# Asian Emotional Body Movement Database: Diverse Intercultural E-Motion Database of Asian Performers (DIEM-A)

Miao Cheng[1], Chia-huei Tseng[1],Ken Fujiwara[2]
Victor Schneider[1], Yoshifumi Kitamura[1]
[1]*Tohoku University, Japan*
[2]*National Chung Cheng University, Taiwan*
cheng@tohoku.ac.jp, tseng@riec.tohoku.ac.jp, psykf@ccu.edu.tw
schneider.victor.pierre.d3@tohoku.ac.jp, kitamura@riec.tohoku.ac.jp

***Abstract*—In this paper, we introduce the Diverse Intercultural E-Motion Database of Asian Performers (DIEM-A), high-quality motion-capture data from 97 Asian professional performers (54 Japanese, 43 Taiwanese). Each performer enacted 12 emotions across 3 self-created scenarios at 3 intensity levels, plus 3 additional neutral scenarios, resulting in a total of 10,767 motion recordings accompanied by videos and textual scenario descriptions. Preliminary evaluations indicated robust overall emotion recognition accuracy and highlighted how scenario context, performer styles, and cultural factors jointly influenced emotional understanding from body motions. We discuss potential interdisciplinary applications of DIEM-A in affective computing, cognitive and social neuroscience, cross-cultural comparison studies, and human-computer interaction by offering rich, ecologically valid data on expressive bodily behavior.**

***Index Terms*—database, emotion, body movement, motion capture, point-light animation, cross-culture**

## I. Introduction

Body movement has become a crucial channel for emotion recognition in affective computing community [1]–[3], as it can convey affective cues that often complement or exceed those expressed through facial and vocal indicators [4]. Subtle variations in posture, kinematics, spatial extent, and temporal patterns can reveal an individual's affective state, illuminating how emotions unfold through everyday actions such as walking, dancing, or playing sports [5]–[8]. However, these cues are inherently challenging to decode because body movements primarily serve functional purposes (e.g grabbing a cup), which make emotional signals less visible and more context-dependent [9]. As a result, this field requires diverse, well-annotated datasets that capture a wide range of cultural backgrounds, personal movement styles, and real-life contexts to train robust recognition algorithms [10]–[13].

In parallel, the motion generation domain has an increased need for emotion-rich datasets. While many publicly available databases offer functional and task-driven movements, there is a relative scarcity of sequences portraying nuanced emotional states [14]. The emerging research on the development of generative models that synthesize authentic, context-aware emotional movements [15] demands quality data capturing the subtle, context-dependent nature of bodily expressions. To advance and grow this field, larger and comprehensive datasets that accurately reflect the complex interplay between motion, emotion, and context is urgently needed [14].

Current emotion-focused body movement databases have a few well-noted limitations. First, many were created to associate with goal-oriented or functional actions, far from the complex movements in actual natural lives. For example, the Emilya database includes emotional expressions in 7 actions (e.g. walking, sitting, and lifting objects), providing a means to compare affective cues embedded in a single action type [8], but it is difficult to construct a natural daily movement sequence with these simple action types only. Secondly, the movements do not always include the whole body. For example, the MPI Emotional Body Expressions Database situates emotions within story-reading narratives by sitting participants [16]. It constraints our understanding to upper body motion only. Thirdly, the same performing scenarios were usually provided to all performers by the experimenters. For example, GEMEP-CS records both bodily cues and vocal components to show how vocalization and movement jointly convey affect [17]. For each recorded affective state, the same 3 scenarios were used for all 10 performers. A more recent database by Zhang et al. [18] collected 22 performers enacting 5 highly recognizable scenarios for 6 basic emotions. In daily situations, affective expressions are common and flow dynamically. A small set of predefined contexts alone is unlikely to represent the richness of affective expressions.

More importantly, most existing databases rely on homogeneous actor pools from Europe and the USA. Research indicates that cultural background and personal experience shape emotional expression and responses [19]–[23], so people may interpret and react differently even to identical situations.

This study was supported by the New Energy and Industrial Technology Development Organization (NEDO [JPNP21004]).

979-8-3315-8015-5/25/$31.00 ©2025 IEEE

TABLE I
OVERVIEW OF DIVERSE INTERCULTURAL E-MOTION DATABASE OF ASIAN PERFORMERS

| Category | Details |
|---|---|
| **Performer Number** | 97 |
| **Performer Distribution** | 54 from Japan, 43 from Taiwan |
| **Takes per performer** | 12 emotions × 3 scenarios × 3 intensity levels + Neutral × 3 scenarios = 111 |
| **Data Format** | Motion capture data: bvh, fbx<br>Video<br>Text: scenario descriptions in English (translated from Japanese and Chinese) |

Therefore, western-centered databases alone may not capture real-world affective diversity. More comprehensive databases are needed to integrate cross-cultural perspectives and personalized induction methods, reflecting how individuals from different backgrounds experience and express emotion.

To fill in the above-mentioned gaps, we propose a new database that emphasizes personalization and cultural diversity. We invited performers to create their own scenarios for each emotion. For multi-cultural comparison, we focus on two sub-Asian populations located in different regions but shared historical and cultural connections: Japan and Taiwan. This database is able to capture subtle expressive variations, contributing to both affective computing applications and theoretical insights into emotional behavior across cultures.

## II. DIEM-A Database Design and Construction

### A. Performers

97 professional performers (50 female, 47 male) participated in motion capture sessions. Their ages ranged from 18 to 69 years (M = 35.3, SD = 13.0), and their performance experience ranged from 1.5 to 50 years (M = 15.5, SD = 11.1).

The performers were recruited from two locations: 54 individuals participated at Tohoku University (TU) in Japan (30 females, 24 males, age 19 to 69 years, mean age 38.4 years), while 43 were involved at National Chung Cheng University (CCU) in Taiwan (20 females, 23 males, age 18 to 60 years, mean age 31.4 years). Recruitment was facilitated through local acting agencies, and all performers were compensated at a professional rate. Each performer signed an informed consent form, acknowledging that their recordings would be used for research. The study received ethical approval from Tohoku University and National Chung Cheng University.

To facilitate effective communication and ensure that performers could express themselves comfortably, the sessions were conducted in their native languages—Japanese for those in Japan and Mandarin for those in Taiwan.

Note on performer selection: We carefully considered whether to recruit amateur or professional performers. We conducted pilot tests with university students and found that amateurs often struggled to portray emotions naturally, resulting in stereotyped and overly simplified performances. In contrast, professional performers were more comfortable conveying emotion authentically in front of the camera. To achieve natural and spontaneous expression often requires years of training. Since our goal is to capture expressive behaviors that closely resemble real-life emotion, we decided to recruit professional performers.

### B. Target Emotions

We included 13 emotions: 7 basic emotions (Joy, Sadness, Anger, Surprise, Fear, Disgust, Contempt), 5 social emotions (Gratitude, Guilt, Jealousy, Shame, Pride), and Neutral. The 7 basic emotions are defined by Ekman and Friesen [24], [25] and are widely recognized as fundamental and universal emotional states [26], [27]. Beyond these basic emotions, we included 5 social emotions: gratitude, guilt, jealousy, shame, and pride. These social emotions were selected based on their prevalence in the literature and their relevance to everyday life scenarios [28]–[30].

### C. Emotion Elicitation and Scenario Preparation

To effectively elicit the target emotions, we implemented a structured protocol to prepare personalized scenarios to provoke each emotion. Before the motion capture session, we provided definitions of the 13 target emotions (see Table II) in the performers' native languages and instructed them to create three distinct scenarios that would evoke each emotion. Performers had full freedom to create their own scenarios. We encouraged performers to send scenarios to us at least one day before the motion capture day, so that they have time to thoughtfully prepare and to rehearse if necessary. Since no director was guiding the acting on the motion capture day, preparation would help ensure smooth performances. In total, each performer prepared 39 scenarios (13 emotions × 3 scenarios). For example, a performer's scenarios for joy included:

- Eating delicious food
- Seeing an old friend for the first time in a long while
- Having intimate physical contact with my partner

All scenarios were prepared in the performers' native languages. We informed the performers that they would enact each scenario at low, medium, and high intensity levels, allowing them the opportunity to rehearse as needed. For the neutral state, they performed at a single intensity level. In total, each performer recorded (12 emotions × 3 scenarios × 3 intensity levels) + (Neutral × 3 scenarios × 1 intensity levels) = 111 performances. This structured preparation ensured that performers could authentically express a wide range of emotions while also accommodating variations in intensity.

### D. Motion capture set-up

At both TU and CCU, we established a Vicon motion capture system (Vicon Motion Systems Ltd., UK) in conjunction with Vicon Shogun Live 1.7 software. This system consisted of 12 optical cameras (Vicon Vero X) positioned to cover a recording area of 4.7 m x 5.8 m at TU (Fig. 1A) and a smaller area of 2.5 m x 3.7 m at CCU (Fig. 1B).

On the day of the motion capture sessions, performers followed the COVID-19 infection control measures, including

TABLE II
EMOTION LIST WITH DEFINITIONS IN ENGLISH, JAPANESE, AND CHINESE

| **Emotion** | **English Definition** | **Japanese** (日本語) | **Chinese** (中文) |
|---|---|---|---|
| Joy | Experiencing an extraordinary feeling of pleasantness, well-being, and sensual delight. | 喜び: 特別な心地よさ，幸福感，官能的な喜びの体験。 | 高興: 特別愉的心情、正向、感官上的喜。 |
| Sadness | Feeling discouraged by the irrevocable loss of a person, thing, or place. | 悲しみ: 人，物，場所などの取り返しのつかない損失に落胆すること。 | 悲傷: 因為無法挽回的人、物、地點，而感到沮喪。 |
| Anger | Extreme displeasure caused by someone's stupid or hostile action. | 怒り: 誰か他の人の愚かな行為や敵対的な行動によって引き起こされる極度の不快感。 | 憤怒: 因為他人的愚蠢行為或是不友善的敵意而感到極度的不。 |
| Surprise | Being abruptly faced with an unexpected and unusual event (without positive or negative connotation). | 驚き: 予期せぬ異常な出来事に突然直面すること（ただし，ポジティブな意味合いもネガティブな意味合いもない）。 | 驚訝: 遇到突如其來、意料之外的事，或是不尋常的事件（不包括正面或負面的意涵）。 |
| Fear | Being faced with an imminent danger that threatens our survival or physical well-being. | 恐れ: 生存または身体の健康を脅かす差し迫った危険に直面すること。 | 害怕: 正面臨迫在眉睫的、會威脅到我們性命或身體安危的危險。 |
| Disgust | Revulsion when faced with an unpleasant object or environment. | 嫌悪: 不快な物体や環境に直面したときの反発。 | 厭惡: 遇到讓人不快的物體或是環境時的反應。 |
| Contempt | Disapproval of the socially or morally reprehensible conduct of another person. | 軽蔑: 他人の社会的または道徳的に非難されるべき行為に対する不支持。 | 輕蔑: 否定他人在社會上或道上應該被譴責的行為。 |
| Gratitude | A sense of thankfulness and happiness in response to receiving a gift, either a tangible benefit (e.g., a present, favor) given by someone or a fortunate happenstance (e.g., a beautiful day). | 感謝: 誰かから与えられた具体的な利益（プレゼントや好意など），または幸運な出来事（晴れた日など）を受け取ったことに対する感謝の気持ちや幸福感。 | 因為偶遇的好事（例如：好天氣）或是收到來自他人的贈予、實質的利益（例如：禮品、幫忙）而感受到的感謝與喜。 |
| Guilt | A person believes or realizes—accurately or not—that they have compromised their own standards of conduct or have violated universal moral standards and bear significant responsibility for that violation. | 罪悪感: 自分自身の行動基準を損なった，あるいは普遍的な道徳基準に違反し非難されるべきだと信じていること。 | 罪惡感: 一個人相信或理解到——即便不見得正確——自己違背了自己的行為標準、觸犯了普世的道標準，並且為該行為背負著顯著的責任。 |
| Jealousy | A negative emotion in which an individual resents a third party for appearing to take away (or being likely to take away) the affections of a loved one. | 嫉妬: 好意を抱いている相手の興味や関心，愛情を奪っている（または奪いそうな）第三者を恨む否定的な感情。 | 嫉妒: 因為第三方的出現並且奪走了（或看似將要奪走）我們所愛之人的關注，而引起的負面情緒。 |
| Shame | Self-esteem shaken by an error or clumsiness for which one feels responsible. | 恥: 自分の責任であると感じるミスや不手際で自尊心が揺らぐこと。 | 丟臉: 因錯誤或笨拙而感到自責，進而動搖了自尊心。 |
| Pride | Feeling of triumph following a success or a personal achievement (one's own or that of someone close). | 誇り: 成功や個人的な達成感（自分自身や親しい人の成功）に伴う勝利の感覚。 | 驕傲: 因成功或個人（自己或是親近的人）的成就，而生的勝利感。 |
| Neutral | Not emotional. | ニュートラル: 感情を出さない。 | 中性無情緒: 沒有任何情緒。 |

temperature checks, the use of hand sanitizer, and wearing face masks. After these precautions, performers received detailed instructions regarding the data collection process and signed informed consent forms.

Performers were well fitted to Vicon whole-body suits, that included a cap, gloves, shoes cover, and top and bottom garments. Their skeleton model was created using the "Production" marker set, which included 57 retro-reflective markers attached to various joints and body segments (Fig. 2).

Calibration was conducted following the two-step procedure outlined in Vicon Shogun Live 1.7. This process began with an A-pose followed by a Range of Motion (ROM) assessment. During the ROM assessment, performers were instructed to move through the full range of motion for each limb and joint designated for capture. During the calibration, the range of motion of all joints was calculated and the length of the bones of the skeleton was determined. The skeleton stayed consistent through all the recordings for the same performer, and this consistency allowed for precise body position measurements and few data cleaning even in the case of marker occlusions (e.g., when a performer was sitting or lying on their back).

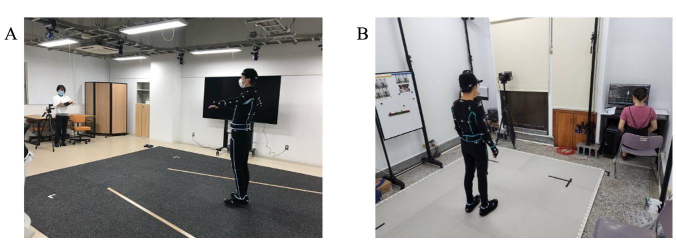

Fig. 1. Motion Capture Set-up in TU (A) and CCU (B).

Before the formal capture session, performers adapted themselves to moving in the motion capture suits and completed a practice recording trial. This preparation ensured that they were comfortable with the equipment and familiar with the data capture procedure.

Due to a laboratory relocation and hardware upgrades, Japanese performers 1–5 were recorded with an OptiTrack system (NaturalPoint Inc., Corvallis, OR, USA) using 24 Flex V100r2 cameras and Motive 2.1.1 software. The skeleton marker set was "Baseline" with 41 markers. All other performers were recorded on the Vicon system, as described above.

### E. Recording and performing procedure

After the practice trial, formal motion capture session began. In each performance, the actors portrayed the assigned emotion, scenario, and intensity through whole-body movement for a duration of 3 to 10 seconds. To have clear starting and ending signals, performers began and ended each portrayal with a T-pose.

Between scenarios, performers were free to take breaks to recalibrate their emotional states before the next portrayal. They were encouraged to perform in their own ways, with no constraints imposed on their expression styles. This flexibility allowed for more authentic and varied expressions. Furthermore, if a performance exceeded 10 seconds, it was still accepted following the "no constraints" principle. In addition to the motion capture system, all performances were recorded from front view using a video camera, and audio was captured via a wireless microphone.

Performers participated in the motion capture sessions individually. Only one performer was present during each session, which ensured that no one could observe another's performance beforehand. This arrangement minimized any potential influence from peers and allowed for a diverse range of expressions and movement patterns.

### F. Interview

After the motion capture session, each performer participated in a semi-structured interview. During this interview, they reviewed all 111 recorded performances with the experimenters. This process allowed performers to elaborate on their scenarios and acting strategies, providing valuable insights into how they expressed different emotions and intensities through body movement. Performers' explanation also highlighted cultural nuances and individual differences in scenario selection and emotional expression styles.

Performers were given the option to retake any recording they were not satisfied with. If a retake occurred, the new recording would replace the previous one. Requests for retakes were rare, totaling 11 scenarios requested by 8 performers.

We observed varying levels of self-awareness among performers regarding their scenario choices and emotional expressions. Some performers articulated their thought processes and feelings clearly, while others relied more on intuition and needed additional time to articulate their experiences. On average, each interview session lasted between 30 to 60 minutes. The whole interview was recorded with a video camera.

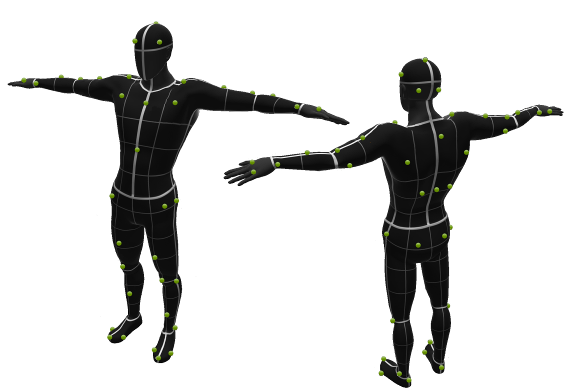

Fig. 2. Vicon "Production" marker set setup, composed of 57 markers.

## III. Post-Process

### A. Motion Data Processing

After the capture, the data was cleaned and processed. Due to the high quality of the capture process, only a few instances of jumping markers were observed per files, caused by occlusion. To automatize the correction on all files, the jumps were automatically detected using the mean amplitude deviation of each marker's acceleration. The motion was then smoothed using cubic spline interpolation between the period before the marker was occluded and after it became visible again. All the algorithm's parameters were fine-tuned through human observations of the results over multiple files. We also removed extra markers that appeared in the view field of motion capture system.

To be able to use all the data together, we needed the data captured in Japan and Taiwan to be consistent in their orientation. To that end, we reoriented the Taiwan motion capture data, so that the directions are the same in all files: Z-axis as the upward direction, Y-axis as the frontward direction, and X-axis as the direction from left to right shoulder.

From the markers and with the data acquired during the calibration process, the motion capture software calculated the position of the bones in the reconstructed performer's avatar. The bones orientation were given using a local rotation system with Euler angles. Since there were different sequences of rotational axes for different bones, we changed them so that all the rotation followed the rotational axes sequence ZYX.

Finally, we exported the data in both .bvh and .fbx formats, two very popular format in computer animation to describe body motion. Our files were named following the format ⟨Country_PerformerID_Emotion_ScenarioNumber_Intensity⟩.

As mentionned in Section II-D, Japanese performers 1 to 5 were recorded using a different system. Since the data of those performers need more pre-processing to make sure that they are consistent with the rest of the other performers' data, they will be released as part of this dataset at a later time.

### B. Scenario Translation

Scenarios were originally provided in Japanese or Chinese. Each scenario was translated into English by three bilingual experts using a multi-step procedure to ensure accuracy and cultural appropriateness. First, one translator produced an initial translation for each scenario. The other two translators reviewed and compared the translated text with the original, having two rounds of cross-checking. In cases where they reached a consensus, no further modifications were made. If the reviewers agreed with the initial translation, no further modifications were made. In cases of disagreement or ambiguity, the translators proposed revisions and conducted follow-up

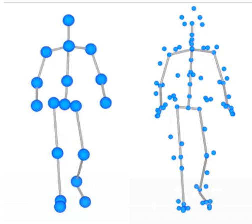

Fig. 3. Examples of point-light animation in partial (18 points) and full (57 points) marker conditions.

discussions to resolve the issues. Any persistent discrepancies were discussed collectively among all three translators until a final consensus was reached. Noteworthy, when the source text contained ambiguous expressions, the interview transcripts were referred to for contextual clarification. This approach helped preserve the nuance of each scenario and capture subtle cultural references. The complete list of translated scenarios is provided with the motion data.

## IV. Evaluation on Expressed Emotion

This database features 97 performers expressing emotions in personalized scenarios. To assess how accurately these emotions can be recognized and to capture the potential variability across performers and scenarios, we conducted an emotion recognition study with a separate group of observers. Because the full DIEM-A database contains more than 10,000 clips, we started with a subset to evaluate emotion recognition accuracy and identify general patterns before scaling up to the entire dataset. In a prior study [9], we examined six Japanese performers portraying 8 basic emotions. For the current evaluation, we followed the same paradigm with 6 Taiwanese performers to compare against the Japanese performers' results.

### *A. Selected Motion Stimuli*

We selected motion data from 6 Taiwanese performers (3 females, 3 males, mean age 38.5 years, mean performance experience 17 years). Clips from eight basic emotions (Joy, Anger, Sadness, Fear, Contempt, Disgust, Surprise, and Neutral) at a mid-level intensity were converted into point-light displays (18 or 57 points, see Fig. 3). We removed T-pose at the beginning and end of each recording.

### *B. Participants and Procedure*

39 university students (30 females, 9 males; all Asian, located in Taiwan and Hong Kong; mean age 22 years) completed an online emotion recognition task using the Labvanced platform (http://www.labvanced.com/). Each participant received a briefing via video conference, provided written informed consent, and accessed the experiment through an individualized link.

During each trial, participants viewed a point-light animation video and selected one of eight possible emotions in response to the question: "Which emotion is the person expressing?". In total, each participant viewed 288 videos, including 8 emotions $\times$ 3 scenarios $\times$ 6 performers $\times$ 1 intensity level (mid-level) $\times$ 2 marker point conditions (partial vs. full). These videos were divided into four blocks of 20–30 minutes each, with three breaks scheduled every 24 trials. The entire session lasted about 1.5–2 hours, and participants were free to take additional breaks as needed.

### *C. Results*

Our analysis focused on whether each emotion was recognized accurately, and whether performance varied across different performers and scenarios. We began with an overall accuracy assessment before conducting deeper analyses at the performer and scenario levels. This approach offers a detailed view of how bodily expression of emotion might be influenced by individual differences and contextual factors.

*1) Recognition Accuracy by Emotion:* We summarized the recognition accuracy of emotional animations with 18 points (partial) and 57 points (full) in Fig. 4. The average emotion recognition was 43.3%, and all emotion categories were detected above chance (12.5% in an 8-forced choice task). A repeated-measures ANOVA with Emotion (8 levels) $\times$ Marker Condition (2 levels) showed significant effects of Emotion ($F$(7, 266)=85.97, $p$ <.001, $\eta^2_{\mathrm{p}}$ =0.69), Marker ($F$(7, 38)=38.48, $p$ <.001, $\eta^2_{\mathrm{p}}$ =0.50), as well as a interaction effect ($F$(7, 266)=4.94, $p$ <.001, $\eta^2_{\mathrm{p}}$ =0.12).

Post-hoc multiple comparisons (Bonferroni-adjusted) indicated that Neutral was recognized most accurately (71.4%, *ps* <.001 against all other emotions). The next most recognizable emotions were Fear (52.2%), Joy (50.1%), Sadness (45.9%), and Anger(45.7%). The worst recognized are Contempt (32.3%), Disgust (24.4%) and Surprise (24.3%).

To examine whether the number of markers influenced recognition accuracy, we compared partial (18 points) vs full (57 points) conditions. Mean recognition accuracy in full points condition (45.6%) was significantly higher than the partial points condition (41.0%). Further post-hoc analyses with Bonferroni adjustments indicated that Fear and Disgust benefited the most from additional markers on the trunk and head. Specifically, the recognition of Fear increased from 47.7% (partial) to 56.7% (full), and Disgust from 19.4% (partial) to 29.3% (full).

*2) Error Analysis:* A confusion matrix (Fig. 5) illustrates how participants misclassified each emotion. Although Neutral achieved the highest overall recognition rate it also contributed to a significant portion of errors for most other emotions, particularly Disgust, Surprise, and Contempt, which were often judged as Neutral. Certain pairs of emotions were also prone to confusion. For example, 19.8% of Surprise clips were labeled as Fear, and 11.4% of Fear clips were labeled as Surprise. Additionally, Disgust (16.0%) and Contempt (11.3%) were frequently mistaken for Anger.

*3) Recognition in Detail: Across Performers and Scenarios:* To explore how recognition accuracy varies by performer and scenario, we analyzed the data at a more granular level (see Supplementary Table 1). Overall, results show wide variability in accuracy both within and across performers. First, scenario context plays a crucial role. Certain scenarios achieved near-perfect recognition, while others fell below 10%, suggesting that the specific context can strongly influence observers' judgments. For example, for Anger scenarios, ranged recognition rates ranged from 1.3% (e.g., "Being cheated on in a romantic relationship") to 97.4% (e.g., "Seeing someone being bullied"). These large discrepancies suggest that scenario content can dramatically affect observer perception.

Second, performing style matters. Even when performers prepared the same scenarios, recognition rates varied. For example, "Eating Delicious Food" for Joy, where recognition rates ranged from as low as 1.3% to as high as 79.5% and 73.1%; "Watching a Horror Movie" for Fear yielded accuracy of 43.6%, 96.2%, and 82.1%. Such discrepancies suggest that while the scenario label itself remains constant, performers' individual differences in acting approach, posture, and movement nuances appear to significantly influence how observers interpret emotional cues. Some performers appear more challenging to recognize. For example, Performer 4 showed lower average accuracy across emotions (Fig. 6), suggesting that certain expressive styles may systematically present fewer or more ambiguous emotional cues.

Additionally, substantial variability emerged within the same performer. One performer might achieve a near-perfect recognition rate for one scenario yet fail to convey another scenario in a way that observers accurately interpret. For example, for Performer 1, recognition accuracy for anger scenarios ranged from 28.2% to 69.2%. Please see more data of emotion recognition accuracy and variance across performers and scenarios in Supplementary Table 1.

Taken together, the above results indicate that even similar emotions can be expressed with widely varying clarity, possibly due to differences in personal experiences, acting styles, or scenario complexity.

## V. Discussion

We introduce the Diverse Intercultural E-Motion Database of Asian Performers (DIEM-A), a collection of high-quality motion capture data featuring emotional expressions by Asian professional performers with a wide range of backgrounds. DIEM-A incorporates diverse emotion triggers and contextual

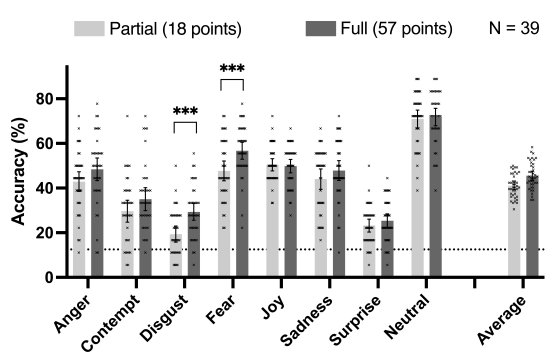


Fig. 4. Mean and 95% CI of emotion recognition accuracy in partial vs. full marker conditions (***, $p < .001$).

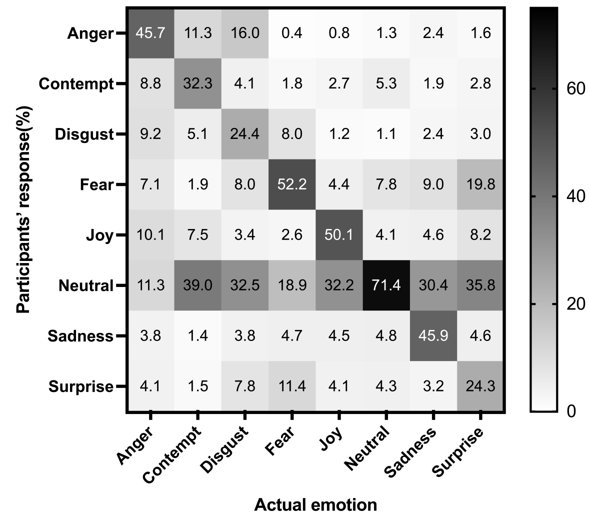


Fig. 5. Confusion matrix of 8 emotions.

cues by allowing each performer to draw on personalized scenarios. The evaluation results illustrated how scenario context and performer style interact, resulting in a full spectrum of ambiguity/clarity: from subtle to highly recognizable expressions. This makes DIEM-A a valuable addition for a broad spectrum of academic and industry applications in cognitive scientists, engineers, and machine learning.

### A. Database Uses

First, DIEM-A is a good resource for computational modeling as it offers whole-body movement 3-D data alongside video and scenario descriptions. It can be used to train and develop models for a range of tasks, including emotion recognition and motion classification. The inclusion of textual descriptions also enables multimodal fusion, allowing researchers to integrate linguistic or audio cues with body motion features. This makes the database well-suited for designing and benchmarking advanced affective systems that perceive, interpret, and respond to users' bodily expressions. Additionally, the diversity of motion data supports synthetic data generation, using 3D body coordinates as a template for scenarios where real-world data collection is challenging.

Second, DIEM-A serves as a valuable stimuli dataset for cognitive and neural sciences. Psychologists and neuroscientists often require carefully crafted stimuli to elicit specific affective states. Existing options typically fall into two categories: naturalistic but noisy video clips [31], [32], or high quality motion data captured in controlled environments with limited contextual variety [33], [34]. DIEM-A bridges these approaches by providing a wide range of context-rich scenarios, while also allowing extraction of pure motion (e.g., biological motion) for isolating the body's role in emotion perception. It also enables analyses of how specific kinematic features and body parts contribute to emotional expression. Researchers can apply methods such as kinematic feature extraction, principal component analysis (PCA), and spatiotemporal

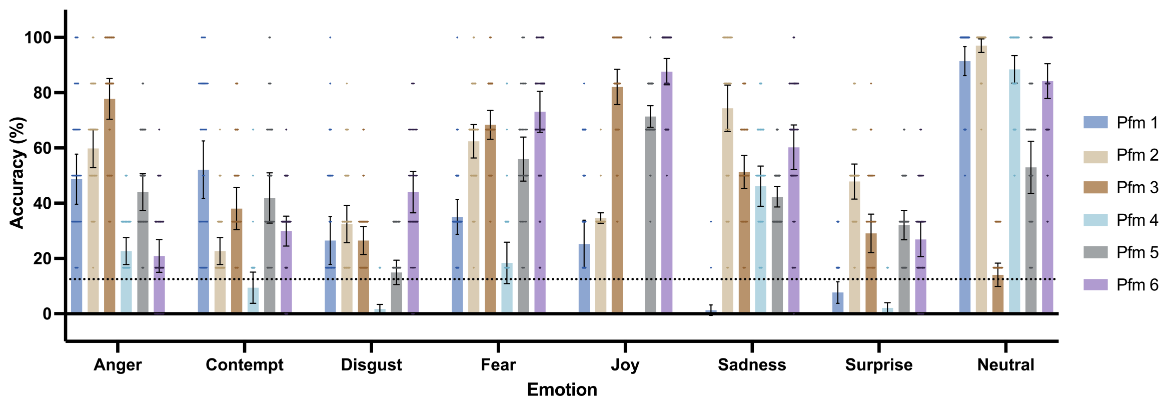


Fig. 6. Mean and 95% CI of recognition accuracy of 8 emotions for 6 performers in the evaluation experiment.

clustering to identify movement patterns relevant to different emotions. Consequently, researchers can better examine behavioral, neural or physiological responses to nonverbal cues and gain deeper insights into how the human express and decode emotional states from bodily expressions.

Third, DIEM-A can contribute to cross-cultural research by providing performances from Japanese and Taiwanese, two Asian subpopulations that share similarities yet also have different social norms. Recently there has been growing attention on Asian focused database (e.g. [18]). Our database expands on such efforts by including two Asian subgroups. Our database enables researchers to investigate subtle cultural variations in contextual triggers, expression style, and perception strategy. This approach provides a nuanced perspective on emotional communication within Asian contexts, facilitating deeper cross-cultural understanding in affective computing, psychology, linguitics, and related disciplines. Additionally, DIEM-A can be integrated into larger international datasets, helping researchers distinguish universal features from culturally specific ones.

Industrial fields can also benefit from the current database. Human-computer interaction (HCI) researchers can integrate DIEM-A data into virtual or augmented reality systems that respond to subtle whole-body expressions. Animators and game developers can leverage high-fidelity motion capture for realistic character movements, while social robotics practitioners can design culturally nuanced gestures that resonate with users.

### *B. Database Access*

The complete DIEM-A dataset will be available upon acceptance. Motion capture files (BVH and FBX formats), textual scenarios, de-identified performer demographics, and preliminary evaluation results will be freely available for noncommercial research purposes. Interested researchers can request access by contacting the authors and signing a licensing agreement.

Materials and results from current evaluation experiment are available on Open Science Framework (OSF link).

### *C. Limitations and Future Work*

The current dataset has a few areas that could be further developed in the future.

First, our current evaluation primarily aims to demonstrate the dataset's initial validity and usefulness for emotion recognition research. Given the large dataset size (over 10,000 clips), a full-scale human evaluation was beyond the scope of this paper. Future researchers can design evaluation/rating studies based on their research questions, such as emotion classification, valence and arousal ratings, Yes/No judgment to identify if an emotion, and etc.

Second, more recent databases extend into interactive, multimodal recordings [35]–[37]. For example, BEAT includes body motion, facial expression, audio, and text from multilingual conversations [36]. NNIME provides audio, video, and electrocardiogram data from Chinese dyadic interactions [37]. Our database does not include facial expressions, since data collection occurred during the COVID-19 pandemic, performers wore masks throughout the recordings. Future extensions of the database could integrate facial expression to better capture the multimodal nature of emotional expression.

Third, the emotion categories used in DIEM-A followed traditional labels commonly employed in emotion research, which may not fully capture the subtleties of everyday emotional experiences. Emotional experiences often reflect social norms and personal interpretations, making fixed labels insufficient. To address this, we plan to invite performers to revisit and freely label their own recorded performances, generating personalized labels that better represent the emotions they genuinely experienced. This approach would allow us to capture additional cultural nuances and enhance the ecological validity of emotion ground truth labeling.

Finally, to broaden the scope and generalizability of DIEM-A, we intend to extend the database to other population, including performers from additional cultural backgrounds. We also encourage collaborations with other research groups for cross-validation, enabling comparative analyses and deeper insights into cultural variations in emotional perception and expression.

## Ethical Impact Statement

All experimental procedures, including motion capture data collection and subsequent evaluation experiments, were reviewed and approved by the Human Research Ethics Committee in Tohoku University and National Chung Cheng University. All methods adhered to the relevant guidelines and regulations. Informed consent was obtained from each performer and participant. All performers and participants in evaluation study were free to pause or withdraw at any time without penalty, minimizing the potential for physical or psychological distress. All provided data in the database will be anonymized.

Although performers recalled personal scenarios to enact each emotion, their expressions are posed rather than spontaneous reactions, and therefore differ from genuine daily-life responses. The evaluation study focused on how observers recognized these posed expressions, thus reflecting the perception of emotional cues embedded in body movement, rather than perception of true affective states. Researchers using this dataset must be cautious when generalizing to real-world scenarios where spontaneous emotional reactions are more common.

The dataset will be made available solely for noncommercial, research-oriented purposes. Users must sign a licensing agreement and comply with the dataset's ethical guidelines, which prohibit any use that could infringe upon performers' privacy or lead to misuse of emotion expression recognition technology. Researchers are encouraged to remain mindful of the inherent limitations of posed expressions when designing studies and interpreting results.

## Acknowledgment

The authors express their sincere gratitude to all the performers who contributed to the database. Special thanks are due to Mr. Jun Hirose and Mr. Kazumasa Fujita from Sendai SCS Musical Institute, Mr. Yasuhiro Takahashi and Ms. Aya Sasaki from Sendai Theater Studio 10-box, Mr. Po-Hsiang Ko from EnjoyToEnjoy (Chiayi), and Mr. Hsin-Hung Lin from TGLB (Tainan), for their valuable support in performer recruitment.

We also thank Mr. Shoi Higashiyama, Ms. I-Chia Lin, Ms. Mei-Wun Wang, Mr. Yi-Chun Chuang, Mr. Shao-Kang Lee, and Mr. Chen-Yuan Hsieh for their assistance during the motion capture sessions. We are grateful to Ms. Sayaka Makabe, Ms. Mika Ono, and Ms. Nobuyo Maejima for their administrative support.

## References

[1] F. Noroozi, C. A. Corneanu, D. Kamińska, T. Sapiński, S. Escalera, and G. Anbarjafari, "Survey on emotional body gesture recognition," *IEEE transactions on affective computing*, vol. 12, no. 2, pp. 505–523, 2018.

[2] H. Zacharatos, C. Gatzoulis, and Y. L. Chrysanthou, "Automatic emotion recognition based on body movement analysis: a survey," *IEEE computer graphics and applications*, vol. 34, no. 6, pp. 35–45, 2014.

[3] M. P. A. Ramaswamy and S. Palaniswamy, "Multimodal emotion recognition: A comprehensive review, trends, and challenges," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 14, no. 6, p. e1563, 2024.

[4] B. De Gelder, A. W. de Borst, and R. Watson, "The perception of emotion in body expressions," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 6, no. 2, pp. 149–158, 2015.

[5] M.-A. Mahfoudi, A. Meyer, T. Gaudin, A. Buendia, and S. Bouakaz, "Emotion expression in human body posture and movement: A survey on intelligible motion factors, quantification and validation," *IEEE Transactions on Affective Computing*, vol. 14, no. 4, pp. 2697–2721, 2022.

[6] D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, and K. Scherer, "Toward a minimal representation of affective gestures," *IEEE Transactions on Affective Computing*, vol. 2, no. 2, pp. 106–118, 2011.

[7] M. Poyo Solanas, M. J. Vaessen, and B. de Gelder, "The role of computational and subjective features in emotional body expressions," *Scientific reports*, vol. 10, no. 1, p. 6202, 2020.

[8] N. Fourati and C. Pelachaud, "Emilya: Emotional body expression in daily actions database," in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 2014, pp. 3486–3493.

[9] M. Cheng, C.-H. Tseng, K. Fujiwara, S. Higashiyama, A. Weng, and Y. Kitamura, "Toward an asian-based bodily movement database for emotional communication," *Behav. Res. Methods*, vol. 57, no. 1, p. 10, 2025.

[10] M. Baradaran, P. Zohari, A. Mahyar, H. Motamednia, D. Rahmati, and S. Gorgin, "Cultural-aware ai model for emotion recognition," in *2024 13th Iranian/3rd International Machine Vision and Image Processing Conference (MVIP)*. IEEE, 2024, pp. 1–6.

[11] Y. Mohamed, M. Abdelfattah, S. Alhuwaider, F. Li, X. Zhang, K. W. Church, and M. Elhoseiny, "Artelingo: A million emotion annotations of wikiart with emphasis on diversity over language and culture," in *2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022*, 2022.

[12] A. Kappas, "Smile when you read this, whether you like it or not: Conceptual challenges to affect detection," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 38–41, 2010.

[13] T. Olugbade, M. Bieńkiewicz, G. Barbareschi, V. D'amato, L. Oneto, A. Camurri, C. Holloway, M. Björkman, P. Keller, M. Clayton *et al.*, "Human movement datasets: An interdisciplinary scoping review," *ACM Computing Surveys*, vol. 55, no. 6, pp. 1–29, 2022.

[14] W. Zhu, X. Ma, D. Ro, H. Ci, J. Zhang, J. Shi, F. Gao, Q. Tian, and Y. Wang, "Human motion generation: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 4, pp. 2430–2449, 2023.

[15] T. Yu, J. Wang, J. Wang, J. Luo, and G. Zhou, "Towards emotion-enriched text-to-motion generation via llm-guided limb-level emotion manipulating," in *Proceedings of the 32nd ACM International Conference on Multimedia*, 2024, pp. 612–621.

[16] E. Volkova, S. De La Rosa, H. H. Bülthoff, and B. Mohler, "The mpi emotional body expressions database for narrative scenarios," *PloS one*, vol. 9, no. 12, p. e113647, 2014.

[17] T. Bänziger, M. Mortillaro, and K. R. Scherer, "Introducing the geneva multimodal expression corpus for experimental research on emotion perception," *Emotion*, vol. 12, no. 5, pp. 1161–1179, Oct. 2012.

[18] M. Zhang, L. Yu, K. Zhang, B. Du, B. Zhan, S. Chen, X. Jiang, S. Guo, J. Zhao, Y. Wang *et al.*, "Kinematic dataset of actors expressing emotions," *Scientific data*, vol. 7, no. 1, p. 292, 2020.

[19] P. Chen, A. Chung-Fat-Yim, T. Guo, and V. Marian, "Cultural background and input familiarity influence multisensory emotion perception." *Cultural Diversity & Ethnic Minority Psychology*, vol. 30, no. 3, p. 487, 2024.

[20] M. H. Immordino-Yang, X.-F. Yang, and H. Damasio, "Cultural modes of expressing emotions influence how emotions are experienced." *Emotion*, vol. 16, no. 7, p. 1033, 2016.

[21] N. Binetti, N. Roubtsova, C. Carlisi, D. Cosker, E. Viding, and I. Mareschal, "Genetic algorithms reveal profound individual differences in emotion recognition," *Proceedings of the National Academy of Sciences*, vol. 119, no. 45, p. e2201380119, 2022.

[22] A. Javanbakht, S. Tompson, S. Kitayama, A. King, C. Yoon, and I. Liberzon, "Gene by culture effects on emotional processing of social cues among east asians and european americans," *Behavioral Sciences*, vol. 8, no. 7, p. 62, 2018.

[23] D. Matsumoto and H. C. Hwang, "Cultural influences on nonverbal behavior," *Nonverbal communication: Science and applications*, pp. 97–120, 2013.

[24] P. Ekman and W. V. Friesen, “A new pan-cultural facial expression of emotion,” *Motiv. Emot.*, vol. 10, no. 2, pp. 159–168, Jun. 1986.
[25] P. Ekman, “An argument for basic emotions,” *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.
[26] D. Matsumoto, “Cultural similarities and differences in display rules,” *Motivation and emotion*, vol. 14, no. 3, pp. 195–214, 1990.
[27] R. W. Levenson, “Basic emotion questions,” *Emotion review*, vol. 3, no. 4, pp. 379–386, 2011.
[28] D. Trampe, J. Quoidbach, and M. Taquet, “Emotions in everyday life,” *PloS one*, vol. 10, no. 12, p. e0145450, 2015.
[29] C. Scrivner, D. Sznycer, A. W Lukaszewski, and L. Al-Shawaf, “Social emotions are governed by a common grammar of social valuation: Theoretical foundations and applications to human personality and the criminal justice system,” 2024.
[30] K. R. Scherer, T. Wranik, J. Sangsue, V. Tran, and U. Scherer, “Emotions in everyday life: Probability of occurrence, risk factors, appraisal and reaction patterns,” *Social Science Information*, vol. 43, no. 4, pp. 499–570, 2004.
[31] S. Sonkusare, M. Breakspear, and C. Guo, “Naturalistic stimuli in neuroscience: critically acclaimed,” *Trends in cognitive sciences*, vol. 23, no. 8, pp. 699–714, 2019.
[32] H. Saarimäki, “Naturalistic stimuli in affective neuroimaging: A review,” *Frontiers in human neuroscience*, vol. 15, p. 675068, 2021.
[33] J. Bachmann, J. Munzert, and B. Krüger, “Neural underpinnings of the perception of emotional states derived from biological human motion: a review of neuroimaging research,” *Frontiers in psychology*, vol. 9, p. 1763, 2018.
[34] S. Lammers, G. Bente, R. Tepest, M. Jording, D. Roth, and K. Vogeley, “Introducing acass: an annotated character animation stimulus set for controlled (e) motion perception studies,” *Frontiers in Robotics and AI*, vol. 6, p. 94, 2019.
[35] L. Piwek, K. Petrini, and F. Pollick, “A dyadic stimulus set of audiovisual affective displays for the study of multisensory, emotional, social interactions,” *Behavior Research Methods*, vol. 48, pp. 1285–1295, 2016.
[36] H. Liu, Z. Zhu, N. Iwamoto, Y. Peng, Z. Li, Y. Zhou, E. Bozkurt, and B. Zheng, “Beat: A large-scale semantic and emotional multi-modal dataset for conversational gestures synthesis,” in *European conference on computer vision*. Springer, 2022, pp. 612–630.
[37] H.-C. Chou, W.-C. Lin, L.-C. Chang, C.-C. Li, H.-P. Ma, and C.-C. Lee, “Nnime: The nthu-ntua chinese interactive multimodal emotion corpus,” in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2017, pp. 292–298.

## Supplementary Table 1. Emotion recognition accuracy across performers and scenarios

| Performer | Scenario (3 per performer) | Emotion | Accuracy (mean) | Accuracy variance (STD) |
|---|---|---|---|---|
| 1 | Bullying at school | Anger | 69.2% | 46.5% |
| 1 | A long wait for Uber Eats compounded by the delivery of the wrong order | Anger | 28.2% | 45.3% |
| 1 | Roommate's dog running into my room | Anger | 48.7% | 50.3% |
| 2 | Bumping with a passerby | Anger | 87.2% | 33.6% |
| 2 | Having a car accident | Anger | 62.8% | 48.6% |
| 2 | Quarreling with a colleague or intimate partner | Anger | 29.5% | 45.9% |
| 3 | Getting bumped into while walking on the street. | Anger | 74.4% | 43.9% |
| 3 | Seeing someone being bullied | Anger | 97.4% | 15.9% |
| 3 | Being wronged | Anger | 61.5% | 49.0% |
| 4 | Being cheated on in a romantic relationship | Anger | 3.8% | 19.4% |
| 4 | The loss of a loved one due to homicide | Anger | 1.3% | 11.3% |
| 4 | Having an argument | Anger | 62.8% | 48.6% |
| 5 | Being misunderstood | Anger | 23.1% | 42.4% |
| 5 | Breaking up because of a betrayal | Anger | 85.9% | 35.0% |
| 5 | Direct reports repeatedly making mistakes at work | Anger | 23.1% | 42.4% |
| 6 | A social movement/street protest | Anger | 48.7% | 50.3% |
| 6 | Being elbowed by someone when playing basketball | Anger | 12.8% | 33.6% |
| 6 | Being cut in line after waiting for a long time | Anger | 1.3% | 11.3% |
| 1 | Hearing a despised teacher talking nonsense on the podium | Contempt | 65.4% | 47.9% |
| 1 | Witnessing the life of a person I dislike getting better and better | Contempt | 65.4% | 47.9% |
| 1 | Seeing someone lying | Contempt | 25.6% | 43.9% |
| 2 | Hearing someone give a wrong answer | Contempt | 2.6% | 15.9% |
| 2 | Arguing with someone from a protest | Contempt | 56.4% | 49.9% |
| 2 | Bullying others | Contempt | 9.0% | 28.8% |
| 3 | Working with a colleague who constantly flatters the boss | Contempt | 41.0% | 49.5% |
| 3 | Being irresponsible | Contempt | 21.8% | 41.6% |
| 3 | Cheating on someone/(having) an affair | Contempt | 51.3% | 50.3% |
| 4 | Parasite singles | Contempt | 14.1% | 35.0% |
| 4 | Littering cigarette butts | Contempt | 7.7% | 26.8% |
| 4 | Forcing someone to buy something | Contempt | 6.4% | 24.7% |
| 5 | Looking down on those who used connections to get desirable positions | Contempt | 37.2% | 48.6% |
| 5 | experiencing discrimination based on abilities | Contempt | 52.6% | 50.3% |
| 5 | Dissing someone | Contempt | 35.9% | 48.3% |
| 6 | Attacking opponents in an election for falsifying their education experience | Contempt | 14.1% | 35.0% |
| 6 | seeing someone walking without a mask on the street | Contempt | 0.0% | 0.0% |
| 6 | Lip syncing | Contempt | 75.6% | 43.2% |
| 1 | Walking into a smelly and dirty bathroom | Disgust | 21.8% | 41.6% |
| 1 | Red bean soup | Disgust | 30.8% | 46.5% |
| 1 | Cleaning up dog poop | Disgust | 26.9% | 44.6% |
| 2 | A mosquito flying nearby | Disgust | 52.6% | 50.3% |
| 2 | Walking into a disgusting room | Disgust | 10.3% | 30.5% |
| 2 | Eating something yucky | Disgust | 34.6% | 47.9% |
| 3 | Being sexually harrassed | Disgust | 66.7% | 47.4% |
| 3 | Getting frightened by a prank | Disgust | 6.4% | 24.7% |
| 3 | Dealing with a rude customer | Disgust | 6.4% | 24.7% |
| 4 | Being in an uncomfortably hot and humid environment | Disgust | 5.1% | 22.2% |
| 4 | Encountering a friend I dislike | Disgust | 0.0% | 0.0% |
| 4 | Food | Disgust | 0.0% | 0.0% |
| 5 | A kid who's hating on vegetables | Disgust | 38.5% | 49.0% |
| 5 | Running into someone I dislike | Disgust | 3.8% | 19.4% |
| 5 | smelling something awful and seeing the dirtiness | Disgust | 2.6% | 15.9% |
| 6 | eating something I dislike | Disgust | 55.1% | 50.1% |
| 6 | Riding my motorcycle behind a garbage truck | Disgust | 55.1% | 50.1% |
| 6 | smelling teenagers' sweat and foot odor | Disgust | 21.8% | 41.6% |
| 1 | Hearing a loud noise while in a focused state | Fear | 16.7% | 37.5% |
| 1 | Seeing a cockroach | Fear | 44.9% | 50.1% |
| 1 | Watching a horror movie | Fear | 43.6% | 49.9% |
| 2 | Getting ready to get an injection/undergo an operation | Fear | 26.9% | 44.6% |
| 2 | Feeling like a monster is approaching | Fear | 67.9% | 47.0% |
| 2 | Being hunted | Fear | 92.3% | 26.8% |
| 3 | Climbing to a high place | Fear | 73.1% | 44.6% |
| 3 | Getting lost | Fear | 35.9% | 48.3% |
| 3 | Watching a horror movie | Fear | 96.2% | 19.4% |
| 4 | Getting an injection | Fear | 11.5% | 32.2% |
| 4 | Losing a loved one | Fear | 11.5% | 32.2% |
| 4 | Encountering ghosts | Fear | 32.1% | 47.0% |
| 5 | Receiving threats | Fear | 74.4% | 43.9% |
| 5 | Being surrounded | Fear | 35.9% | 48.3% |
| 5 | Touching something scary | Fear | 57.7% | 49.7% |

| | | | | |
|---|---|---|---|---|
| 6 | Walking alone on a suspension bridge | Fear | 61.5% | 49.0% |
| 6 | Watching a horror movie | Fear | 82.1% | 38.6% |
| 6 | Being chased by a wild dog while riding a bicycle | Fear | 75.6% | 43.2% |
| 1 | Eating delicious food | Joy | 2.6% | 15.9% |
| 1 | Going on a date | Joy | 32.1% | 47.0% |
| 1 | Being in the midst of nature | Joy | 41.0% | 49.5% |
| 2 | Eating delicious food | Joy | 1.3% | 11.3% |
| 2 | Seeing an old friend for the first time in a long while | Joy | 100.0% | 0.0% |
| 2 | Having intimate physical contact with my partner | Joy | 2.6% | 15.9% |
| 3 | Winning a prize | Joy | 70.5% | 45.9% |
| 3 | Traveling | Joy | 96.2% | 19.4% |
| 3 | Eating delicious food | Joy | 79.5% | 40.6% |
| 4 | Watching funny TV shows | Joy | 0.0% | 0.0% |
| 4 | Eating sweets | Joy | 0.0% | 0.0% |
| 4 | Winning the Uniform Invoice lottery. | Joy | 0.0% | 0.0% |
| 5 | Encouraging a child to play with other children | Joy | 92.3% | 26.8% |
| 5 | Winning championships and receiving gifts | Joy | 23.1% | 42.4% |
| 5 | Hanging out with my friends | Joy | 98.7% | 11.3% |
| 6 | Going to a concert | Joy | 100.0% | 0.0% |
| 6 | Winning the biggest prize at Weiya | Joy | 89.7% | 30.5% |
| 6 | Eating delicious food | Joy | 73.1% | 44.6% |
| 1 | Meditating | Neutral | 98.7% | 11.3% |
| 1 | Using a singing bowl | Neutral | 96.2% | 19.4% |
| 1 | Sleeping | Neutral | 79.5% | 40.6% |
| 2 | Meditating | Neutral | 93.6% | 24.7% |
| 2 | Reading a book | Neutral | 98.7% | 11.3% |
| 2 | Drinking water | Neutral | 98.7% | 11.3% |
| 3 | Encountering a sudden storm. | Neutral | 0.0% | 0.0% |
| 3 | Spotting a huge hole in the road ahead. | Neutral | 0.0% | 0.0% |
| 3 | Queuing to get on the bus | Neutral | 42.3% | 49.7% |
| 4 | Practicing meditation | Neutral | 97.4% | 15.9% |
| 4 | Taking a walk | Neutral | 82.1% | 38.6% |
| 4 | Reading a book | Neutral | 85.9% | 35.0% |
| 5 | Listening to music | Neutral | 39.7% | 49.3% |
| 5 | Washing my hands | Neutral | 53.8% | 50.2% |
| 5 | Memorizing vocabulary and writing drafts | Neutral | 65.4% | 47.9% |
| 6 | Waking up and brushing my teeth | Neutral | 92.3% | 26.8% |
| 6 | Scrolling on my phone | Neutral | 85.9% | 35.0% |
| 6 | Taking notes | Neutral | 74.4% | 43.9% |
| 1 | The passing of a close friend | Sadness | 1.3% | 11.3% |
| 1 | Going through a breakup | Sadness | 1.3% | 11.3% |
| 1 | The missing case of someone I care about | Sadness | 1.3% | 11.3% |
| 2 | Being rejected for a job application | Sadness | 65.4% | 47.9% |
| 2 | Experiencing the loss of a loved one/mourning | Sadness | 84.6% | 36.3% |
| 2 | Watching a heartbreaking scene in a series | Sadness | 73.1% | 44.6% |
| 3 | A family member passing away | Sadness | 32.1% | 47.0% |
| 3 | Getting sick | Sadness | 33.3% | 47.4% |
| 3 | Going through a breakup | Sadness | 88.5% | 32.2% |
| 4 | Failing an exam | Sadness | 67.9% | 47.0% |
| 4 | Seeing an ex's belongings | Sadness | 15.4% | 36.3% |
| 4 | The passing of a loved one | Sadness | 55.1% | 50.1% |
| 5 | Mourning the loss of a good friend | Sadness | 26.9% | 44.6% |
| 5 | feeling helpless | Sadness | 1.3% | 11.3% |
| 5 | Going through a heartbreak | Sadness | 98.7% | 11.3% |
| 6 | The passing of a family member | Sadness | 69.2% | 46.5% |
| 6 | Breaking my favorite glass | Sadness | 41.0% | 49.5% |
| 6 | My favorite team losing a game | Sadness | 70.5% | 45.9% |
| 1 | Finding out that my ex-girlfriend is with my ex-boyfriend | Surprise | 5.1% | 22.2% |
| 1 | Learning that my student had a baby | Surprise | 0.0% | 0.0% |
| 1 | Seeing someone running naked in the street | Surprise | 17.9% | 38.6% |
| 2 | Seeing something fly by | Surprise | 60.3% | 49.3% |
| 2 | Noticing someone is behind me while using my phone | Surprise | 74.4% | 43.9% |
| 2 | Realizing the building in front is about to collapse as I walk towards it | Surprise | 9.0% | 28.8% |
| 3 | Seeing a friend lose weight successfully | Surprise | 24.4% | 43.2% |
| 3 | Witnessing a car accident | Surprise | 29.5% | 45.9% |
| 3 | Running into a friend on the road | Surprise | 33.3% | 47.4% |
| 4 | Discovering a friend's secret | Surprise | 0.0% | 0.0% |
| 4 | Running into an acquaintance | Surprise | 6.4% | 24.7% |
| 4 | Realizing my family background is different from what I had initially believed | Surprise | 0.0% | 0.0% |
| 5 | Getting scammed | Surprise | 7.7% | 26.8% |
| 5 | Hearing a secret | Surprise | 30.8% | 46.5% |
| 5 | Being frightened | Surprise | 57.7% | 49.7% |
| 6 | Pregnancy test | Surprise | 29.5% | 45.9% |
| 6 | First time watching a magic show | Surprise | 16.7% | 37.5% |
| 6 | Learning about the sudden death of a classmate whom I haven't talked to for a lon | Surprise | 34.6% | 47.9% |