

Development and Validation of the Emotion Laden Sentences Toolbox for Emotion Recognition (ELSTER)

Rohit Verma¹, Navkiran Kalsi², Neha Priya Shrivastava³, Anita Sheerha⁴, Ishita Dhyani⁴ and Shubham Narnoli⁵

ABSTRACT

Introduction: Emotion recognition plays a crucial role in our social interactions and overall well-being. The present cross-sectional study aimed to develop and validate Emotion Laden Sentences Toolbox for Emotion Recognition (ELSTER), that utilizes emotion-laden sentences as stimuli to assess individuals' ability to perceive and identify emotions conveyed through written language.

Methods: In Phase I, a comprehensive set of emotion-laden sentences in English language were validated by 25 (eight males and 17 females) qualified mental health professionals (MHPs). In Phase II, the sentences that received high interrater agreement in Phase I were selected and then a Hindi version of the same sentences was also developed. The English and Hindi database was then validated among 50 healthy individuals (30 males and 20 females).

Results: The percentage hit rate for all the emotions after exclusion of contempt was 84.3% with a mean kappa for emotional

expression being 0.67 among MHPs. The percentage hit rate of all emotion-laden sentences across the database was 81.43% among healthy lay individuals. The mean hit rate percentage for English sentences was similar to Hindi sentences with a mean kappa for emotional expression being 0.63 for the combined English and Hindi sentences.

Conclusion: The ELSTER database would be useful in the Indian context for researching textual emotion recognition. It has been validated among a group of experts as well as healthy lay individuals and was found to have high inter-rater reliability.

Keywords: Emotion, recognition, emotion-laden, textual emotion, sentence, affect

Key Message: Social communication is dependent upon the ability of individuals to interpret the emotional expression of others. The recognition of emotional expressions is based on both verbal and

non-verbal stimuli. While there are databases existing for emotion recognition of non-verbal stimuli, none exist purely based on textual form of emotion expression. Emotion Laden Sentences Toolbox for Emotion Recognition (ELSTER) database would be useful in the context of researching emotion recognition for English and Hindi reading populations.

Emotions are increasingly being studied in domains such as psychology, neurology, human-computer interaction, linguistics, and so on. They play a profound role in a human's life. Emotions have aided our ancestors in adaptation and survival; they are a powerful means of communication; and they have a considerable influence on our cognitive functions, such as attention, memory, and decision-making.¹ Being able to recognize these emotions helps increase self-awareness as well as stronger connections and more meaningful interactions

¹Additional Professor, All India Institute of Medical Sciences (AIIMS), New Delhi, Delhi, India. ²Post Doctoral Fellow, All India Institute of Medical Sciences (AIIMS), New Delhi, Delhi, India. ³Research Officer, All India Institute of Medical Sciences (AIIMS), New Delhi, Delhi, India. ⁴Research Assistant, All India Institute of Medical Sciences (AIIMS), New Delhi, Delhi, India. ⁵Senior Resident, Brain Mapping Lab, Department of Psychiatry, All India Institute of Medical Sciences (AIIMS), New Delhi, Delhi, India.

HOW TO CITE THIS ARTICLE: Verma R, Kalsi N, Shrivastava NP, Sheerha A, Dhyani I and Narnoli S. Development and Validation of the Emotion Laden Sentences Toolbox for Emotion Recognition (ELSTER). *Indian J Psychol Med.* 2024;46(3):253–259.

Address for correspondence: Rohit Verma, All India Institute of Medical Sciences, Ansari Nagar, New Delhi, Delhi 110029, India.
E-mail: rohit.aiims@gmail.com

Submitted: 09 Aug. 2023
Accepted: 21 Nov. 2023
Published Online: 31 Jan. 2024



Copyright © The Author(s) 2024

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution- NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-Commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the Sage and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

ACCESS THIS ARTICLE ONLINE
Website: journals.sagepub.com/home/szj
DOI: 10.1177/02537176231219980

with others. It also assists individuals in efficiently managing their own emotions. Therefore, emotion recognition is a valuable skill to possess.

When discussing emotion recognition, it is critical to mention the theory of mind (ToM). ToM involves the capacity to comprehend the thoughts, beliefs, intentions, and desires of other individuals, as well as to acknowledge that these mental states can vary from one's own.² The conventional understanding of ToM is that it can be separated into two systems: cognitive and affective. Affective ToM refers to understanding the other person's emotions, feelings, and affective states. On the other hand, cognitive ToM, indicates an understanding of the other person's cognitive states, beliefs, and intentions.³ While few researchers have stated that the affective component of ToM declines with age^{4,5} others have opposed this notion suggesting that age has no effect on affective ToM.^{6,7}

Failure to recognize emotions can impact an individual's social and emotional functioning. Understanding and interpreting social cues, such as others' emotional states and facial expressions, can be particularly difficult for individuals with autism,^{8,9} schizophrenia,¹⁰⁻¹³ depression,^{14,15} and social anxiety disorder.^{16,17}

There are several tasks that are frequently used for assessing emotion recognition, each of which concentrates on a distinct type of emotional signal. While most of these measures are based on recognizing emotions in non-verbal stimuli, particularly facial expressions,¹⁸ there are a few databases that evaluate other forms of non-verbal expressions. The Reading the Mind in the Eyes Test (RMET), evaluates the capacity to attribute emotions by looking at photos of people's eyes.¹⁹ Geneva Emotion Recognition Test (GERT) assesses the ability to identify emotions using a variety of modalities, such as body language, voice cues, and facial expressions.²⁰ The Body Action and Posture (BAP) test evaluates videos or pictures of body language and identifies the emotions being conveyed.²¹ The Montreal Affective Voices (MAV) database assesses emotion recognition from vocal cues.²² The Emotion Attribution Task (EAT) asks participants to assign emotional states after listening to written stories of emotional situations read aloud by the researcher.²³ Although

assessments that use verbal stimuli do exist, none of them focus on reading and comprehension.

Furthermore, previous studies in the field of emotion research have primarily focused on two key dimensions: valence and arousal. Arousal pertains to the degree of activation within the appetitive or defensive motivation system that is triggered by an emotion. On the other hand, valence refers to the classification of emotions as either positive or negative.²⁴ Regardless of their emotional valence, it has been noted that words associated with emotions undergo faster processing compared to neutral words. This occurrence is commonly known as the 'emotion effect'.^{25,26}

The emotion word type was neglected in a large proportion of earlier studies on emotion word processing.^{27,28} Recently, there has been a greater focus on comparing emotion-label words with emotion-laden words. Emotion-laden words, such as 'successful', 'failed', elicit an individual's emotions through the word's implications, as opposed to emotion-label terms, which directly explain or describe one's affective states like 'upset' or 'happy'.^{28,29} Compared to emotion-laden words, emotion-label words have a stronger priming effect and faster response times.³⁰ It was also observed that emotion-label words had a processing advantage over emotion-laden words regardless of whether it was a first or second language used.³¹ But in everyday usage, it is not uncommon that emotion-laden words or sentences are used to attribute emotions. In the ever-growing world of social media, such words or sentences may have a larger implication for recognizing emotional states.

There can be several advantages to assessing emotion recognition using textual stimuli. It will allow a standardised presentation of emotions and multiple people can be assessed at the same time. The use of language may also eliminate any kind of visual bias. Furthermore, using textual stimuli will help us tap into language-based processing of emotion recognition, which has not been studied much. Thus, the development of a text-based emotion recognition tool will provide an alternative perspective on how individuals recognize and identify emotions.

The primary objective of this study is to create a comprehensive assessment database that measures the ability to

recognize emotions based on written/textual stimuli. Specifically, the focus is on evaluating individuals' proficiency in understanding and perceiving emotions conveyed through emotion-laden sentences. By developing such a tool, researchers aim to provide a reliable and valid means of assessing individuals' competence in recognizing and interpreting emotions expressed through written language. The tool will enable researchers and practitioners to gain insights into individuals' emotional perception abilities and potentially identify any deficits or variations in emotion recognition skills.

Methodology

Development of the Database

The data presented in the current article on the development of Emotion Laden Sentences Toolbox for Emotion Recognition (ELSTER) database is part of a larger study comprising the development of a larger toolbox for varying facets of emotion recognition.^{18,32} This cross-sectional study was conducted from March 2019 to August 2021 at the Department of Psychiatry, All India Institute of Medical Sciences, New Delhi, India. The Institute Ethics Committee approved the study. Based on the previous literature, eight emotion expressions were selected—neutral, happiness, anger, sadness, disgust, contempt, fear, and surprise.¹⁸

The emotion-laden sentences were initially selected in English language from freely available online books, novels, and other textual literature. Ten sentences depicting each emotional state were selected keeping the length to a maximum of 40 words in proper grammatical format by three researchers. Thereafter, two other researchers re-assessed the sentences and reduced the number to six sentences for each emotion keeping the sentence length to a maximum of 25 words in proper grammatical format. Thus, a total of 48 emotion-laden sentences were developed.

Validation of the Database

Participants

In Phase I, the database was validated by 25 qualified mental health professionals (MHPs) from the Department of Psychiatry of a tertiary care hospital (eight

males and 17 females). All raters had a normal or corrected-to-normal vision and volunteered for no-cost participation. All participants were proficient in the English language and were educated with at least an undergraduate degree.

In Phase II, the sentences that received high interrater agreement in Phase I were selected and then a Hindi version of the same sentences was also developed by three researchers independently. Thereafter, two other researchers reassessed the Hindi sentences and a final Hindi set with concurrence was reached at. The English and Hindi database was then validated among 50 healthy individuals presenting as caregivers to patients attending the psychiatry outpatient clinic of a tertiary care hospital (30 males and 20 females). All participants were educated at least up to higher secondary and consented to participation in the study.

Procedure

Raters were presented with randomised sentences of different emotions on a 27-inch computer screen. They were seated 50–70 cm away from the screen and proceeded at their own pace in the presence of a researcher (NPS or NK). Raters were asked to recognize the expression that was the best fit for the emotion depicted in the sentence, choosing one of the eight response categories in Phase I: happiness, anger, sadness, neutral, surprise, fear, contempt, disgust and others. In Phase II, the ‘other’ and ‘contempt’ response category was removed based on Phase I findings.

Analysis

The emotion expression recognition of the sentences was evaluated by hit rate percentage with standard deviation (SD), where the proportion of raters who agree with the intended expression was calculated for each emotion. Initially, for each sentence, hit rates were calculated for emotion recognition. Then, five sentences of each emotion with the highest hit rates were selected and the mean hit rate for the emotion was calculated from these sentence sets. In addition, we also calculated Fleiss’ kappa, which is a chance-corrected measure of agreement between the intended expression and the raters’ labels. The agreement interpretation for Fleiss’ kappa was taken as: <0: poor, 0.0–0.20: slight, 0.21–0.40: fair, 0.41–0.60: moderate, 0.61–0.80: substantial, and 0.81–1.0: almost perfect.³³

Results

Phase I

A total of 48 emotion-laden sentences were presented to MHPs. For each sentence, we calculated how many participants identified the correct emotion. The overall mean hit rate percentage of all emotions across the database was 77.92% (SD = 25.11). The percentage hit rates for each emotion are depicted in **Table 1**. The hit rate for contempt was very low—36% (SD = 20.70); hence, the contempt emotion was removed from further analysis. The new mean hit rate percentage of all the emotions after the exclusion of contempt was 84.3% (SD = 8.67). We calculated the hit rates for each

sentence and finally selected five sentences with the highest hit rates for each emotion to be included in the database—ELSTER. The hit rate of each selected sentence for individual emotion is provided in supplementary **Table S1**.

There was substantial agreement (Fleiss Kappa = 0.67) between the raters ($n = 25$) for overall emotion recognition of the sentences. The agreement for individual emotion was also found to be substantial (0.59–0.78) except for contempt (Kappa = 0.34) as shown in **Table 2**. The kappa for sad emotion sentence recognition was at almost perfect agreement.

Phase II

After removing sentences having low hit rates (less than 75%) for each emotion and removing contempt-expressing sentences, the remaining 35 emotion-laden sentences (five sentences for each of the seven emotions—happiness, anger, sadness, neutral, surprise, fear and disgust) were presented to healthy lay individuals.

The overall mean hit rate percentage of all emotion-laden sentences across the database was 81.43% (SD = 10.38). The mean hit rate percentage for English sentences was 84.28% (SD = 13.83) and for Hindi sentences was 80.32% (SD = 11.46). The percentage hit rates for each emotion are depicted in **Table 1**. The hit rate of each selected sentence for individual emotion is provided in supplementary **Table S2**.

There was substantial agreement (Fleiss Kappa = 0.63) between the raters

TABLE 1.

Hit Rate for Emotion Recognition in Phase I and Phase II.

Emotion	Phase I MHPs for English ELS Average Hit Rate% (SD)	Phase II Healthy Lay Individuals for Overall ELS Average Hit Rate% (SD)	Phase II Healthy Lay Individuals for English ELS Average Hit Rate% (SD)	Phase II Healthy Lay Individuals for Hindi ELS Average Hit Rate% (SD)
Anger	83.33 (15.05)	77.20 ± 5.93	77.14 ± 10.59	77.22 ± 4.96
Contempt	36.00 (20.70)	–	–	–
Disgust	80.67 (26.09)	69.20 ± 11.36	64.28 ± 13.36	71.11 ± 13.40
Fear	75.33 (32.53)	81.60 ± 8.64	85.71 ± 7.14	80.01 ± 10.28
Happiness	92.00 (4.38)	89.20 ± 3.34	94.28 ± 3.19	87.22 ± 5.04
Neutral	77.33 (21.41)	92.00 ± 5.09	94.28 ± 7.82	91.11 ± 5.34
Sadness	94.00 (8.29)	82.00 ± 11.91	94.28 ± 5.97	77.22 ± 15.26
Surprise	84.67 (14.84)	78.80 ± 8.08	80.01 ± 14.63	78.33 ± 13.37

ELS, emotion-laden sentences; MHPs, mental health professionals; SD, standard deviation.

TABLE 2.

Agreement on Individual Emotion Categories in Phase I with MHPs.

Rating Category	Kappa	Asymptotic			Asymptotic 95% Confidence Interval	
		Standard Error	z	Sig.	Lower Bound	Upper Bound
1	.68	.008	81.83	<.001	.66	.69
2	.29	.008	35.67	<.001	.28	.31
3	.71	.008	84.68	<.001	.68	.72
4	.72	.008	86.09	<.001	.70	.73
5	.75	.008	89.52	<.001	.73	.76
6	.59	.008	71.01	<.001	.57	.60
7	.79	.008	94.61	<.001	.77	.80
8	.72	.008	86.70	<.001	.70	.73
9	.06	.008	6.63	<.001	.03	.07

MHPs-Mental health professionals.

TABLE 3.

Agreement on Individual Emotion Categories in Phase II.

Emotion Category	Combined ELS			English ELS			Hindi ELS		
	Kappa	95% CI		Kappa	95% CI		Kappa	95% CI	
		Lower Bound	Upper Bound		Lower Bound	Upper Bound		Lower Bound	Upper Bound
Anger	0.61	0.60	0.61	0.64	0.61	0.67	0.59	0.57	0.60
Disgust	0.59	0.58	0.60	0.53	0.50	0.57	0.62	0.60	0.63
Fear	0.65	0.64	0.66	0.68	0.64	0.71	0.65	0.63	0.66
Happy	0.68	0.67	0.69	0.76	0.72	0.79	0.67	0.65	0.68
Sad	0.63	0.62	0.64	0.80	0.76	0.83	0.57	0.56	0.59
Surprise	0.56	0.55	0.57	0.54	0.51	0.58	0.58	0.57	0.59
Neutral	0.66	0.65	0.67	0.78	0.75	0.82	0.62	0.61	0.63

CI, confidence interval; ELS, emotion laden sentences; All Kappa values were significant at $p < .001$. Standard Error value for English, Hindi and combined sentences was 0.018, 0.007 & 0.005, respectively.

($n = 50$) for overall emotion recognition for the combined English and Hindi sentences. The agreement for individual emotion was also found to be moderate to substantial (0.56–0.68) as shown in **Table 3**.

Among the participants, the distribution of English emotion-laden sentences ($n = 14$) and Hindi emotion-laden sentences ($n = 36$), was similar among the genders (Males vs. Females: 30 vs. 20, Chi-square = 0.15, $p = .70$). The mean age was also comparable among the genders (Male: 36.93 \pm 11.98 years, Female: 34.25 \pm 12.64 years, $t = 0.76$, $p = .45$).

When assessing inter-rater reliability individually for language-dependent performance, there was a substantial agreement for both English and Hindi sentences for overall emotion recognition (Fleiss Kappa = 0.68 & 0.61, respectively). The agreement for individual emotion was also found to be moderate to substantial for both English (0.53–0.80) and Hindi

(0.57–0.65) language emotion-laden sentences (**Table 3**).

Discussion

This study proposes an assessment toolbox that measures textual emotion recognition using emotion-laden sentences. The current database contains emotion-laden sentences in English and Hindi language. A total of 35 emotion-laden sentences were finalised, five from each of the seven emotions (happiness, anger, sadness, neutral, surprise, fear and disgust). Fleiss' kappa and hit rates were calculated to evaluate the tool's reliability and validity. These sentences were validated initially by MHPs in Phase I and subsequently by healthy individuals in Phase II, for the intent of standardisation.

Like Ekman's research, the initial phase of the study included the emotion of 'contempt'.³⁴ However, just like Ekman,

this category along with 'others', was eliminated due to low accuracy rates in phase I. Similarly, our findings from the previous study on emotion recognition of facial emotion in All India Institute of Medical Sciences (AIIMS) Facial Toolbox for Emotion Recognition (AFTER) database,¹⁸ also found that the contempt emotion had poor hit rates among the study population. The same was also true for the Dynamic Emotion Laden Inventory of Videos for Emotion Recognition (DELIVER) database that assessed emotion recognition through videos of emotion expressions.³² Thus, it appears that contempt may not be recognizable as an easily identifiable emotion, particularly for the Indian population. This may be true for other cultures too as most other emotion databases also do not include contempt.

Phase I MHPs had a higher average hit rate for most emotion categories compared

to Phase II healthy lay individuals, who scored a higher hit rate than MHPs in fear and neutral emotion categories only. It may suggest a potential influence of professional training on emotion recognition abilities for most emotions.

Among the healthy individuals, there was substantial agreement for overall emotion recognition for the combined as well as individual English ELS and Hindi ELS. There was generally a good agreement among raters for categorising emotions, with moderate to almost perfect Kappa values observed for most emotion categories. This indicates a reliable and consistent categorization of emotions within each category across different conditions. The Kappa values were all statistically significant, indicating a high level of agreement beyond chance. Except for disgust and surprise emotion, the hit rates were higher for English ELS than for Hindi ELS. This may indicate that language does play a role in emotion recognition.

Numerous studies suggest that there is a difference in language processing in a native language and second language.³⁵⁻³⁷ It is observed that there's a difference in terms of affective processing in those who acquired a second language later in life and use it as a foreign language. The second language appears to be processed in a semantic way rather than affectively. This distinction is responsible for the reduction of framing biases observed in the processing of the second language.³⁸

In the current study, gender was not found to have any notable impact on emotion recognition. Many studies have underlined the gender advantage in decoding emotions with females showing better recognition over males.^{39,40} However, research does not always confirm the maintenance of gender differences across the age span suggesting larger differences for teenagers and young adults.^{39,41} Studies have reported that females seem to show significantly higher scores in anger recognition and also seem to show more ability than males to recognize paired emotions. Studies in affective valence processing have shown sex-specificity with females displaying hypersensitivity to negative valence.⁴² But these studies were conducted using facial expression databases, thus limiting their generalizability for emotion recognition across different stimuli.

As previously mentioned, several tests and assessments are available to evaluate emotion recognition. Emotion recognition from face tasks essentially assesses the capacity to discriminate facial expressions. The tasks that use auditory stimuli test prosodic information discrimination. The suggested ELSTER database uses emotion-laden sentences, which can provide a higher level of linguistic complexity. Furthermore, ELSTER corresponds to real-world communication situations. Moreover, the database is available in a bilingual language unlike any other to provide a more comprehensive assessment, particularly for the Hindi reading population. No such database is currently available.

The Emotion Attribution Task (EAT) was not immune to bias. Even though the stimuli used were written language, the task required the examiner to read those stories out loud to the participants.²³ This defeats the point of using written language. Furthermore, this does not guarantee standard and consistent presentation of stimuli each time as there are studies that suggest the role of auditory prosody in recognition or appreciation of emotion too. ELSTER will eliminate these biases by making the participant use their own cognitive and language processing skills.

Also, using emotion-laden words instead of emotion-label words also has major advantages. Recognition of emotion using emotion-laden sentences would require conceptual knowledge of the emotions. It would also require an understanding of semantics. Whereas emotion-label sentences may be easier to identify or recall because emotion-label words would be a part of explicit memory. Therefore, researchers also suggest that emotion-label words are recognised faster as compared to emotion-laden words.³⁰

With changing times, as human-machine interaction is increasing and becoming more common, research is being done in the field of natural language processing (NLP). Affective computing refers to techniques for detecting, recognizing, and predicting human emotions with the goal of adapting computational systems to these states.⁴³ The toolkit EmoTxt has been shown to achieve high accuracy in emotion recognition from text.⁴⁴ Colnerič & Demšar⁴⁵

used a collection of tweets to explore the use of deep learning for emotion detection on Twitter. Mohammad & Kiritchenko⁴⁶ studied emotion-word hashtags in tweets using them as emotion labels to generate a large lexicon of word-emotion associations from these hashtags. The authors also suggested improving emotion classification accuracy in a different non-tweet's domain using the lexicon. The traditional machine learning can only learn from a fixed-size vector of features and thus features are commonly built upon bag-of-words. Even though computational techniques such as deep learning have yielded considerable performance improvements for a variety of tasks in NLP, naive network architectures struggle with the task of emotion recognition.⁴⁷

To the best of our knowledge, there is no existing database that provides a reliable and validated emotion recognition database for textual stimuli. The English as well as Hindi version of the ELSTER database has high hit rates and good inter-rater reliability both among a cohort of experts and healthy individuals for emotion recognition. The emotion-laden sentences from this database can be used freely for emotion research.

The development of the ELSTER database utilizing emotion-laden sentences represents a significant advancement in the field of emotion research. The database provides a standardized and psychometrically sound instrument for assessing individuals' ability to recognize and interpret emotions conveyed through written language. The database has significant prospects in both research and clinical contexts. It offers valuable utility in evaluating emotion recognition abilities among clinical populations, including individuals diagnosed with autism spectrum disorders or mood disorders. Moreover, it can contribute to comprehending emotional processing and its correlation with social functioning and interpersonal communication. Future research should focus on further validation and refinement of ELSTER, as well as exploring its utility in various contexts and populations.

Conclusion

The ELSTER database would be useful in the context of conducting research in

the field of textual emotion recognition. ELSTER has been validated in a cohort of experts and healthy lay individuals. Based on the hit rates and the good inter-rater reliability, it might be concluded that the ELSTER database offers a valid set of affective stimuli for recognising emotions.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The study is sponsored through the Cognitive Science Research Initiative (CSRI) grant of Department of Science and Technology (DST) (D.O. No. DST/CSRI/2017/186 dated April 12, 2018) to Dr Rohit Verma.

Supplemental Material

Supplemental material for this article is available online.

ORCID iD

Rohit Verma  <https://orcid.org/0000-0001-7418-2511>

References

- LeBlanc VR, McConnell MM, Monteiro SD. Predictable chaos: a review of the effects of emotions on attention, memory and decision making. *Adv Health Sci Educ* 2015; 20: 265–282.
- Premack D, Woodruff G. Does the chimpanzee have a theory of mind? *Behav Brain Sci* 1978; 1(4): 515–526.
- Shamay-Tsoory SG, Aharon-Peretz J. Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia* 2007; 45(13): 3054–3067.
- Mahy CE, Vetter N, Kühn-Popp N, et al. The influence of inhibitory processes on affective theory of mind in young and old adults. *Aging Neuropsychol Cogn* 2014; 21(2): 129–145.
- Grainger SA, Henry JD, Phillips LH, et al. Age deficits in facial affect recognition: the influence of dynamic cues. *J Gerontol B Psychol Sci Soc Sci* 2017; 72(4): 622–632.
- Castelli I, Baglio F, Blasi V, et al. Effects of aging on mindreading ability through the eyes: an fMRI study. *Neuropsychologia* 2010; 48(9): 2586–2594.
- Yıldırım E, Soncu Büyükişcan E, Gürvit H. Affective theory of mind in human aging: is there any relation with executive functioning? *Aging Neuropsychol Cogn* 2020; 27(2): 207–219.
- Humphreys K, Minshew N, Leonard GL, et al. A fine-grained analysis of facial expression processing in high-functioning adults with autism. *Neuropsychologia* 2007; 45(4): 685–695.
- Ashwin E, Chapman E, Colle L, et al. Impaired recognition of negative basic emotions in autism: a test of the amygdala theory. *Soc Neurosci* 2006; 1: 349–363.
- Edwards J, Jackson HJ, Pattison PE. Emotion recognition via facial expression and affective prosody in schizophrenia: a methodological review. *Clin Psychol Rev* 2002; 22(6): 789–832.
- Barkhof E, de Sonneville LM, Meijer CJ, et al. Specificity of facial emotion recognition impairments in patients with multi-episode schizophrenia. *Schizophr Res* 2015; 2(1): 12–19.
- Leppänen JM, Niehaus DJ, Koen L, et al. Emotional face processing deficit in schizophrenia: a replication study in a South African Xhosa population. *Schizophr Res* 2006; 84(2–3): 323–330.
- Catalan A, Gonzalez de Artaza M, Bustamante S, et al. Differences in facial emotion recognition between first episode psychosis, borderline personality disorder and healthy controls. *PLoS One* 2016; 11(7): e0160056.
- Demenescu LR, Kortekaas R, den Boer JA, et al. Impaired attribution of emotion to facial expressions in anxiety and major depression. *PLoS one* 2010; 5(12): e15058.
- Kohler CG, Hoffman LJ, Eastman LB, et al. Facial emotion perception in depression and bipolar disorder: a quantitative review. *Psychiatry Res* 2011; 188(3): 303–309.
- Tseng HH, Huang YL, Chen JT, et al. Facial and prosodic emotion recognition in social anxiety disorder. *Cogn Neuropsychiatry* 2017; 22(4): 331–345.
- Maoz K, Eldar S, Stoddard J, et al. Angry-happy interpretations of ambiguous faces in social anxiety disorder. *Psychiatry Res* 2016; 241: 122–127.
- Verma R, Kalsi N, Shrivastava NP et al. Development and Validation of the AIIMS Facial Toolbox for Emotion Recognition (AFTER). *Indian J Psychol Med* 2022; 45(5): 471–475.
- Baron-Cohen S, Wheelwright S, Hill J, et al. The “Reading the Mind in the Eyes” test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry* 2001; 42(2): 241–251.
- Schlegel K, Grandjean D, Scherer KR. Introducing the Geneva emotion recognition test: an example of Rasch-based test development. *Psychol Assess* 2014; 26(2): 666.
- Dael N, Mortillaro M, Scherer KR. The body action and posture coding system (BAP): development and reliability. *J Nonverbal Behav* 2012; 36: 97–121.
- Belin P, Fillion-Bilodeau S, Gosselin F. The Montreal Affective Voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behav Res Methods* 2008; 40(2): 531–539.
- Prior M, Marchi S, Sartori G. Social cognition and behavior. A tool for assessment. Upsilon Domenighini Editore, Padova, 2003.
- Lang PJ, Simons RF, Balaban MT. Attention and orienting: sensory and motivational processes. Mahwah, NJ: Lawrence Erlbaum Associates; 1997.
- Kousta ST, Vinson DP, Vigliocco G. Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 2009; 112(3): 473–481.
- Vinson D, Ponari M, Vigliocco G. How does emotional content affect lexical processing? *Cogn Emot* 2014; 28(4): 737–746.
- Altarriba J, and Basnight-Brown DM. The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *Int J Biling* 2011; 15(3): 310–328.
- Zhang J, Wu C, Meng Y, et al. Different neural correlates of emotion-label words and emotion-laden words: An ERP study. *Front Hum Neurosci* 2017; 11: 455.
- Pavlenko A. Emotion and emotion-laden words in the bilingual lexicon. *Biling: Lang Cogn* 2008; 11(2): 147–164.
- Kazanas SA, Altarriba J. The automatic activation of emotion and emotion-laden words: Evidence from a masked and unmasked priming paradigm. *Am J Psychol* 2015; 128(3): 323–336.
- Tang D, Fu Y, Wang H, et al. The embodiment of emotion-label words and emotion-laden words: Evidence from late Chinese–English bilinguals. *Front Psychol* 2023; 14: 1143064.
- Verma R, Kalsi N, Sheerha A, et al. Dynamic emotion laden inventory of videos for emotion recognition (DELIVER): development and validation. *Indian J Psychiatry* 2023; 65(Suppl 1): S138.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977: 159–174.
- Ekman P, Friesen WV, O’Sullivan M, et al. Universals and cultural differences in the judgments of facial expressions of emotion. *J Pers Soc Psychol* 1987; 53(4): 712.
- Harris CL. Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *J Multiling Multicult Dev* 2004; 25(2–3): 223–247.

36. Pavlenko A. Affective processing in bilingual speakers: Disembodied cognition? *Int J Psychol* 2012; 47(6): 405–428.
37. Caldwell-Harris CL. Emotionality differences between a native and foreign language: theoretical implications. *Front Psychol* 2014; 5: 1055.
38. Keysar B, Hayakawa SL, An SG. The foreign-language effect: thinking in a foreign tongue reduces decision biases. *Psychol Sci* 2012; 23(6): 661–668.
39. Olderbak S, Wilhelm O, Hildebrandt A, et al. Sex differences in facial emotion perception ability across the lifespan. *Cogn Emot* 2019; 33(3): 579–588.
40. Abbruzzese L, Magnani N, Robertson IH, et al. Age and gender differences in emotion recognition. *Front Psychol* 2019; 10: 2371.
41. Thompson AE, Voyer D. Sex differences in the ability to recognise non-verbal displays of emotion: a meta-analysis. *Cogn Emot* 2014; 28(7): 1164–1195.
42. Singh V, Kumar S, Jha A, et al. Exploring sex-specific affect processing and depression-anxiety comorbidity in a gender-inequitable country. *Minerva Psichiatr* 2023; 63:00. (First published online)
43. Picard R. Affective computing. Cambridge, Massachusetts Institute of Technology.
44. Calefato F, Lanubile F, Novielli N. Emotxt: a toolkit for emotion recognition from text. In 2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) 2017 Oct 23 (pp. 79–80). IEEE.
45. Colnerić N, and Demšar J. Emotion recognition on twitter: comparative study and training a unison model. *IEEE Trans Affect Comput* 2018; 11(3): 433–446.
46. Mohammad SM, Kiritchenko S. Using hashtags to capture fine emotion categories from tweets. *Comput Intell* 2015; 31(2): 301–326.
47. Kratzwald B, Ilić S, Kraus M, et al. Deep learning for affective computing: text-based emotion recognition in decision support. *Decis Support Syst* 2018; 115: 24–35.