scientific reports

OPEN

Check for updates

Development of the RIKEN database for dynamic facial expressions with multiple angles

Shushi Namba^{1,2⊠}, Wataru Sato^{1⊠}, Saori Namba², Hiroki Nomiya³, Koh Shimokawa⁴ & Masaki Osumi⁴

The development of facial expressions with sensing information is progressing in multidisciplinary fields, such as psychology, affective computing, and cognitive science. Previous facial datasets have not simultaneously dealt with multiple theoretical views of emotion, individualized context, or multi-angle/depth information. We developed a new facial database (RIKEN facial expression database) that includes multiple theoretical views of emotions and expressers' individualized events with multi-angle and depth information. The RIKEN facial expression database contains recordings of 48 Japanese participants captured using ten Kinect cameras at 25 events. This study identified several valence-related facial patterns and found them consistent with previous research investigating the coherence between facial movements and internal states. This database represents an advancement in developing a new sensing system, conducting psychological experiments, and understanding the complexity of emotional events.

Developing a new facial database will contribute to progress in many domains, such as psychology, affective computing, and cognitive science. Recent studies have reported on developing several facial databases well suited for many situations and research purposes, including deception detectione¹, free speec³hes², group discussion², spontaneous tears⁴, pain-related face^{5,6}, and social stigma⁷. The most enthusiastic among them has been the development of the facial expression database that conveys emotions (see the systematic survey⁸; review⁹; and meta-database⁷). These studies are the basis for extensive applied research^{10,11}.

This paper introduces a new facial database called the RIKEN facial expression database with the potential to achieve multiple goals depending on the individual research aims for facial expressions of emotions. For example, the user can explore the relationship between facial movements and annotated information derived from the multiple theoretical views of emotion (emotional labels, valence, and appraisal components).

There are three issues in developing a new database for facial expressions of emotion: the deficit in multiple theoretical views of emotion, insufficient description of individualized events that induce affective responses, and the lack of multi-angle and depth information.

(a) In the scientific study of emotion, three overarching traditions can be identified: basic emotion theory, the theory of constructed emotion, and appraisal theory¹². Each of these theories can distinguish itself by the psychological states they rely upon for emotional expression. Basic emotion theory involves emotional labels, the theory of constructed emotion encompasses valence and arousal values¹³, and appraisal theory centers on the appraisal dimensions of emotional events (novelty check)¹⁴. Many databases rely on basic emotion theory (i.e., six basic emotion categories¹⁵⁻²⁴). Therefore, there is a lack of facial expression data based on other theoretical models of emotion. Available databases that apply the theory of constructed emotion are limited compared to those that apply the basic emotion theory or simplified emotional labels, and many existing databases sometimes use observers' ratings of valence and arousal (AffectNet²⁵; AFEW-VA²⁶) rather than the expresser's view (but see the Stanford Emotional Narratives Dataset²⁷). Furthermore, the component process model (CPM¹⁴) derived from appraisal theory assumes that the results of each appraisal check drive the dynamics of emotion sequentially. The only facial database that relies on this is the actor database developed by ²⁸. In the theoretical discussion of emotion, a recent scholar recommended the multiple ways to define emoion²⁹. A database with annotations of various theoretical backgrounds would be desirable because it can be used flexibly according to the research purpose or practice.

¹RIKEN, Psychological Process Research Team, Guardian Robot Project, Kyoto 6190288, Japan. ²Department of Psychology, Hiroshima University, Hiroshima 7398524, Japan. ³Faculty of Information and Human Sciences, Kyoto Institute of Technology, Kyoto 6068585, Japan. ⁴KOHINATA Limited Liability Company, Osaka 5560020, Japan. [⊠]email: nashushi@hiroshima-u.ac.jp; wataru.sato.ya@riken.jp

(b) Contextual information is important to understand emotional expressions and perceptions³⁰⁻³². Le Mau et al.³³ emphasize the role of contextual information when inferring internal states through facial movements. Each emotional instance can be considered a loose concept³⁴. Different persons perceive the same event differently. Being insulted may cause one person to feel anger while another may feel contempt or fear. An individual's developmental history, including cultural learning, changes facial movements associated with affect^{35,36}. Therefore, when expanding a facial database, it is important to address individualized eliciting contexts or conditions. A new database is expected to collect various perspectives on specific and personal events with several evaluations (labels, valence and arousal, or appraisal checks) rather than the same standardized situations.

(c) Many facial stimuli have been created using only two-dimensional (2D) images or clips^{15,16,19-21,37,38}. Additionally, there are a limited number of facial databases with comparable multi-angle and depth information^{28,39–41}. However, multi-angle data and depth information have several advantages in individual research practices. Psychologists have indicated that the angle of the face (e.g., frontal or profile view) is important in studying face perception. Guo and Shaw⁴² showed that profile faces have significantly decreased perceived intensity compared to frontal faces. If angles influence our facial perception, a database with multiple angles contributes to their psychological research. Moreover, multi-angle information is gaining attention in computer science. From multiangle images, the state-of-the-art algorithm can generate volumetric radiance representation⁴³. Echoing volumetric representation and 4D information, which adds dynamic information to 3D faces, have been the focus of increasing attention in psychology research, such as research on face perception⁴⁴. For example, Chelnokova and Laeng⁴⁵ showed that 3D faces could be recognized better than 2D faces. Scholars have developed new databases that directly measure depth information using tools such as Kinect to enrich science using facial databases^{46,47}. Therefore, multi-angle recording and obtaining depth information have increasingly become standard in affective computing^{28,48-50}. Multi-angle and depth information are expected to be useful in reconstructing a face with many features and extracting detailed facial movements. Collecting this information is important for conducting psychological experiments and training or developing automated sensing systems.

This study developed a new facial database that includes individualized contexts and multiple theoretical views of emotions with multi-angle and depth information. We aimed to create a facial database based on 25 individual events corresponding to valence and arousal⁵¹. Furthermore, we obtained free description labeling data⁵² and rating values associated with appraisal dimensions^{53,54} from the 25 events prepared by the participants. We are currently performing manual facial action coding using the Facial Action Coding System (FACS), a comprehensive, anatomically based system for describing all visually discernible facial movement⁵⁵. This database makes these annotated facial movements publicly available data (8 people are already annotated and available). There are 29 types of manually annotated facial Action Units (AU). Please complete the following form (https://forms.gle/XMYiXaXHhfszCb4c6) to request access to the RIKEN facial expression database. People who want to use the RIKEN facial expression database must agree to the end-user license agreement.

Here, our main purpose was to report the characteristic of the RIKEN facial expression database. To understand the nature of emotional events within this database, we performed a text analysis of word frequency for all the events. Next, this study investigated the relationships between the appraisal dimensions, including valence and arousal, for the events. We then provided an overview of the characteristics of the events tagged with each emotion label. This was achieved by assessing the frequency of labels assigned by participants and examining the mean values of all the evaluative elements of the events. These can provide insight into the emotional events the database targets to create facial reactions.

Finally, we aimed to elucidate the nature of this database as a "facial" database. We quantified facial expressions using an automated FACS analysis and analyzed their annotated information, such as valence and arousal. This approach was designed to be a practical use case. As a well-established relationship, there is an association between AU 4 (brow lowerer) and negative valence, AU 12 (lip corner puller) and positive valence. We predicted these relationships because several psychophysiological studies recording facial electromyography showed that activity in the corrugator supercilii (related to AU 4) and zygomatic major muscle (related to AU 12) is negatively and positively associated with subjective valence experiences, respectively⁵⁶⁻⁶¹. This study also conducted an exploratory investigation into the relationship underlying arousal and each facial muscle. The focus of this paper remains the demonstration of use cases.

Methods

Participants

Forty-eight Japanese adults (22 female and 26 male) aged between 20 and 30 years (mean = 23.33; SD = 3.65) participated in the recording sessions. The participants were recruited from a local human resource center in Kyoto. Individuals were informed about the purpose of the study, methodology, risks, right to withdraw, handling of individual information, and the voluntary nature of participation. All participants gave written informed consent before recording the facial movements. Informed consent included whether the participants agreed to their videos being shown for academic purposes, including psychological experiments and affective computing. Each participant was paid 13,000 JPY for participation and database creation. The Ethics Committee of the RIKEN (Protocol number: Wako3 2020-21) approved the experimental procedure and study protocol. The study was conducted according to the Declaration of Helsinki. Our main purpose was not to estimate population indices for effect sizes. Therefore, power analyses were not available.

Procedures

All participants were instructed to remember and write down 25 events that occurred in their lives, with five valences (strongly unpleasant, unpleasant, neutral, pleasant, and strongly pleasant feelings) and five levels of arousal (very low arousal or sleepiness, low arousal, middle arousal, high arousal, and very high arousal) one

week before the recording session to obtain individualized emotional events (Fig. 1). Qualtrics was applied as the platform for collecting the events. Participants were instructed to describe a single event corresponding to each cell in Fig. 1. The order of events (i.e., each cell in Fig. 1) for the valence and arousal combinations was randomized. Participants also rated appraisal checks from 1 (strongly disagree) to 5 (strongly agree) for novelty (predictable: "the event was predictable"; familiar: "the event was common"), goal significance ("the event was important to you") and coping potential ("the event could have been controlled and avoided if you had taken appropriate actions") for each event described. These appraisal checks were derived from a previously reported facial database that relied on the CPM as its theoretical basis²⁸. The participants were also asked to describe the possible labels for each event freely.

On the day of the facial clip recording, the participants were given a further explanation of the experiment. They were transferred to the recording location (the first basement floor of the Advanced Telecommunications Research Institute International). Figure 2 displays the recording environment. The participants sat in chairs with their faces fixed in a steady position. We set up three photographic lights (AL-LED-SQA-W: Toshiba) and illuminated each participant's face from the upper right, left, and lower sides to make clear their faces and remove shadows.

We then asked the participants to remove their masks and glasses. We set up an environment to record facial movements using ten Azure Kinect DK1880 cameras with (Microsoft; $2D:1920 \times 1080$; depth: 640×576) pixel resolution and 30 frames per second to record the participants' facial movements as video clips. The interval between the left and right horizontal cameras was 22.5°, and the cameras were used at 22.5°, 45°, and 90° (skipping 67.5° in this database). Images were taken to avoid interference between multiple depth cameras, with each camera shifted by 160 microseconds in timing. A Software Development Kit program was used to create a program to record facial movements. The depth information was limited to eight cameras to reduce the processing load and avoid equipment errors: one upper and two lower cameras, and front-facing, two left, and two right cameras; a green carpet covered as much of the background as possible (Fig. 2).

The experimenter verbally narrated the individual event descriptions collected one week before the recording session for each expression. Participants were instructed to vividly reexperience their emotions and practice expressing them through facial expressions using a hand mirror before the recording. Participants were allowed to remember the events and practice their facial expressions with no time restrictions. When participants felt ready, they sounded a bell to initiate the recording and the experimenter verbally narrated the events again. The recording process was structured into distinct segments. The timing of each segment was indicated by beep sounds (onset:880 Hz; peak:1174 Hz; offset:880 Hz) produced by the speaker system to control the participants in producing their expressions according to the time course. The models were instructed to express an emotional expression rooted in pre-described events for the initial 1 s, maintain the intended emotional expression for 2 s, and then return to a neutral expression for one second. The order of events was also randomized.

Data analysis

We extracted 17 facial movements to evaluate the pattern of facial expressions using OpenFace⁶²: AU 1 (inner brow raiser), AU 2 (outer brow raiser), AU 4 (brow lowerer), AU 5 (upper lid raiser), AU 6 (cheek raiser), AU 7 (lid tightener), AU 9 (nose wrinkler), AU 10 (upper lip raiser), AU 12 (lip corner puller), AU 14 (dimpler), AU 15 (lip corner depressor), AU 17 (chin raiser), AU 20 (lip stretcher), AU 23 (lip tightener), AU 25 (lips parts), AU 26 (jaw drop), and AU 45 (blink). Among the automated facial movement detection systems, OpenFace had a relatively good performance⁶³. Since Namba, Sato, and Yoshikawa⁶⁴ also found that facial images from the front



Figure 1. Events referred to the affect grid⁵¹.



Figure 2. (A) The setup of the apparatus. The camera settings illustrated by Autodesk Fusion 360. Three lights were shone on the face from under the feet and from above on the left and right sides, and facial expressions were captured with one upper and two lower cameras, and front-facing, two left, and two right cameras. (B) Showcasing samples of the images. A green carpet covered as much of the background as possible.

.....

view have the highest accuracy in OpenFace, only facial expressions from the front-view camera were targeted in this study. Given the procedure's nature, in which facial combinations' intensity is expected to be maximal during the apex beep sound, we mainly focused on the middle frame (i.e., 61 frames).

We used R⁶⁵ for statistical analysis. We used the tm and openxlsx packages to perform text mining for each event^{66–68}. The psych package⁶⁹ was used to check the correlation between several appraisal dimensions. We used the nnTensor package to reduce dimension for data extracted by OpenFace⁷⁰. We used the tidyverse package for data visualization⁷¹. Based on ample psychophysiological evidence, we predicted and analyzed the relationships between valence and AUs 4/12 using hierarchical linear regression modeling using the lmerTest package⁷². The results were considered significant at p < 0.05. To elucidate the relationship between arousal and AUs, we used Bayesian Lasso regression⁷³, treating arousal as the dependent variable and utilizing all AUs as independent variables with the tuning parameter set at a degree of freedom of 1⁷⁴. The AU data were standardized, and only results that did not encompass zero within the 95% confidence interval were reported. All codes are available on the Gakunin RDM (https://dmsgrdm.riken.jp:5000/uphvb/). The design and analysis of this study were not pre-registered.

Ethics declarations

The Ethics Committee of the RIKEN (Protocol number: Wako3 2020-21) approved all experimental procedures and protocols. This research was conducted according to the Declaration of Helsinki.

Consent to participate

All participants provided written informed consent to participate before the beginning of the experiment.

Results

The detail of the events. As indicated in the "Methods" section, we obtained 1,200 events (48 participants × 5 valences × 5 arousals). All Japanese events were translated and back-translated into English using TEXT (https:// www.text-edit.com/english-page/). Table 1 depicts the top 3 frequently used English words for each event obtained by text mining. In the obtained database, words that appeared to be common events (frequency of 10/48 or more) occurred in valence 4 * arousal 4 events (friend) and valence 5 * arousal 5 events (passing the university entrance exam). The latter, in particular, shows that university entrance exams greatly affected emotional events because this research was limited to young participants.

The correspondence between the valence, arousal, and appraisal ratings is also presented in Table 2. Valence and arousal appeared to be positively associated with the appraisal of importance for each event (rs > 0.20). The

V1A1		371.4.2		V1A2		VIAA		VIAE	
VIAI		VIA2		VIAS		VIA4		VIA5	
Sleep, exam	4	Sleep, wait	3	Person	9	Fail, bike	4	Someone	5
Bed, person	4	Small, get	3	Dislike, got	4	Find, time	3	Person, hit	4
Depress, bad	3	Tired, want	2	Parttime, job	4	Bad, fell	3	Fell, head	3
V2A1		V2A2		V2A3		V2A4		V2A5	
Morning	7	Day, rain	5	Want	4	Work	7	Game	4
Sleep	6	Start, intend	3	Dislike, favorite	3	Take	4	Someone, make	3
Day, wake	3	Something, well	3	Well, hours	3	Custom, train	3	Mistake, get 3	
V3A1		V3A2		V3A3		V3A4		V3A5	
Watch, class	4	Train	5	Day, lunch	4	Take, time	4	News	5
Take, time	4	Ride, watch	4	meal, without	4	School, friend	4	Day	4
Space, sleep	3	Bike, take	3	Spend, time	3	Room, make	3	Time, watch	3
V4A1		V4A2		V4A3	V4A4 V4A5				
Sleep	7	Good	6	Favorite, watch	5	Friend	11	Game	5
Bed	5	Watch	5	Buy, new	3	Favorite	6	Friend	4
Play, home	3	Holiday, movie	4	Good, job	3	Game, make	4	Favorite, time	3
V5A1	V5A2 V5A3 V5A4 V5A		V5A5						
Sleep	7	Message, won	4	Friend	7	Friend, favorite	7	Exam, pass	16
Work, hot	4	Finish, work	4	Favorite	6	Play	6	Entrance	13
Holiday, coffee	4	Tire, good	3	Time, watch	5	Watch	5	University	10

Table 1. Frequent words in each individualized event. "…" means that there is more than one word with a frequency equivalent to the words listed in the table. V represents valence, and A represents arousal. The word 'get/got' was removed as it is used as a 'be' verb in Japanese.

	Predict	Similarity	Important	Control
Similarity	0.451			
Important	0.147	-0.062		
Control	0.294	0.176	0.033	
Valence	0.215	0.079	0.237	-0.020
Arousal	-0.054	-0.194	0.208	-0.033

Table 2. Correspondence between valence, arousal, and appraisal ratings.

results also revealed that high predictability increased the valence of the event (r=0.22). Additionally, the more unfamiliar the event, the higher the arousal (r=-0.19). For the correlations between appraisal dimensions, positive correlations were found between predictability and familiarity, and predictability and controllability (rs>0.29).

As the participants were asked to freely describe the possible labels for each event, each event had an emotional term that the participant subjectively labelled. To provide information labelled as an individual event, Table 3 lists the most frequently used labels of emotions using the free description data. Only the top 18 modes (N = 730/1200) are listed. Positive emotional labels, such as joy, happiness, and fun, indicated high valence; negative emotional labels, such as anger, sadness, and unpleasantness, indicated low valence. Arousal was high with surprise and impatience. Although there were other interesting correspondences between the controllability component and frustration, predictability, and fun, these were not examined as they went beyond the study's purpose of overviewing the events in our database.

The detail of facial movements

The facial data for some events are missing due to camera malfunctions and participant problems, although the events themselves were recorded as stated above. The available number of frames was 142,865. When only the peak frame was extracted, there were 1,190 frames. Six events for male and four for female participants were missing expressions. Ultimately, 1,190 data points were available for analysis. Figure 3 shows the facial patterns of all the individual events associated with valence and arousal. Visual inspection revealed that AU4 (lower brow) and AU7 (lid tightener) were strongly expressed during negative events (V1-V2). Positive events (V4-V5) induced AU6 (cheek raiser), AU7, AU10 (upper lip raiser), AU12 (lip corner puller), and AU14 (dimpler), which can be considered strong smiling expressions. The intensity of facial movements may be relatively low in neutral events (V3) compared with the two valenced events. Moreover, in positive events, Fig. 3 indicates that higher arousal was associated with more mouth-opening movements (AU25: lip parts and AU26: jaw drop). In the peak intensity frame, we also checked the correlation between the estimated AUs and appraisal dimensions (Table 4).

Emotional label	N	Valence	Arousal	Predict	Similarity	Important	Control
Joy (喜び)	119	4.45 (0.59)	3.21 (1.38)	2.83 (1.17)	2.92 (1.17)	3.91 (1.03)	2.99 (1.26)
Anger (怒り)	89	1.45 (0.52)	3.48 (1.39)	2.34 (1.15)	2.73 (1.20)	3.38 (1.25)	2.94 (1.34)
Happy (嬉しい)	59	4.46 (0.57)	3.51 (1.32)	3.03 (1.27)	2.69 (1.19)	3.93 (0.96)	2.37 (1.14)
Sadness (悲しみ)	54	1.43 (0.57)	3.02 (1.45)	2.54 (1.30)	2.56 (1.21)	3.70 (1.25)	3.15 (1.45)
None (無)	39	2.97 (0.35)	2.28 (1.07)	3.46 (1.41)	4.00 (1.08)	2.41 (1.23)	3.18 (1.41)
Relief (安心)	37	3.86 (0.75)	2.46 (1.24)	3.76 (0.93)	3.59 (1.19)	3.76 (0.93)	3.54 (1.04)
Fun (楽しい)	36	4.17 (0.74)	3.56 (1.25)	4.17 (0.97)	3.89 (1.09)	4.14 (1.13)	3.64 (1.22)
Surprise (驚き)	36	2.72 (0.97)	4.08 (1.18)	1.58 (0.94)	2.19 (1.04)	2.61 (1.32)	2.11 (1.24)
Unpleasant (不快)	29	1.48 (0.51)	2.59 (1.12)	2.69 (1.49)	3.41 (1.09)	3.14 (1.27)	3.31 (1.39)
Disappointment (落胆)	28	2.00 (1.02)	2.61 (1.55)	2.82 (1.09)	2.82 (1.12)	3.21 (1.34)	2.93 (1.44)
Accomplishment (達成感)	27	4.26 (0.86)	3.26 (1.46)	3.70 (0.91)	3.52 (0.89)	3.78 (0.93)	3.63 (0.84)
Satisfaction (満足)	27	4.22 (0.85)	3.19 (1.30)	3.59 (1.15)	2.93 (1.38)	4.04 (0.81)	3.07 (1.17)
Disgust (嫌悪)	26	1.50 (0.65)	2.81 (1.27)	2.69 (1.05)	3.19 (0.90)	2.69 (0.97)	2.85 (0.97)
Impatience (焦り)	26	2.04 (0.66)	4.12 (1.07)	2.27 (1.15)	2.69 (1.35)	3.50 (1.21)	3.54 (1.27)
Pleasure (快感)	25	4.52 (0.71)	2.76 (1.36)	3.48 (1.33)	3.56 (1.29)	3.80 (1.29)	2.96 (1.40)
Happiness (幸せ)	25	4.76 (0.44)	2.60 (1.58)	3.24 (1.36)	3.36 (1.19)	4.08 (1.04)	3.24 (1.42)
Frustrated (悔しい)	24	1.71 (0.55)	2.75 (1.48)	2.92 (1.21)	3.08 (0.93)	3.25 (1.11)	3.88 (1.23)
Expectation (期待)	24	4.04 (0.55)	3.25 (1.29)	3.29 (1.33)	3.42 (1.02)	3.92 (0.72)	2.67 (1.46)

 Table 3. Eighteen emotional labels and corresponding valence, arousal, and appraisal checks.



Figure 3. Peak intensities of each facial movement; V represents valence, and A represents arousal.

	Valence	Arousal	Predict	Similarity	Important	Control
AU01_r	-0.03	0.02	-0.06	-0.02	0.02	-0.03
AU02_r	0.03	0.00	0.01	0.03	0.01	-0.02
AU04_r	-0.21	0.01	-0.06	-0.11	-0.02	0.06
AU05_r	0.02	0.01	-0.05	-0.02	0.07	-0.09
AU06_r	0.30	0.15	0.07	-0.12	0.21	0.03
AU07_r	0.01	0.08	-0.03	-0.07	0.08	0.04
AU09_r	0.06	-0.03	0.00	0.00	0.02	0.03
AU10_r	0.28	0.18	0.08	-0.09	0.19	0.01
AU12_r	0.49	0.18	0.11	-0.09	0.25	0.00
AU14_r	0.34	0.09	0.07	-0.07	0.16	-0.01
AU15_r	0.00	0.03	0.01	0.00	0.03	-0.03
AU17_r	0.02	-0.02	-0.01	-0.02	0.04	0.03
AU20_r	0.07	-0.03	0.03	0.00	0.04	0.06
AU23_r	0.02	-0.02	0.00	-0.01	0.08	0.03
AU25_r	0.13	0.14	-0.02	-0.08	0.11	0.02
AU26_r	0.02	0.02	-0.02	-0.06	0.01	0.01
AU45_r	0.03	-0.04	0.00	0.00	0.00	-0.05

Table 4. Correspondence between action units and valence and arousal values and appraisal ratings.

Compared to the correlations between valence and some facial movements, such as AU12 (lip corner puller: r = 0.49), the combinations of all facial movements and other appraisal dimensions were relatively low (|r|s < 0.25).

A hierarchical linear regression model examined the relationship between valence/arousal and the AUs. Consistent with our predictions, the result indicated that the valence values significantly predicted the intensity of AU4 (brow lowerer) negatively ($\beta = -0.11$, t = 6.38, p < 0.001) and that of AU12 (lip corner puller) positively ($\beta = 0.28$, t = 12.64, p < 0.001). Besides, the arousal values significantly predicted the intensity scores of AU 12 ($\beta = 0.10$, t = 9.19, p < 0.001). Furthermore, post-hoc sensitivity power analysis using the simr package⁷⁵ indicated that the current sample size (i.e., N = 1190) was sufficient to detect all coefficients in the hierarchical linear regression models with a significance level of $\alpha = 0.05$ and 99% power.

To explore the new relationship between arousal and the AUs, we also used the Bayesian Lasso regression. Action Units 12 (lip corner puller) and 25 (opening the mouth) were found to predict arousal (β s = 0.13, 95% Credible Intervals [0.01, 0.26] and 0.08, 95% Credible Intervals [0.00, 0.17]). However, none of the other predictors predicted arousal performance, resulting from 95% CIs that included zero.

We confirmed the dynamics of the facial expressions obtained in this database by applying non-negative matrix factorization to reduce dimensionality and extract spatiotemporal features ⁷⁶. This approach can identify dynamic facial patterns⁷⁷⁻⁷⁹. The factorization rank was determined using cophenetic coefficients⁸⁰ and the dispersion index⁸¹. Information on factorization rank is available on the Gakunin RDM (https://dmsgrdm.riken. jp:5000/uphvb/).

Figure 4 displays the AU profiles of the top four components. We interpreted Component 1 as a Duchenne marker (AU6, 7), Component 2 as blinking and other facial movements (AU1, 14, 17, 45), Component 3 as a lower brow (AU4), and Component 4 as smiling (AU6, 10, 12, 14) by visually inspecting the relative contribution of each AU to the independent components. These results were also consistent with the peak intensities of each facial movement (Fig. 3).

Figure 5 lists how the spatial components changed over time for each valence and arousal combination. Visual inspection of component 1 (Duchenne marker) revealed that negative (V1-V2) and positive (V4-V5) events showed larger movements (e.g., V1A1 and V5A5). This result is consistent with the finding that eye constriction is systematically associated with the facial expressions of negative and positive emotions⁸². Component 2 (blinking and other facial movements) can be interpreted as the relaxation movement of tension associated with the expression of deliberate facial manipulation or noise unrelated to the main emotional expression because this movement increases during the offset duration (frames = 91–120) after the peak duration (frames = 31–90). For Component 3 (lower brow), negative expressions (V1–V2) produced more intense facial changes than other expressions (V3–V5). Component 4 (smiling) occurred more frequently during positive events (V4–V5) than others (V1–V3).

In summary, blinking and other facial movements, such as raising the inner eyebrow and chin, were (i.e., Component 2) peculiar to the offset of deliberate facial expressions in naive Japanese participants. More interestingly, the results clarified that smiling is related to the positive (Component 4), lowering of eyebrows is related to the negative (Component 3), and eye constriction (Component 1) corresponds to both values.

As a supplementary analysis and an example of the potential uses of the database, it may be useful to visualize dynamic changes rather than correlations in the peak frame (Table 4) as the relationship between one appraisal dimension and one facial pattern. According to Scherer's theory, the appraisals (and the corresponding AUs) appear sequencetially. Figure 6 shows the relationship between one appraisal dimension (important) and one



Figure 4. Heatmap of each component's loadings for facial expressions of all events. Value colors represent each facial movement's contribution to component scores.



Figure 5. Temporal changes in the four components for the facial expressions of 25 events.





component (AU6, 10, 12, 14). This indicates that as the appraisal of the importance of an event increase, more smiles are seen in response to the event.

Discussion

This study developed a new facial database with expresser annotations such as individualized emotional events, appraisal checks, and free description labels with multi-angle and depth information. The results (Table 3) indicate that the words for each event had few matches, implying that the database has a large variance in emotional events. A database with various events and individual evaluations can be verified for academic purposes. For example, researchers can investigate issues such as the typical elements of events labeled as anger and the appraisal components that constitute them in a data-driven manner or as a starting point.

According to the analysis of front-view facial expressions, facial movements related to pleasant and unpleasant valences were observed. For example, lowering the brow was related to a negative valence, whereas pulling the lip corner was related to a positive valence. These results are consistent with previous findings investigating the coherence between valence and facial muscle electrical activity. Moreover, the Bayesian lasso analysis reported that mouth movements such as AU12 (pulling the lip corner) and AU25 (opening the mouth) were also associated with arousal. Opening the mouth has been shown to increase arousal attribution from observers⁴⁰, which corresponds to ratings on the part of the expressers. This contributed to understanding the relationship between specific facial action and arousal. In addition to the data provided here, we are currently performing manual facial action coding by certified FACS coders. We will open the annotation data in the future (now, data for 8 people is already annotated and available in the same database. There are 29 types of manually annotated facial actions). In recent years, amidst the controversy about emotion⁸³, there have been increasing efforts to extract facial movements⁸⁴. The opening of databases, including manual FACS annotations that include in-depth information, can prime how research in affective computing can be further developed.

While this study provides a new facial database on emotions, certain limitations exist. First, the number of participants was small, given the diversity of facial movements and emotional events. In particular, the database only includes recordings of Japanese participants, which may limit its generalizability to other populations. Future research using similar environments, as represented in Fig. 2, will create additional databases for young and older adult participants and extend to other cultures or ethnicities beyond the Japanese population. Second, this study dealt with only facial responses to emotional expression. However, other aspects such as vocal or physiological responses would be important for understanding emotional communication^{85–87}. Expansion of those modalities could provide a useful database to understand emotion further. Finally, we did not investigate how depth or infrared information can be used, and the lighting conditions do not influence this information compared to 2D color images. This database will be an important foundation for developing a robust sensing system for facial movements in room conditions. Using these databases, we provide an internal state estimation algorithm via an Application Programming Interface combined with smartphones and other devices. Furthermore, we would like to utilize this technology to develop solutions for people with difficulty communicating.

The database, including the expressers' events, labels, and appraisal checking intensity, is available as a RIKEN facial expression database for academic purposes. The notable features of this database are as follows: (a) availability of multiple theoretical views for emotion (valence and arousal, appraisal dimensions, and free emotion label), (b) variety of events, and (c) rich information taken from 10 multi-angle and depth cameras.

Data availability

The RIKEN facial expression database is freely available to the research community (https://dmsgrdm.riken.jp: 5000/uphvb/). An End User License Agreement (EULA) must be produced to access the database.

Received: 8 August 2023; Accepted: 5 December 2023 Published online: 08 December 2023

References

 Lloyd, E. P. et al. Miami University deception detection database. Behav. Res. Methods 51, 429–439. https://doi.org/10.3758/s13428-018-1061-4 (2019).

- Şentürk, Y. D., Tavacioglu, E. E., Duymaz, İ, Sayim, B. & Alp, N. The Sabancı University Dynamic Face Database (SUDFace): Development and validation of an audiovisual stimulus set of recited and free speeches with neutral facial expressions. *Behav. Res. Methods* 55, 1–22. https://doi.org/10.3758/s13428-022-01951-z (2022).
- Girard, J. M. et al. Sayette group formation task (gft) spontaneous facial expression database. Proc. Int. Conf. Autom. Face Gesture Recognit. https://doi.org/10.1109/FG.2017.144 (2017).
- Küster, D., Baker, M. & Krumhuber, E. G. PDSTD—The Portsmouth dynamic spontaneous tears database. *Behav. Res. Methods* 54, 1–15. https://doi.org/10.3758/s13428-021-01752-w (2021).
- Lucey, P., Cohn, J. F., Prkachin, K. M., Solomon, P. E. & Matthews, I. Painful data: The UNBC-McMaster shoulder pain expression archive database. In 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG). 57–64 https://doi.org/10. 1109/FG.2011.5771462 (2011).
- Fernandes-Magalhaes, R. et al. Pain Emotion Faces Database (PEMF): Pain-related micro-clips for emotion research. Behav. Res. Methods 55, 3831–3844. https://doi.org/10.3758/s13428-022-01992-4 (2023).
- Workman, C. I. & Chatterjee, A. The Face Image Meta-Database (fIMDb) & ChatLab Facial Anomaly Database (CFAD): Tools for research on face perception and social stigma. *Methods Psychol.* 5, 100063. https://doi.org/10.1016/j.metip.2021.100063 (2021).
- 8. Dawel, A., Miller, E. J., Horsburgh, A. & Ford, P. A systematic survey of face stimuli used in psychological research 2000-2020. Behav. Res. Methods 54, 1-13. https://doi.org/10.3758/s13428-021-01705-3 (2021).
- Krumhuber, E. G., Skora, L., Küster, D. & Fou, L. A review of dynamic datasets for facial expression research. *Emot. Rev.* 9, 280–292. https://doi.org/10.1177/1754073916670022 (2017).
- Li, S. & Deng, W. Deep facial expression recognition: A survey. IEEE Trans. Affect. Comput. 13, 1195–1215. https://doi.org/10. 1109/TAFFC.2020.2981446 (2022).
- Ekundayo, O. S. & Viriri, S. Facial expression recognition: A review of trends and techniques. *IEEE Access* 9, 136944–136973. https://doi.org/10.1109/ACCESS.2021.3113464 (2021).
- Gendron, M. & Feldman Barrett, L. Reconstructing the past: A century of ideas about emotion in psychology. *Emot. Rev.* 1, 316–339. https://doi.org/10.1177/1754073909338877 (2009).
- Barrett, L. F. & Russell, J. A. The structure of current affect: Controversies and emerging consensus. *Curr. Dir. Psychol. Sci.* 8, 10–14. https://doi.org/10.1111/1467-8721.00003 (1999).
- Scherer, K. R. Appraisal considered as a process of multi-level sequential checking. In (eds Scherer, K. R., Schorr, A. & Johnstone, T.) Appraisal Processes in Emotion: Theory, Methods, Research. 92–120. (Oxford University Press, 2001).
- Calvo, M. G. & Lundqvist, D. Facial expressions of emotion (KDEF): Identification under different display-duration conditions. Behav. Res. Methods 40, 109–115. https://doi.org/10.3758/BRM.40.1.109 (2008).
- Chung, K. M., Kim, S. J., Jung, W. H. & Kim, V. Y. Development and validation of the Yonsei Face Database (Yface DB). Front. Psychol. 10, 2626. https://doi.org/10.3389/fpsyg.2019.02626 (2019).
- Ebner, N., Riediger, M. & Lindenberger, U. FACES—A database of facial expressions in young, middle-aged, and older women and men: development and validation. *Behav. Res. Methods* 42, 35–362. https://doi.org/10.3758/BRM.42.1.351 (2010).
- Holland, C. A. C., Ebner, N. C., Lin, T. & Samanez-Larkin, G. R. Emotion identification across adulthood using the dynamic FACES database of emotional expressions in younger, middle aged, and older adults. *Cogn. Emot.* 33, 245–257. https://doi.org/10.1080/ 02699931.2018.1445981 (2019).
- Langner, O. et al. Presentation and validation of the Radboud Faces Database. Cogn. Emot. 24, 1377–1388. https://doi.org/10.1080/ 02699930903485076 (2010).
- LoBue, V. & Thrasher, C. The child affective facial expression (CAFE) set: Validity and reliability from untrained adults. Front. Psychol. 5, 1532. https://doi.org/10.3389/fpsyg.2014.01532 (2015).
- Van Der Schalk, J., Hawk, S. T., Fischer, A. H. & Doosje, B. Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion* 11, 907–920. https://doi.org/10.1037/a0023853 (2011).
- Mavadati, S. M., Mahoor, M. H., Bartlett, K., Trinh, P. & Cohn, J. F. Disfa: A spontaneous facial action intensity database. *IEEE Trans. Affect. Comput.* 4, 151–160. https://doi.org/10.1109/T-AFFC.2013.4 (2013).
- Sneddon, I., McRorie, M., McKeown, G. & Hanratty, J. The Belfast induced natural emotion database. *IEEE Trans. Affect. Comput.* 3, 32–41. https://doi.org/10.1109/T-AFFC.2011.26 (2011).
- 24. Zhang, X. *et al.* Bp4d-spontaneous: A high-resolution spontaneous 3d dynamic facial expression database. *Image Vis. Comput.* **32**, 692–706. https://doi.org/10.1016/j.imavis.2014.06.002 (2014).
- Mollahosseini, A., Hasani, B. & Mahoor, M. H. Affectnet: a database for facial expression, valence, and arousal computing in the wild. *IEEE Trans. Affect. Comput.* 10, 18–31. https://doi.org/10.1109/TAFFC.2017.2740923 (2017).
- Kossaifi, J., Tzimiropoulos, G., Todorovic, S. & Pantic, M. AFEW-VA database for valence and arousal estimation in-the-wild. *Image Vis. Comput.* 65, 23–36. https://doi.org/10.1016/j.imavis.2017.02.001 (2017).
- Ong, D. C. et al. Modeling emotion in complex stories: The Stanford Emotional Narratives Dataset. IEEE Trans. Affect. Comput. 12, 579–594. https://doi.org/10.1109/TAFFC.2019.2955949 (2019).
- Seuss, D., Dieckmann, A., Hassan, T., Garbas, J. U., Ellgring, J. H., Mortillaro, M. & Scherer, K. Emotion expression from different angles: A video database for facial expressions of actors shot by a camera array. In 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII). 35–41 https://doi.org/10.1109/ACII.2019.8925458 (2019).
- 29. Scarantino, A. How to define emotions scientifically. Emot. Rev. 4, 358-368. https://doi.org/10.1177/1754073912445810 (2012).
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M. & Pollak, S. D. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest* 20, 1–68. https://doi.org/10.1177/1529100619832930 (2019).
- Barrett, L. F., Mesquita, B. & Gendron, M. Context in emotion perception. Curr. Dir. Psychol. Sci. 20, 286–290. https://doi.org/10. 1177/0963721411422522 (2011).
- Chen, Z. & Whitney, D. Tracking the affective state of unseen persons. Proc. Natl. Acad. Sci. 116, 7559–7564. https://doi.org/10. 1073/pnas.1812250116 (2019).
- Le Mau, T. et al. Professional actors demonstrate variability, not stereotypical expressions, when portraying emotional states in photographs. Nat. Commun. 12, 1–13. https://doi.org/10.1038/s41467-021-25352-6 (2021).
- Fehr, D. & Russell, J. A. Concept of emotion viewed from a prototype perspective. J. Exp. Psychol. Gen. 113, 464–486. https://doi. org/10.1037/0096-3445.113.3.464 (1984).
- 35. Barrett, L. F. How Emotions are Made: The Secret Life of the Brain. 448 (Pan Macmillan, 2017).
- Griffiths, P. E. What Emotions Really Are: The Problem of Psychological Categories. Vol. 293. https://doi.org/10.7208/chicago/97802 26308760.001.0001 (University of Chicago Press, 1997).
- Dawel, A., Miller, E. J., Horsburgh, A. & Ford, P. A systematic survey of face stimuli used in psychological research 2000–2020. Behav. Res. Methods 54, 1–13. https://doi.org/10.3758/s13428-021-01705-3 (2022).
- Tottenham, N. et al. The NimStim set of facial expressions: Judgments from untrained research participants. Psychiatry Res. 168, 242–249. https://doi.org/10.1016/j.psychres.2008.05.006 (2009).
- Cudeiro, D., Bolkart, T., Laidlaw, C., Ranjan, A. & Black, M. J. Capture, learning, and synthesis of 3D speaking styles. In Proceedings of the IEEE Computer Society Conference on Computer Vision Pattern Recognition 10101–10111. https://doi.org/10.48550/arXiv. 1905.03079 (2019).
- Fujimura, T. & Umemura, H. Development and validation of a facial expression database based on the dimensional and categorical model of emotions. *Cogn. Emot.* 32, 1663–1670. https://doi.org/10.1080/02699931.2017.1419936 (2018).

- Ueda, Y., Nunoi, M. & Yoshikawa, S. Development and validation of the Kokoro Research Center (KRC) facial expression database. Psychologia 61, 221–240. https://doi.org/10.2117/psysoc.2019-A009 (2019).
- Guo, K. & Shaw, H. Face in profile view reduces perceived facial expression intensity: An eye-tracking study. Acta. Psychol. 155, 19–28. https://doi.org/10.1016/j.actpsy.2014.12.001 (2015).
- Mihajlovic, M., Bansal, A., Zollhoefer, M., Tang, S. & Saito, S. KeypointNeRF: Generalizing image-based volumetric avatars using relative spatial encoding of keypoints. *Eur. Conf. Comput. Vis.* https://doi.org/10.1007/978-3-031-19784-0_11 (2022).
- Burt, A. L. & Crewther, D. P. The 4D space-time dimensions of facial perception. Front. Psychol. 11, 1842. https://doi.org/10.3389/ fpsyg.2020.01842 (2020).
- Chelnokova, O. & Laeng, B. Three-dimensional information in face recognition: An eye-tracking study. J. Vis. 11, 27. https://doi. org/10.1167/11.13.27 (2011).
- Aly, S., Trubanova, A., Abbott, A. L., White, S. W. & Youssef, A. E. VT-KFER: A Kinect-based RGBD+ time dataset for spontaneous and non-spontaneous facial expression recognition. In 2015 International Conference on Biometrics (ICB). 90–97 https://doi.org/ 10.1109/ICB.2015.7139081 (2015).
- Boccignone, G., Conte, D., Cuculo, V. & Lanzarotti, R. AMHUSE: a multimodal dataset for HUmour SEnsing. In Proceedings of the 19th ACM International Conference on Multimodal Interaction. 438–445 https://doi.org/10.1145/3136755.3136806 (2017).
- Cheng, S., Kotsia, I., Pantic, M. & Zafeiriou, S. 4dfab: A large scale 4d database for facial expression analysis and biometric applications. In Proceedings of the IEEE Conference on Computer Vision on Pattern Recognition. 5117–5126 https://doi.org/10.1109/CVPR. 2018.00537 (2018).
- Li, X. et al. 4DME: A spontaneous 4d micro-expression dataset with multimodalities. IEEE Trans. Affect. Comput. https://doi.org/ 10.1109/TAFFC.2022.3182342 (2022).
- Matuszewski, B. J. et al. Hi4D-ADSIP 3-D dynamic facial articulation database. Image Vis. Comput. 30, 713–727. https://doi.org/ 10.1016/j.imavis.2012.02.002 (2012).
- Russell, J. A., Weiss, A. & Mendelsohn, G. A. Affect grid: A single-item scale of pleasure and arousal. J. Pers. Soc. Psychol. 57, 493–502. https://doi.org/10.1037/0022-3514.57.3.493 (1989).
- Haidt, J. & Keltner, D. Culture and facial expression: Open-ended methods find more expressions and a gradient of recognition. Cogn. Emot. 13, 225–266. https://doi.org/10.1080/026999399379267 (1999).
- Scherer, K. R. Profiles of emotion-antecedent appraisal: Testing theoretical predictions across cultures. Cogn. Emot. 11, 113–150. https://doi.org/10.1080/026999397379962 (1997).
- Scherer, K. R., Mortillaro, M., Rotondi, I., Sergi, I. & Trznadel, S. Appraisal-driven facial actions as building blocks for emotion inference. J. Pers. Soc. Psychol. 114, 358–379. https://doi.org/10.1037/pspa0000107 (2018).
- 55. Ekman, P., Friesen, W. V. & Hager, J. C. Facial Action Coding System 2nd edn. (Research Nexus eBook, 2002).
- Bradley, M. M. & Lang, P. J. Affective reactions to acoustic stimuli. Psychophysiology 37, 204215. https://doi.org/10.1111/1469-8986. 3720204 (2000).
- 57. Greenwald, M. K., Cook, E. W. & Lang, P. J. Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *J. Psychophysiol.* **3**, 51–64 (1989).
- Lang, P. J., Greenwald, M. K., Bradley, M. M. & Hamm, A. O. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x (1993).
- Larsen, J. T., Norris, C. J. & Cacioppo, J. T. Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology* 40, 776–785. https://doi.org/10.1111/1469-8986.00078 (2003).
- Sato, W., Fujimura, T., Kochiyama, T. & Suzuki, N. Relationships among facial mimicry, emotional experience, and emotion recognition. *PLoS One* 8, e57889. https://doi.org/10.1371/journal.pone.0057889 (2013).
- Sato, W., Kochiyama, T. & Yoshikawa, S. Physiological correlates of subjective emotional valence and arousal dynamics while viewing films. *Biol. Psychol.* 157, 107974. https://doi.org/10.1016/j.biopsycho.2020.107974 (2020).
- Baltrusaitis, T., Zadeh, A., Lim, Y. C. & Morency, L. P. Openface 2.0: Facial behavior analysis toolkit. In 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). 59–66 https://doi.org/10.1109/FG.2018.00019 (2018).
- Namba, S., Sato, W., Osumi, M. & Shimokawa, K. Assessing automated facial action unit detection systems for analyzing crossdomain facial expression databases. Sensors 21, 4222. https://doi.org/10.3390/s21124222 (2021).
- 64. Namba, S., Sato, W. & Yoshikawa, S. Viewpoint robustness of automated facial action unit detection systems. *Appl. Sci.* 11, 11171. https://doi.org/10.3390/app112311171 (2021).
- 65. R Core Team. R: A Language and Environment for Statistical Computing. http://www.R-project.org/ (R Foundation for Statistical Computing, 2021).
- 66. Feinerer I. & Hornik, K. tm: Text Mining Package. R package version 0.7-10. https://CRAN.R-project.org/package=[™] (2022).
- 67. Feinerer, I., Hornik, K. & Meyer, D. Text mining infrastructure in R. J. Stat. Softw. 25, 1–54. https://doi.org/10.18637/jss.v025.i05 (2008)
- Schauberger P. & Walker, A. openxlsx: Read, Write and Edit xlsx Files. R package version 4.2.5.1. https://CRAN.R-project.org/packa ge=openxlsx (2022).
- 69. Revelle, W. psych: Procedures for Personality and Psychological Research. https://CRAN.R-project.org/package=psych (2022)
- Tsuyuzaki, K., Ishii, M. & Nikaido, I. nnTensor: Non-Negative Tensor Decomposition. R package version 1.1.9. https://github.com/ rikenbit/nnTensor (2022).
- 71. Wickham, H. et al. Welcome to the tidyverse. J. Open Source Softw. 43, 1686. https://doi.org/10.21105/joss.01686 (2019).
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. ImerTest Package: Tests in linear mixed effects models. J. Stat. Softw. 82, 1-26 https://doi.org/10.18637/jss.v082.i13 (2017).
- Park, T. & Casella, G. The Bayesian lasso. J. Am. Stat. Assoc 103, 681–686. https://doi.org/10.1198/016214508000000337 (2008).
- Bürkner, P. C. Advanced Bayesian multilevel modeling with the R Package brms. *R J.* 10, 395–411. https://doi.org/10.32614/RJ-2018-017 (2018).
- Green, P. & MacLeod, C. J. SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods Ecol. Evol.* 7, 493–498. https://doi.org/10.1111/2041-210X.12504 (2016).
- Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. Nature 401, 788–791. https://doi.org/ 10.1038/44565 (1999).
- Delis, I., Panzeri, S., Pozzo, T. & Berret, B. A unifying model of concurrent spatial and temporal modularity in muscle activity. J. Neurophysiol. 111, 675–693. https://doi.org/10.1152/jn.00245.2013 (2014).
- Perusquía-Hernández, M., Dollack, F., Tan, C. K., Namba, S., Ayabe-Kanamura, S. & Suzuki, K. Smile action unit detection from distal wearable electromyography and computer vision. In 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021). 1–8. https://doi.org/10.1109/FG52635.2021.9667047 (2021).
- Namba, S., Nakamura, K. & Watanabe, K. The spatio-temporal features of perceived-as-genuine and deliberate expressions. *PLoS One* 17, e0271047. https://doi.org/10.1371/journal.pone.0271047 (2022).
- Brunet, J. P., Tamayo, P., Golub, T. R. & Mesirov, J. P. Metagenes and molecular pattern discovery using matrix factorization. *Proc. Natl. Acad. Sci.* 101, 4164–4169. https://doi.org/10.1073/pnas.0308531101 (2004).
- Kim, H. & Park, H. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics* 23, 1495–1502. https://doi.org/10.1093/bioinformatics/btm134 (2007).

- Mattson, W. I., Cohn, J. F., Mahoor, M. H., Gangi, D. N. & Messinger, D. S. Darwin's Duchenne: Eye constriction during infant joy and distress. *PLoS One* 8, e80161. https://doi.org/10.1371/journal.pone.0080161 (2013).
- Cordaro, D., Fridlund, A. J., Keltner, D., Russell, J. A. & Scarantino, A. Debate: Keltner and Cordaro vs. Fridlund vs. Russell. http:// emotionresearcher.com/the-great-expressions-debate/ (2015).
- Cohn, J. F., Ertugrul, I. O., Chu, W. S., Girard, J. M., Jeni, L. A. & Hammal, Z. Affective facial computing: Generalizability across domains. In *Multimodal Behavior Analysis in the Wild* 407–441 https://doi.org/10.1016/B978-0-12-814601-9.00026-2 (2019).
 - Cowen, A., Sauter, D., Tracy, J. L. & Keltner, D. Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychol. Sci. Public Interest* 20, 69–90 (2019).
- Koelstra, S. *et al.* Deap: A database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput* 3, 18–31 (2011).
 Rueff-Lopes, R., Navarro, J., Caetano, A. & Silva, A. J. A Markov chain analysis of emotional exchange in voice-to-voice commu-
- nication: Testing for the mimicry hypothesis of emotional contagion. Hum. Commun. Res. 41, 412–434 (2015).

Author contributions

Sh.N and Sa.N. carried out the experiments. Sh.N analyzed the data. All authors designed the plan of this study and wrote the manuscript. All authors reviewed the manuscript. Participants provided written informed consent for publication of their aggregate data before the beginning of the experiment.

Funding

This research was supported by The Telecommunications Advancement Foundation (to SN) and Japan Science and Technology Agency-Mirai Program (Grant No. JPMJMI20D7 to WS).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.N. or W.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023