 600–605. doi: 10.3758/BF03193031
- Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714
- Salikova, D., and Kosonogov, V. (2025). Complex emotional experiences: theoretical significance, ways of induction and therapeutic potential. *Curr. Psychol.* 44, 1962–1975. doi: 10.1007/s12144-024-07207-7
- Schaefer, A., Nils, F., Sanchez, X., and Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. *Cogn. Emot.* 24, 1153–1172. doi: 10.1080/02699930903274322
- Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., et al. (2017). Measuring aesthetic emotions: a review of the literature and a new assessment tool. *PLoS One* 12:e0178899. doi: 10.1371/journal.pone.0178899
- Schirmer, A., and Adolphs, R. (2017). Emotion perception from face, voice, and touch: comparisons and convergence. *Trends Cogn. Sci.* 21, 216–228. doi: 10.1016/j.tics.2017.01.001
- Schmidt, L. A., and Trainor, L. J. (2001). Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cogn. Emot.* 15, 487–500. doi: 10.1080/02699930126048
- Schöne, B., Kisker, J., Sylvester, R. S., Radtke, E. L., and Gruber, T. (2023). Library for universal virtual reality experiments (luvre): a standardized immersive 3D/360° picture and video database for VR based research. *Curr. Psychol.* 42, 5366–5384. doi: 10.1007/s12144-021-01841-1
- Singh, P., Budhiraja, R., Jalote, P., Kumar, M., and Singh, P. (2025). Translating emotions to annotations: a participant perspective of physiological emotion data collection. *PACMHCI*. 9:CSC195. doi: 10.1145/3711093
- Skaramagkas, V., Ktistakis, E., Manousos, D., Kazantzaki, E., Tachos, N. S., Tripoliti, E., et al. (2023). eSEE-d: emotional state estimation based on eye-tracking dataset. *Brain Sci.* 13:589. doi: 10.3390/brainsci13040589
- Soleymani, M., Caro, M. N., Schmidt, E. M., Sha, C. Y., and Yang, Y. H.. (2013). 1000 songs for emotional analysis of music. In: Proceedings of the 2nd ACM international workshop on Crowdsourcing for multimedia. New York, NY, USA: ACM; p. 1–6. doi: 10.1145/2506364.2506365
- Styliadis, C., Ioannides, A. A., Bamidis, P. D., and Papadelis, C. (2014). Amygdala responses to valence and its interaction by arousal revealed by MEG. *Int. J. Psychophysiol.* 93, 121–133. doi: 10.1016/j.ijpsycho.2013.05.006
- Suissa-Rocheleau, L., Benning, S. D., and Racine, S. E. (2019). Associations between self-report and physiological measures of emotional reactions to food among women with disordered eating. *Int. J. Psychophysiol.* 144, 40–46. doi: 10.1016/j.ijpsycho.2019.08.004
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* 168, 242–249. doi: 10.1016/j.psychres.2008.05.006
- Tully, L. M., Blendermann, M., Fine, J. R., Zakskorn, L. N., Fritz, M., Hamlett, G. E., et al. (2024). The SocialVidStim: a video database of positive and negative social evaluation stimuli for use in social cognitive neuroscience paradigms. *Soc. Cogn. Affect. Neurosci.* (SCAN) 19. doi: 10.1093/scan/nsae024
- Vrana, S. R., Spence, E. L., and Lang, P. J. (1988). The startle probe response: a new measure of emotion? *J. Abnorm. Psychol.* 97, 487–491. doi: 10.1037//0021-843x.97.4.487
- Vytal, K., and Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based Meta-analysis. *J. Cogn. Neurosci.* 22, 2864–2885. doi: 10.1162/jocn.2009.21366
- Zentner, M., Grandjean, D., and Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion* 8, 494–521. doi: 10.1037/1528-3542.8.4.494
- Zhang, Z., Peng, Y., Jiang, Y., and Chen, T. (2023). The pictorial set of emotional social interactive scenarios between Chinese adults (ESISCA): development and validation. *Behav. Res. Methods* 56, 2581–2594. doi: 10.3758/s13428-023-02168-4



## Glossary

<b>IAPS</b> - International Affective Picture System	<b>IAVRS</b> - International Affective Virtual Reality System
<b>IADS</b> - International Affective Digitized Sounds	<b>GEMS</b> - Geneva Emotional Music Scale
<b>VR</b> - Virtual Reality	<b>EmoMusic</b> - Emotionally Labeled Music Database
<b>KAPODI</b> - Searchable Database of Emotional Stimuli Sets	<b>MIAS</b> - Mood Inductive Audio Stories.
<b>DEVO</b> - Database of Emotional Videos from Ottawa	<b>EMOVO Corpus</b> - Italian Emotional Speech Database
<b>EmoMadrid</b> - Emotional Pictures Database for Affect Research	<b>ANEW</b> - Affective Norms for English Words
<b>EPSS</b> - Empathy for Pain Stimuli System	<b>SAM</b> - Self-Assessment Manikin
<b>ESISCA</b> - The Pictorial Set of Emotional Social Interactive Scenarios between Chinese Adults	<b>SCR</b> - Skin Conductance Response
<b>SocialVidStim</b> - Social Video Stimuli	<b>EEG</b> - Electroencephalography
<b>luVR</b> - Library for Universal Virtual Reality Experiments	<b>MEG</b> - Magnetoencephalography
<b>AVDOS-VR</b> - Affective Video Database with Physiological Signals and Continuous Ratings Collected Remotely in VR	<b>fMRI</b> - Functional Magnetic Resonance Imaging
	<b>ERP</b> - Event-Related Potentials