scientific data

OPEN

DATA DESCRIPTOR

Lights, Camera, Emotion: REELMO's 1060 Hours of Affective Reports to Explore Emotions in Naturalistic Contexts

Erika Sampaolo[1], Giacomo Handjaras[2], Giada Lettieri[3,4] ✉ & Luca Cecchetti[1,4]

Emotions are central to human experience, yet their complexity and context-dependent nature challenge traditional laboratory studies. We present REELMO (REal-time EmotionaL responses to MOvies), a novel dataset bridging controlled experiments and naturalistic affective experiences. REELMO includes 1,060 hours of moment-by-moment emotional reports across 20 affective states collected during the viewing of 60 full-length movies, along with additional measures of personality traits, empathy, movie synopses, and overall liking from 161 participants. It also features fMRI data from 20 volunteers recorded while watching the full-length movie Jojo Rabbit. Complemented by visual and acoustic features as well as semantic content derived from deep-learning models, REELMO provides a comprehensive platform for advancing emotion research. Its high temporal resolution, rich annotations, and integration with fMRI data enable investigations into the interplay between sensory information, narrative structures, and contextual factors in shaping emotional experiences, as well as the study of affective chronometry, mixed-valence states, psychological trait influences, and machine learning applications in affective (neuro)science.

Background & Summary

Literature, performing arts, and modern media, such as movies and television, are highly effective in evoking complex and vivid emotional experiences. This may explain why people engage with them daily[1]. Furthermore, the connection between visual storytelling and emotion is deeply rooted in human culture, originating in early theater and evolving naturally into modern cinema. In this progression, directors have inherited the challenging role of playwrights in discovering the best ways to create scenes that evoke suspense, contempt, tenderness, or relief.

Interestingly, affective sciences have neither fully utilized the extensive knowledge from media research on how to induce and control emotions[2] nor drawn on film art creations proven highly effective in making people laugh or cry. Instead, psychologists and neuroscientists have traditionally relied on stimuli that are simple, static, and limited in contextual information, such as standardized sets of images[3,4], to induce emotions. Although the effectiveness of movies in evoking affective reactions is acknowledged[5,6], researchers have largely overlooked the role of contextual information, a crucial element in emotion elicitation that is well understood by successful screenwriters. Instead, they have favored the use of relatively brief movie excerpts designed to elicit prototypical emotional states, such as sadness or joy[7–10], thereby sacrificing complexity for the sake of standardization.

While existing paradigms have provided valuable insights into the experience of emotions, they may not fully capture the true nature of affective states and their nuances. First, subjective experiences are often complex and multifaceted, characterized by blends of emotions[11], even contradictory ones, resulting in states such as bittersweetness[12]. This complexity, central to many films and genres, is difficult to replicate with brief stimuli. Second, emotional experiences unfold over time, exhibiting dynamic patterns[13] and transitions between states[14]. Studying these dynamics requires longer-form stimuli that can capture the temporal evolution of affective

[1]Social and Affective Neuroscience (SANe) group, MoMiLab, IMT School for Advanced Studies Lucca, Lucca, Italy. [2]Methods for Advanced Biosignal Analysis (MABA) group, MoMiLab, IMT School for Advanced Studies Lucca, Lucca, Italy. [3]Affective Physiology and Interoception (API) group, MoMiLab, IMT School for Advanced Studies Lucca, Lucca, Italy. [4]These authors contributed equally: Giada Lettieri, Luca Cecchetti. ✉e-mail: giada.lettieri@imtlucca.it

responses, which is challenging with short excerpts. Last, the role of context in emotional experience is significant[15]. The meaning and affective impact of a given stimulus can dramatically shift based on the surrounding narrative and accumulated contextual information[16]. For example, an isolated scene might appear neutral, such as a protagonist drinking from a cup. However, if the plot reveals that the cup has been poisoned, transforming this scene into a tragic finale, it will likely elicit strong negative affect. These three factors contribute to determining the ecological validity of emotion reports collected in laboratory settings, a critical consideration in affective science. Notably, the same challenges that are inherent to behavioral research are equally relevant when the goal is to establish where and how emotions arise in the brain[17].

One possible strategy to address these concerns is to present participants with full-length movies, narratives, or musical pieces during laboratory sessions or brain scans, mirroring activities they typically engage in during daily life. This approach, known as naturalistic stimulation, has been successfully employed over the last two decades[18] to investigate several aspects of cognition, such as memory[19], language[20], and perceptual organization[21,22], and has motivated several data-sharing efforts[23–26]. To fully exploit naturalistic stimulation, physiological, behavioral, and brain data are typically complemented by detailed stimulus descriptions. These annotations relate to physical properties, such as scene brightness or soundtrack loudness, or to informational content, such as the presence of specific objects in a scene[22], types of social interactions[25], or word representations derived from embeddings[27]. In the context of affective neuroscience, brain activity can also be paired with reports of portrayed emotions[28,29] or subjective feelings[30,31]. In this regard, we previously developed a method that allows individuals to report their feelings in real-time while viewing full-length movies, both in terms of emotion categories, ranging from a few[32] to as many as 15[33], and affective dimensions[34]. In spite of the success of naturalistic stimulation in advancing our understanding of cognition, this approach has only recently begun to attract significant attention from emotion researchers[17,30,35,36].

Here, we present REELMO (REal-time EmotionaL responses to MOvies), a dataset comprising 637 behavioral experimental sessions in which 152 participants provided real-time reports of their subjective experience using 20 emotion labels while watching 60 full-length movies across 18 genres. This amounts to a total of 1,060 hours of annotations. The emotion reports are complemented by a range of stimulus descriptions, including visual and acoustic features, subtitles for extracting semantic content, movie synopsis recorded after viewing the film, overall liking, as well as the personality and trait empathy of participants. Additionally, for one of these movies (Jojo Rabbit; Waititi, 2019), we provide measures of brain activity collected from 20 individuals during two 1-hour functional magnetic resonance imaging (fMRI) sessions (totaling 40 hours of data). The neuroimaging data and accompanying metadata are organized in a standardized BIDS format, ensuring adherence to best practices for data sharing and interoperability. The dataset is available on Figshare, with comprehensive documentation detailing the experimental procedures and annotation protocols.

REELMO has the potential to drive significant advancements in emotion research and neuroscience. It enables investigations into how low-level sensory information influences felt emotions, as well as the role of narrative structures and contextual information in shaping emotional responses. The dataset also facilitates the study of complex co-occurrences of affective states, addressing longstanding questions, such as the existence of mixed-valence states. Researchers can explore affective chronometry by leveraging the dataset's high temporal resolution and continuous nature, investigating the duration of emotional experiences or differences between emotion categories in terms of their temporal profiles. Additionally, the dataset provides opportunities to study whether between-participant similarities in psychological traits, such as personality or empathy, explain synchronization in emotional reports. By leveraging movie synopses and overall ratings of enjoyment, researchers can study the relationship between the intensity and nature of emotional experiences and memory or appreciation. Importantly, by equipping neuroscientists with a high-quality fMRI dataset, these questions can also be examined through the lens of brain function. Lastly, the scale and richness of REELMO make it suitable for machine learning applications, including affective computing and predictive analyses. We envision REELMO as a collaborative resource that will evolve through contributions from the research community, such as the addition of peripheral physiological measures or eye-tracking data. By fostering such collaborative efforts, the dataset aims to serve as a platform for advancing the study of emotion and its interplay with cognition and perception.

Methods

Participants. A total of 161 Italian young adults (80 females; mean age = 26 years, standard deviation [SD] = 3 years) contributed to the dataset. Of these, 152 individuals (73 females; mean age = 26 years, SD = 3 years) participated in at least one behavioral experiment, during which they watched a full-length movie in the laboratory, annotated their felt emotions in real time, recorded a synopsis of the movie plot, and rated their overall liking of the film. Participants received a reimbursement of €20 per session for their contribution.

Twenty individuals (10 females; mean age = 27 years, SD = 3 years) underwent fMRI scanning while watching one of the 60 movies—Jojo Rabbit—retrospectively reported the emotions they felt during the session, and recorded a synopsis of the movie plot. Of these, nine volunteers took part only in the fMRI experiment, while eleven also contributed to the behavioral dataset by participating in at least one session involving movies other than Jojo Rabbit. Participants in the fMRI experiment received compensation of €50 per session for their contribution.

All participants reported no history of psychiatric or neurological disorders and no use of psychoactive medications. Informed consent was obtained in accordance with the Declaration of Helsinki and approved by the local ethics committee (Comitato Etico Area Vasta Nord Ovest [CEAVNO], Protocol No. 1485/2017). The consent procedure included permission to share anonymized data and the right to withdraw from the study at any time without providing a reason. The study complied with data protection regulations, including the EU General Data Protection Regulation (GDPR, 2016/679) and the Italian legislative decree D.Lgs 196/2003.

Stimuli. We selected 60 movies based on several criteria: duration (at least 1 hour), genre (e.g., comedy, drama, action), recognition (e.g., Oscar nominations, Golden Palm at Cannes), and ratings from both critics and audience (e.g., IMDb.com). All selected movies were original full-length motion pictures dubbed in Italian, with durations ranging from 80 to 148 minutes. A complete list of the movies –including release year, director(s), duration, genre(s)– is provided in Table 1.

To collect real-time affective reports during behavioral experimental sessions, each movie was divided into runs lasting 20 to 30 minutes, allowing participants to take short breaks between runs. Movies were segmented using iMovie software (MacBook Air, Retina, 13-inch, 2019), ensuring that cuts did not disrupt ongoing scenes to preserve narrative continuity. The original movie length, excluding credits, was retained, with an additional fade-in and fade-out applied to each movie segment to ensure smooth transitions (see https://github.com/giacomohandjaras/REELMO/tree/main/movie_frames/movie_frames_v1.0.pdf for details). Each film part was exported in MP4 format, preserving the original aspect ratio.

From the initial sample of 60 full-length movies, we selected Jojo Rabbit for the fMRI experiment. This film portrays the events of World War II from the unique perspective of a German boy who befriends a Jewish girl, offering a poignant exploration of themes such as the absurdity and tragedy of war, as well as the transformative power of love and connection. Jojo Rabbit adeptly interweaves moments of profound grief that evoke tears with scenes of humor and levity, creating an emotionally rich narrative that elicits a wide spectrum of feelings. For the in-scanner viewing sessions, the movie was segmented into eight runs using iMovie software (MacBook Air, Retina, 13-inch, 2019). Each segment included an 8-second overlap (equivalent to 4 TRs, with dummy volumes discarded) from the end of the preceding run, except for the first run, which began with an 8-second black screen (4 TRs discarded). The last run concluded with an 8-second black screen. Movie segments were exported in MP4 format, maintaining the original aspect ratio but adjusted to a resolution of 800×600 to ensure compatibility with the stimulation apparatus.

Procedure. After eligibility screening, participants completed four self-report questionnaires online via the Qualtrics platform (https://www.qualtrics.com). The questionnaires were divided into two separate links, with each requiring participants to provide consent for data processing before access. Responses were anonymized using unique alphanumeric codes assigned to each participant. The first link included three psychometric assessments in their Italian versions: the Interpersonal Reactivity Index[37] (IRI), the Empathy Index[38] (EI), and the Affective and Cognitive Measure of Empathy[39] (ACME), which assessed various dimensions of empathy. The second link contained the Italian version of the Big Five Inventory-2[40] (BFI-2; https://www.situationslab.com/big-five-inventory-2), a tool for evaluating personality traits.

Before each behavioral and brain imaging experimental session, participants' mood was assessed using the Positive and Negative Affect Schedule (PANAS)[41]. Each session took place in a soundproof experimental room, where participants sat alone in a comfortable setup. Detailed instructions about the task were provided before data collection began. Participants were asked to report their feelings in real-time at any moment during the movie (10-Hz sampling rate) by selecting from 20 emotion labels, balanced between positive and negative affective states. To mitigate cognitive load and ensure usability, we adopted a series of intuitive design strategies. Emotion labels were grouped by valence: positive on the left, negative on the right, and mixed or ambivalent in the center of the screen. Within each group, labels were arranged alphabetically. This layout encouraged a simple heuristic: upon detecting a change in emotional state, participants first judged its valence (pleasant, unpleasant, or ambiguous), then scanned the appropriate screen region, and finally selected the matching label based on alphabetical order. To support this process and ensure a shared understanding of emotion labels, participants were informed about the structure of the interface and also given a glossary with definitions of all emotional terms during the familiarization phase. Also, each participant watched the selected movie for the first time during the experimental session to maintain consistency and avoid the influence of expectations on their emotional experience[1].

Movies were displayed on the upper portion of a 24-inch HP monitor connected to a MacOS 10.13.6 laptop. The emotion labels were shown on the bottom of the screen, arranged as evenly spaced bars along the horizontal axis. Participants wore headphones and used an external QWERTY keyboard to interact with the interface. Using the keyboard's arrow keys, participants navigated between emotion labels, while specific keys ("q" to increase, "a" to decrease) adjusted the intensity of their emotional response. They could assign three levels of intensity to each label –0% (not felt), 50% (moderate intensity), or 100% (high intensity)– and select multiple emotions simultaneously. Selected labels were outlined in red, with the fill color of the corresponding bar dynamically indicating the chosen intensity.

To familiarize participants with the procedure, a training phase was conducted before the movie session. This included 39 practice trials where participants selected one or two labels displayed on the screen in a random order, using the same controls required for the main task. The task was implemented using in-house software designed for tagging events during naturalistic stimulation, based on Psychtoolbox[42] (version 3.0.16). This software provides a versatile interface for capturing real-time emotional responses, as demonstrated in previous studies[32–34].

After the movie, participants rated their overall enjoyment on a scale from 0 to 100 and provided a detailed synopsis of the plot using Audacity software (version 2.4.2). For each movie, emotional reports were collected in real-time over the film's entire duration, from ten participants balanced for gender. For Jojo Rabbit, selected for the fMRI experiment, emotional reports were collected from 47 participants. Depending on the movie's duration, each behavioral experimental session lasted between 120 and 150 minutes.

For the fMRI experimental sessions, participants watched the full-length movie Jojo Rabbit split into eight parts over two 1-hour sessions while lying in the scanner. Consistent with previous studies employing naturalistic stimulation[23], participants were instructed to remain still and simply enjoy the movie. Stimuli were delivered

MOVIE TITLE	YEAR	DIRECTOR	LENGTH (mins)	IMDb GENRES (January 2022)
10 Cloverfield Lane	2016	D. Trachtenberg	105	Drama, Horror, Science-Fiction
500 Days of Summer	2009	M. Webb	95	Comedy, Drama, Romance
About Time	2013	R. Curtis	124	Comedy, Drama, Fantasy
African Cats	2011	A. Fothergill, K. Scholey	89	Adventure, Documentary
American Pie	1999	P. Weitz, C. Weitz	95	Comedy
Autumn in New York	2000	J. Chen	103	Drama, Romance
Bad Tales	2020	D., F. D'Innocenzo	98	Drama, Mystery, Thriller
Borat	2006	L. Charles	84	Comedy
Borat 2	2020	J. Woliner	95	Comedy
Bruce Almighty	2003	T. Shadyac	91	Comedy, Fantasy
Carnage	2011	R. Polanski	79	Comedy, Drama
Children of Men	2006	A. Cuaron	110	Action, Drama, Science-Fiction
Dallas Buyers Club	2013	J.M. Vallée	117	Biography, Drama
Deadpool	2016	T. Miller	108	Action, Adventure, Comedy
Dogman	2018	M. Garrone	120	Crime, Drama, Thriller
Due Date	2010	T. Phillips	95	Comedy, Drama
Elephant	2003	G. Van Sant	81	Crime, Drama, Thriller
Everything Is Illuminated	2005	L. Schreiber	105	Comedy, Drama
Ex Machina	2015	A. Garland	108	Drama, Science-Fiction, Thriller
Extremely Loud & Incredibly Close	2011	S. Daldry	129	Adventure, Drama, Mystery
Funny Games	2007	M. Haneke	111	Crime, Drama, Thriller
Hachi: A Dog's Tale	2009	L. Hallstrom	93	Biography, Drama, Family
Hot Fuzz	2007	E. Wright	121	Action, Comedy, Mystery
I, Tonya	2017	C. Gillespie	121	Biography, Comedy, Drama
Isle of Dogs	2018	W. Anderson	101	Adventure, Animation, Comedy
Jojo Rabbit	2019	T. Waititi	108	Comedy, Drama, War
Joker	2019	T. Phillips	123	Crime, Drama, Thriller
Kingsman	2014	M. Vaughn	129	Action, Adventure, Comedy
Klaus	2019	S. Pablos	96	Adventure, Animation, Comedy
Little Miss Sunshine	2006	J. Dayton, V. Faris	101	Comedy, Drama
Lost in Translation	2003	S. Coppola	102	Comedy, Drama
Mama	2013	A. Muschietti	100	Fantasy, Horror, Science-Fiction
Masterminds	2015	J. Hess	93	Biography, Comedy, Crime
Meet the Parents	2000	J. Roach	108	Comedy, Romance
Midsommar	2019	A. Aster	148	Drama, Horror, Mystery
Monkey Kingdom	2015	M. Linfield, A. Fothergill	82	Documentary
Oldboy	2003	P. Chan-wook	120	Action, Drama, Mystery
Pan's Labirynth	2006	G. Del Toro	118	Drama, Fantasy, War
School of Rock	2003	R. Linklater	108	Comedy, Music
Shaun of the Dead	2004	E. Wright	100	Comedy, Horror
Snowpiercer	2013	J. Bong	126	Action, Drama, Romance
Soul	2020	P. Docter, K. Powers	90	Adventure, Animation, Comedy
Swiss Army Man	2016	D. Kwan, D. Scheinert	97	Comedy, Drama, Fantasy
The 40 Year Old Virgin	2005	J. Apatow	116	Comedy, Romance
The Cabin in the Woods	2012	D. Goddard	95	Horror, Mystery, Thriller
The Dictator	2012	L. Charles	83	Comedy
The Father	2020	F. Zeller	97	Drama, Mystery
The Gentlemen	2020	G. Ritchie	113	Action, Comedy, Crime
The Hangover	2009	T. Phillips	100	Comedy
The Intouchables	2011	O. Nakache, E. Toledano	112	Biography, Comedy, Drama
The Killing of a Sacred Deer	2017	Y. Lanthimos	116	Drama, Horror, Mystery
The King's Speech	2010	T. Hooper	118	Biography, Drama, History
The Lobster	2015	Y. Lanthimos	118	Comedy, Drama, Romance
The Platform	2019	G. Gaztelu-Urrutia	94	Horror, Science-Fiction, Thriller
The Wind Rises	2013	H. Miyazaki	126	Animation, Biography, Drama
Three Billboards Outside Ebbing, Missouri	2017	M. McDonagh	115	Comedy, Crime, Drama

Continued

MOVIE TITLE	YEAR	DIRECTOR	LENGTH (mins)	IMDb GENRES (January 2022)
Toc Toc	2017	V. Villanueva	90	Comedy, Family, Fantasy
Up	2009	P. Docter, B. Peterson	96	Adventure, Animation, Comedy
Wolfwalkers	2020	T. Moore, R. Stewart	103	Adventure, Animation, Family
Zoolander	2001	B. Stiller	89	Comedy

Table 1. List of the 60 full-length movies, including title, year of production, director(s), length in minutes, genres according to the IMDB database as of January 2022.

using MR-compatible LCD goggles and headphones (VisualStim Resonance Technology). The goggles provided a resolution of 800 × 600 pixels at 60 Hz, with a visual field of 30° × 22° on a 5-inch display. The headphones offered 30 dB noise attenuation and a frequency response range of 40 Hz to 40 kHz. Movie clips were presented using VLC Videolan (version 3.0.12).

To minimize participant discomfort, an interleaved acquisition design was employed. Each session began with the first participant entering the scanner to complete anatomical scans, followed by viewing the first three or four movie parts, depending on their compliance. While the first participant took a break, the second participant underwent the same protocol. Subsequently, the first participant re-entered the scanner to complete additional anatomical scans for co-registration and watch the remaining five or four parts of the movie. The second participant followed the same sequence. Importantly, as in the behavioral experiments, participants viewed the Jojo Rabbit movie for the first time during the fMRI sessions, ensuring untainted emotional responses.

During the break and after completing the fMRI scans, participants reviewed scenes from the movie clips they had just watched. They reported their subjective feelings using the 20 emotion categories applied in the behavioral experimental sessions, specifying the onset of emotional instances. At the conclusion of the two fMRI sessions, participants recorded a detailed plot synopsis using Audacity software (version 2.4.2). The entire fMRI experiment lasted approximately three hours per participant.

Demographics, stimuli, and questionnaires' scoring. Participants completed the IRI, a self-report questionnaire consisting of 28 items divided into four 7-item subscales. These subscales assess distinct dimensions of empathy: Fantasy (FS), Perspective-Taking (PT), Empathic Concern (EC), and Personal Distress (PD). Following the IRI, participants completed the EI, which includes two subscales: Emotional Contagion (EMOCONT) and Behavioral Contagion (BEHACONT). Each subscale contains 7 items, for a total of 14 items. Participants also filled out the ACME, which evaluates three dimensions of empathy across three 12-item subscales: Cognitive Empathy (COG), Affective Resonance (RES), and Affective Dissonance (DIS). In a separate link, participants completed the BFI-2, which measures five personality traits across subscales for Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Negative Emotionality (N).

All participant-related information, including ID, gender, age, education level, performance at the training task, completed experimental sessions, and scores from the questionnaires, is available in the repository within the dataset-description.tsv file. Additionally, we stored a comprehensive descriptive file for the selected movies, movies.tsv, which includes: title, director(s), year of release, genre classification based on IMDb (as of January 2022), full-length duration (hh:mm:ss), frames per second (fps), length of each experimental run (hh:mm:ss), total number of runs, and total duration of the experiment. The participant-movie-watchlist.tsv file is structured as a matrix with participant IDs as rows and movie titles as columns. Each cell contains a value of "1" if the participant viewed the corresponding movie and "0" otherwise. Each movie folder also includes a text file listing all participants who watched the movie. This file contains additional details, such as their PANAS scores collected before the experiment and relevant session information.

Emotion features. The data output from each experimental session, encompassing all recorded parameters, is stored in the emotion-features file. For each run, we collected the participant id, the movie title, the list of the 20 tagging categories, the rating sampling frequency, and the timeseries of affective ratings, structured as the number of categories by the length of the run, downsampled by a factor of 10. Additionally, for each run, we aggregated the downsampled affective ratings provided by participants to generate a group-level report of subjective feelings. This aggregated data is stored in the subject-avg file.

Visual properties. We extracted visual features using VGG-19[43], a pre-trained Convolutional Neural Network (CNN) comprising 16 convolutional layers and three fully connected layers. To prepare the training data, we preprocessed images in accordance with VGG-19 methodology. Using the FFMPEG tool (https://www.ffmpeg.org/), we selected the central frame from each second of the 60 movies. For non-isometric images, we adjusted the aspect ratio to 4:3 by cropping a square at the center with dimensions equal to the stimulus height, followed by resampling to a resolution of 224 × 224 pixels. For isometric images, we directly resized them to the same resolution. The processed images were then fed into VGG-19, where a series of convolutional layers acted as filters to extract complex features in a feed-forward manner. Max-pooling layers were used to reduce spatial information while transforming the images into feature maps. For our analyses, we used VGG-19 as a feature extractor, focusing on the first fully connected layer (fc6). In this layer, all inputs from the preceding layers were connected to each activation unit of the subsequent layer, producing a final matrix of 4,096 dimensions for each image. To derive global visual features, we transformed RGB color values into CIE coordinates and extracted lightness, contrast, chroma, and hue. This was achieved by remapping the movie frames into the JzAzBz perceptually uniform color space[44].

Acoustic properties. For the extraction of acoustic features, we used VGG-ish[45], a pre-trained deep-learning model designed for audio classification, which consists of six convolutional layers and three fully connected layers. We preprocessed the audio stream of the movies to match the input characteristics used to train the classifier by computing the log-mel spectrogram for each second of the stimuli. Next, the acoustic signals were converted into input images, which were resized to 96×64 pixels to be fed into the network. Similar to the visual features, we used VGG-ish to extract acoustic features from the first fully connected layer (fc1-1), resulting in a total of 4,096 dimensions for each second of sound. Additionally, we estimated the power spectrum[46] and generated a vector of 449 values representing the power of the acoustic signal across frequencies from 0 Hz to 15 kHz, distributed in 33.5 Hz bands.

Plot recording. The MP3 files containing the movie plot, as recorded by participants at the end of each experimental session, were manually transcribed and translated into English using ChatGPT (version 4), with prompts to ensure fidelity to the original.

Semantic properties. We obtained both English and Italian subtitles for all 60 movies and synchronized them with the corresponding audio streams using the freely available subtitle editing software Jubler (https://www.jubler.org/). To extract semantic features, we processed the English subtitles using GPT-3. Specifically, we transformed the sentences into numerical vectors by applying the Curie embedding model, which generates a 4,096-dimensional array for each sentence. These vectors were then resampled at the temporal level, creating a feature vector for each second of the movie. For sentences spanning multiple seconds, the corresponding arrays were replicated to cover the entire duration of the sentence.

fMRI data. Brain activity was recorded from 20 participants over two 1-hour sessions using a Philips 3 T Ingenia scanner equipped with a 32-channel head coil. Functional images were collected across eight runs (totaling 3,087 timepoints per participant) using gradient recall echo planar imaging (GRE-EPI) with the following parameters: TR = 2,000 ms, TE = 30 ms, FA = 75°, FOV = 240 mm, in-plane resolution = 80×80, slice thickness = 3 mm, acquisition voxel size = $3 \times 3 \times 3$ mm, and 39 slices (full brain coverage including the cerebellum). The slices were acquired in an alternating plus direction, denoted as *alt* + *z* in AFNI[47] notation. For each of the two 1-hour sessions, a high-resolution three-dimensional T1-weighted anatomical image was captured using a magnetization-prepared rapid gradient echo (MPRAGE) sequence and the following parameters: TR = 7 ms; TE = 3.2 ms; FA = 9°; FOV = 224, acquisition matrix = 224×224; slice thickness = 1 mm; voxel size = $1 \times 1 \times 1$ mm; 180 sagittal slices. In addition, for all participants except sub-03, T2-weighted anatomical images (TR = 2,500 ms; TE = 251 ms; FA = 90°; FOV = 224, reconstructed matrix = 224×224; slice thickness = 1 mm; voxel size = $1 \times 1 \times 1$ mm; 196 sagittal slices) and Fluid-Attenuated Inversion Recovery (FLAIR) images (TR = 8,000 ms; TE = 360 ms; FA = 90°; FOV = 224, reconstructed matrix = 480×480; slice thickness = 0.550 mm; voxel size = $0.550 \times 0.533 \times 0.533$ mm; 332 sagittal slices) were also acquired.

Preprocessing. MRI data were converted from DICOM to NIfTI format (dcm2nii v1.0.20211006) and preprocessed using AFNI[47] v22.1.10, FSL[48] v6.0.5.1 and ANTS[49] v2.3.5–126. The preprocessing of anatomical data was conducted using a robust pipeline that ensures high-quality alignment and normalization across participants. First, all anatomical images were defaced and cropped to focus on the brain region only (FSL's *robustfov*). Subsequently, the images were resampled to a resolution of 1 mm isotropic using sinc interpolation with a Hanning window of 7 voxels (FSL's *flirt*). Initial brain extraction was performed on all anatomical scans using FSL's BET tool, applying an inclusive threshold (f = 0.3) to retain all relevant brain structures. The T1-weighted image acquired at the beginning of the first MRI session was used as a reference for all the other anatomical images. For each participant, this was aligned to the anterior commissure-posterior commissure (ACPC) plane. The alignment was achieved by calculating the transformation matrix between the brain-extracted image and the MNI152 2009c symmetric template using affine registration (*flirt*). The affine transformation matrix was then converted to a rigid-body matrix using FSL's *aff2rigid*, and the transformation was applied to the original T1 image using sinc interpolation (7 voxels Hanning window). A more refined brain extraction was performed on the ACPC-aligned T1-weighted image using ANTs' antsBrainExtraction.sh with the OASIS template. The resulting binary brain mask was used to generate the final brain-extracted image by multiplying the mask with the ACPC-aligned image using AFNI's *3dcalc*. Bias field inhomogeneity correction was then applied to the brain-extracted image using ANTs' antsAtroposN4.sh. The same correction was also applied to the non-brain-extracted T1-weighted image by dividing it by the bias field with *3dcalc*. The other acquired anatomical images, such as T2-weighted and FLAIR images (when available), were co-registered to the ACPC-aligned T1-weighted image using *flirt* with six degrees of freedom and sinc interpolation (7 voxels Hanning window). The binary brain mask generated from the ACPC-aligned T1-weighted image was then applied to these anatomical images to isolate the brain area. For participants with multiple T1-weighted images, an average image was created in the ACPC space using AFNI's *3dMean* tool, generating both brain-extracted and non-brain-extracted versions. Then, the brain-extracted average T1-weighted image was linearly and non-linearly registered to the MNI152 2009c symmetric template using AFNI's *3dQwarp* with a minimum patch of 21 mm, applying a 3 mm Gaussian blur to the image and no blur to the template. The resulting linear and non-linear transformations were applied to all ACPC-aligned anatomical images using AFNI's *3dNWarpApply* with sinc interpolation and a Hanning window (5 mm).

Functional preprocessing involved several steps for each participant and functional run. To adjust for slice-dependent acquisition delays, we performed slice timing correction (AFNI's *3dTShift*) using Fourier interpolation and selecting the first slice of the volume as a reference. Brain extraction was carried out with BET to

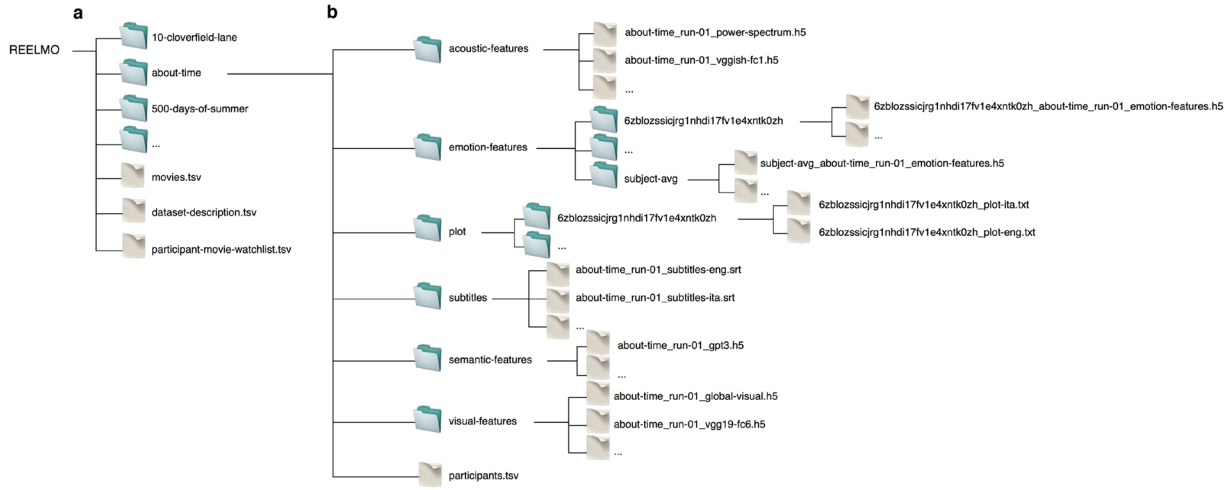

Fig. 1 Structure of the behavioral dataset. (**a**) Overview of the general directory structure, organized into folders corresponding to each of the 60 full-length movies. (**b**) Detailed contents of each movie folder, including computational features and semantic annotations derived from the films.

remove non-brain tissue and applied to the entire 4-dimensional dataset (-F option). Head motion was corrected by aligning each volume to the midpoint of the time series using rigid-body transformations with six degrees of freedom (AFNI's *3dvolreg*). Motion parameters and transformation matrices were inspected, and average voxel intensity images were calculated (AFNI's *3dTstat*). EPI-to-anatomical alignment was achieved using both linear (AFNI's align_epi_anat.py, giant move, and lpc cost function) and nonlinear (*3dQwarp*) registration to account for EPI's geometric distortions. Functional data were subsequently transformed into the MNI152 standard space in a single interpolation step by catenating the linear transformation matrices and non-linear deformation fields obtained during the EPI-to-anatomical and anatomical-to-template registrations (*3dNwarpApply*, wsinc5 interpolation). Standard-space functional images retained the original 3 mm isotropic resolution, and brain masks were similarly transformed using nearest-neighbor interpolation. Functional images were smoothed to achieve a 6 mm full width at half maximum using AFNI's *3 dBlurToFWHM*. Signal intensity scaling was performed to express BOLD activity as percentage changes (*3dcalc*). A conjunction brain mask was generated by combining masks from all runs, ensuring complete brain coverage across all sequences (AFNI's *3dmask_tool*). Deconvolution analysis was used to regress out motion parameters and slow signal trends while accounting for temporal autocorrelation (AFNI's *3dDeconvolve* with the *-polort A* option and *3dREMLfit*). For each participant, all functional runs were provided as input to *3dDeconvolve* and *3dREMLfit*, which automatically concatenated the time series across runs. Regressors of no interest (i.e., head motion) were concatenated accordingly and passed to the same functions, ensuring alignment with the full duration of the combined time series. Since the fitted model included only nuisance regressors, the *3dDeconvolve* and *3dREMLfit* functions were instructed to output the residual time series (*-errts* option), which were then used in subsequent analyses (i.e., encoding and intersubject correlation analyses). Group-level averages of brain-extracted EPI and anatomical images, as well as EPI brain masks, were generated using *3dMean*.

Data Records

Data are available from the Figshare repository at https://doi.org/10.6084/m9.figshare.28255745[50]. README files within the repository provide detailed descriptions of the available content. Data are organized as.tar files, each corresponding to a specific movie and its related data. The organization of the behavioral and fMRI data is illustrated in Figs. 1, 2, respectively.

While we are unable to share the raw movie stimuli due to copyright restrictions, we provide all necessary information to enable other researchers to obtain and reproduce the same stimuli independently. Specifically, the file *movie_frames_v1.0.pdf*, available on the project's GitHub page (https://github.com/giacomohandjaras/REELMO/tree/main/movie_frames), includes the 10-digit Amazon Standard Identification Number (ASIN) or the media platform associated with each movie, allowing for unambiguous identification and access. This file also includes the timing and screenshots of the first and last fully visible frames (excluding fade-ins and fade-outs) for each run and movie. This information allows others to apply the same trimming procedure to the original source material, ensuring consistency in stimulus presentation.

Participants' demographics and questionnaires' scoring. Location: dataset-description.tsv, participant-movie-watchlist.tsv

File formats: tab-separated value

Description: The dataset-description.tsv file includes participants' demographic information—specifically, sex, age, and years of education—as well as their accuracy scores at the training, and their scores on empathy and personality questionnaires. The data is organized with one row per participant and one column per variable

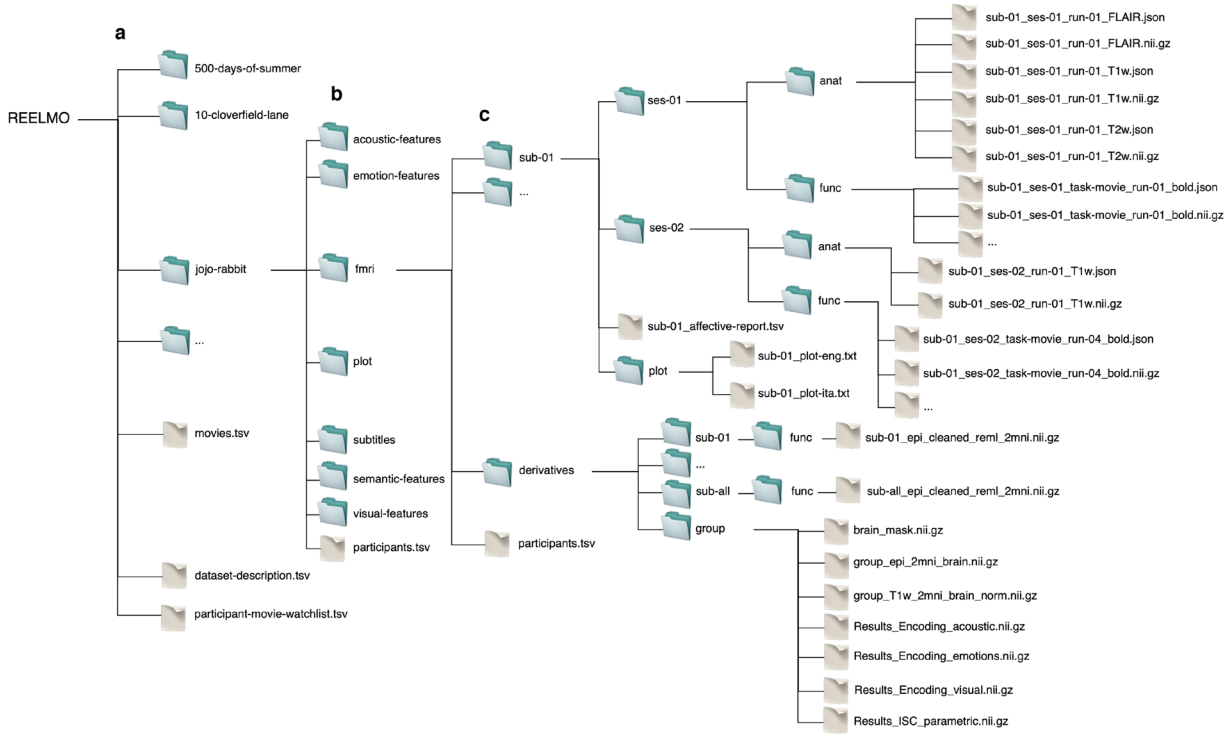

Fig. 2 Structure of the fMRI dataset. (**a**) Overview of the general directory structure. (**b**) Detailed contents of the "jojo-rabbit" folder, including the "fmri" subfolder. (**c**) Organization of subject-specific directories, showing functional and anatomical data along with derivatives.

or questionnaire subscale. Another file, participant-movie-watchlist.tsv, lists which participants annotated their emotions for each movie and, for those who also took part in the fMRI experiment, provides the corresponding identifier used in the jojo-rabbit/fmri folder. This file is similarly organized with one row per participant and one column per movie.

Participants in the behavioral experimental session. Location: <movie>/participants.tsv

File format: tab-separated value

Description: Participants who reported their felt emotions while watching a specific movie. This file includes their unique ID, mood, overall rating of the movie, indication of whether the plot was recorded, and information on any exclusions and their reasons. The data is structured with one row per participant and each variable as a separate column.

Stimuli. Location: movies.tsv

File formats: tab-separated value

Description: A summary of the movies included in the database, detailing their title, identification code, release year, director(s), genre(s), duration (in hh:mm:ss format), frame rate (frames per second), number of runs, duration of each run (in hh:mm:ss format), total experiment duration, and any additional notes. The data is organized with one row per movie and each variable as a separate column.

Emotion features. Location: <movie>/emotion-features/<participant-ID>/<participant-ID>_<movie>_run-0[1-n]_emotion-features.h5, <movie>/emotion-features/subject-avg/subject-avg_<movie>_run-0[1-n]_emotion-features.h5

Files format: Hierarchical Data Format

Description: Data from a single run and participant, in which felt emotions were annotated while watching a specific movie. The HDF5 (.h5) file includes the following elements: the movie title (/Movie), a matrix of emotion annotations (/Ratings), the sampling frequency of the annotations (/RatingsSamplingFrequency), the participant's identification number (/Subject), and the list of emotion categories used for annotation (/TaggingCategories).

Data from one run, with emotion annotations averaged across participants. The HDF5 (.h5) file contains: the movie title (/Movie), a matrix of emotion annotations averaged across individuals (/Ratings), the sampling frequency of the annotations (/RatingsSamplingFrequency), an indication that the data is averaged across participants (/Subject), and the list of emotion categories used for annotation (/TaggingCategories).

Visual properties. Location: <movie>/visual-features/<movie>_run-0[1-n]_vgg19-fc6.h5, <movie>/visual-features/<movie>_run-0[1-n]_global-visual.h5

Files format: Hierarchical Data Format

Description: Visual features extracted from the fully connected layer 6 of the VGG network for a specific movie segment. The HDF5 (.h5) file contains: the sampling frequency at which the activations were recorded (/SamplingFrequency), and the corresponding layer activations (/fc6).

Global visual features computed for a specific movie segment. The HDF5 (.h5) file includes: the sampling frequency of feature extraction (/SamplingFrequency), along with measures of chroma (/chroma), contrast (/contrast), hue (/hue), and lightness (/lightness) at each sampling point.

Acoustic properties. Location: <movie>/acoustic-features/<movie>_run-0[1-n]_vggish-fc1.h5, <movie>/acoustic-features/<movie>_run-0[1-n]_power-spectrum.h5

Files format: Hierarchical Data Format

Description: Acoustic features extracted from the fully connected layer 1 of the VGGish network for a specific movie segment. The HDF5 (.h5) file includes: the sampling frequency at which the activations were recorded (/SamplingFrequency), and the corresponding layer activations (/fc1).

Data related to the sound power of a specific movie segment. The HDF5 (.h5) file contains: the sampling frequency of the sound power computation (/SamplingFrequency), the frequency bands used to segment the audio spectrum (/frequency_bands), and the power values associated with each band (/power).

Plot recording. Location: <movie>/plot/<participant-ID>/<participant-ID>_plot-eng.txt, <movie>/plot/<participant-ID>/<participant-ID>_plot-ita.txt

Files format: text file (UTF-8)

Description: The movie plot transcription provided by a specific participant and translated into English. The data is stored as a single line of text.

The movie plot transcription provided by a specific participant in Italian. The data is stored as a single line of text.

Subtitles. Location: <movie>/subtitles/<movie>_run-0[1-n]_subtitles-eng.srt, <movie>/subtitles/<movie>_run-0[1-n]_subtitles-ita.srt

Files format: SubRip Subtitle

Description: English subtitles for a specific movie segment. Each entry includes a sequential subtitle ID, the start and end times (formatted as hh:mm:ss,MMM– > hh:mm:ss,MMM), and the subtitle text in English. Italian subtitles for a specific movie segment. Each entry includes a sequential subtitle ID, the start and end times (formatted as hh:mm:ss,MMM– > hh:mm:ss,MMM), and the subtitle text in Italian.

Semantic features. Location: <movie>/semantic-features/<movie>_run-0[1-n]_gpt3.h5

File format: Hierarchical Data Format

Description: Semantic features for a specific movie segment, extracted using the GPT-3 model. The HDF5 (.h5) file contains: the sampling frequency at which sentence embeddings were generated (/SamplingFrequency), the embeddings themselves (/gpt3), and the corresponding subtitle sentence for each sampling point (/subtitles).

fMRI metadata. Location: jojo-rabbit/fmri/dataset_description.json, jojo-rabbit/fmri/participants.json, jojo-rabbit/fmri/participants.tsv, jojo-rabbit/fmri/README, jojo-rabbit/fmri/CHANGES, jojo-rabbit/fmri/task-movie_bold.json, jojo-rabbit/fmri/sub-[01-20]/ses-01/anat/sub-[01-20]_ses-01_run-01_[T2w-FLAIR/T1w/T2w].json,
jojo-rabbit/fmri/sub-[01-20]/ses-01/func/sub-[01-20]_ses-01_task-movie_run-0[1-n]_bold.json,
jojo-rabbit/fmri/sub-[01-20]/ses-02/anat/sub-[01-20]_ses-02_run-01_T1w.json,
jojo-rabbit/fmri/sub-[01-20]/ses-02/func/sub-[01-20]_ses-02_task-movie_run-0[n + 1-8]_bold.json,
jojo-rabbit/fmri/sub-[01-20]/plot/sub-[01-20]_plot-eng.txt, jojo-rabbit/fmri/sub-[01-20]/plot/sub-[01-20]_plot-ita.txt, jojo-rabbit/fmri/sub-[01-20]/sub-[01-20]_affective-report.tsv

Files format: JavaScript Object Notation, text file (UTF-8), tab-separated value

Description: Metadata accompanying the fMRI data collected while participants watched the full-length movie Jojo Rabbit in the scanner. The dataset_description.json file includes essential information such as the dataset name, BIDS version, authors, DOI, and license. The participants.json file describes the variables recorded in participants.tsv, which include the participant identification number, sex, age, years of education, whether the participant provided a brief summary of their emotional experience after the fMRI session, whether they gave a verbal summary of the movie plot, whether the participant was excluded, and any issues related to data acquisition. In participants.tsv, the data is organized with one row per participant, with each variable represented in a separate column.

The README file provides a concise description of the fMRI dataset along with detailed information on the data structure and file contents. The CHANGES file lists all dataset versions along with their respective creation dates. The task-movie_bold.json file contains parameters related to the movie task functional acquisition, including the scanner vendor and model, the strength of the static magnetic field, slice thickness, echo time, repetition time, flip angle, and the software used to convert DICOM files to NIfTI format.

Each participant's folder (sub-[01–20]) contains individual JSON files that store acquisition parameters for the anatomical ([T2w-FLAIR/T1w/T2w]) and functional scans (task-movie_run-0[1–8]_bold) for each session ([ses-01/ses-02]). Additionally, the dataset includes transcriptions of the movie plot in both Italian (sub-[01–20]_plot-ita.txt) and English (sub-[01–20]_plot-eng.txt), as well as emotion annotations (sub-[01–20]_affective-report.tsv).

The plot transcriptions are stored as plain text, with one line per file. The affective reports are organized as a matrix, with each row corresponding to a run and each column to a specific emotion category. Each cell contains the timing (in hh:mm:ss format) at which the participant retrospectively reported experiencing a given emotion. Multiple instances of the same emotion are separated by semicolons, while the absence of a reported emotion is indicated by "n/a".

fMRI data. Location: jojo-rabbit/fmri/sub-[01-20]/ses-01/anat/sub-[01-20]_ses-01_run-01_[T2w-FLAIR/T1w/T2w].nii.gz,
jojo-rabbit/fmri/sub-[01-20]/ses-01/func/sub-[01-20]_ses-01_task-movie_run-0[1-n]_bold.nii.gz,
jojo-rabbit/fmri/sub-[01-20]/ses-02/anat/sub-[01-20]_ses-02_run-01_T1w.nii.gz,
jojo-rabbit/fmri/sub-[01-20]/ses-02/func/sub-[01-20]_ses-02_task-movie_run-0[n + 1-8]_bold.nii.gz
Files format: Neuroimaging Informatics Technology Initiative
Description: Functional and structural magnetic resonance imaging data after conversion from DICOM to compressed NIFTI format. Each participant's folder (sub-[01–20]) stores individual compressed NIFTI files relative to the anatomical ([T2w-FLAIR/T1w/T2w]) and functional scans (task-movie_run-0[1–8]_bold) for each session ([ses-01/ses-02]).

fMRI derivatives. Location: jojo-rabbit/fmri/derivatives/sub-[01-20]/func/sub-[01-20]_epi_cleaned_reml_2mni.nii.gz, jojo-rabbit/fmri/derivatives/sub-all/func/sub-all_epi_cleaned_reml_2mni.nii.gz,
jojo-rabbit/fmri/derivatives/group/brain_mask.nii.gz, jojo-rabbit/fmri/derivatives/group/group_epi_2mni_brain.nii.gz, jojo-rabbit/fmri/derivatives/group/group_T1w_2mni_brain_norm.nii.gz, jojo-rabbit/fmri/derivatives/group/Results_Encoding_acoustic.nii.gz, jojo-rabbit/fmri/derivatives/group/Results_Encoding_emotions.nii.gz, jojo-rabbit/fmri/derivatives/group/Results_Encoding_visual.nii.gz,
jojo-rabbit/fmri/derivatives/group/Results_ISC_parametric.nii.gz
Files format: Neuroimaging Informatics Technology Initiative
Description: This folder contains derivative files generated from the processing of the (f)MRI data using the code available at https://github.com/giacomohandjaras/REELMO. Each participant's folder (sub-[01–20]) includes compressed NIfTI files with voxelwise time series of brain activity recorded during movie watching. These time series are concatenated across runs, preprocessed to remove physiological and scanner-related artifacts, and normalized to the MNI152 standard space (sub-[01–20]_epi_cleaned_reml_2mni.nii.gz). The sub-all folder includes group-level voxelwise data, averaged across all participants (sub-all_epi_cleaned_reml_2mni.nii.gz). The group folder, instead, contains several group-level files: a binary brain mask used to distinguish brain voxels from non-brain tissue (brain_mask.nii.gz), a 3D group EPI image generated by averaging functional images across time points and participants (group_epi_2mni_brain.nii.gz), and a 3D group T1-weighted image created by averaging participants' individual T1w scans (group_T1w_2mni_brain_norm.nii.gz). All of these files are registered to the MNI152 standard space. The group folder also includes results from encoding analyses conducted on acoustic (Results_Encoding_acoustic.nii.gz), visual (Results_Encoding_visual.nii.gz), and emotional (Results_Encoding_emotions.nii.gz) features. Each of these 4D NIfTI files contains the coefficient of determination (R^2) for the full model, log-transformed uncorrected p-values (log_pval-raw_parametric), FDR-corrected p-values (log_pval-fdr95_parametric), and the regression coefficients (including the intercept and one slope per regressor in the encoding matrix). Last, the folder includes the results from the intersubject correlation analysis (Results_ISC_parametric.nii.gz), also in MNI152 standard space. This 4D file stores, for each voxel, the ISC value averaged across all participant pairs, along with the log-transformed uncorrected and FDR-corrected p-values (log_pval-raw_parametric and log_pval-fdr95_parametric, respectively).

Technical Validation

Emotional labels. The 20 emotion labels used by participants to describe their moment-by-moment feelings during movie watching (eight positive, nine negative, and three with no clear valence) were selected from the affective taxonomy we recently proposed[51]. This taxonomy was derived from an analysis of spoken and written descriptions of 265 affective states, emphasizing the role of language in shaping how emotions are conceptualized. The 20 labels span diverse semantic clusters across the emotional spectrum and are: admiration, amusement, excitement, joy, amazement, contentment, relief, tenderness, compassion, confusion, surprise, agitation, anguish, disappointment, uneasiness, disgust, contempt, fear, anger, and sadness, along with their Italian translations (ammirazione, divertimento, eccitazione, gioia, meraviglia, soddisfazione, sollievo, tenerezza, compassione, perplessità, sorpresa, agitazione, angoscia, delusione, disagio, disgusto, disprezzo, paura, rabbia, tristezza). This selection aligns with recent research on affective taxonomies[11,52,53], which highlights a consensus around distinct emotional states rooted in varying appraisals and with specific antecedents.

The analysis of the training data indicate that the task was manageable for the vast majority of participants: the median accuracy was 93.10%, with the 25th percentile at 86.21% and the 5th percentile at 72.41%. The inclusion of these metrics in the dataset-description.tsv allows future users of the dataset to eventually apply their own inclusion criteria based on training performance or to incorporate this information in data analysis.

Data aggregation and analysis. Single-participant emotion reports for each movie segment (i.e., run) were concatenated to generate an overall emotional profile for the full film. Figure 3a displays the emotion annotations from three representative participants for Jojo Rabbit, with the color scale indicating the presence and intensity of emotional experiences over time. To derive group-level annotations for each movie—illustrated in Fig. 3b for Jojo Rabbit—individual ratings were first binarized to indicate only the presence or absence of each emotion at each timepoint, regardless of intensity, and then summed across participants. The resulting emotion-by-time

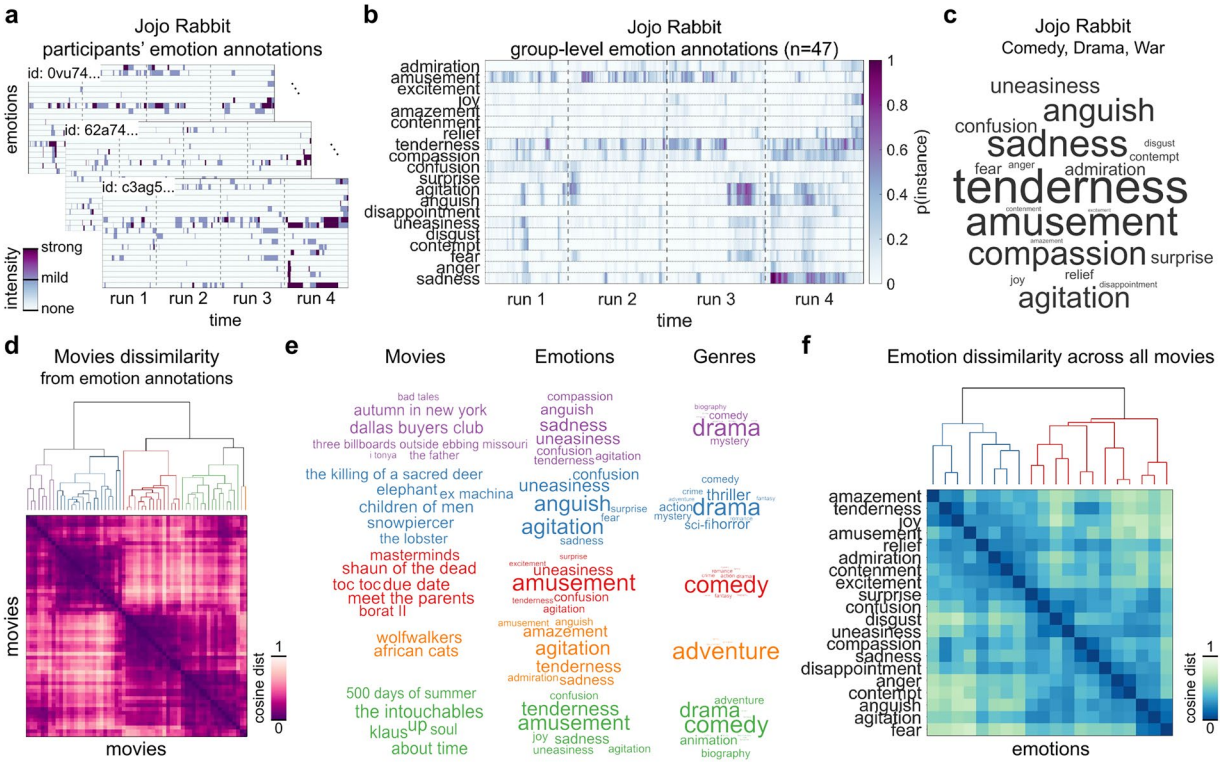

Fig. 3 Panel (**a**) shows Jojo Rabbit emotion annotations at the individual level, with each emotion-by-time matrix representing one participant's report. Color indicates emotion intensity, and dashed lines mark timepoints where the movie was split into shorter runs to reduce fatigue. Panel (**b**) presents group-level annotations, obtained by binarizing individual data (emotion present/absent) and summing across participants. The resulting matrix was scaled by its global maximum for comparability. Panel (**c**) displays the frequency of each reported emotion during the movie, derived from the group matrix and scaled by total occurrences. Panel (**d**) shows pairwise dissimilarity between all movies, based on cosine distances computed from movie-by-emotion matrices. Brighter colors reflect greater emotional divergence. Hierarchical clustering (average linkage) and the gap statistic identified the optimal number of clusters. Panel (**e**) summarizes key features of each cluster, including representative movies, dominant emotions, and common genres. Movie title size reflects distance from the cluster centroid; emotion and genre size indicate average frequency within the cluster. Panel (**f**) illustrates emotion co-occurrence across movies. Cosine dissimilarities between emotion categories were computed and clustered, revealing a valence-based structure: positive (blue) and negative (red) emotions formed distinct clusters, with finer sub-groupings such as (1) anguish–agitation–fear, (2) sadness–compassion, (3) anger–contempt, (4) confusion–disgust–uneasiness, and (5) joy–tenderness–amazement.

matrix was scaled by its global maximum to facilitate comparisons across films. Each cell in the matrix reflects the degree of consensus among participants in experiencing a specific emotion at a given timepoint. These matrices, thus, capture the dynamic evolution of emotional responses during movie viewing while preserving variability due to individual differences. Importantly, shared affective responses emerged during specific narrative moments, underscoring the collective emotional resonance of key scenes.

To summarize the overall emotional content of each film, we computed the average occurrence of each emotion across timepoints, producing a 20-element vector where each value reflects the relative prevalence of a given emotion. This summary is visualized using word clouds, as shown in Fig. 3c for Jojo Rabbit, in which tenderness, amusement, sadness, and compassion stand out. These results highlight the film's emotionally complex and bittersweet tone.

Applying the same procedure to all 60 movies yielded a movie-by-emotion matrix, from which we computed the cosine distance between films (Fig. 3d). We then applied hierarchical clustering with average linkage and determined the optimal number of clusters using MATLAB's evalclusters function and the gap statistic (globalMaxSE), considering solutions from one to fifteen clusters. The analysis revealed five distinct groupings of films based on their emotional profiles (Fig. 3e). The first (i.e., purple) includes primarily drama films such as Dallas Buyers Club, Autumn in New York, and Three Billboards Outside Ebbing Missouri, which are marked by high levels of sadness, uneasiness, and anguish. The second (i.e., blue) comprises a mix of dramas, thrillers, and horror films, including Children of Men, Snowpiercer, and The Killing of a Sacred Deer, and is characterized by anguish, agitation, and uneasiness. A third group (i.e., red) contains mostly comedies such as Due Date, Shaun of the Dead, and Meet the Parents, which primarily evoke amusement alongside a sense of uneasiness. A fourth group (i.e., green) includes emotionally layered films often labeled by IMDb as both drama and comedy, such as

Up, The Intouchables, and About Time. These films typically elicit tenderness, amusement, and sadness. Last, a fifth (i.e., orange) and smaller group comprises two adventure films—Wolfwalkers and African Cats—which are characterized by a combination of tenderness, agitation, amazement, and sadness.

To examine the relationships among the 20 emotion categories across the film set, we transposed the movie-by-emotion matrix used in the previous analysis. This resulted in an emotion-by-movie matrix reflecting the average occurrence of each emotion across films. Using pairwise cosine distances, we constructed an emotion-by-emotion dissimilarity matrix and applied hierarchical clustering following the same approach. As shown in Fig. 3f, the resulting structure reveals a broad organization of emotional states based on valence, with positive emotions forming one cluster (i.e., blue) and negative emotions forming another (i.e., red). Within this overarching division, more nuanced groupings emerged. For instance, anguish, agitation, and fear clustered together, as did sadness and compassion; anger and contempt formed a separate grouping, while confusion, disgust, and uneasiness aligned closely. On the positive side, joy, tenderness, and amazement were found to co-occur. These results provide insight into the affective structure that underlies emotional responses to naturalistic movie stimuli.

Inter-subject correlation and encoding analysis. To evaluate the reliability of fMRI responses elicited by Jojo Rabbit, we conducted inter-subject correlation[18] (ISC) and encoding[32] analyses.

Regarding ISC, we extracted the preprocessed time series of brain activity from each voxel and calculated the average Pearson's correlation coefficient (r) across all possible pairs of participants. The average ISC (Fig. 4a) highlights brain areas showing high synchronization (e.g., superior temporal cortex, occipital lobe) and reveals effect sizes comparable to other datasets of the highest quality[25].

For the encoding analysis, we evaluated the ability of affective reports and low-level acoustic and visual features to explain the brain activity elicited by Jojo Rabbit. Specifically, for the subjective reports of felt emotions, the predictor matrix was derived from emotion annotations aggregated across 47 participants who watched Jojo Rabbit outside the scanner (as shown in Fig. 3b). This approach resulted in an emotion-by-timepoint (20-by-3,087) matrix which was downsampled to the fMRI resolution of 2 s per point and convolved using a double-gamma function in SPM[54], with default parameters, to account for the standard hemodynamic response. In this matrix each value reflects the probability (i.e., the normalized number of participants) of a given emotional instance occurring at a particular time, along with its associated conceptual representation (i.e., emotion label). This approach yields a probabilistic model of shared emotional experiences during the movie—a model that we believe offers unique insights into how the brain interprets and categorizes affective states. It is therefore important to note that the emotion encoding analysis is centered on emotion conceptualization, rather than intensity or salience as it happens instead in typical emotion elicitation paradigms. Prior research using movie excerpts has already demonstrated the effectiveness of film in eliciting strong emotional responses[55] and in mapping the intensity of those responses onto brain connectivity–particularly of the amygdala[56]. However, we believe that one of the added values of full-length movies lies in their ability to engage processes related to emotion conceptualization and meaning-making—constructs that are less studied with traditional elicitation paradigms and for which the ventromedial prefrontal cortex plays a key role[57]. For each voxel, the emotion matrix served as a set of predictors in a mass-univariate encoding analysis aimed at explaining brain activity[32], averaged across the 20 participants who watched Jojo Rabbit inside the MRI scanner.

The results (coefficient of determination, R^2) indicated an effect size of 9% explained variance in the ventromedial prefrontal cortex ($R^2 = 0.093$; $x = +14$, $y = +44$, $z = -14$) based on the affective reports (Fig. 4b), aligning with previous findings obtained from similar models[33]. Among the regions where the affective model explained a greater proportion of variance there are also the amygdala ($R^2 = 0.050$; $x = +22$, $y = -2$, $z = -14$), and the subgenual anterior cingulate cortex ($R^2 = 0.071$; $x = +9$, $y = +37$, $z = -9$)–all regions consistently implicated in emotion processing. To assess the psychological relevance of this pattern, we conducted a meta-analytic, correlation-based decoding analysis[58], the results of which are presented in Fig. 4c. Specifically, we correlated the unthresholded affect encoding map with each of 50 unthresholded topic maps derived from a Latent Dirichlet Allocation (LDA) analysis of the Neurosynth database[59] through NiMARE[60] (v0.4.2 - 2025-03-11). Each topic was labeled by GPT-4o based on its most heavily weighted terms. Among the ten topics most strongly associated with the affect encoding map were "theory of mind and social cognition" (Topic 8; $r = 0.216$), "social interaction, empathy, and moral cognition" (Topic 28; $r = 0.174$), "face perception and emotion recognition" (Topic 40; $r = 0.116$), and "emotion processing and affective regulation" (Topic 26; $r = 0.115$).

For the encoding analysis of acoustic features, we extracted a single-column model based on sound energy, calculated by summing the power spectrum intensities of the acoustic stimulation across all frequency bands[46]. Similar to the previous approach, sound energy was convolved with a double-gamma function and used as a predictor of brain activity. As expected, the strongest association was measured in the bilateral temporal cortices (Fig. 4d; $R^2 = 0.145$; $x = -51$, $y = -15$, $z = +9$) and area 55b ($R^2 = 0.070$; $x = -52$, $y = 0$, $z = +56$). The corresponding decoding results (Fig. 4e) identified highly relevant topics such as "auditory and speech perception" (Topic 6; $r = 0.617$) and "language processing and reading" (Topic 37; $r = 0.340$).

For visual properties, we defined an encoding model using four low-level features (contrast, lightness, chroma, and hue), which were also convolved with a double-gamma function and used as predictors in a general linear model. The results (Fig. 4f) revealed the mapping of these features in the calcarine sulcus ($R^2 = 0.025$; $x = +16$, $y = -104$, $z = -2$), fusiform gyrus ($R^2 = 0.014$; $x = -48$, $y = -52$, $z = -10$), parahippocampal gyrus ($R^2 = 0.025$; $x = +26$, $y = -46$, $z = -10$), and right posterior superior temporal sulcus ($R^2 = 0.015$; $x = +56$, $y = -44$, $z = +4$). For the visual R^2 map, the results of the correlation-based decoding (Fig. 4g) reveal associations with meta-analytic maps of "semantic representations and object knowledge" (Topic 38; $r = 0.439$), "multisensory integration and sensory modalities" (Topic 42; $r = 0.402$), "face perception and emotion recognition" (Topic 40; $r = 0.330$), "mental imagery, scene construction, and navigation" (Topic 41; $r = 0.300$), and "motion perception and visual processing" (Topic 45; $r = 0.254$).

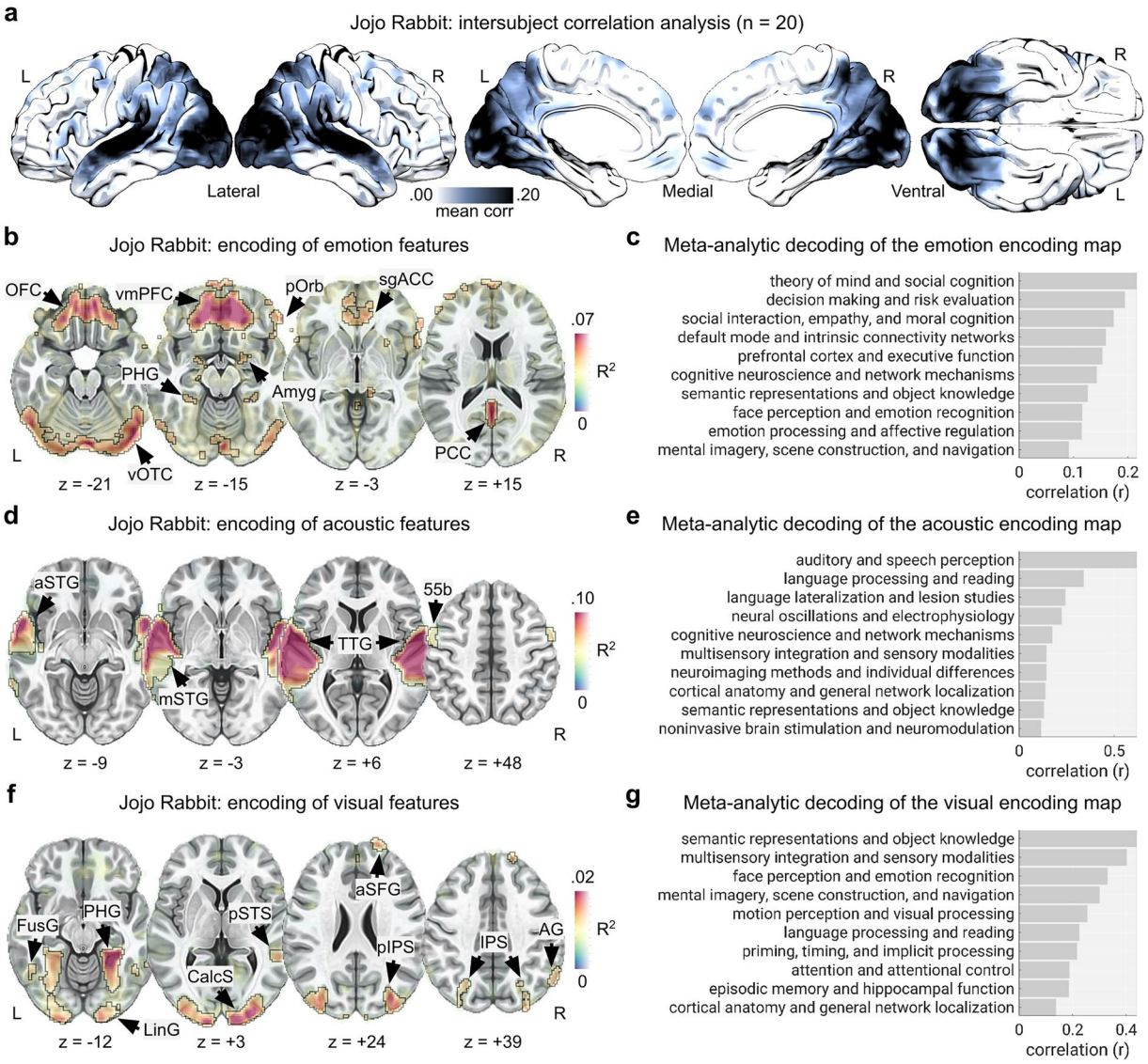

Fig. 4 Panel (**a**) shows the unthresholded intersubject correlation (ISC) map computed during the viewing of Jojo Rabbit, with strongest synchronization in auditory and visual areas, and additional effects in transmodal regions such as the precuneus, parietal lobules, and medial prefrontal cortex. Panel (**b**) presents results from the affective encoding analysis: group-level emotion ratings (n = 47) were used to predict fMRI activity in an independent sample (n = 20). Color intensity reflects R^2 values (thresholded at 0.04), with notable peaks in emotion-related regions such as the ventromedial prefrontal cortex ($R^2 = 0.093$), amygdala ($R^2 = 0.050$), and subgenual anterior cingulate cortex ($R^2 = 0.071$). Panel (**c**) shows results from meta-analytic decoding of this map, based on correlations with 50 topic maps from Neurosynth (LDA-based, GPT-4o-labeled). Top associated topics include "theory of mind and social cognition" (r = 0.216), "social interaction, empathy, and moral cognition" (r = 0.174), and "emotion processing and affective regulation" (r = 0.115). To validate sensory encoding, we repeated the analysis using acoustic and visual features. Panel (**d**) shows encoding of sound energy ($R^2 > 0.04$), with peaks in auditory regions such as the transverse temporal gyrus ($R^2 = 0.145$) and area 55b ($R^2 = 0.070$). Corresponding decoding results (panel e) highlight topics like "auditory and speech perception" (r = 0.617) and "language processing and reading" (r = 0.340). Panel (**f**) presents encoding of visual features ($R^2 > 0.01$), with highest values in visual and multisensory areas including the calcarine sulcus ($R^2 = 0.025$), fusiform gyrus ($R^2 = 0.014$), parahippocampal gyrus ($R^2 = 0.025$), and posterior superior temporal sulcus ($R^2 = 0.015$). Meta-analytic decoding (panel g) confirms relevance to topics such as "semantic representations and object knowledge" (r = 0.439), "multisensory integration and sensory modalities" (r = 0.402), and "motion perception and visual processing" (r = 0.254).

Therefore, using the REELMO fMRI data, behavioral annotations, and stimulus descriptions, we were able to demonstrate that brain regions known to process emotional, acoustic, and visual features are engaged in their respective domains.

Code availability

The code for the REELMO dataset is publicly available at the https://github.com/giacomohandjaras/REELMO repository. The code includes the fMRI preprocessing scripts and the ISC and encoding analysis.

Received: 5 February 2025; Accepted: 8 May 2025;
Published online: 15 May 2025

References

1. Grall, C. & Finn, E. S. Leveraging the power of media to drive cognition: A media-informed approach to naturalistic neuroscience. *Social Cognitive and Affective Neuroscience* **17**(6), 598–608 (2022).
2. Bordwell D., Thompson K., Smith J. Film Art: An Introduction (Eleventh Edition). New York: McGraw-Hill Education (2016).
3. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention* **1**(39-58), 3 (1997).
4. Marchewka, A., Żurawski, Ł., Jednoróg, K. & Grabowska, A. The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior research methods* **46**, 596–610 (2014).
5. Westermann, R., Spies, K., Stahl, G. & Hesse, F. W. Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of social psychology* **26**(4), 557–580 (1996).
6. Boğa, M., Koyuncu, M., Kaça, G. & Bayazıt, T. O. Comparison of emotion elicitation methods: 3 methods, 3 emotions, 3 measures. *Current Psychology* **42**(22), 18670–18685 (2023).
7. Philippot, P. Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and emotion* **7**(2), 171–193 (1993).
8. Gross, J. J. & Levenson, R. W. Emotion elicitation using films. *Cognition & emotion* **9**(1), 87–108 (1995).
9. Schaefer, A., Nils, F., Sanchez, X. & Philippot, P. Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and emotion* **24**(7), 1153–1172 (2010).
10. Maffei, A. & Angrilli, A. E-MOVIE-Experimental MOVies for Induction of Emotions in neuroscience: An innovative film database with normative data and sex differences. *Plos one* **14**(10), e0223124 (2019).
11. Cowen, A. S. & Keltner, D. Semantic space theory: A computational approach to emotion. *Trends in Cognitive Sciences* **25**(2), 124–136 (2021).
12. Miyamoto, Y., Uchida, Y. & Ellsworth, P. C. Culture and mixed emotions: co-occurrence of positive and negative emotions in Japan and the United States. *Emotion* **10**(3), 404 (2010).
13. Kuppens, P. & Verduyn, P. Emotion dynamics. *Current Opinion in Psychology* **17**, 22–26 (2017).
14. Thornton, M. A. & Tamir, D. I. Mental models accurately predict emotion transitions. *Proceedings of the National Academy of Sciences* **114**(23), 5982–5987 (2017).
15. Greenaway, K. H., Kalokerinos, E. K. & Williams, L. A. Context is everything (in emotion research). *Social and Personality Psychology Compass* **12**(6), e12393 (2018).
16. Hogan, P. C. What literature teaches us about emotion. Cambridge University Press (2011).
17. Saarimäki, H. Naturalistic stimuli in affective neuroimaging: A review. *Frontiers in human neuroscience* **15**, 675068 (2021).
18. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *science* **303**(5664), 1634–1640 (2004).
19. Baldassano, C. *et al.* Discovering event structure in continuous narrative perception and memory. *Neuron* **95**(3), 709–721 (2017).
20. Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E. & Gallant, J. L. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **532**(7600), 453–458 (2016).
21. Popham, S. F. *et al.* Visual and linguistic semantic representations are aligned at the border of human visual cortex. *Nature neuroscience* **24**(11), 1628–1636 (2021).
22. Setti, F. *et al.* A modality-independent proto-organization of human multisensory areas. *Nature Human Behaviour* **7**(3), 397–410 (2023).
23. Hanke, M. *et al.* A studyforrest extension, simultaneous fMRI and eye gaze recordings during prolonged natural stimulation. *Scientific data* **3**(1), 1–15 (2016).
24. Nastase, S. A., Liu, Y. F., Hillman, H., Norman, K. A. & Hasson, U. Leveraging shared connectivity to aggregate heterogeneous datasets into a common response space. *NeuroImage* **217**, 116865 (2020).
25. Visconti di Oleggio Castello, M., Chauhan, V., Jiahui, G. & Gobbini, M. I. An fMRI dataset in response to "The Grand Budapest Hotel", a socially-rich, naturalistic movie. *Scientific Data* **7**(1), 383 (2020).
26. Aliko, S., Huang, J., Gheorghiu, F., Meliss, S. & Skipper, J. I. A naturalistic neuroimaging database for understanding the brain using ecological stimuli. *Scientific Data* **7**(1), 347 (2020).
27. Goldstein, A. *et al.* Shared computational principles for language processing in humans and deep language models. *Nature neuroscience* **25**(3), 369–380 (2022).
28. Labs, A. *et al.* Portrayed emotions in the movie "Forrest Gump". *F1000Research* **4**, 92 (2015).
29. Camacho, M. C. *et al.* Large-scale encoding of emotion concepts becomes increasingly similar between individuals from childhood to adolescence. *Nature neuroscience* **26**(7), 1256–1266 (2023).
30. Chang, L. J. *et al.* Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. *Science Advances* **7**(17), eabf7129 (2021).
31. Camacho, M. C. *et al.* EmoCodes: A standardized coding system for socio-emotional content in complex video stimuli. *Affective Science* **3**(1), 168–181 (2022).
32. Lettieri, G. *et al.* Emotionotopy in the human right temporo-parietal cortex. *Nature communications* **10**(1), 5568 (2019).
33. Lettieri, G. *et al.* Dissecting abstract, modality-specific and experience-dependent coding of affect in the human brain. *Science Advances* **10**(10), eadk6840 (2024).
34. Lettieri, G. *et al.* Default and control network connectivity dynamics track the stream of affect at multiple timescales. *Social cognitive and affective neuroscience* **17**(5), 461–469 (2022).
35. Vaccaro, A. G. *et al.* Neural patterns associated with mixed valence feelings differ in consistency and predictability throughout the brain. *Cerebral Cortex* **34**(4), bhae122 (2024).
36. Finn, E. S., Corlett, P. R., Chen, G., Bandettini, P. A. & Constable, R. T. (2018). Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. *Nature communications*, **9**(1), 2043.
37. Albiero, P., Ingoglia, S. & Lo Coco, A. Contributo all'adattamento italiano dell'Interpersonal Reactivity Index. *Testing Psicometria Metodologia* **13**(2), 107–125 (2006).
38. Jordan, M. R., Amir, D. & Bloom, P. Are empathy and concern psychologically distinct? *Emotion* **16**(8), 1107 (2016).
39. Vachon, D. D. & Lynam, D. R. Fixing the problem with empathy: Development and validation of the affective and cognitive measure of empathy. *Assessment* **23**(2), 135–149 (2016).
40. Soto, C. J. & John, O. P. The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of personality and social psychology* **113**(1), 117 (2017).

41. Terraciano, A., McCrae, R. R. & Costa, P. T. Jr Factorial and construct validity of the Italian Positive and Negative Affect Schedule (PANAS). *European journal of psychological assessment* **19**(2), 131 (2003).
42. Kleiner, M., Brainard, D. & Pelli, D. What's new in Psychtoolbox-3? (2007).
43. Simonyan, K. & Zisserman, A. Two-stream convolutional networks for action recognition in videos. *Advances in neural information processing systems*, **27** (2014).
44. Safdar, M., Cui, G., Kim, Y. J. & Luo, M. R. Perceptually uniform color space for image signals including high dynamic range and wide gamut. *Optics express* **25**(13), 15131–15151 (2017).
45. Hershey, S. *et al*. CNN architectures for large-scale audio classification. *In 2017 ieee international conference on acoustics, speech and signal processing (icassp)* (pp. 131–135). IEEE (2017, March).
46. de Heer, W. A., Huth, A. G., Griffiths, T. L., Gallant, J. L. & Theunissen, F. E. The hierarchical cortical organization of human speech processing. *Journal of Neuroscience* **37**(27), 6539–6557 (2017).
47. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research* **29**(3), 162–173 (1996).
48. Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. *Fsl. Neuroimage*, **62**(2), 782–790 (2012).
49. Avants, B. B., Tustison, N. & Song, G. Advanced normalization tools (ANTS). *Insight j* **2**(365), 1–35 (2009).
50. Sampaolo, E., Handjaras, G., Lettieri, G. & Cecchetti, L. Lights, Camera, Emotion: REELMO's 1060 Hours of Affective Reports to Explore Emotions in Naturalistic Contexts. *Figshare* https://doi.org/10.6084/m9.figshare.28255745 (2025).
51. Lettieri, G. *et al*. A network-based hierarchical taxonomy of affect from language. *Affective Science* **3**, 208–222 (2022).
52. Cordaro, D. T. *et al*. Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion* **18**(1), 75 (2018).
53. Cowen, A. S. & Keltner, D. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the national academy of sciences* **114**(38), E7900–E7909 (2017).
54. Friston, K. J. *et al*. Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping* **2**(4), 189–210 (1994).
55. Fernández-Aguilar, L., Navarro-Bravo, B., Ricarte, J., Ros, L. & Latorre, J. M. How effective are films in inducing positive and negative emotional states? A meta-analysis. *PloS one* **14**(11), e0225040 (2019).
56. Wang, L. *et al*. Arousal modulates the amygdala-insula reciprocal connectivity during naturalistic emotional movie watching. *Neuroimage* **279**, 120316 (2023).
57. Roy, M., Shohamy, D. & Wager, T. D. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends in cognitive sciences* **16**(3), 147–156 (2012).
58. Poldrack, R. A. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron* **72**(5), 692–697 (2011).
59. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nature methods* **8**(8), 665–670 (2011).
60. Salo, T. *et al*. NiMARE: Neuroimaging Meta-Analysis Research Environment. *Aperture Neuro* **3**, 1–32 (2022).

Acknowledgements

The authors would like to thank Ms. Agnese Scarpellini and Ms. Chiara Fabbro for their help in the acquisition of the data. This work is part of the PRIN 2022 "Bittersweet, Ancient and Modern: Mixed Affect in Emotional Experience from Plato to Brain Imaging" project granted to Luca Cecchetti (protocol 2022CS3XR3_001) and funded by Unione europea - Next Generation EU, Missione 4 Componente 2 Inv. 1.1 CUP J53D23007980006.

Author contributions

E.S.: Investigation, Data curation, Visualization, Writing - Original Draft, Writing - Review & Editing. G.H.: Methodology, Software, Formal analysis, Data curation, Visualization, Writing - Original Draft, Writing - Review & Editing. G.L.: Conceptualization, Methodology, Data curation, Supervision, Writing - Original Draft, Writing - Review & Editing. L.C.: Conceptualization, Methodology, Visualization, Supervision, Project administration, Funding acquisition, Writing - Original Draft, Writing - Review & Editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.L.

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.