Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 47 (2026), No. 6, pp. 2465–2489
DOI : 10.47974/JIOS-2226

Validation of video stimuli for basic emotion elicitation : A multimodal approach using self-reports and physiological signals

Vilas Machhi *
Apurva Shah †
Department of Computer Science & Engineering
The M S University of Baroda
Vadodara 390001
Gujarat
India

Abstract

Dependable emotion generation is an essential requirement for enhancing affective computing studies, especially for the development and assessment of strong machine learning models. This study tackles the practical issue of finding suitable, already- existing video stimuli that reliably trigger particular emotional responses. Our main aim was to confirm a selected collection of video clips that can evoke the six fundamental emotions (happiness, sadness, anger, fear, disgust, surprise) along with a neutral condition. In this research, 24 individuals from Vadodara, Gujarat, India, watched the chosen video clips. Emotional reactions were assessed via self-report ratings on specific emotions and a 5-point scale for valence-arousal, and confirmed by physiological indicators (GSR and PPG). The findings reveal an overall emotion elicitation accuracy of 77.38% & 69.64% according to self-reports, and a Cohen's Kappa value of 0.736 & 0.645 respectively for Session-1 & Session-2, signifying significant alignment between intended and perceived emotions. Valence-arousal mapping validated anticipated emotional distributions. Additionally, substantial differences ($p<0.001$) were noted in physiological indicators (SCR Peaks, Heart Rate Range, Heart Rate Jumps) among the triggered emotions, offering objective validation. This study recognizes and confirms a collection of impactful video stimuli for inducing emotions, providing a readily accessible resource for the affective computing field without requiring the creation of new datasets, thus simplifying future experimental setups and improving the consistency of emotion-related data gathering.

Subject Classification: *68T01, 68U35, 92C55.*

Keywords: *Affective computing, Emotion elicitation, Video stimuli, Basic emotions, Self-report, Physiological signals, Valence-arousal.*

* *E-mail:* **vilas.machhi-cse@msubaroda.ac.in (Corresponding Author)**
† *E-mail:* **apurva.shah-cse@msubaroda.ac.in**

1. Introduction

Affective computing is an interdisciplinary field dedicated to recognizing, interpreting, processing, and simulating human emotions [1]. It has increasing relevance in human- computer interface (HCI), robotics, education, healthcare, entertainment, and other areas [2, 3]. For instance, computers aware of emotions can facilitate tailored and motivated learning experiences [4], assist with empathy in healthcare professions [5], and support advanced IoMT-based physiological data analysis for anxiety status prediction using cloud computing [32], while also improve engagement in user entertainment [6]. Facilitating intelligent systems requires understanding human emotional states and responding accurately. This raisesthe need to create datasets rich in emotions, which is vital for training and validating the underlying machine learning frameworks [7]. Thus, the methods used to elicit emotions impact the accuracy and dependability of the data collected, and in turn, the robustness and usefulness of the affective computing systems developed [8].

Despite its critical importance, consistently and genuinely eliciting specific human emotions in controlled experimental settings remains a significant challenge [9]. Traditional methods for emotion induction often include exposing participants to static images, playing music, or asking them to recall autobiographical memories [10, 11]. Although these techniques have advantages, they often face drawbacks like insufficient dynamic context, possible subjective interpretation, or issues related to privacy and the reliability of personal recollection [12]. The static nature of images can limit the depth and duration of emotional responses, while music's emotional impact can be highly subjective and culturally dependent [13]. Moreover, purely cognitive elicitation through memory recall can be inconsistent and difficult to verify objectively [14].

Conversely, video stimuli offer a strong and ecologically relevant approach to evoking emotions [15]. Videos provide a vibrant, dynamic, and multisensory experience by combining visual narratives, auditory cues (speech, music, sound effects), and time-based development [16]. This comprehensive sensory stimulus can more closely mimic actual emotional experiences, leading to stronger, enduring, and authentic emotional reactions in people [17]. The structured nature of video content offers greater experimental control and replicability compared to more abstract or internal elicitation methods [18]. Consequently, videos are often considered the ideal format for generating the varied and nuanced emotional information needed by sophisticated affective computing systems [19].

Although the benefits of using video for eliciting emotions are evident, a significant challenge exists in pinpointing and confirming specific video clips that consistently trigger distinct emotional responses [20]. Researchers frequently encounter the challenge of either dedicating substantial resources to develop new, validated video datasets—a labor-intensive and intricate process [21]—or resorting to existing, non- validated clips with unpredictable emotional effects. This study tackles this gap by concentrating on confirming the effectiveness of existing video clips as material for evoking emotions. Our main goal is to discover and verify a collection of video clips that can consistently evoke the six fundamental emotions (Happy, Sad, Angry, Disgust, Fear, and Surprise), in addition to a neutral state, as recognized and reported by human participants via self-reports and corroborated by physiological responses. This research specifically aims to provide a readily usable and validated resource for the affective computing community, without necessitating the creation of a new, large-scale video dataset. For this study, 24 participants viewed and rated the emotional impact of the selected video clips.

The remainder of this paper is organized as follows: Section 2 presents the related work in emotion elicitation and affective computing. Section 3 provides the Foundations of emotions and measurement. Section 4 details the methodology employed for participant recruitment, video stimulus selection, experimental procedure, and emotion measurement. Section 5 presents the results of our validation study, highlighting the efficacy of the identified video clips. Finally, Section 6 discusses the implications, limitations, and future directions of this research, followed by a conclusion in Section 7.

2. Related Work

The area of emotion elicitation is essential to affective computing, supplying the required data for training and assessing systems that can comprehend and react to human feelings. This segment examines current methodologies and key contributions in this field, especially emphasizing video-centric approaches, and underscores the gap that our research seeks to fill.

2.1 *Traditional and Common Emotion Elicitation Methods*

Historically, various methods have been employed to induce emotional states in experimental settings. These methods range from simple static stimuli to complex immersive experiences. Common techniques include:

- Picture viewing: The use of emotionally impactful static images, like those from the International Affective Picture System (IAPS), is a commonly employed approach because of its standardization and user-friendliness [12, 29]. Nevertheless, still images generally provoke fleeting and weaker emotional reactions when contrasted with moving stimuli.
- Music induction: Carefully selected musical pieces have been shown to reliably induce specific emotions or affective dimensions (e.g., happiness, sadness, relaxation) [13, 26]. The effectiveness, however, can be highly subjective and culturally dependent.
- Autobiographical recall: Requesting participants to remember and relive previous emotional experiences can elicit intense, authentic feelings [14]. Although powerful, this approach is challenging to manage among individuals, can provoke privacy issues, and might result in varying emotional responses.
- Olfactory or gustatory stimuli: Specific aromas or flavors can provoke particular feelings, especially revulsion, although their use in extensive emotion research is restricted [30].

These techniques, although useful, frequently struggle to represent the complete complexity and fluid nature of actual emotional experiences, which restricts the ecological validity of the gathered data for training advanced affective computing models.

2.2 *Video-Based Emotion Elicitation and Existing Datasets*

Video stimuli have emerged as a powerful alternative due to their ability to combine visual narratives, auditory cues, and temporal progression, closely mimicking natural emotional encounters [15, 17]. This multimodal richness allows for more profound and sustained emotional induction compared to static images or isolated auditory stimuli. Consequently, numerous research efforts have focused on creating and utilizing video-based emotion elicitation databases for affective computing research:

- Film-clip based datasets: Many studies leverage excerpts from feature films or documentaries to induce emotions [10, 16]. Examples include sets of film clips standardized for emotion induction, often validated through self-report measures [16, 20]. Although effective, the overall accessibility and specific emotional impact of individual segments from commercial movies can be challenging for other researchers to determine or replicate because of copyright and access restrictions.
- Multimodal Affective Databases: A number of extensive databases combine video stimuli with physiological signals and behavioral labels to offer diverse, multi-faceted emotional data. Significant instances consist of:
  - DEAP (Database for Emotion Analysis using Physiological Signals): This dataset includes physiological signals and frontal face video from 32 subjects, who viewed 40 one-minute music videos, assessing their arousal, valence, dominance, and liking [19].
  - MAHNOB-HCI: This dataset contains multimodal data (EEG, eye tracking, audio, video, peripheral physiological signals) from individuals engaging with videos and various stimuli, marked for emotional responses [18].
  - GEMEP (Geneva Multimodal Emotion Portrayals): Although mainly centered on acted portrayals, GEMEP provides a thoroughly validated database of emotional expressions, featuring video content, suitable for elicitation research [21].
  - DREAMER: A multimodal database containing EEG and ECG data captured during emotional elicitation through audio-visual stimuli, along with self- reported ratings for valence, arousal, and dominance [31].

Despite the existence of these valuable resources, a common challenge persists: many existing datasets are either too large for quick, targeted validation, culturally specific, or the precise, individual clip efficacy for eliciting discrete basic emotions (without significant co-occurrence of unintended emotions) is not always transparently reported or rigorously validated in a way that directly supports the selection of a minimal, highly effective set of stimuli.

2.3 *The Contribution of This Research*

Drawing from the groundwork established by earlier studies, our research seeks to meet the practical demand for a concise, thoroughly validated collection of pre- existing video clips selected for their effectiveness in evoking the six fundamental emotions and a neutral state. In contrast to extensive dataset creation initiatives, our emphasis is on delivering an immediately usable and verified resource for researchers needing dependable emotional stimuli for their experiments or model training, without the burden of extensive stimulus development or preliminary testing. Utilizing a dual-modality validation strategy (self-report and physiological signals) within a particular cultural framework, we provide a distinct contribution by pinpointing a collection of videos with established subjective and objective emotional effects, thus facilitating future research in affective computing.

3. Foundations of Emotion and Measurement

3.1. *Basic Emotions*

The study of emotions is a crucial element of psychology and neuroscience, providing valuable understanding for affective computing. Among various theories, the concept of fundamental emotions has achieved significant popularity, largely owing to Paul Ekman's contributions [22]. This theory suggests that specific emotions are basic, common to all human cultures, and have unique physiological, behavioral, and experiential traits. The six fundamental emotions that are widely acknowledged are: happiness, sadness, anger, fear, disgust, and surprise. Every one of these emotions is thought to relate to distinct facial expressions that are similarly recognized and generated among various groups, indicating an inherent biological foundation [23].

Every one of these emotions is thought to relate to distinct facial expressions that are similarly recognized and generated among various groups, indicating an inherent biological foundation [23]. The inclusion of a neutral state is equally crucial, serving as a baseline against which emotional deviations can be measured and ensuring that the system can differentiate between an absence of strong emotion and the presence of a specific one.

3.2 *Emotion Measurement Paradigms*

Measuring human emotion is a complex endeavor, often approached through a multi-component perspective, encompassing subjective experience, behavioral expression, and physiological responses [9, 24].

- Subjective/Experiential Measures: These reflect a person's aware emotional condition. The predominant approach consists of self-report surveys, in which participants directly assess their emotional experiences, usually employing Likert scales to rate intensity or specific emotion identifiers [25]. Instruments such as the Self-Assessment Manikin (SAM) enable individuals to evaluate emotions across continuous scales like valence (pleasant-unpleasant) and arousal (intense- calm) [11]. Although self-reports provide immediate access to personal emotions, they are vulnerable to intentional biases, social expectations, and the difficulty of accurately expressing internal experiences. Regardless of these constraints, self- report continues to be crucial for confirming the perceived emotional influence of stimuli, especially in research centered on stimulus effectiveness.
- Behavioral Measures: This entails monitoring and evaluating visible signs of emotion. Notable instances consist of facial expressions, which can be objectively categorized through systems such as the Facial Action Coding System (FACS) or identified by automated facial recognition software [7]. Vocal elements (e.g., tone, pitch, speech speed) along with body language (e.g., posture, gestures) offer valuable behavioral signals of emotional conditions [26]. Behavioral metrics provide observable information, usually less susceptible to intentional alteration than self-reports, yet their meaning can remain unclear without additional context.
- Physiological Measures: These objective metrics reflect the activity of the autonomic nervous system linked to emotional arousal. Frequently utilized physiological signals consist of Electrocardiography (ECG) for monitoring heart rate and heart rate variability, Galvanic Skin Response (GSR) or electrodermal activity (EDA) for alterations in sweat gland function, and Electroencephalography (EEG) for recording brain electrical activity [19]. These methods offer immediate, subconscious signals of emotional strength and value. Although obtaining them necessitates specialized tools and correlating them to specific emotions can be intricate, they provide a valuable, objective addition to subjective accounts for comprehensive emotion evaluation.

For the purpose of this study, which aims to validate the perceived emotional impact of video stimuli, self-report measures were chosen as the primary method of assessment. Physiological signals were recorded and analysed to provide a more thorough validation, offering objective confirmation of the induced emotional states.

4. Methodology

This segment describes the experimental framework and methods used to assess the emotion-evoking potential of existing video stimuli through self-report assessments and physiological signal monitoring involving 24 participants.

4.1 *Participant Recruitment*

A total of 24 participants (12 males, 12 females), aged between 18 and 30 years (M = 23.5, SD = 2.8), were recruited from the local university community in Vadodara, Gujarat, India, through voluntary participation. All participants reported normal vision and hearing, and no history of neurological or psychological disorders. Before taking part, informed consent was acquired from every participant, detailing the study's objectives, methods, possible risks, and their entitlement to exit at any moment without consequence.

4.2 *Video Stimulus Selection Process*

The selection process for video stimuli focused on identifying short, impactful clips from publicly available sources likely to evoke one of the six basic emotions (happiness, sadness, anger, fear, disgust, surprise) or a neutral emotion. The main objective was to confirm the emotional effectiveness of these clips, rather than to develop a new dataset.

A preliminary selection of around 50 video clips (lasting between 1 and 3 minutes) was assembled. These clips originated from various content types, including film trailers, brief documentary segments, popular online videos, and nature scenes. Important factors for the first selection involved:

- Distinct Emotional Elements: Videos were selected for their compelling storyline and visual/auditory signals that represent a particular target emotion.
- Limited Confounding Variables: Preference was given to clips featuring sparse dialogue, low cultural specificity (when feasible for universal feelings), and distinct emotional trajectories to minimize the chances of mixed or unclear emotional reactions.

- Technical Quality: Each chosen video possessed adequate resolution and clear sound.

The ultimate collection for the experiment included 14 distinct video clips: two clip representing each of the six primary emotions (happiness, sadness, anger, fear, disgust, surprise) and two clips for the neutral condition. This decision aimed to identify specific, significant instances for every targeted emotion. Details regarding these video stimuli are available in Table 1.

Table 1

Details on Validated Video Stimuli

Video ID	Intended Emotion	Original Source (Title, Year, Creator / Platform)	Duration	Key Emotional Content Description
Session-1				
V1	Happy	Dhammal (2007), Comedy Film, "Jaldi Ka Kaam Shaitan Ka" scene	0:01:02	A funny and cheerful segment showcasing comedic lines and gestures meant to provoke laughter.
V2	Sad	YouTube (2022), "Son Leaving His Mother in an Old Age Home"	0:03:19	A touching and heartfelt moment showing a son parting with his aged mother at a Old age house, accompanied by somber music and faces of sorrow.
V3	Fear	YouTube (2020), "Horror scene from Diet"	0:02:18	A gripping and thrilling horror snippet featuring sudden scares and unsettling sound effects, aimed at provoking emotions of fear.
V4	Anger	YouTube (2021), "Triggering OCD part1"	0:02:18	A segment showing events or actions that are typically viewed as aggravating or bothersome, aimed at provoking anger or irritation.

Contd...

V5	Surprise	YouTube (2019), "Japanese people precision waking pattern"	0:00:30	A brief, unforeseen segment highlighting an unusual or exact action, aimed at generating authentic amazement and curiosity.
V6	Disgust	YouTube (2023), "Cleaning an OPEN diabetic ulcer"	0:02:01	A disturbing and graphic video depicting a medical procedure on an open sore, designed to provoke intense feelings of disgust.
V7	Neutral	YouTube (2018), "Serene Duck on a Quiet Pond"	0:01:00	A serene and simple video showcasing a duck gliding smoothly on a quiet pond, accompanied by soothing sounds, intended to inspire a neutral emotional response
Session-2				
V1	Happy	YouTube (2019), "Babies Laughing"	0:01:43	Best Babies Laughing Video Compilation
V2	Sad	YouTube (2017), "Searching for Mother: A Lonely Cub's Journey"	0:02:30	A small child bear is seen wandering alone through a dense forest, his movements slow and uncertain. He pauses often, sniffing the air and looking around, clearly searching for his missing mother.
V3	Fear	YouTube (2016), "The moonlight man"	0:02:11	A young woman finishes her tasks and walks alone toward her parked car in a quiet, dimly lit parking garage. As she approaches her car, she senses something strange—a subtle movement in the shadows.

Contd...

V4	Anger	YouTube (1993), "Schindler's List Immolation Scene"	0:02:09	A barren, snow-covered field becomes the site of horror. Nazi soldiers begin burning the bodies of murdered Jews, their lifeless forms stacked in massive piles. Thick black smoke rises into the pale sky, while flames consume the remains.
V5	Surprise	Up (2009)," flying house scene"	0:01:30	Carl ties thousands of balloons to his house, and it lifts off the ground, soaring into the sky. As the house floats peacefully above the city, Carl begins his long-dreamed adventure, leaving behind his old life in a magical, uplifting moment.
V6	Disgust	YouTube (2006), "Disgusting video compilation"	0:01:55	A severely infected fingernail is removed, releasing thick pus. The scene then shows an internal organ with decaying tissue and bacterial infection, oozing with fluid and rot.
V7	Neutral	YouTube (2005), "Meditation Music"	0:01:00	A tranquil 5minute meditation clip featuring serene ocean visuals set to calming ambient music.

4.3 *Experimental Procedure*

To keep outside distractions to a minimum, participants did the experiment alone in quiet, dimly lighted environment. Every participant was positioned roughly 60 cm away from a 24-inch screen and equipped with noise-canceling headphones to guarantee uniform audio transmission. The experimental software, created with a tailored Python script, managed the display of video stimuli. Following the video viewing, the participant must complete a self-report ratings form, and while watching the video, the participant is required to wear a wearable device to gather their physiological data.

To ensure comprehensive data collection and manage potential fatigue, the experiment was structured into two separate sessions. Each session consisted of 7 distinct video clips, drawn from the initial pool of 50 selected clips. The videos within each session were presented in a fully randomized order for every participant. The order of the two sessions was also counterbalanced across participants to mitigate any potential order effects.

The procedure for each participant was as follows:

1. Instructions: Participants were told to focus on the videos and to assess their own experienced emotions right after each clip. The focus was on sharing their true emotional experience rather than the feelings displayed by the characters in the video.
2. Video Presentation: The 14 video clips were shown in a fully randomized sequence for every participant to mitigate possible order effects, fatigue, or desensitization.
3. Post-Video Evaluation: Right after every video, the display shifted to a rating interface for self-assessment.
4. Inter-Trial Interval: A 30-second black screen appeared between consecutive video presentations. This break was introduced to help emotional conditions revert nearer to a baseline, minimizing carry-over impacts from one video to another.

4.4. *Emotion Elicitation Measurement*

Emotional responses were captured through two modalities: self-report questionnaires and physiological signal recording.

4.4.1. Self-Report Measures

Following every video segment, participants were asked to evaluate their present emotional condition using a specific scale for Valence-Arousal ratings and to identify their specific emotion:

- Discrete Emotion Rating: We have utilized seven target emotions: Happiness, Sadness, Anger, Fear, Disgust, Surprise, and Neutral. The participant must identify their specific emotion.
- Valence-Arousal Evaluation: Participants assessed their emotional reactions utilizing a 5-point Likert scale for valence (1 = negative, 5

= positive) and arousal (1 = relaxed, 5 = stimulated). This offered a dimensional viewpoint on the evoked feelings.

4.4.2. Physiological Signal Recording

Physiological data were consistently monitored during the experiment to obtain objective signs of emotional arousal. GSR (Galvanic Skin Response) and PPG (Photoplethysmography) signals were obtained through a wearable device that incorporated both a GSR sensor and a MAX30102 sensor. The unprocessed physiological signals were refined to identify essential characteristics linked to emotional conditions:

- SCR Peaks (GSR): Count of Skin Conductance Response peaks, representing activation of the sympathetic nervous system and levels of arousal.
- Heart Rate Range (PPG): The variability range of heart rates during the video clip, indicating the activity of the autonomic nervous system.
- Heart Rate Spikes (PPG): The occurrence of notable rises in heart rate, possibly signifying stress, shock, or enthusiasm.

5. Results

This section details the extensive results from the emotion elicitation validation study, examining both the self-report ratings and physiological responses of the 24 participants for the chosen video clips. The goal was to find out which video clips best and most precisely brought forth the primary emotions of happiness, sadness, anger, fear, disgust, surprise and a neutral condition.

5.1. *Emotion Elicitation Accuracy (Self-Reports)*

We examined participant self-reports to determine the efficacy of the selected video stimuli in eliciting target emotions. Twenty-four participants viewed both the session consist of 14 video clips designed to evoke emotions (two for each of Happy, Sad, Fear, Anger, Surprise, Disgust, Neutral), generating 336 responses (24 participants × 14 clips). Every participant indicated their felt emotion following each clip.

The accuracy of the elicitation was determined by comparing each video's intended emotion with the highest emotion reported by the participant. Table 2 & 3 displays the confusion matrix for session-1 and session-2 obtained from all participant responses, showing the frequency with which, each intended emotion was accurately recognized as the primary felt emotion and how frequently it was mistaken for other emotions.

Table 2

Confusion Matrix (Intended vs Reported Emotions) for Session-1

Intended \ Reported	Happy	Sad	Fear	Anger	Surprise	Disgust	Neutral	Accuracy (%)
Happy	24	0	0	0	0	0	0	100%
Sad	1	23	0	0	0	0	0	95.83%
Fear	0	0	17	2	2	0	3	70.83%
Anger	4	1	1	12	1	1	4	50%
Surprise	2	0	0	0	15	0	7	62.5%
Disgust	0	1	1	0	0	19	3	79.16%
Neutral	4	0	0	0	0	0	20	83.33%

Table 3

Confusion Matrix (Intended vs Reported Emotions) for Session-2

Intended \ Reported	Happy	Sad	Fear	Anger	Surprise	Disgust	Neutral	Accuracy (%)
Happy	24	0	0	0	0	0	0	100%
Sad	0	20	2	0	2	0	0	83.33%
Fear	0	0	16	0	3	1	4	66.66%
Anger	0	10	3	7	0	2	2	29.16%
Surprise	12	0	0	0	10	0	2	41.66%
Disgust	0	0	1	1	0	21	1	87.5%
Neutral	5	0	0	0	0	0	19	79.16%

The total precision for all emotions was determined by adding the correctly identified emotions (diagonal values) and dividing by the total responses:

Overall Precision for Session-1 = (24 + 23 + 17 + 12 + 15 + 19 + 20) / 168 ≈ 77.38% and Overall Precision for Session-2 is 69.64%.

This outcome reflects a strong alignment between the desired emotional message of the videos and the participants' self-reported feelings.

5.2 *Cohen's Kappa Agreement Score*

We used Cohen's Kappa (κ) to assess how well participants' intended emotions matched their self-reported feelings, taking into account random chance. The estimated score is: $\kappa = 0.736$ and $\kappa = 0.645$ for session-1 and session-2.

Based on Landis and Koch's well-known scale, this score indicates a high level of agreement [27]. This additionally reinforces the dependability and consistency of the video-based elicitation protocol, validating that participants' reactions were not random but significantly matched with the intended emotional categories.

5.3 *Valence–Arousal Mapping*

Participants assessed their emotional reactions to video stimuli utilizing a 5-point Likert scale for valence (1 = unpleasant, 5 = pleasant) and arousal (1 = calm, 5 = excited). Table 4 shows the mean ± standard deviation of valence and arousal ratings for each emotion, averaged across all participants for the respective video clip.

Table 4 shows the mean ± SD for valence and arousal ratings (scale: 1-5).

Table 4

Emotion	Mean Arousal (SD)	Mean Valence (SD)
Happy	4 (0.88)	4.375 (0.49)
Sad	3 (1.14)	2.79 (0.97)
Angry	3.05 (0.99)	1.83 (1.09)
Fear	3.54 (0.93)	1.62 (0.71)
Disgust	3.29 (0.99)	1.45 (0.72)
Surprise	3.75 (0.89)	4 (0.72)
Neutral	2.25 (0.84)	3.75 (0.94)

These findings confirm that the video clips successfully elicited emotional states consistent with the circumplex model of affect [28], which depicts emotions in a 2D framework defined by valence and arousal.

- Happy and Surprise emotions exhibited strong scores in both valence and arousal, indicating positive and vigorous affective conditions.
- Fear, anger, and disgust had low valence scores but moderate arousal levels, suggesting they are unpleasant yet stimulating experiences.

- Sadness had low valence and moderate arousal, characteristic of low-energy negative feelings
- Neutral was rated moderate-to-high on valence and low on arousal, supporting its intended emotionally stable tone.

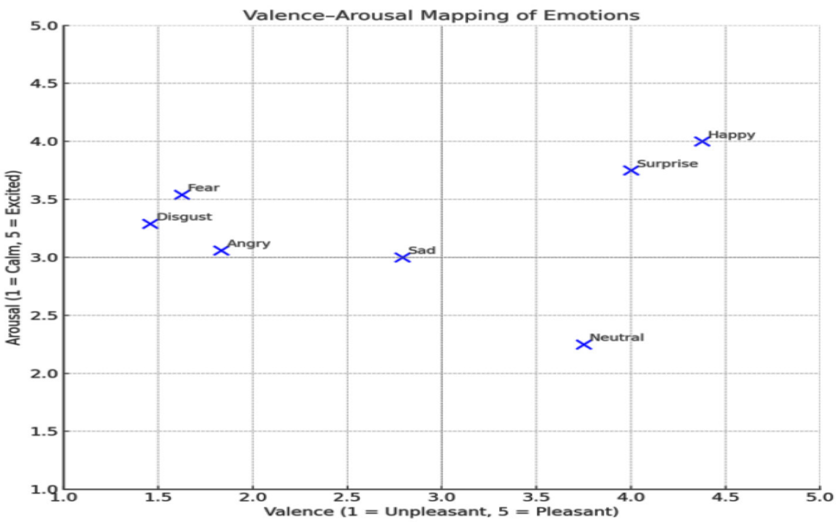

Figure 1

2d plot of Valence-Arousal mapping

Figure 1 illustrates the visual representation of how each emotion is placed in the valence-arousal space according to participant ratings.

5.4 *Physiological Signal Responses*

We analyzed three primary biosignal features obtained from the recorded GSR and PPG signals to investigate physiological changes linked to various emotional states: SCR Peaks (GSR), Heart Rate Range (PPG), and Heart Rate Jumps (PPG). Everyfeature was evaluated within the seven main emotion categories: Happy, Neutral, Disgust, Angry, Fear, Sad, and Surprise.

These physiological metrics under different emotional states will be illustrated using boxplots in the provided Figures 2-4.

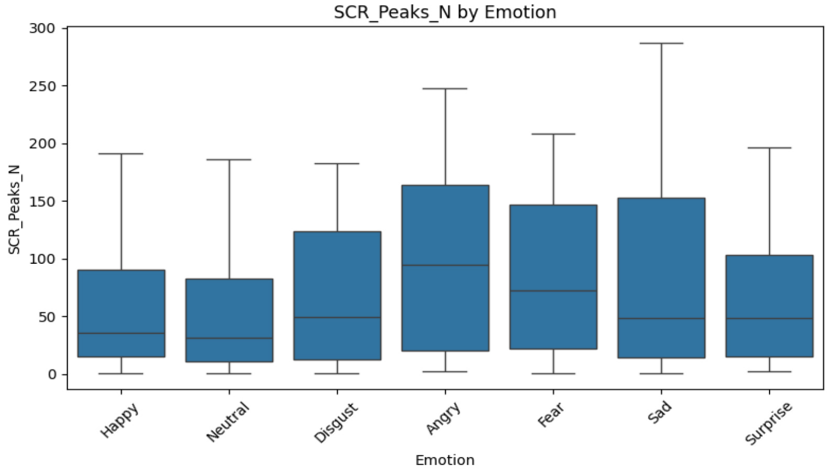

Figure 2

Illustrates that Angry, Fear, and Sad provoked greater SCR peak counts than Neutral and Happy, signifying increased sympathetic arousal.

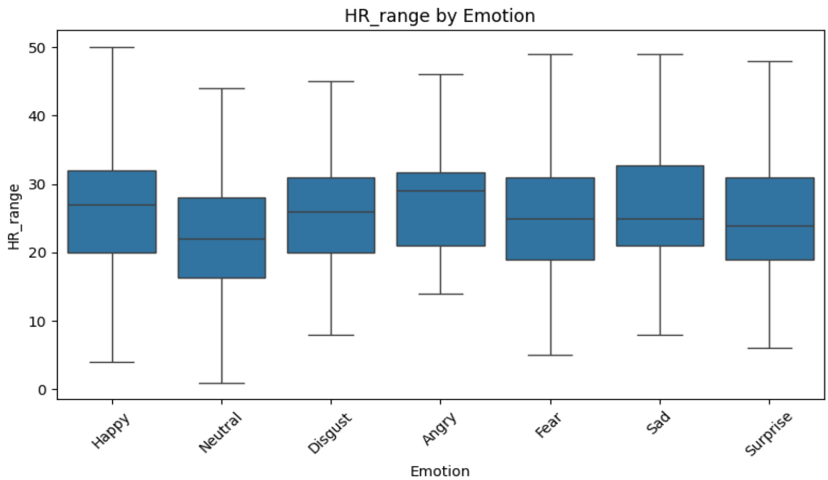

Figure 3

Shows the range of heart rates. Emotions such as Anger and Sadness showed greater HR variability, potentially indicating activation of the autonomic nervous system.

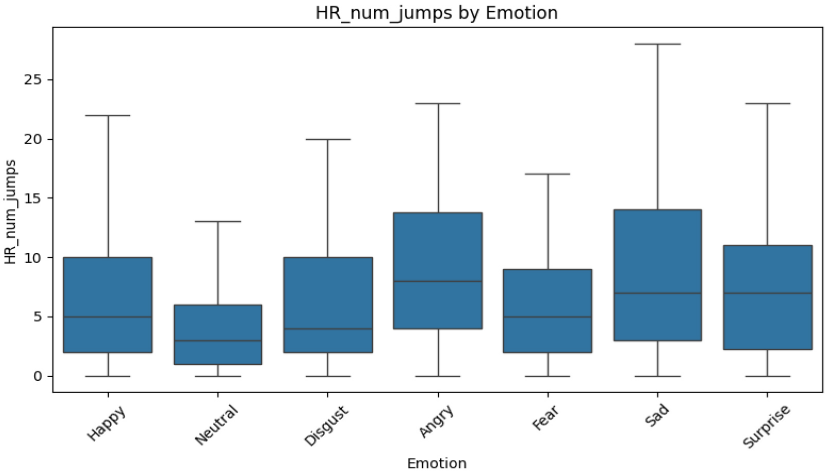

Figure 4

Displays the frequency of heart rate increases. Feelings like Sadness, Anger, and Surprise exhibited the most frequent HR spikes, indicating reactions of stress or excitement.

Statistical Analysis

A one-way ANOVA was conducted for each feature to statistically confirm if physiological responses vary significantly across emotions. Table 5 displays the F- statistics and p-values derived from these analyses.

Table 5

Results of One-Way ANOVA for Physiological Characteristics

Feature	F-Statistic	p-value
SCR_Peaks_N	9.82	1.59e-10
HR_range	6.06	3.00e-06
HR_num_jumps	7.46	7.97e-08

Table 5 demonstrates that the p-values for every feature are well below the typical significance level ($p<0.05$), suggesting that the variations in physiological responses among emotions are statistically meaningful. This offers objective proof supporting the subjective self-reports and verifying that the video stimuli triggered quantifiable physiological alterations aligned with specific emotional states.

6. Discussion

This research thoroughly examined emotion triggering through multimedia stimuli, evaluating both self-reported and physiological reactions. The main conclusions are outlined as follows:

- Emotion Elicitation Accuracy: The validation study demonstrated consistent efficacy in emotion elicitation, with an Overall Precision of 77.38% for Session 1and 69.64% for Session 2 (Cohen's κ=0.736 & κ=0.645). While both sessions achieved a substantial level of agreement between intended and self-reported emotions, the slight decrease in precision in Session 2 warrants further consideration. This difference could potentially be attributed to factors such as participant fatigue, habituation to the experimental procedure, or inherent variations in the emotional intensity of the video clips presented in the second session compared to the first. Nevertheless, the results from both sessions collectively confirm the effectiveness of the selected video stimuli in reliably eliciting their target emotions, providing a valuable resource for future affective computing research.
- Valence–Arousal Distribution: Emotions were effectively represented in the 2D valence–arousal framework using a 5-point scale. As anticipated, feelings such Happy and Surprise occupied the high valence–high arousal quadrant, while Sad, Fear, and Disgust clustered around low valence–moderate arousal regions. Neutral focused on moderate valence and low arousal, validating its intended emotionally steady tone.
- Physiological Responses: Notable distinctions were noted in physiological indicators among various emotions. SCR Peaks, Heart Rate Range, and Heart Rate Jumps showed considerable variation ($p<0.001$) across different emotions, as validated by one-way ANOVA.
    - Anger, Fear, and Sadness produced increased GSR responses (more SCR peaks).
    - Sadness and anger exhibited greater heart rate variations and an increased number of heart rate fluctuations, indicating activation of the sympathetic nervous system.
- Validation of Stimuli: The video stimuli employed were confirmed to be effective at provoking emotions. Their design yielded uniform

subjective and physiological reactions that matched the intended emotional states, offering strong dual-modality validation.

The overall analysis of self-reports and physiological data confirms the strength of the selected stimuli and endorses the multimodal strategy for emotion recognition in affective computing studies.

6.1 *Limitations of the Study*

Even with its significant contributions, this research has multiple limitations that deserve attention:

- Dependence on Self-Report as Main Source and Physiological as Supportive Evidence: Although our research combined both self-report and physiological assessments, the in-depth examination for specificity mainly centered on self-reports. Although physiological data offered objective validation of arousal levels, a deeper investigation into the distinct physiological signatures of individual emotions (beyond general arousal) was not the main emphasis. Future research might investigate the inter-modal connection in greater depth.
- Sample Size and Demographics: The research included a modest group of 24 participants from a particular cultural setting (Vadodara, Gujarat, India). Although this ensured uniformity within the sample, the applicability of these results to larger populations, especially those from diverse cultures, might be constrained. Cultural factors can shape emotional reactions to specific stimuli [20].
- Particularity of Video Selection: Although attempts were made to choose clips rich in emotional content, the inherent characteristics of existing video material make it difficult to fully isolate a singular emotion. As noted with anger and surprise (refer to Table 2 & 3), videos frequently provoke mixed or simultaneous emotions, showcasing the intricacy of actual emotional experiences. This research did not examine the subtleties of these mixed feelings thoroughly.
- Insufficient Evaluation of Long-Term Effects: The research concentrated on short-term emotional reactions. It did not evaluate the length of the induced emotions or the possibility of habituation with repeated encounters, which might be important for long-term experimental designs.

- Restricted Influence on Video Material: Our method utilized existing videos instead of creating tailored emotional stimuli as seen in other research. This resulted in reduced detailed control over particular visual, auditory, and narrative aspects that could either enhance or hinder emotional elicitation.

6.3 *Future Directions*

Leveraging the findings from this validation study, multiple potential directions for future research come to light:

- Enhanced Multimodal Examination: Future research should delve deeper into the intricate relationship between self-reported feelings and physiological reactions. This may include machine learning methods to link physiological characteristics to specific emotions, or explore the time-based patterns of emotional reactions across different modalities.
- Cultural Replication and Validation: Conducting this research again with larger and more diverse groups of participants from different cultural backgrounds is crucial to evaluate the general applicability or cultural particularity of the emotional effects of these video clips. This would be vital for creating affective computing systems that can be applied worldwide.
- Identifying Complex and Blended Emotions: Future studies may focus on recognizing and validating video stimuli that represent more intricate and blended emotional states (e.g., awe, empathy, nostalgia, contempt), which are becoming more pertinent in advanced HCI contexts.
- Dynamic and Adaptive Elicitation: Exploring methods for real-time adjustments of video content based on a user's current emotional condition may result in more compelling and tailored affective experiences. This would entail closed-loop systems in which emotion recognition informs stimulus presentation, possibly even creating content.
- Ethical Concerns and Privacy: As technologies for eliciting and recognizing emotions progress, forthcoming studies must consistently emphasize ethical concerns, including data privacy, the potential abuse of emotion recognition, and ensuring the welfare of participants, particularly when addressing sensitive emotional topics.

- AI-Driven Content Analysis for Elicitation: Creating tools powered by AI to examine video content and forecast its emotional influence could significantly help in the automated choice or even creation of impactful emotion-evoking stimuli, minimizing manual curation work.

7. Conclusion

This study presented a structured review of pre-existing video stimuli for consistently eliciting the six essential emotions (happiness, sadness, anger, fear, disgust, and surprise) and a neutral condition in human subjects. Using a detailed experimental design with 24 participants and self-report measurements, we successfully identified certain video parts that dependably induced their intended emotional reactions with great precision. The findings demonstrate the usefulness of selected video snippets as fast and simply accessible approaches for eliciting emotions in affective computing investigations. This study provides a confirmed set of stimuli that aids in the creation of superior emotional datasets, improves the assessment of affective systems, and simplifies experimental frameworks for future studies. Although there are limitations mainly tied to dependence on self-reporting and the demographics of the sample, this study offers a valuable resource and establishes a foundation for more thorough, multimodal, and culturally varied inquiries into video-based emotion elicitation, thus progressing the area of affective computing.

References

[1] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press (1997).

[2] J. Tao and T. Tan, “Affective computing: A review,” in *Proc. Int. Conf. Machine Learning and Cybernetics*, vol. 1, pp. 151–156 (2005).

[3] R. A. Calvo and S. K. D’Mello, “Affective computing and education: Learning about emotions,” *IEEE Trans. Affective Computing*, vol. 1, no. 1, pp. 18–30 (2010).

[4] B. Kort, R. Reilly, and R. W. Picard, “An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion,” in *Proc. IEEE Int. Conf. Advanced Learning Technologies (ICALT)*, pp. 6–9 (2001).

[5] K. Hone and J. Patrick, "Affective computing for mental health," *Int. J. Human-Computer Studies*, vol. 70, no. 6, pp. 337–340 (2012).

[6] J. Gratch and S. Marsella, "A social appraisal framework for emotion," *Autonomous Agents and Multi-Agent Systems*, vol. 8, no. 3, pp. 173–197 (2004).

[7] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1424–1445 (2000).

[8] K. R. Scherer, "Psychological models of emotion," in *The Neuropsychology of Emotion*, J. C. Borod, Ed. Oxford, U.K.: Oxford Univ. Press, pp. 137–162 (2000).

[9] J. A. Coan and J. J. B. Allen, *Handbook of Emotion Elicitation and Assessment*. Oxford, U.K.: Oxford Univ. Press (2007).

[10] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition & Emotion*, vol. 9, no. 1, pp. 87–107 (1995).

[11] M. M. Bradley and P. J. Lang, *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*. Gainesville, FL, USA: Center for Research in Psychophysiology, Univ. of Florida (1999).

[12] G. Stemmler, M. Heldmann, A. Pauls, and K. R. Scherer, "Constraints on emotion elicitation by affective pictures," *Psychophysiology*, vol. 38, no. 2, pp. 295–306 (2001).

[13] S. Vieillard, I. Peretz, and P. Boucher, "A study of the emotional impact of music on physiological arousal," *Psychophysiology*, vol. 47, no. 4, pp. 629–638 (2010).

[14] H. C. Lench, S. A. Flores, and S. W. Bench, "Discrete emotions predict behavior and judgment: A meta-analytic review," *Psychological Bulletin*, vol. 137, no. 5, pp. 834–855 (2011).

[15] P. Philippot, "The effect of emotion on the production of facial expression: A review of the literature," *Int. J. Psychophysiology*, vol. 14, no. 3, pp. 171–182 (1993).

[16] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot, "Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion research," *Behavior Research Methods*, vol. 42, no. 4, pp. 932–942 (2010).

[17] J. Rottenberg, R. D. Ray, and J. J. Gross, "Emotion elicitation using films," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan

and J. J. B. Allen, Eds. Oxford, U.K.: Oxford Univ. Press, pp. 9–28 (2007).

[18] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition in naturalistic environments," *IEEE Trans. Affective Computing*, vol. 3, no. 1, pp. 47–63 (2012).

[19] S. Koelstra *et al.*, "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affective Computing*, vol. 3, no. 1, pp. 18–31 (2012).

[20] J. Lu, C. Ma, and J. Zhang, "Challenges and advancements in emotion recognition," *J. Visual Communication and Image Representation*, vol. 45, pp. 107–118 (2017).

[21] J. Wagner, K. R. Scherer, and T. Bänziger, "The Geneva Multimodal Emotion Portrayals (GEMEP) corpus: An emotion-eliciting video database," *IEEE Trans. Affective Computing*, vol. 5, no. 2, pp. 177–189 (2014).

[22] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, vol. 6, no. 3–4, pp. 169–200 (1992).

[23] P. Ekman and W. V. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA, USA: Consulting Psychologists Press (1978).

[24] I. B. Mauss and M. D. Robinson, "Measures of emotion: A review," *Cognition & Emotion*, vol. 23, no. 2, pp. 209–237 (2009).

[25] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The PANAS scales," *J. Personality and Social Psychology*, vol. 54, no. 6, pp. 1063–1070 (1988).

[26] P. N. Juslin and K. R. Scherer, Eds., *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford, U.K.: Oxford Univ. Press (2008).

[27] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174 (1977).

[28] J. A. Russell, "A circumplex model of affect," *J. Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178 (1980).

[29] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*. Gainesville, FL, USA: Univ. of Florida (2008).

[30] E. T. Rolls, *Emotion and Decision-Making Explained*. Oxford, U.K.: Oxford Univ. Press (2014).

[31] C. D. Katsis, Y. Goletsis, and D. I. Fotiadis, "An integrated approach for emotion recognition using physiological signals and self-assessment," *IEEE Trans. Information Technology in Biomedicine*, vol. 15, no. 3, pp. 370–377 (2011).

[32] B. Ramani, W. Patel, and K. Solanki, "Stress Ocare: An advance IoMT based physiological data analysis for anxiety status prediction using cloud computing," *J. Discrete Mathematical Sciences and Cryptography*, vol. 25, no. 4, pp. 1019–1029 (2022). doi: 10.1080/09720529.2022.2072426.

Received October, 2025