International Journal of Psychophysiology 226 (2026) 113414

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Registered Report Stage I

Registered Report Stage I: A Registered Report on transfer effects of emotion recognition training: Impacts on Visual attention and emotional reactivity in healthy adults

Nils R. Sommer [a], Nayla Sokhn [b], Roberto Caldara [b], Katja Schlegel [a,*]

[a] Institute of Psychology, University of Bern, Bern, Switzerland
[b] Department of Psychology, University of Fribourg, Fribourg, Switzerland

ARTICLE INFO

Keywords:
Emotion recognition training
Emotional intelligence
Visual attention
Emotional reactivity
Eye-tracking

ABSTRACT

Emotion recognition ability (ERA) – the capacity to accurately infer emotions from nonverbal facial, vocal, and bodily expressions – is a fundamental socio-emotional skill linked to improved social functioning, clinical counseling efficacy, workplace performance, and interpersonal relationships. While numerous ERA training programs for healthy adults have demonstrated improvements in ERA as measured through standardized tests, evidence of transfer effects to real-world socio-emotional processes remains limited. This study investigates whether a short multimodal ERA training (Training of Emotion Recognition Ability, TERA) impacts visual attention to emotional stimuli and emotional reactivity compared to a control attention and memory training in a randomized controlled trial (RCT). Participants (planned sample of $N = 230$) will be randomly assigned to either the TERA or control group. Visual attention will be assessed using eye-tracking while participants view videos of autobiographical emotional accounts. Emotional reactivity will be measured through self-reported emotional congruence (alignment of participants' emotions with those of video protagonists) and sympathy for protagonists. We expect that the TERA group will exhibit longer dwell time on protagonists' eye regions, greater emotional congruence and sympathy, and more accurate judgments of protagonists' displayed emotions compared to controls. Additionally, mediation analyses will investigate whether heightened attention to the eyes explains the training's effects on emotion recognition accuracy, emotional congruence, and sympathy. This study addresses critical gaps in the literature by examining the transfer effects of ERA training on attentional and emotional processes using ecologically valid stimuli. Findings could inform the development of targeted interventions to enhance socio-emotional functioning in diverse populations.

1. Introduction

Emotion recognition ability (ERA) – the ability to accurately infer emotions in others from nonverbal expressions in their face, voice, and body (Bänziger, 2016) – is a crucial skill for various social and emotional outcomes and is measured with standardized performance tests such as the Diagnostic Analysis of Nonverbal Accuracy (DANVA; Nowicki and Duke, 1994) or the Geneva Emotion Recognition Test (GERT; Schlegel et al., 2014). Higher ERA is associated with better social functioning (Hall et al., 2009a; Schmid Mast and Hall, 2018), greater effectiveness in clinical counseling (Abargil and Tishby, 2022; Hall et al., 2015; Hall et al., 2009b) and workplace success (Kranefeld and Blickle, 2021; Momm et al., 2015). Furthermore, higher ERA is linked to better negotiation outcomes (Elfenbein et al., 2007; Schlegel et al., 2018), teacher-student interactions and learning (Bernieri, 1991; Kurkul, 2007), as well as friendship quality and peer status in children (Wang et al., 2019). Given these positive links, researchers often advocate ERA training to boost social functioning and job performance in healthy individuals (Abargil and Tishby, 2022; Blanch-Hartigan et al., 2012; Ruben et al., 2020; Schmid Mast and Hall, 2018).

Over the past fifteen years, numerous ERA training programs for healthy adults have been developed and evaluated (Blanch-Hartigan, 2012; Chege et al., 2025; Döllinger et al., 2021, 2023; Hurley, 2012; Matsumoto and Hwang, 2011; Ruben et al., 2015; Schlegel et al., 2017). However, their effectiveness has primarily been assessed using standardized ERA tests, often with stimuli similar to those used in training.

---

* Corresponding author at: University of Bern, Institute of Psychology, Fabrikstrasse 8, 3012, Bern, Switzerland.
E-mail address: katja.schlegel@unibe.ch (K. Schlegel).

https://doi.org/10.1016/j.ijpsycho.2026.113414
Received 20 December 2025; Received in revised form 1 April 2026; Accepted 12 May 2026

For instance, Blanch-Hartigan (2012) found that students performed best on multimodal ERA tests when trained on analogous items. Similarly, Döllinger et al. (2021, 2023) observed that students and psychotherapists improved most on ERA tests using the same stimulus category as in training (e.g., multimodal stimuli vs. facial micro-expressions). While such improvements in ERA test performance in healthy adults are well documented (e.g., Döllinger et al., 2021; Schlegel et al., 2017), they may reflect test-specific learning rather than a genuine enhancement of the underlying socio-emotional ability.

To our knowledge, only two studies in healthy adults have examined transfer effects of ERA training on real-life ERA-related outcomes: Matsumoto and Hwang (2011) reported that training in recognizing micro-expressions improved third-party ratings of participants' social and communicative job skills. Schlegel (2021) found that multimodal ERA training led to more egalitarian economic outcomes and higher self- and other-ratings of collaboration and positive affect during negotiations. While these findings are promising, the mechanisms underlying these effects remain unclear, and empirical evidence of transfer remains limited. For instance, the positive outcomes reported by Matsumoto and Hwang (2011) and Schlegel (2021) may be explained through enhanced attention to others' emotional expressions, but these and other possible explanations have yet to be tested empirically.

This randomized controlled trial (RCT) will investigate the effects of ERA training on emotional attention and reactivity. Using eye-tracking, we will measure visual attention to emotional autobiographical videos (Wieck et al., 2022) and compare participants receiving a short multimodal ERA training with those undergoing a control attention and memory training. Additionally, we will assess and compare self-reported emotional reactivity – specifically, emotional congruence (the extent to which participants adopt the emotions expressed by video protagonists) and sympathy (the degree to which participants feel moved by and sympathetic toward the protagonists' accounts) – between the two groups. This study is the first to explore the transfer effects of ERA training on visual attention to emotional stimuli, as well as on emotional congruence and sympathy.

1.1. *ERA and emotional attention*

Theoretical frameworks suggest that ERA is closely linked to emotional attention. Fiori et al. (2022, 2023) argue that high emotional intelligence (EI), with ERA as a core component, involves both greater emotion knowledge (crystallized ability, measured by EI/ERA tests) and more effective emotion processing (fluid ability, assessed via tasks targeting selective attention and inhibition for emotional stimuli). Empirical support comes from behavioral experiments showing that individuals with high EI exhibit heightened attention to emotional stimuli, as demonstrated in emotional Stroop, Go/NoGo, and dot-probe tasks (Elfenbein et al., 2017; Fiori et al., 2022; Nicolet-dit-Félix et al., 2023).

Eye-tracking research indicates that longer fixations on others' eyes in emotional contexts are related to more accurate perception of emotions. The eye region is considered a “diagnostic area” for recognizing emotions like anger, sadness, and fear (Beaudry et al., 2014; Blais et al., 2017; Eisenbarth and Alpers, 2011; Schurgin et al., 2014; Smith et al., 2005). However, previous studies have not consistently linked individual differences in ERA test performance to specific gaze patterns in neurotypical populations (e.g., Yitzhak et al., 2021, 2022). For instance, Yitzhak et al. (2022) found no correspondence between gaze patterns and ERA using static and dynamic facial expressions. Nevertheless, Soker-Elimaliah et al. (2024) associated higher ERA with increased dwell time on the eyes of emotional facial expressions in a free viewing task. In another study, Vaidya et al. (2014) showed that fixation patterns did not predict accuracy for high-intensity posed expressions, but for subtle, daily-life-like expressions, eye fixations strongly predicted recognition accuracy across all emotions.

The study by Vaidya et al. (2014) suggests that inconsistent findings on gaze allocation to the eyes in ERA tasks may stem from the low ecological validity of commonly used stimuli. Most prior research has relied on static images (Beaudry et al., 2014; Eisenbarth and Alpers, 2011; Pollux et al., 2014; Smith et al., 2005), brief stimulus presentations (Blais et al., 2017), or a limited range of easily recognizable posed emotions (Beaudry et al., 2014; Yitzhak et al., 2022). As a result, such tasks often provide limited discrimination between individuals with higher versus lower underlying ERA due to ceiling effects (e.g., Yitzhak et al., 2022).

Another indicator for the role of eye-focused visual attention in emotion recognition comes from atypical gaze patterns in individuals with autism spectrum disorder (ASD). Compared to neurotypical controls, those with ASD often exhibit reduced attention to the eyes when viewing faces and emotional expressions (Riddiford et al., 2022). Concurrently, individuals with ASD – or neurotypical individuals with high ASD-like traits – tend to perform worse on ERA tests, though findings are mixed (Yeung, 2022). In healthy adults, further evidence suggests that the ERA advantage often found in women (Hall et al., 2025) might be due to more eye-focused gaze patterns (Hall et al., 2010).

Taken together, the above literature suggests that individuals with higher ERA may be more likely to focus on others' eyes when processing ecologically valid stimuli, such as longer, naturalistic videos of emotional expressions. This implies that ERA training could increase attention to the eyes. In trainings that use ecologically valid and multimodal stimuli such as the TERA (Schlegel et al., 2017), participants learn to pay attention to and accurately interpret the most important cues for various emotional expressions. Many of the learned cues appear around the eye region, such as the typical wrinkles around the eyes when expressing joy or amusement or the wide-open eyes in surprise or fear. The only previous ERA training study that investigated visual attention supports this: Pollux et al. (2014) found that both children and adults increased their fixation on the eyes after ERA training, with adults focusing more on the eyes in sadness stimuli and children doing so across all emotions. However, this study was limited by a small sample, lack of a randomized control group, and reliance on static facial expressions of only four emotions.

1.2. *ERA and emotional reactivity*

ERA has also been linked to stronger reactions to emotional stimuli. Multiple reviews discuss that ERA inherently integrates both cognitive decoding and emotional contagion processes (Elfenbein, 2014; Niedenthal and Brauer, 2012; Schurz et al., 2021; Zaki and Ochsner, 2011). Emotion simulation theories suggest that perceiving emotional expressions triggers unconscious mimicry (mimicking nonverbal behaviors of others) and emotional contagion/ experience sharing (taking over and experiencing emotions of others; e.g., Gallese and Goldman, 1998; for an overview, see Niedenthal and Brauer, 2012). In the EI field, it is hypothesized that EI, with ERA as a core component, is linked to a “hypersensitivity” to emotions (Davis and Nichols, 2016; Fiori et al., 2023; Fiori and Ortony, 2021; Schlegel, 2020). These authors argue that high ERA may increase awareness and attunement to emotions in one's surroundings, which would amplify emotion contagion and one's own emotional reactions. This emotional reactivity may have both positive and potentially negative intrapersonal consequences for individuals with higher ERA.

Research shows that ERA is linked to both subjectively experienced and physiological reactivity to positive and negative emotions. Individuals with higher ERA rate happy stimuli as more positive (Fiori et al., 2024), exhibit greater facial mimicry, perceive emotional stimuli as more intense, and report higher empathic concern (Drimalla et al., 2019). However, higher ERA is also linked to increased stress susceptibility and stronger physiological stress reactions (Bechtoldt and Schneider, 2016), more negative attributions of adverse events (Scherer, 2020), higher intensity of participants' own negative emotions and empathic concern for a video protagonist's negative emotions (Katzorreck and Kunzmann, 2018), greater avoidance coping (Sommer et al., 2025), heightened arousal in response to anger (Fiori et al., 2024), and stronger reactions to anger induction (Fiori and Ortony, 2016) in high ERA individuals. Thus, higher ERA may amplify emotional reactivity to both positive and negative emotions in one's surroundings.

Accordingly, it can be assumed that ERA training may also heighten subjective emotional reactivity. Increased attention to emotions and the relevant cues, as described earlier, may intensify emotional responses, including emotional congruence (i.e., “the vicarious experience of another person's emotional state”; Wieck et al., 2022, p. 76) and sympathy (feeling moved by another person). However, to our knowledge, the effects of ERA training on emotional reactivity remain unexplored. We note that, because previous evidence linking ERA to emotional reactivity is largely correlational, increased emotional reactivity may occur without a corresponding increase in decoding accuracy, and vice versa. This is consistent with theoretical models that distinguish mental state attribution, which is more closely related to accuracy, from experience sharing, which is more closely related to reactivity, as partially dissociable processes (see Zaki and Ochsner, 2011). At the same time, theory suggests that the increased perceptual readiness for, and heightened attention to, emotional stimuli fostered by ERA training may promote both more accurate decoding and stronger emotional reactivity (e.g., Fiori et al., 2023).

1.3. *The present study*

This randomized controlled trial (RCT) examines how emotion recognition ability (ERA) training influences visual attention and emotional reactivity to ecologically valid emotional expressions, with the aim of clarifying attentional processes involved in everyday emotion recognition. Participants will be randomly assigned to either a multimodal ERA training, the TERA (Schlegel et al., 2017), which teaches the recognition of 14 emotions using multimodal stimuli (i.e., videos with sound), or to a control attention and memory training that uses the same stimuli but focuses on the target actors' appearance. This control condition was chosen to rule out possible effects on fixation patterns that might arise merely from repeated exposure to videos of emotional expressions during training.

After training, participants will watch videos of autobiographical accounts of positive and negative emotional experiences (Wieck et al., 2022). After each video, they will rate both the perceived emotions of the protagonist and their own felt emotions using the 21 emotion items from Wieck et al. (2022), and they will additionally complete three sympathy items. Visual attention will be measured using eye tracking with an EyeLink Portable Duo (SR Research Ltd., 2024) and analyzed both in terms of dwell time on regions of interest (ROIs) and with a more data-driven fixation-mapping approach using iMap4 software (Lao et al., 2017). Emotional reactivity will be assessed following the standard procedure of Wieck et al. (2022), namely through emotional congruence, defined as the intraclass correlation between participants' own emotion ratings and the protagonists' reported emotions, and sympathy, assessed with the additional sympathy items. Emotion recognition accuracy in the autobiographical videos task will likewise be indexed by an intraclass correlation between the protagonists' reported emotions and participants' judgments of the protagonists' emotions.

Because the present study aims to examine transfer to a more naturalistic task and to assess multiple consequences of ERA training, including gaze behavior and emotional reactivity, we deliberately did not choose an outcome measure that fully mirrors the TERA in stimulus type and emotion labels. Instead, we expect that the enhanced perceptual readiness for, and knowledge about, diagnostic nonverbal cues fostered by the TERA will provide complementary information that contributes measurably to participants' understanding of the autobiographical videos, and thereby to their emotion recognition accuracy, affective reactivity, and gaze behavior, despite the videos' rich verbal content. This is consistent with the notion that in naturalistic settings, the two aspects of communication (verbal and nonverbal) are strongly interlocked (e.g., Hall et al., 2009a, 2009b; Jones & LeBaron, 2002). The two tasks were therefore selected to overlap substantially in their main primary emotion categories, without matching completely in all trained and rated emotions. This partial overlap allows us to test whether training effects generalize beyond highly similar laboratory tasks.

It is expected that participants in the ERA-training group show higher emotion recognition accuracy in the autobiographical videos task (H1) and higher dwell time on the eye region of the protagonists of the videos, measured with ROI (H2), compared to the control group. The ROI results will be complemented using fixation mapping analysis. Although the TERA trains multiple expressive cues across modalities rather than explicitly directing attention to the eyes, the eye region appears to be especially diagnostic for emotion recognition (e.g., Blais et al., 2017; Schurgin et al., 2014; Vaidya et al., 2014), and adults viewing spoken faces continue to devote substantial gaze to the eyes (Võ et al., 2012; Hessels, 2020). We therefore expect increased eye looking as one plausible and measurable transfer mechanism, while acknowledging that transfer may also occur through improved use of vocal and contextual cues.

Further, it is expected that participants in the ERA-training group show higher emotional congruence (H3) with and sympathy (H4) toward the protagonists of the autobiographical videos, compared to the control group. Furthermore, in a secondary set of hypotheses, mediation effects of higher dwell time on the eyes measured with ROI are expected, potentially explaining the effect of the ERA training on emotion recognition accuracy (H5), emotional congruence (H6) and sympathy (H7).

2. Methods

2.1. *Procedure*

At the beginning of the session, participants will be briefed, provide informed consent, and complete a demographics questionnaire on Qualtrics (https://www.qualtrics.com), followed by baseline state affect and ERA measures (for explorative purposes). Participants will be randomly assigned to either ERA training or control attention and memory training, stratified by gender (male, female, other) using Qualtrics randomization, with the experimenter blind to condition assignment. After completing the assigned 40-min training, participants will complete a post-training state affect measure. Following eye-tracker introduction, calibration, and validation, participants will watch eight autobiographical videos (Wieck et al., 2022) in random order using PsychoPy (v2025.1.1; Peirce et al., 2019), while gaze behavior is recorded. After each video, participants will complete brief questionnaires measuring perceived and felt emotions. A two-minute break occurs after the first four videos, with repeated eye-tracking calibration and validation. Drift checks precede each video, with recalibration if needed. After the last video and questionnaire, participants will complete a brief post-training ERA test (outcome-neutral manipulation check), then receive a debriefing. Additional questionnaires on self-reported response confidence for the baseline ERA test and the autobiographical videos task will also be administered but will not be used in the present analyses. The session duration is approximately 1 h and 45 min. An overview of the timeline and procedure is shown in Fig. 1.

The study was approved by the ethics committee of the Faculty of Human Sciences at the University of Bern (reference number 2025-07-06) in accordance with the Declaration of Helsinki. Data collection is expected to take approximately 6 to 9 months following Stage 1 approval, and the completed manuscript will be resubmitted within 12 months.

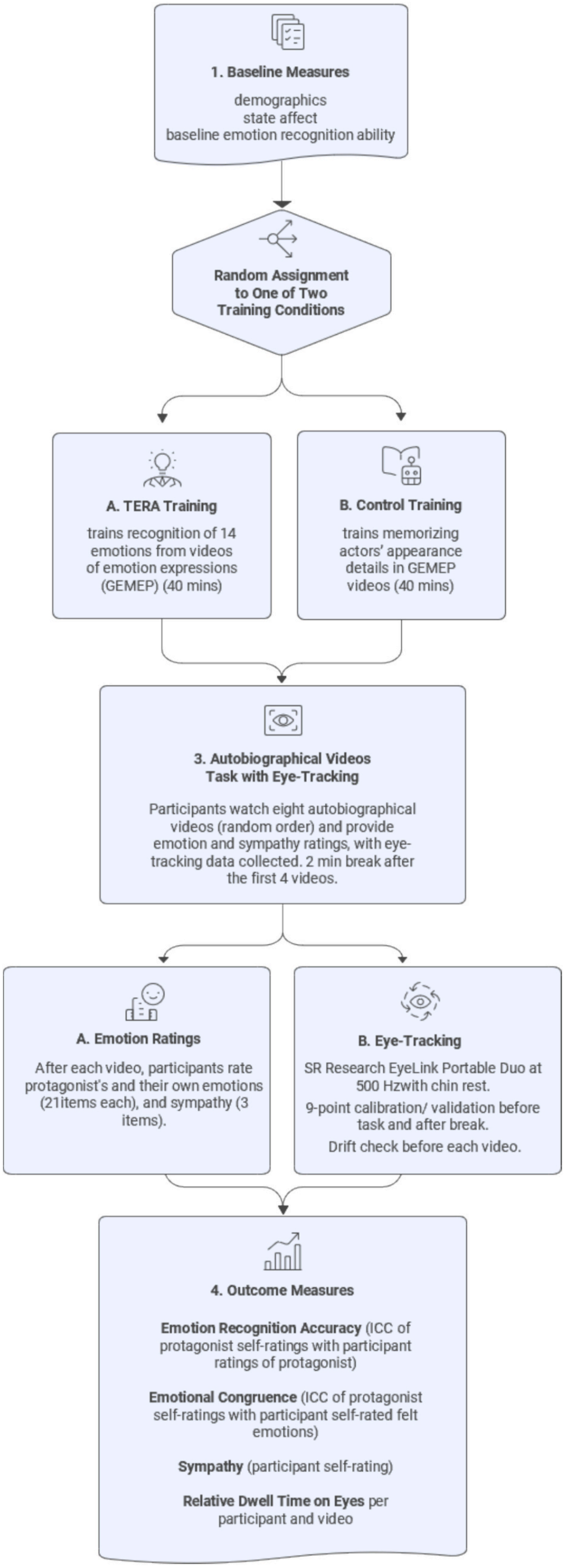

Fig. 1. Overview of the experimental procedure.

2.2. Sample

The sample will include psychology students recruited through the first author's department participant pool (receiving course credit) and young adults recruited by masters students writing their thesis in the project, with a planned minimum of 210 participants (for a-priori power analyses see Section 2.5.5). Inclusion criteria are: age 18–35 years, German native language, and contact lens use (not glasses) for vision correction if myopic or hyperopic. In a first step, 240 participants will be recruited. If after exclusions (either due to tracking loss or due to eye-tracking calibration failure, see Section 2.5.2) less than 210 participants remain in the data, another 30 participants are recruited. No analysis except the calculation of tracking loss will be conducted before this decision is made; exclusions due to failed calibration will be continually recorded in the experiment protocol.

2.3. Materials

2.3.1. Emotion recognition training

As the ERA training, we selected the self-administered 40-minute Training of Emotion Recognition Ability (TERA; Schlegel et al., 2017), as it encompasses a wide range of positive and negative emotions and integrates multiple sensory modalities (facial, vocal, and bodily expressions). The TERA has demonstrated improvements in ERA in both healthy adults and clinical populations, including patients with depression (Preis et al., 2025) and anorexia nervosa (Preis et al., 2020), and increased egalitarian economic outcomes and cooperative ratings in negotiations (Schlegel, 2021). It consists of two parts: an instruction part and a practice part.

In the instruction part, participants view an AI-generated instructor video describing the training procedure. For the original TERA (Schlegel et al., 2017) a real human experimenter provided instructions, but here an AI-avatar created using HeyGen (https://www.heygen.com/) ensures comparability with the control training. For each of 14 emotions, participants view the AI-instructor explaining the emotion's meaning (e.g., “People typically experience pride after a personal success or accomplishment…”) and describing typical facial, postural, gestural, and vocal cues (e.g., “You may recognize pride through the head tilted back; expansive posture…”), followed by a written cue list. Participants then observe these cues in two example videos from the Geneva Multimodal Emotion Portrayals database (GEMEP; Bänziger et al., 2012). The GEMEP items are short videos (1–3 s) of 10 different actors (varying gender and age) displaying emotions through facial, postural, gestural, and vocal cues (pseudo-language), corresponding to the respective instructions (e.g., pride). Participants can replay videos up to three times while rereading cues below each video. This procedure covers 14 emotions: six positive (amusement, interest, joy, pleasure, pride, relief), seven negative (anger, anxiety, despair, disgust, fear, irritation, sadness), and one neutral (surprise).

In the practice part, participants watch 42 additional GEMEP clips and identify which of the 14 emotions was displayed after each clip. Feedback (correct/incorrect) is provided after each choice, with one opportunity to re-watch and correct. After a second incorrect answer, the correct emotion is displayed before proceeding. The training (including instruction and practice) takes approximately 40 min. Detailed TERA descriptions and materials, including example videos, are available on OSF (https://osf.io/x3pj9/overview?view_only=656b90cfe8be4a17baace019c7c795c9).

2.3.2. Control attention and memory training

The control condition consists of an attention and memory training using the same GEMEP video stimuli as the TERA, created specifically to rule out viewing behavior changes due to repeated exposure to emotional expression videos. Instead of identifying emotions, participants memorize actor appearance details and answer multiple-choice questions after each clip (adapted from the familiarity control condition described in Schlegel et al., 2017). Participants are told this training increases attention to detail and short-term memory skills, which appeared plausible during piloting.

Participants sequentially complete three modes of these video and question pair items: 1) a direct mode, where questions are asked right after the clip (total of 30 items in one block), 2) a "1-back" mode, where the questions are asked only after the next clip (i.e., the questions need to be answered about the previous clip, and not the clip right before the question; total of 40 items in 8 blocks) and 3) a "2-back" mode, where the respective questions are asked only after the next two clips (total of 30 items in 6 blocks). Each mode is explained in an instruction video (using the same AI avatar as in the TERA) and participants complete practice items to get used to the procedure. Feedback (correct vs. incorrect with the correct choice displayed) is provided after each question. The training (including instruction and practice) takes approximately 40 min.

2.3.3. Autobiographical videos task

During the eye-tracking section, participants will watch eight videos from Wieck et al. (2022) showing German-speaking protagonists recounting real positive or negative workplace experiences (e.g., a teacher witnessing severe bullying). Videos range from 1:30 to 2:00 min and were selected to balance protagonist gender (male/female) and story valence (positive/negative), elicit moderate emotional reactivity, and ensure high verbal comprehensibility based on Wieck et al.'s (2022) validation statistics.

After each video, participants complete the ratings described by Wieck et al. (2022): (1) the perceived intensity of the protagonist's emotions and (2) the intensity of their own emotions while watching the video, using 21 adjectives rated on a scale from 1 (not at all) to 5 (extremely). These adjectives represent four negative primary emotion categories (anger, disgust, fear, and sadness) and three positive primary emotion categories (happiness, pride, and relaxation), with three adjectives per category: angry, mad, furious; sad, downhearted, grieved; afraid, alarmed, worried; disgusted, nauseated, sickened; happy, glad, delighted; proud, productive, satisfied; and relaxed, calm, easygoing. In addition, participants rate three sympathy adjectives (sympathetic, moved, and compassionate) using the same response format.

Based on the protagonists' and participants' ratings, the following scores are calculated for each participant and video: emotional congruence (intraclass correlation coefficient [ICC] between protagonists' and participants' own emotion intensity ratings), sympathy (mean of participants' ratings on the three sympathy adjectives), and emotion recognition accuracy (ICC between protagonists' own emotion intensity ratings and participants' ratings of the protagonists' emotion intensity). Following Wieck et al. (2022), emotional congruence and emotion recognition accuracy are calculated using two-way random, consistency, single-rating ICCs. In the psych package in R, this corresponds to the Shrout and Fleiss ICC(3,1) coefficient and will be computed using the ICC function (Revelle, 2026). For inferential analyses, ICC values will be Fisher z-transformed to approximate normality (Fisher, 1954). For descriptive reporting, back-transformed ICC values will be used to facilitate interpretation.

2.3.4. Baseline and post-training emotion recognition ability

The Perceiving AI Generated Emotions test (Weidmann and Xu, 2025) will be used to assess baseline ERA for exploratory purposes and to characterize the sample. In this brief test, participants are presented with 35 images of facial expressions, each corresponding to one of 20 emotion words shown in random order, and are asked to select the matching emotion from six response options. The target and distractor emotions span a broad range of positive and negative as well as basic and complex emotional states (for a complete list and all stimuli, see the materials on OSF: https://osf.io/x3pj9/overview?view_only=656b90cfe8be4a17baace019c7c795c9). Responses are scored dichotomously (0 = incorrect, 1 = correct), and a mean accuracy score across all items is computed for each participant.

Post-training ERA will be assessed with a brief 20-item version of the short Geneva Emotion Recognition Test (GERT-S; Schlegel and Scherer, 2016). The GERT-S is a performance-based emotion recognition test in which participants view short video clips of actors expressing 14 different emotions. These clips are drawn from the same GEMEP stimulus database (Bänziger et al., 2012) as the materials used in the TERA and the control attention and memory training. After each clip, participants select which of the 14 emotions, corresponding to the emotions trained in the TERA, was expressed by the actor. Responses are scored dichotomously (0 = incorrect, 1 = correct), and a mean accuracy score is computed across all items for each participant. The brief 20-item version was developed for and has been used in a large randomized controlled trial with young adults (ECoWeB; Watkins et al., 2024a, 2024b).

2.3.5. Affect

Participant's affect before and after training (baseline and post-training affect, respectively) is assessed with the 10-item short form of the Daniels five-factor measure of affective well-being (D-FAW-S; Russell and Daniels, 2018) for exploratory purposes. The D-FAW-S represents the two dimensions of valence and activation of affect as proposed by Russell (2003). The positive and negative activated affect scoring method (Russell and Daniels, 2018) combines adjectives "anxious," "at ease" (reversed), "annoyed," and "calm" (reversed) into negative affect, and "happy," "gloomy" (reversed), "active," "tired" (reversed), "motivated," and "bored" (reversed) into positive affect. Participants rate each adjective on a six-point Likert scale from 1 (not at all) to 6 (very much) indicating how they feel right now. The D-FAW-S demonstrates convergent validity, with Russell and Daniels (2018) finding high correlations between D-FAW-S subscales and the Positive and Negative Affect Scale (PANAS; Watson et al., 1988).

2.4. Apparatus

An SR-Research Eyelink Portable Duo (SR Research Ltd., 2024) collects eye-tracking data. Participants will sit at a table with a height-adjustable chair and with head movements restricted by a chin rest. The distance from the monitor to the eyes is 950 mm, the distance from the Eyelink camera to the eyes is 550 mm. The autobiographical videos are displayed on a 520 × 325 mm, 1920 × 1200 pixel monitor (16:10 aspect ratio), which corresponds to a 32.18° horizontal and 20.16° vertical visual angle of the full screen. The videos will be displayed at the center of the screen in a resolution of 1708 × 960 pixel(27.87° horizontal, 15.35° vertical visual angle) using PsychoPy (v2025.1.1; Peirce et al., 2019) with a gray background. Nine-point calibration and validation procedures will be used. Both eyes will be tracked if possible; if one eye shows low accuracy, only the high-accuracy eye is tracked (determined by calibration and validation). Drift checks will precede each video, and recalibration and validation will be performed if deviation from the drift check target is excessive. The eye-tracker samples at 500 Hz.

2.5. Data analysis

All data preparation and analyses will be performed in R version 4.4.2 (R Core Team, 2024), except for the preprocessing of eye-tracking data and the complementary fixation duration maps analysis, which will be performed in Matlab (version 9.13.0; The MathWorks Inc., 2022) and iMap4 (Lao et al., 2017).

2.5.1. Eye-tracking data preprocessing

In a preparatory step, all frames of the eight autobiographical videos were manually annotated with pixel coordinates (vertical and horizontal) of the protagonist's left eye, right eye, and mouth centers relative to the presentation screen. This step was necessary due to the dynamic format of the stimuli, and the coordinates of these points are later used for the creation of ROI and for the alignment procedure across videos. ROIs around left and right eye regions were created with 100-pixel radii around the annotated eye points for each frame (see Fig. 2). This preparatory step has already been completed to test preprocessing and analysis procedures using eye-tracking data from five pilot participants.

After data collection, sample reports will be exported from participants' eye-tracking recordings using Eyelink Data Viewer software (SR Research Ltd., 2025), including gaze locations (vertical and horizontal pixel coordinates relative to the presentation screen), sample allocation to specific frames (averaging 16–17 samples per video frame; 30 fps video, 500 Hz tracker), and automatic fixation detection (using Data Viewer's built-in algorithm). Next, fixations across all videos will be aligned to the first frame of one reference video using the three manually annotated reference points (left eye, right eye, mouth) and Matlab's built-in transformation functions *fitgeotrans* and *transformPointsForward* to ensure comparable gaze locations across all eight videos. *fitgeotrans* computes geometric transformations based on corresponding points across frames, while *transformPointsForward* applies this transformation to fixation coordinates, mapping them into the reference video's first frame coordinate system.

In a last step, fixation duration within the specified left and right eye ROIs for each participant and video is calculated and divided by total fixation duration during the respective video, representing relative dwell time on the eyes (ROI). For iMap4 analyses (Lao et al., 2017), the gaze positions aligned to the reference video will be used. For each fixation detected by Data Viewer, positions will be averaged between both eyes and durations are summed across all samples.

2.5.2. *Data quality checks*

Due to forced response settings in Qualtrics and Psychopy, no missing questionnaire or test data are expected. To ensure data quality and exclude careless responders, participants performing at or below 16.7% guessing probability on the baseline ERA test (PAGE; Weidmann and Xu, 2025) will be excluded from further analysis. Internal consistency of baseline ERA, baseline and post-training affect, emotion recognition accuracy, emotional congruence, and sympathy in the autobiographical videos task, as well as relative fixation duration to the eyes will be assessed using Cronbach's alpha, with values >0.70 considered acceptable. Lower values will be discussed as limitations.

Eye-tracking data quality is ensured by applying a nine-point calibration and validation procedure twice during the eye-tracking section of the laboratory session, once before the first video, and once after a two-minute break after the first 4 videos. The Eyelink-validation procedure is only accepted if the average deviation from the validation points (i.e., accuracy) is below 1°, and maximum error below 1.5° visual angle. Additionally, before each trial, an Eyelink drift-check procedure is performed: Participants fixate a central target, and the system only proceeds to the video if gaze falls within 2° of the target. In the case of a higher deviation, a new nine-point calibration and validation procedure are performed. If for a participant the calibration and validation procedure fails three times, and if that is not due to short-term inattention (which typically can be alleviated by re-instruction and re-calibration), participants are given a dummy-version of the autobiographical videos task, in which they follow the same procedure but their gaze is not tracked. These data will still be used for the analyses to test H1, H3, and H4, but not the ones for H2, H5, H6, and H7. Furthermore, trials with more that 50% tracking loss and participants with more than 50% average tracking loss across videos will be excluded from all analyses on eye-tracking data. Sensitivity analyses will be conducted excluding trials and participants with more than 20% tracking loss and will be reported in supplementary materials. These thresholds were chosen based on common practice (e.g., Povedano-Montero et al., 2025) and suggestions by Nyström et al. (2025).

Fig. 2. ROI of the left and right eyes.
Note. ROI for the left and right eyes are created with a radius of 100 pixels (in red) around the centers of the eyes (annotated in each frame of the video). This frame of one of the autobiographical videos is depicted with permission of the video protagonist and the original authors (Wieck et al., 2022). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5.3. *Descriptive statistics*

Basic descriptive analyses (mean values, standard deviations, skewness and kurtosis) and bivariate intercorrelations will be reported for all variables.

2.5.4. *Manipulation check*

As a manipulation check, performance in the post-training GERT-S is compared between the two conditions using a robust Welch's *t*-test. If the TERA group shows significantly higher scores than the control group, the manipulation will be considered successful.

2.5.5. *Hypothesis tests*

2.5.5.1. Primary hypotheses H1, H2, H3, and H4. To test whether participants in the TERA condition show higher emotion recognition accuracy (H1), higher dwell time on the eyes (ROI; H2), higher emotional congruence (H3), and higher sympathy (H4) compared to the control condition, simple linear mixed analyses will be conducted. Condition (TERA vs. control training) will be dummy-coded and included with control variables gender and age, as level-2 predictors. No level-1 predictors will be specified, and participant identity is added as a random intercept to account for variability within participant. These analyses will be conducted with the R-package *lme4* (Bates et al., 2015). Two-sided *t*-tests will assess whether the regression coefficient of condition differs significantly from zero using the package *lmerTest* (Kuznetsova et al., 2026). As measures of effect size, marginal R-squared will be computed for each model, as well as standardized regression coefficients and partial eta-squared for each predictor. Results will be reported with and without control variables. The *p*-values of the regression coefficients of condition will be corrected for multiple testing using the Hochberg method (Hochberg, 1988). Linearity, homoscedasticity, normality and independence of residuals, normality of random effects, and influential outliers will be assessed visually for all linear mixed models with the *check_model* function from the *performance* package in R (Lüdecke et al., 2026). Visually determined deviations from the assumptions will be discussed as limitations. To enhance transparency, sensitivity analyses will be conducted by estimating all models using robust linear mixed models from the R package *robustlmm* (Koller, 2026). These sensitivity analyses, as well as plots for visual assumption checks will be reported in the supplementary materials.

A-priori power analyses were calculated with the package *mlmpower* (Enders et al., 2025). The following effect sizes from previous individual studies were taken into account, as there are no meta-analytic effect sizes available: For H1, Schlegel et al. (2017) found effect sizes of $r =$ 0.22–0.59 (calculated from the originally stated d and η^2) for the effect of the TERA on multimodal and facial ERA test performance across multiple studies. For H2, Pollux et al. (2014) found an effect size of ERA training on dwell time on the eye region of r ~ 0.32 (calculated from the originally stated $\eta^2 = 0.1–0.13$). For H3 and H4, the effects of ERA-training on self- and other-rated social skills (r ~ 0.28; Matsumoto and Hwang, 2011) and competitiveness and positive emotions (r ~ 0.20–0.23; Schlegel, 2021) were relied upon. The lowest of these effect sizes, r ~ 0.20 (recalculated to R2 = 0.04; Schlegel, 2021), an assumed intra-class-correlation for repeated measures of ICC = 0.5 (see Arend and Schäfer, 2019) and eight measurements per participant (the eight videos in the autobiographical videos task) were entered. The simulation analysis with 1000 replications indicated a required sample size of at least 150 participants to reliably reach a power of 1-β = 0.90 with a significance level of α = 0.05.

2.5.5.2. *Complementary analysis for H2.*

As a complementary analysis to study whether participants in the TERA condition show higher dwell time on the eyes compared to the control condition, the procedure for comparative statistical fixation duration maps using iMap4 as described in Lao et al. (2017) will be applied. Fixation durations (coordinates aligned across all eight videos) are smoothed with a two-dimensional Gaussian kernel function at 1° of visual angle, using the *estimated* option in iMap4. Maps are then normalized by each video's sum duration. iMap4 conducts pixel-wise linear mixed models on smoothed and normalized fixation maps with fixation duration as the dependent variable, training condition as the independent variable, and participant identity as the random effects structure. Multiple comparisons correction uses a bootstrap spatial clustering method (*cluster size* option in iMap).

Fig. 3 shows an example fixation duration map based on eye-tracking data across all videos of five pilot participants (research assistants; four female, one male; age range 24–28), of which three completed the TERA and two the control attention and memory training. In this example, control training participants looked significantly longer at the nose area (indicated by black-edged area). This example uses no multiple comparisons correction to illustrate the test; with correction, no significant difference would be found with only five participants. Results from this analysis will be used to complement and visually interpret the results from H2, but not as a formal hypothesis test.

2.5.5.3. *Secondary hypotheses H5, H6, and H7.*

To test mediation effects of dwell time on the eye region in explaining the effects of training condition on emotion recognition accuracy (H5), emotional congruence (H6), and sympathy (H7), linear mixed mediation analyses will be conducted using the R-package *mediation* (Tingley et al., 2014). The function *mediate* in this package requires two models as its input: one model to predict the mediator (age, gender, and condition as level-2 predictors, a random intercept for participant identity, and dwell time on the eyes as the dependent variable) and one model to predict the outcome (age, gender, and condition as level-2 predictors, dwell time on the eyes as a level-1 predictor, a random intercept for participant identity, and either emotion recognition accuracy [H5], emotional congruence [H6], and sympathy [H7] as the dependent variable). The function then computes the indirect effect, direct effect, and total effect for the combined models using the Monte Carlo method with 1000 draws to determine the *p*-value and 95% confidence intervals. Results will again be reported with and without control variables. It is acknowledged that with the present study design, it will not be possible to interpret the mediation effect as strictly causal, as the mediator and dependent variables are both measured during the autobiographical videos task.

Current software does not support A-priori power analyses specific for the linear mixed mediations proposed here. The following approximations for the secondary hypotheses were conducted instead: First, Monte Carlo simulations with the Shiny app by Schoemann et al. (2017) were conducted using ordinary linear regression, with the same effect sizes as for H1-H4. Additionally, the effect size of the mediator predicting the dependent variables was drawn from Vaidya et al. (2014), who found $r = 0.24–0.40$ associations between dwell time on the eyes and ERA. This analysis suggested a minimum required sample size of 210 participants to reach a power of 1-β = 0.90 with a significance level of α = 0.05. We would expect this number to be significantly lower if multiple measurements could be integrated into the calculation (due to enhanced precision and thereby power; see, e.g. Vickers, 2003), which is not possible in the shiny app to date. Second, we ran power analyses (using the mlmpower package Keller, 2026) on the two separate linear mixed models that form the input for the *mediate* function in R. Similar to the power analyses for hypotheses H1-H4, using the lowest expected effect size of $r = 0.20$, the simulation analysis with 1000 replications indicated a required sample size of at least 150 participants to reliably reach a power of 1-β = 0.90 with a significance level of α = 0.05. Taken together, we decided to use the conservative approach and plan to recruit a sample of at least 210 participants.

CRediT authorship contribution statement

Nils R. Sommer: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Nayla Sokhn:** Writing – review & editing, Visualization, Methodology, Formal analysis. **Roberto Caldara:** Writing – review & editing, Supervision, Resources, Methodology. **Katja Schlegel:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Registration statement

After stage I in principle acceptance, the authors will register the approved protocol on the Open Science Framework (https://osf.io/).

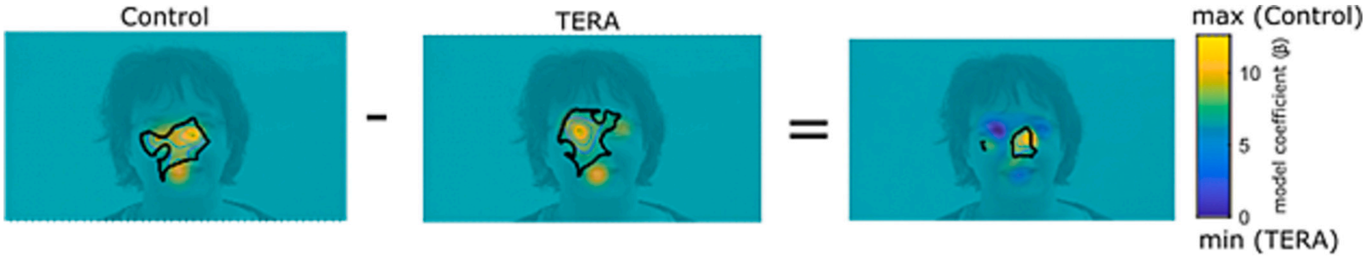

Fig. 3. Example comparative fixation map.

Note. Eye-tracking data of five pilot participants were used for this analysis. Blue areas indicate areas where participants in the TERA group fixated longer compared to participants in the control group, and yellow areas vice versa. The areas with black edges indicate areas of statistically significant differences between conditions (for illustrative purposes without correction of multiple comparisons). Data of all eight videos are collapsed and displayed on top of a single frame of the reference video. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Permission to reproduce materials

The authors have received permission to use the autobiographical video task created by Wieck et al. (2022) and to reproduce extracts such as example images.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors used NotebookLM (https://notebooklm.google/) to support an initial overview and synthesis of previous research and Le Chat (Mistral Medium 3.1, https://chat.mistral.ai/chat) to improve readability and concision of the text. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. The authors used HeyGen (https://www.heygen.com/) to create instruction videos for the ERA and control trainings. The contents of these videos were created by the authors or adapted from previous work, and HeyGen was solely used to make the instructors in the videos comparable between the training conditions.

Funding information

This research is supported by funding from the Swiss National Science Foundation awarded to Katja Schlegel (10000469).

Declaration of competing interest

The authors declare no conflicts of interest.

References

Abargil, M., Tishby, O., 2022. How therapists' emotion recognition relates to therapy process and outcome. Clin. Psychol. Psychother. 29 (3), 1001–1019. https://doi.org/10.1002/cpp.2680.

Arend, M.G., Schäfer, T., 2019. Statistical power in two-level models: a tutorial based on Monte Carlo simulation. Psychol. Methods 24 (1), 1–19. https://doi.org/10.1037/met0000195.

Bänziger, T., 2016. Accuracy of judging emotions. In: Hall, J.A., Schmid Mast, M., West, T.V. (Eds.), The Social Psychology of Perceiving Others Accurately. Cambridge University Press, pp. 23–51.

Bänziger, T., Mortillaro, M., Scherer, K.R., 2012. Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. Emotion 12 (5), 1161–1179. https://doi.org/10.1037/a0025827.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–48. https://doi.org/10.18637/jss.v067.i01.

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., Tapp, R., 2014. Featural processing in recognition of emotional facial expressions. Cognit. Emot. 28 (3), 416–432. https://doi.org/10.1080/02699931.2013.833500.

Bechtoldt, M.N., Schneider, V.K., 2016. Predicting stress from the ability to eavesdrop on feelings: emotional intelligence and testosterone jointly predict cortisol reactivity. Emotion 16 (6), 815–825. https://doi.org/10.1037/emo0000134.

Bernieri, F.J., 1991. Interpersonal sensitivity in teaching interactions. Personal. Soc. Psychol. Bull. 17 (1), 98–103. https://doi.org/10.1177/0146167291171015.

Blais, C., Fiset, D., Roy, C., Saumure Régimbald, C., Gosselin, F., 2017. Eye fixation patterns for categorizing static and dynamic facial expressions. Emotion 17 (7), 1107–1119. https://doi.org/10.1037/emo0000283.

Blanch-Hartigan, D., 2012. An effective training to increase accurate recognition of patient emotion cues. Patient Educ. Couns. 89 (2), 274–280. https://doi.org/10.1016/j.pec.2012.08.002.

Blanch-Hartigan, D., Andrzejewski, S.A., Hill, K.M., 2012. The effectiveness of training to improve person perception accuracy: a meta-analysis. Basic Appl. Soc. Psychol. 34 (6), 483–498. https://doi.org/10.1080/01973533.2012.728122.

Chege, M., Yang, F., Lovell, S., Al-Ramahi, N., Shiu, S., Hu, A., Stojanoski, B., Stevenson, R.A., 2025. Emotion recognition training across sensory modalities: comparing the effects of multisensory and unisensory training (65xzp_v1). PsyArXiv. https://doi.org/10.31234/osf.io/65xzp_v1.

Davis, S.K., Nichols, R., 2016. Does emotional intelligence have a "dark" side? A review of the literature. Front. Psychol. 7. https://doi.org/10.3389/fpsyg.2016.01316.

Döllinger, L., Laukka, P., Högman, L.B., Bänziger, T., Makower, I., Fischer, H., Hau, S., 2021. Training emotion recognition accuracy: results for multimodal expressions and facial micro expressions. Front. Psychol. 12. https://doi.org/10.3389/fpsyg.2021.708867.

Döllinger, L., Högman, L.B., Laukka, P., Bänziger, T., Makower, I., Fischer, H., Hau, S., 2023. Trainee psychotherapists' emotion recognition accuracy improves after training: emotion recognition training as a tool for psychotherapy education. Front. Psychol. 14. https://doi.org/10.3389/fpsyg.2023.1188634.

Drimalla, H., Landwehr, N., Hess, U., Dziobek, I., 2019. From face to face: the contribution of facial mimicry to cognitive and emotional empathy. Cognit. Emot. 33 (8), 1672–1686. https://doi.org/10.1080/02699931.2019.1596068.

Eisenbarth, H., Alpers, G.W., 2011. Happy mouth and sad eyes: scanning emotional facial expressions. Emotion 11 (4), 860–865. https://doi.org/10.1037/a0022758.

Elfenbein, H.A., 2014. The many faces of emotional contagion: an affective process theory of affective linkage. Organ. Psychol. Rev. 4 (4), 326–362. https://doi.org/10.1177/2041386614542889.

Elfenbein, H.A., Foo, M.D., White, J., Tan, H.H., Aik, V.C., 2007. Reading your counterpart: the benefit of emotion recognition accuracy for effectiveness in negotiation. J. Nonverbal Behav. 31 (4), 205–223. https://doi.org/10.1007/s10919-007-0033-7.

Elfenbein, H.A., Jang, D., Sharma, S., Sanchez-Burks, J., 2017. Validating emotional attention regulation as a component of emotional intelligence: a Stroop approach to individual differences in tuning in to and out of nonverbal cues. Emotion 17 (2), 348–358. https://doi.org/10.1037/emo0000145.

Enders, C.K., Keller, B.T., Woller, M.P., 2025. A simple Monte Carlo method for estimating power in multilevel designs. Psychol. Methods 30 (5), 980–996. https://doi.org/10.1037/met0000614.

Fiori, M., Ortony, A., 2016. Are emotionally intelligent individuals hypersensitive to emotions? Testing the curse of emotion. Acad. Manag. Proc. 2016 (1), 10023. https://doi.org/10.5465/ambpp.2016.10023abstract.

Fiori, M., Ortony, A., 2021. Initial evidence for the hypersensitivity hypothesis: emotional intelligence as a magnifier of emotional experience. J. Intell. 9 (2), 24. https://doi.org/10.3390/jintelligence9020024.

Fiori, M., Udayar, S., Vesely Maillefer, A., 2022. Emotion information processing as a new component of emotional intelligence: theoretical framework and empirical evidence. Eur. J. Personal. 36 (2), 245–264. https://doi.org/10.1177/08902070211007672.

Fiori, M., Vesely-Maillefer, A.K., Nicolet-Dit-Félix, M., Gillioz, C., 2023. With great sensitivity comes great management: how emotional hypersensitivity can be the superpower of emotional intelligence. J. Intell. 11 (10), 198.

Fiori, M., Gillioz, C., Nicolet-dit-Félix, M., 2024. Emotional intelligence and emotional reactivity: understanding the hypersensitivity hypothesis. Personal. Individ. Differ. 230, 112792. https://doi.org/10.1016/j.paid.2024.112792.

Fisher, R.A., 1954. Statistical Method for Research Workers, 12th ed. Oliver and Boyd. https://doi.org/10.1002/qj.49708235130.

Gallese, V., Goldman, A., 1998. Mirror neurons and the simulation theory of mind-reading. Trends Cogn. Sci. 2 (12), 493–501. https://doi.org/10.1016/S1364-6613(98)01262-5.

Hall, J.A., Andrzejewski, S.A., Yopchick, J.E., 2009a. Psychosocial correlates of interpersonal sensitivity: a meta-analysis. J. Nonverbal Behav. 33 (3), 149–180. https://doi.org/10.1007/s10919-009-0070-5.

Hall, J.A., Roter, D.L., Blanch, D.C., Frankel, R.M., 2009b. Nonverbal sensitivity in medical students: implications for clinical interactions. J. Gen. Intern. Med. 24 (11), 1217–1222. https://doi.org/10.1007/s11606-009-1107-5.

Hall, J.K., Hutton, S.B., Morgan, M.J., 2010. Sex differences in scanning faces: does attention to the eyes explain female superiority in facial expression recognition? Cognit. Emot. 24 (4), 629–637. https://doi.org/10.1080/02699930902906882.

Hall, J.A., Ship, A.N., Ruben, M.A., Curtin, E.M., Roter, D.L., Clever, S.L., Smith, C.C., Pounds, K., 2015. Clinically relevant correlates of accurate perception of patients' thoughts and feelings. Health Commun. 30 (5), 423–429. https://doi.org/10.1080/10410236.2013.865507.

Hall, J.A., Gunnery, S.D., Schlegel, K., 2025. Gender and accuracy in decoding affect cues: a meta-analysis. J. Intell. 13 (3), 38. https://doi.org/10.3390/jintelligence13030038.

Hessels, R.S., 2020. How does gaze to faces support face-to-face interaction? A review and perspective. Psychon. Bull. Rev. 27 (5), 856–881. https://doi.org/10.3758/s13423-020-01715-w.

Hochberg, Y., 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75 (4), 800–802. https://doi.org/10.1093/biomet/75.4.800.

Hurley, C.M., 2012. Do you see what I see? Learning to detect micro expressions of emotion. Motiv. Emot. 36 (3), 371–381. https://doi.org/10.1007/s11031-011-9257-2.

Jones, S.E., LeBaron, C.D., 2002. Research on the relationship between verbal and nonverbal communication: emerging integrations. J. Commun. 52 (3), 499–521. https://doi.org/10.1111/j.1460-2466.2002.tb02559.x.

Katzorreck, M., Kunzmann, U., 2018. Greater empathic accuracy and emotional reactivity in old age: the sample case of death and dying. Psychol. Aging 33 (8), 1202–1214. https://doi.org/10.1037/pag0000313.

Keller, B.T., 2026. mlmpower: power analysis and data simulation for multilevel models (version 1.0.11). https://cran.r-project.org/web/packages/mlmpower/index.html.

Koller, M., 2026. robustlmm: robust linear mixed effects models (version 3.4-2). https://cran.r-project.org/web/packages/robustlmm/index.html (computer software).

Kranefeld, I., Blickle, G., 2021. Emotion recognition ability and career success: assessing the roles of GMA and conscientiousness. Personal. Individ. Differ. 168, 110370. https://doi.org/10.1016/j.paid.2020.110370.

Kurkul, W.W., 2007. Nonverbal communication in one-to-one music performance instruction. Psychol. Music 35 (2), 327–362. https://doi.org/10.1177/0305735607070385.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., Jensen, S.P., 2026. lmerTest: tests in linear mixed effects models (version 3.2-1). https://cran.r-project.org/web/packages/lmerTest/index.html.

Lao, J., Miellet, S., Pernet, C., Sokhn, N., Caldara, R., 2017. iMap4: an open source toolbox for the statistical fixation mapping of eye movement data with linear mixed modeling. Behav. Res. Methods 49 (2), 559–575. https://doi.org/10.3758/s13428-016-0737-x.

Lüdecke, D., Makowski, D., Ben-Shachar, M.S., Patil, I., Waggoner, P., Wiernik, B.M., Thériault, R., Arel-Bundock, V., Jullum, M., Bacher, E., Luchman, J., 2026. performance: assessment of regression models performance (version 0.16.0). https://cran.r-project.org/web/packages/performance/index.html.

Matsumoto, D., Hwang, H.S., 2011. Evidence for training the ability to read microexpressions of emotion. Motiv. Emot. 35 (2), 181–191. https://doi.org/10.1007/s11031-011-9212-2.

Momm, T., Blickle, G., Liu, Y., Wihler, A., Kholin, M., Menges, J.I., 2015. It pays to have an eye for emotions: emotion recognition ability indirectly predicts annual income. J. Organ. Behav. 36 (1), 147–163. https://doi.org/10.1002/job.1975.

Nicolet-dit-Félix, M., Gillioz, C., Mortillaro, M., Sander, D., Fiori, M., 2023. Emotional intelligence and attentional bias to emotional faces: evidence of hypersensitivity towards emotion information. Personal. Individ. Differ. 201, 111917. https://doi.org/10.1016/j.paid.2022.111917.

Niedenthal, P.M., Brauer, M., 2012. Social functionality of human emotion. Annu. Rev. Psychol. 63, 259–285. https://doi.org/10.1146/annurev.psych.121208.131605.

Nowicki, S., Duke, M.P., 1994. Individual differences in the nonverbal communication of affect: the Diagnostic Analysis of Nonverbal Accuracy Scale. J. Nonverbal Behav. 18 (1), 9–35. https://doi.org/10.1007/BF02169077.

Nyström, M., Hooge, I.T.C., Hessels, R.S., Andersson, R., Hansen, D.W., Johansson, R., Niehorster, D.C., 2025. The fundamentals of eye tracking part 3: how to choose an eye tracker. Behav. Res. Methods 57 (2), 67. https://doi.org/10.3758/s13428-024-02587-x.

Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J.K., 2019. PsychoPy2: experiments in behavior made easy. Behav. Res. Methods 51 (1), 195–203. https://doi.org/10.3758/s13428-018-01193-y.

Pollux, P.M.J., Hall, S., Guo, K., 2014. Facial expression training optimises viewing strategy in children and adults. PLoS One 9 (8), e105418. https://doi.org/10.1371/journal.pone.0105418.

Povedano-Montero, F.J., Bernardez-Vilaboa, R., Trillo, J.R., González-Jiménez, R., Otero-Currás, C., Martínez-Florentín, G., Cedrún-Sánchez, J.E., 2025. Integrating eye-tracking and artificial intelligence for quantitative assessment of visuocognitive performance in sports and education. Photonics 12 (12). https://doi.org/10.3390/photonics12121167.

Preis, M.A., Schlegel, K., Stoll, L., Blomberg, M., Schmidt, H., Wünsch-Leiteritz, W., Leiteritz, A., Brockmeyer, T., 2020. Improving emotion recognition in anorexia nervosa: an experimental proof-of-concept study. Int. J. Eat. Disord. 53 (6), 945–953. https://doi.org/10.1002/eat.23276.

Preis, M.A., Schlegel, K., Rehbein, S., Lorenz, K., Brockmeyer, T., 2025. Training emotion recognition in depression—an experimental study. Br. J. Clin. Psychol. 64 (3), 817–828. https://doi.org/10.1111/bjc.12540.

R Core Team, 2024. R: a language and environment for statistical computing (version 4.4.2). R Foundation for Statistical Computing. https://www.r-project.org/ (computer software).

Revelle, W., 2026. psych: procedures for psychological, psychometric, and personality research (version 2.6.1). https://cran.r-project.org/web/packages/psych/index.html (computer software).

Riddiford, J.A., Enticott, P.G., Lavale, A., Gurvich, C., 2022. Gaze and social functioning associations in autism spectrum disorder: a systematic review and meta-analysis. Autism Res. 15 (8), 1380–1446. https://doi.org/10.1002/aur.2729.

Ruben, M.A., Hall, J.A., Curtin, E.M., Blanch-Hartigan, D., Ship, A.N., 2015. Discussion increases efficacy when training accurate perception of patients' affect. J. Appl. Soc. Psychol. 45 (6), 355–362. https://doi.org/10.1111/jasp.12301.

Ruben, M.A., Blanch-Hartigan, D., Hall, J.A., 2020. Communication skills to engage patients in treatment. In: Hadler, A., Sutton, S., Osterberg, L. (Eds.), The Wiley Handbook of Healthcare Treatment Engagement. John Wiley & Sons, Ltd, pp. 274–296. https://doi.org/10.1002/9781119129530.ch15.

Russell, J.A., 2003. Core affect and the psychological construction of emotion. Psychol. Rev. 110 (1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145.

Russell, E., Daniels, K., 2018. Measuring affective well-being at work using short-form scales: implications for affective structures and participant instructions. Hum. Relat. 71 (11), 1478–1507. https://doi.org/10.1177/0018726717751034.

Scherer, K.R., 2020. Evidence for the existence of emotion dispositions and the effects of appraisal bias. Emotion 21, 1224–1238. https://doi.org/10.1037/emo0000861.

Schlegel, K., 2020. Inter- and intrapersonal downsides of accurately perceiving others' emotions. In: Sternberg, R.J., Kostić, A. (Eds.), Social Intelligence and Nonverbal Communication. Springer International Publishing, pp. 359–395. https://doi.org/10.1007/978-3-030-34964-6_13.

Schlegel, K., 2021. The effects of emotion recognition training on interpersonal effectiveness. Basic Appl. Soc. Psychol. 43 (2), 141–153. https://doi.org/10.1080/01973533.2021.1883021.

Schlegel, K., Scherer, K.R., 2016. Introducing a short version of the Geneva Emotion Recognition Test (GERT-S): psychometric properties and construct validation. Behav. Res. Methods 48 (4), 1383–1392. https://doi.org/10.3758/s13428-015-0646-4.

Schlegel, K., Grandjean, D., Scherer, K.R., 2014. Introducing the Geneva Emotion Recognition Test: an example of Rasch-based test development. Psychol. Assess. 26 (2), 666–672. https://doi.org/10.1037/a0035246.

Schlegel, K., Vicaria, I.M., Isaacowitz, D.M., Hall, J.A., 2017. Effectiveness of a short audiovisual emotion recognition training program in adults. Motiv. Emot. 41 (5), 646–660. https://doi.org/10.1007/s11031-017-9631-9.

Schlegel, K., Mehu, M., van Peer, J.M., Scherer, K.R., 2018. Sense and sensibility: the role of cognitive and emotional intelligence in negotiation. J. Res. Pers. 74, 6–15. https://doi.org/10.1016/j.jrp.2017.12.003.

Schmid Mast, M., Hall, J.A., 2018. The impact of interpersonal accuracy on behavioral outcomes. Curr. Dir. Psychol. Sci. 27 (5), 309–314. https://doi.org/10.1177/0963721418758437.

Schoemann, A.M., Boulton, A.J., Short, S.D., 2017. Determining power and sample size for simple and complex mediation models. Soc. Psychol. Personal. Sci. 8 (4), 379–386. https://doi.org/10.1177/1948550617715068.

Schurgin, M.W., Nelson, J., Iida, S., Ohira, H., Chiao, J.Y., Franconeri, S.L., 2014. Eye movements during emotion recognition in faces. J. Vis. 14 (13), 14. https://doi.org/10.1167/14.13.14.

Schurz, M., Radua, J., Tholen, M.G., Maliske, L., Margulies, D.S., Mars, R.B., Sallet, J., Kanske, P., 2021. Toward a hierarchical model of social cognition: a neuroimaging meta-analysis and integrative review of empathy and theory of mind. Psychol. Bull. 147 (3), 293–327. https://doi.org/10.1037/bul0000303.

Smith, M.L., Cottrell, G.W., Gosselin, F., Schyns, P.G., 2005. Transmitting and decoding facial expressions. Psychol. Sci. 16 (3), 184–189. https://doi.org/10.1111/j.0956-7976.2005.00801.x.

Soker-Elimaliah, S., Rajpersaud, A., Bragerton-Nasert, S., Soker, D., Wagner, J.B., 2024. Visual attention and cardiac response to faces in relation to emotion recognition. J. Psychophysiol. 38 (3), 147–160. https://doi.org/10.1027/0269-8803/a000336.

Sommer, N.R., Carrard, V., Bourquin, C., Berney, A., Schlegel, K., 2025. Social support and avoidance mediate positive and negative effects of emotion recognition ability on mental health in medical students. Sci. Rep. 15 (1), 16910. https://doi.org/10.1038/s41598-025-02025-8.

SR Research Ltd., 2024. EyeLink Portable Duo User Manual (Version 1.0.10). SR Research Ltd.

SR Research Ltd., 2025. EyeLink Data Viewer 5.1.1. SR Research Ltd. (computer software).

The MathWorks Inc., 2022. MATLAB version: 9.13.0 (R2022b). The MathWorks Inc.. https://www.mathworks.com (computer software).

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., Imai, K., 2014. mediation: R package for causal mediation analysis. J. Stat. Softw. 59, 1–38. https://doi.org/10.18637/jss.v059.i05.

Võ, M.L.-H., Smith, T.J., Mital, P.K., Henderson, J.M., 2012. Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. J. Vis. 12 (13), 3. https://doi.org/10.1167/12.13.3.

Vaidya, A.R., Jin, C., Fellows, L.K., 2014. Eye spy: the predictive value of fixation patterns in detecting subtle and extreme emotions from faces. Cognition 133 (2), 443–456. https://doi.org/10.1016/j.cognition.2014.07.004.

Vickers, A.J., 2003. How many repeated measures in repeated measures designs? Statistical issues for comparative trials. BMC Med. Res. Methodol. 3 (1), 22. https://doi.org/10.1186/1471-2288-3-22.

Wang, Y., Hawk, S.T., Tang, Y., Schlegel, K., Zou, H., 2019. Characteristics of emotion recognition ability among primary school children: relationships with peer status and friendship quality. Child Indic. Res. 12 (4), 1369–1388. https://doi.org/10.1007/s12187-018-9590-z.

Watkins, E.R., Warren, F.C., Newbold, A., Hulme, C., Cranston, T., Aas, B., Bear, H., Botella, C., Burkhardt, F., Ehring, T., Fazel, M., Fontaine, J.R.J., Frost, M., Garcia-Palacios, A., Greimel, E., Hößle, C., Hovasapian, A., Huyghe, V.E.I., Karpouzis, K., Taylor, R.S., 2024a. Emotional competence self-help app versus cognitive behavioural self-help app versus self-monitoring app to prevent depression in young adults with elevated risk (ECoWeB PREVENT): an international, multicentre, parallel, open-label, randomised controlled trial. Lancet Dig. Health 6 (12), e894–e903. https://doi.org/10.1016/S2589-7500(24)00148-1.

Watkins, E.R., Warren, F.C., Newbold, A., Hulme, C., Cranston, T., Aas, B., Bear, H., Botella, C., Burkhardt, F., Ehring, T., Fazel, M., Fontaine, J.R.J., Frost, M., Garcia-Palacios, A., Greimel, E., Hößle, C., Hovasapian, A., Huyghe, V.E.I., Karpouzis, K., Taylor, R.S., 2024b. Emotional competence self-help mobile phone app versus cognitive behavioural self-help app versus self-monitoring app to promote mental wellbeing in healthy young adults (ECoWeB PROMOTE): an international, multicentre, parallel, open-label, randomised controlled trial. Lancet Dig. Health 6 (12), e904–e913. https://doi.org/10.1016/S2589-7500(24)00149-3.

Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. 54, 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063.

Weidmann, B., Xu, Y., 2025. Measuring emotion perception ability using AI-generated stimuli: development and validation of the PAGE test. J. Intell. 13 (9). https://doi.org/10.3390/jintelligence13090116.

Wieck, C., Scheibe, S., Kunzmann, U., 2022. Development and validation of film stimuli to assess empathy in the work context. Behav. Res. Methods 54 (1), 75–93. https://doi.org/10.3758/s13428-021-01594-6.

Yeung, M.K., 2022. A systematic review and meta-analysis of facial emotion recognition in autism spectrum disorder: the specificity of deficits and the role of task characteristics. Neurosci. Biobehav. Rev. 133, 104518. https://doi.org/10.1016/j.neubiorev.2021.104518.

Yitzhak, N., Pertzov, Y., Aviezer, H., 2021. The elusive link between eye-movement patterns and facial expression recognition. Soc. Personal. Psychol. Compass 15 (7), e12621. https://doi.org/10.1111/spc3.12621.

Yitzhak, N., Pertzov, Y., Guy, N., Aviezer, H., 2022. Many ways to see your feelings: successful facial expression recognition occurs with diverse patterns of fixation distributions. Emotion 22 (5), 844–860. https://doi.org/10.1037/emo0000812.

Zaki, J., Ochsner, K., 2011. Reintegrating the study of accuracy into social cognition research. Psychol. Inq. 22 (3), 159–182. https://doi.org/10.1080/1047840X.2011.551743.