Social Sciences & Humanities Open 13 (2026) 102387

Contents lists available at ScienceDirect

Social Sciences & Humanities Open

journal homepage: www.sciencedirect.com/journal/social-sciences-and-humanities-open

Review Article

Recognition of emotions and non-verbal communication characteristics in human-robot interactions - the experience of a systematic review

Balázs Őrsi [a], Péter Korondi [b], Csilla Csukonyi [a], Dávid Papp [a], Panna Márkus [a], Bence Tamás Selejó Joó [a], Husam A. Neamah [b], Karolina Eszter Kovács [a,*]

[a] Institute of Psychology, University of Debrecen, 4032, Debrecen, Hungary
[b] Department of Vehicle Engineering and Mechatronics, Department of Electrical Engineering and Mechatronics, Faculty of Engineering, University of Debrecen, 4032, Debrecen, Hungary

ARTICLE INFO

Keywords:
User experience
Human-robot interaction
Robot psychology
Non-verbal communication
Systematic review

ABSTRACT

The aim of this review is to analyse previous research focusing on the non-verbal communication patterns of human-robot interaction, and related study characteristics, HRI-related specificities, methodologic background and main findings. The field of robotics is increasingly recognizing the significance of incorporating affective expression and social interaction. Many non-anthropomorphic mobile robots are currently utilised in various fields. However, these robots lack the ability to project facial expressions and cannot be modified to support affective expressions explicitly, which presents a significant challenge in terms of facilitating naturalistic human-robot interaction with appearance-constrained robots. The current paper introduces the results of a systematic review that follows the PRISMA guidelines. The EBSCO Discovery Service Search Engine was used for a systematic search, yielding 114 records. By a multiple-step analysis, seven studies met the criteria. Research indicates that the physical gestures of robots can be as identifiable as those of humans, highlighting the importance of nonverbal communication in human-robot interaction (HRI). Certain personality traits of the participants appeared to significantly influence the recognition of emotions, particularly evident in terms of confidence and valence. The findings suggest that robots should not only be designed to express emotions clearly but also to enhance these expressions in ways that boost users' confidence in their understanding. Improving the quality of emotional expression and non-verbal cues in human-robot interaction may play a crucial role in Industry 5.0. Future research should focus on the identified underrepresented themes to deepen understanding of emotion recognition in HRI.

1. Introduction

Nonverbal communication has a significant relevance in our everyday lives. Nonverbal cues allow humans to express their inner states and intentions to others, and they have the ability to interpret and understand the inner states and intentions of others from these nonverbal cues (Argyle, 1988; Depaulo & Friedman, 1998). The human face conveys significant information: it identifies the speaker, helps listeners understand speech through lip movements, expresses attitudes, adjusts meanings, provides feedback with head nods, and directs attention with gaze (Skantze, 2017).

Non-verbal communication, through gestures, posture, gaze, and proxemics, plays a crucial role in conveying emotion and intention in both human-human and human-robot interaction. As robots become more socially integrated, understanding how users interpret such cues is key to designing emotionally responsive and believable agents (Sheikholeslami et al., 2017). These signals may also be critical in human-robot interaction (HRI), as nonverbal communication plays a crucial role in achieving effective and natural collaboration. Nonverbal cues, such as body language, facial expressions, and eye contact, are essential tools for robots to understand human intentions and emotions, as well as to communicate their own internal states and intentions (Bethel & Murphy, 2008; Breazeal, 2003; Damiano & Dumouchel, 2018). These signals not only enhance the user experience but also increase trust and acceptance of robots (Dautenhahn, 2007). Research shows that robots capable of subtle nonverbal communication collaborate more effectively with humans and are adapt better to dynamic interactions (Fong et al., 2003). Therefore, integrating nonverbal

* Corresponding author.
E-mail addresses: kovacs.karolina@arts.unideb.hu, karolina9.kovacs@gmail.com (K.E. Kovács).

https://doi.org/10.1016/j.ssaho.2025.102387
Received 20 March 2025; Received in revised form 16 December 2025; Accepted 17 December 2025
Available online 6 January 2026

communication into robotic systems is key to achieving intuitive and harmonious HRI and to significantly contributing to the widespread adoption of technology in everyday life.

1.1. Theoretical models of non-verbal behaviour

Theoretical models of emotions and nonverbal behaviour are based on different scientific foundations and explain the nature of emotions, their expression, and their perception by others in various ways. Ekman's classic theory begins with the biological determination of emotions, positing universal basic emotions (joy, sadness, anger, disgust, fear, and surprise) that are expressed with similar facial expressions across cultures (Ekman, 1972). This theory had a significant impact on emotion research, especially in the coding of facial expressions (e.g., FACS – Facial Action Coding System). Still, critics say it ignores cultural and contextual differences. In contrast, Scherer's Component Process Model describes emotions as complex, dynamic systems that arise from the interaction of several components (cognitive evaluation, physical responses, expressive behavior, subjective experience, and behavioral tendencies) (Scherer, 2001). Emotions are thus not fixed categories, but rather the result of context-dependent interpretative processes that may vary from individual to individual.

Mehrabian adds to this line of thought by stating that when verbal and nonverbal signals contradict each other in the expression of emotions, people tend to consider nonverbal channels, such as tone of voice and facial expressions, to be more authentic, while words themselves contribute only to a small extent to the meaning (Mehrabian, 1972). This suggests that nonverbal signals play a central role in conveying attitudes and emotions. Russell's multidimensional approach, the Circumplex Model of Affect, describes emotions not as separate categories but along two main dimensions: valence (positive–negative) and arousal (Russell, 1980). This model offers a simple and easily representable form for mapping emotional states, especially in psychometric studies. Still, it is less able to capture the social and cognitive complexity of emotions. Fridlund's interactionist model, however, breaks with views that emphasize biological determinism (Fridlund, 1994). According to this theory, emotional expression is not necessarily an automatic reflection of emotional states, but rather communicative act that serves social purposes and are adapted to the presence and behaviour of others. Together, these theories contribute to a more nuanced understanding of the relationship between emotions and nonverbal behaviour, especially in fields such as education, health communication, artificial intelligence, and human-machine interaction.

1.2. Non-verbal communication in HRI

Regarding humanoid robots, in addition to the ability to communicate verbally, it is crucial to equip them with the capacity to interpret nonverbal cues from participants and comprehend their intentions. Several social robots meet this requirement by utilising face detection and speech recognition (e.g. Bohus & Horvitz, 2009; Li et al., 2015). Substantial advancements have been achieved in the development of service robots with social capabilities, enabling them to engage in one-on-one interactions effectively within controlled settings. However, these robots have limited proficiency in perceiving social cues and predicting intentions, which are essential for their deployment in unrestricted environments for natural interactions with multiple individuals.

The concepts of affective computing and embodied interaction provide an essential theoretical foundation for understanding the nonverbal dimensions of HRI more deeply. Affective computing refers to technologies designed to detect, interpret, and simulate human emotions, often through multimodal input such as facial expressions, body posture, and vocal tone (Picard, 1997). This paradigm emphasises the importance of designing robots that are not only responsive to human affect but also capable of expressing emotions in a way that users can interpret clearly and confidently (Ajibo et al., 2025; Corrales-Paredes et al., 2023; Kolomaznik et al., 2024).

Simultaneously, the notion of embodied interaction, which highlights the significance of physical presence, spatial behaviour, and sensorimotor coordination in communication, also correlates with the findings on gesture, posture, and movement in HRI (Dourish, 2001). Embodied interaction frames emotional signalling not as abstract information exchange, but as meaning emerging from the co-regulated dynamics between the human and the physically present robot (Belhassein et al., 2022; Chevalier et al., 2020; Kopp & Wachsmuth, 2010). Subtle variations in robot gaze, timing, or posture can influence human interpretation and emotional engagement, underscoring the situated and embodied nature of affective communication (El-Raheb et al., 2025; Fiore et al., 2013; Schellen et al., 2021).

Recent progress has been made in robot-mediated multiparty conversations, focusing on rule-based management of dialogue engagement and modelling for specific interactions (Mirnig et al., 2013; Zhai & Wibowo, 2023). For instance, Skantze (2017) developed a robot head named Furhat, that utilizes both head pose and eye movements to control its gaze. This design helps mitigate the 'Mona Lisa effect', a phenomenon in which a face (like a painting) appears to gaze at the observer regardless of the observer's position, leading to ambiguity about where the robot is looking. The robot also had animated lip movements, which improved speech comprehension under noisy conditions (Al Moubayed et al., 2013). Moon et al. (2013) stated that a human can recognise the hesitation gestures of a robot when both attempt to reach for the same object simultaneously. Ende et al. (2011) found that gestures referring to objects ("this one" and "from here to there") and those signalling termination ("stop" or "no") are easily identifiable across all types of robots. The study by Haddadi et al. (2013) highlights the importance of finger gestures for understanding robots' instructional gestures.

In social HRI scenarios, interactive robots must possess social intelligence to communicate effectively with humans in natural, bidirectional ways. Social intelligence enables robots to exchange information, establish connections, comprehend, and interact with individuals in environments centred around humans. By incorporating social intelligence, robots can facilitate more successful and engaging interactions, leading to greater acceptance among users (Breazeal, 2004; de Kervenoael et al., 2020; McColl & Nejat, 2014). The challenge lies in developing interactive robots that can perceive and recognise intricate human social behaviours and respond with their own behaviours, utilising a combination of natural communication methods like speech, facial expressions, paralanguage, and body language. The conveying of human intent is more effectively achieved through non-verbal communication, encompassing body language, facial expressions, and vocal intonation, compared to verbal expressions, particularly when expressing affection changes (Abdulghafor et al., 2022). Previous studies have primarily concentrated on identifying human emotions through facial expressions and vocal intonation, or a combination of the two, while giving minimal attention to directly recognizing emotions from body language, although it plays a crucial role in communicating human emotions during interpersonal social interactions.

Numerous recent studies have addressed the development of communicative robots capable of conveying social cues through gestures, recognizing their significance in facilitating smooth, successful interactions between humans and robots (Baddoura & Venture, 2015; Breazeal, 2004). The importance of gestures in HRI has been highlighted, showing that when robots use gestures, people experience more meaningful social exchanges and derive greater enjoyment than in interactions devoid of gestures. Furthermore, individuals tend to exhibit higher levels of engagement, anthropomorphise the robot to a greater extent, view it as more likeable, and express a stronger desire for future contact when gestures are incorporated into the interaction (Chi et al., 2024). In the context of elderly care centres, research suggests that even simple social gestures, such as daily greetings performed by a robot, can have positive effects on the well-being of the older people, bringing them pleasure, comfort, and interest while also playing a vital role in maintaining social connections (Sabelli et al., 2011).

1.3. Purpose of the review

Although emotion recognition has become an increasingly important aspect of HRI, existing reviews typically focus on technical approaches such as machine learning algorithms, sensor technologies, or facial recognition systems. However, little is known about the role of non-verbal communication cues, such as posture, gesture, proxemics, and body movement, in how humans perceive and interpret robot-expressed emotions. The lack of facial expressions in many robot designs and the limitation regarding emotional expressiveness remains an unsolved problem. Moreover, current literature often treats emotion recognition as a unidirectional process (i.e., robots recognizing human emotions), whereas the bidirectional and relational dynamics of emotional exchange between humans and robots remain underexplored. This is particularly crucial as robots transition from industrial tools to social partners in education, healthcare, and public service domains. Therefore, the purpose of the present systematic review is to address a specific gap in the literature: how non-verbal communication cues influence emotion recognition processes in human-robot interaction, particularly from the human user's perspective, and analyse previous research focusing on the non-verbal communication patterns of HRI. Specifically, the current paper aims to analyse the recognition of emotions and non-verbal communication characteristics in human-robot interactions and review how non-verbal communication and emotional expression affect the quality of human-robot interaction. Such a review may contribute to the development of robots and enhance the quality of human-robot interaction in various areas mentioned above. Based on this aim, we formulated the following research question:

1. How do specific non-verbal cues expressed by robots influence human users' ability to recognise emotional states during HRI?
2. What methodological approaches are commonly applied in empirical studies focusing on non-verbal communication in HRI, and can findings be generalised?
3. Which dimensions of non-verbal communication in HRI remain underrepresented in the current literature, particularly in relation to embodied and bidirectional interactions?

By focusing on non-verbal cues as mediators of emotion recognition, this review moves beyond general discussions of affective computing to highlight embodied, relational, and interactional dimensions of HRI that may contribute to designing and evaluating strategies for emotionally responsive robotic systems.

2. Methods

This systematic literature review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2015).

2.1. Literature review

EBSCO Discovery Service Search Engine was used for systematic search, which includes 85 databases. The keywords, using the Boolean search string, were as follows:

(TI OR AB OR SU) (robot* NEAR/3 (gesture* OR "body language" OR posture* OR movement* OR nonverbal* OR "non-verbal*"))

AND.

(TI OR AB OR SU) ("emotion* expression*" OR "affective expression*" OR express* NEAR/3 motion* OR "emotional display*")

AND.

(TI OR AB OR SU) ("human perception" OR "emotion* recognition" OR user* NEAR/3 interpret* OR "perceived emotion*")

The searches were performed in December 2023. Our systematic searches resulted in a total of 114 records. After double filtering, one record was excluded. Unscreened articles were listed in Zotero (V6.0.22, Roy Rosenzweig Center for History and New Media, George Mason University, Washington DC). After title and abstract screening, which was conducted between January and March 2024, 65 records were excluded. Therefore, 49 papers were sent for full-text screening, which led to the inclusion of seven papers in the qualitative synthesis that was carried out between April and June 2024. To strengthen and validate the completeness of the search strategy, supplementary searches were conducted in OpenAlex in November 2025 using the same Boolean logic. These additional searches returned high numbers of initial records; however, after title and abstract screening, none of the results met the predefined inclusion criteria. To further ensure that no relevant studies were missed, backward snowballing procedures were applied by screening the reference lists of all included studies. This procedure produced one additional eligible study.

2.2. Inclusion and exclusion criteria

The following inclusion criteria were set, following the PICOS format (P: Population, I: Interventions, C: Comparisons, O: Outcomes, S: Study designs):

- *Population*: human participants (children, adults, or older adults) involved in direct interaction with a robot in a social, service, or collaborative context; studies must involve humans interpreting or responding to robot-expressed emotions
- *Intervention*: original empirical research published in a peer-reviewed journal; interaction with a robot designed to express emotion through non-verbal communication, such as gesture, posture, movement, spatial behaviour or body orientation;
- *Comparison*: no comparison group required;
- *Outcomes*: qualitative or quantitative assessment of how humans perceive, recognise, or interpret robot emotions, specifically as influenced by non-verbal cues, any kind of psychological impact of the Human-Robot Interaction, Human-Robot Collaboration or Human-Robot Cooperation investigated;
- *Study design*: empirical studies using quantitative, qualitative, or mixed methods approaches, including experiments, observational studies and surveys.

These criteria were defined to focus the review on a specific intersection within HRI research: how human users emotionally interpret non-verbal communication from robots. This conceptual focus was prioritised to support analytic clarity, recognizing that it necessarily limits the breadth of included literature. Papers also must be written in English language and in the disciplines of psychology, social sciences, humanities and engineering. The language restriction was applied due to resource limitations and the unavailability of translation services, which could have decrease the accuracy and consistency of data extraction and quality assessment from non-English sources. While this may introduce a language bias, the decision was made to ensure methodological rigour and reliability in the review process. No date-related restrictions were given in the searching process. Studies focusing on participants were excluded if they were reviews, commentaries, letters to the editor, conference papers, books, book chapters, dissertations, or newspaper articles. This decision was made to ensure a consistent level of methodological rigour and reporting quality across the included studies. In the field of human–robot interaction, conference publications often present preliminary or pilot findings, are subject to stricter length limitations, and frequently lack detailed descriptions of study design, participant characteristics, or analytical procedures. Additionally, conference papers are often later extended into journal articles using overlapping datasets, raising the risk of duplicate inclusion. Restricting the review to journal publications, therefore, reduced heterogeneity and supported a more reliable synthesis of empirical evidence.

Also, the review intentionally applied narrow and conceptually focused inclusion criteria to examine the role of non-verbal communication cues in emotion recognition during human-robot interaction, specifically from the human user's perspective. The aim was to synthesize research that addresses the embodied, interactional, and relational dynamics of non-verbal emotional signalling in HRI contexts. This decision was made to preserve conceptual clarity and ensure that the findings meaningfully address the targeted research question. During the review, no specific theoretical frameworks or modelling techniques were systematically analysed or coded.

2.3. Data extraction and assessment of methodological quality

We performed a multistage screening process to select studies which met the inclusion criteria. The authors independently searched the literature and then reviewed study titles and abstracts. In the next stage, the authors screened the titles and abstracts of all identified records (PM, BS, DP, BŐ, KEK), and fifty per cent of all titles and abstracts were independently assessed by a second review author (BŐ, KEK). All studies whose adequacy was questionable were taken forward to the full-text screening at this stage. Disagreements were resolved through discussion and, where necessary, adjudicated by a third reviewer. In the next step, full-text screening was performed, in which the authors (PM, BS, DP, BŐ, KEK) independently screened all full texts. In cases of uncertainty, the other authors also checked the decision. To assess inter-rater reliability during the full-text inclusion stage, Cohen's kappa coefficient was calculated, yielding a value of $\kappa = 0.78$, indicating substantial agreement. Screening and selection were conducted in Zotero and in a shared spreadsheet using predefined criteria.

For data extraction, an Excel spreadsheet and Data Extraction Forms were applied. We included full article citation, study objectives, study design, how the study attempted to avoid bias, participant characteristics and numbers, the manifestation of Human-Robot Interaction/ Human-Robot Cooperation/Human-Robot Collaboration, the characteristics of communication investigated, the instruments applied during the research, the results/outcome and comments related to study quality.

The risk of bias and the quality of the studies were evaluated by the Joanna Briggs Institute (JBI) critical appraisal tool (randomised controlled trials and non-randomised controlled trials followed by Barker et al. (Barker et al., 2024) and cross-sectional studies followed by Moola et al. (2015). Papers were evaluated according to the appropriate tool on a 4-point scale (yes/no/unclear/not applicable).

While overall quality was moderate to high, several studies lacked transparent reporting on participant recruitment, potential confounding variables, or researcher reflexivity. Also, two studies did not adequately describe how non-verbal cues were operationalised or measured, limiting interpretability. These quality issues should be considered when considering the strength of the evidence and the consistency of the findings (see Table 1) (see Fig. 1).

3. Results

3.1. Characteristics of people involved in studies

Fig. 2 and Table 2 present the most relevant sociodemographic background data of the participants involved in the papers. Given the characteristics and demographic data of the samples involved, vital information needed for the results to be generalisable and comparable is not always provided. Out of the seven published experiments, all had fewer than 100 participants, with Law et al. (2021) having the highest number (N = 84) and Obo and Takizawa (2022) having the lowest number (N = 16). While all included studies reported the number of participants and general age ranges, there was considerable inconsistency in the reporting of demographic characteristics. Only a few studies specified participants' gender or nationality, and none included information on socioeconomic background. This lack of standardized demographic reporting limits the ability to interpret how user

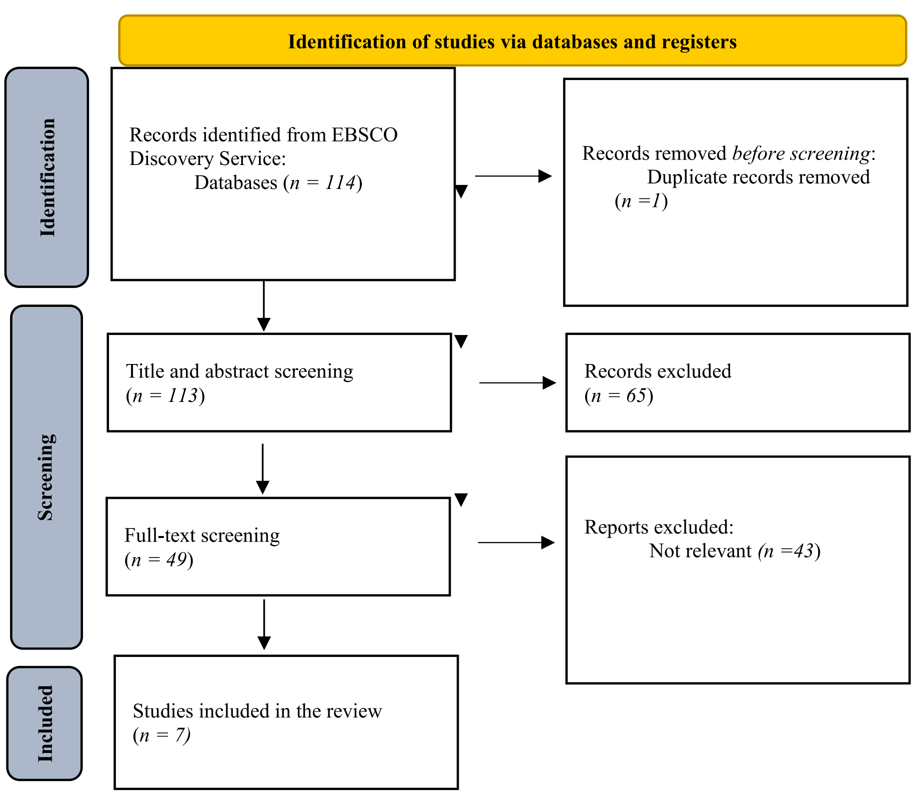

Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram.

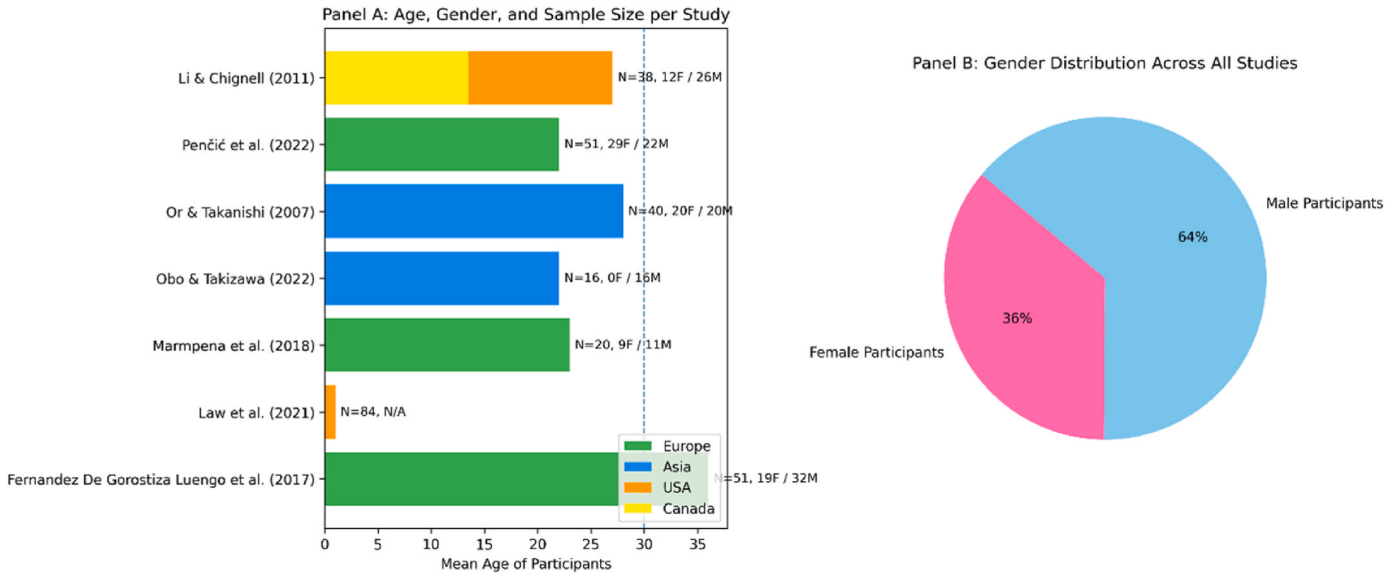

Fig. 2. Sample characteristics of the papers analysed.

Table 1
The evaluation of the risk of bias of the papers based on the JBI tools.

	Law et al. (2021)	Marmpena et al. (2018)	Obo and Takizawa (2022)	J. Li and Chignell (2011)		Fernandez De Gorostiza Luengo et al., 2017	Or and Takanishi (2007)	Penčić et al. (2022)
1. Were the criteria for inclusion in the sample clearly defined?	No	No	No	Yes	**1. Was the sample frame appropriate to address the target population?**	Yes	Yes	Yes
2. Were the study subjects and the setting described in detail?	Yes	Yes	Yes	Yes	**2. Were study participants sampled in an appropriate way?**	N/A	Yes	N/A
3. Was the exposure measured in a valid and reliable way?	Yes	Yes	Yes	Yes	**3. Was the sample size adequate?**	No	No	No
4. Were objective, standard criteria used for measurement of the condition?	Yes	Yes	Yes	Yes	**4. Were the study subjects and the setting described in detail?**	No	Yes	Yes
5. Were confounding factors identified?	No	No	No	Yes	**5. Was the data analysis conducted with sufficient coverage of the identified sample?**	Yed	Yes	Yes
6. Were strategies to deal with confounding factors stated?	No	No	No	Yes	**6. Were valid methods used for the identification of the condition?**	Yes	Yes	Yes
7. Were the outcomes measured in a valid and reliable way?	Yes	Yes	Yes	Yes	**7. Was the condition measured in a standard, reliable way for all participants?**	Yes	Yes	Yes
8. Was appropriate statistical analysis used?	Yes	Yes	Yes	Yes	**8. Was there appropriate statistical analysis?**	Yes	Yes	Yes
				Yes	**9. Was the response rate adequate, and if not, was the low response rate managed appropriately?**	Yes	Yes	Yes

characteristics may have influenced emotional responses to robots or perceptions of non-verbal cues.

Regarding *gender distribution*, Marmpena et al. (2018), Or and Takanishi (2007), Penčić et al. (2022), and Li & Chignell, 2011) reported a relatively equal distribution between males and females, while Fernandez De Gorostiza Luengo et al. (2017) reported twice as many males as females in their sample. Obo and Takizawa (2022) presented an experiment in which only males participated, while Law et al. (2021) did not report gender information about their sample.

Regarding the *age of participants*, Penčić et al. (2022) report the lowest mean, a value of 21 years. Marmpena et al. (2018) report a somewhat higher mean of 23 years, while Fernandez De Gorostiza Luengo et al. (2017) report the highest mean of age (38 for females and 37 for males). Li and Chignell (2011) report the widest age range (18–65 years). Obo and Takizawa (2022) and Or and Takanishi (2007) do not report mean values as they report only ranges, while Law et al. (2021) again fail to present age data about the participants. Therefore, all the research involved young people.

Table 2
Participant demographics in the papers involved in the systematic review and their most important characteristics.

Paper	Country	Sample size	Gender distribution	Age
Fernandez De Gorostiza Luengo et al. (2017)	Spain	51	19 female 32 male	Females: mean = 38, SD = 10.1 % Males = mean = 37, SD = 10.0 %
Law et al. (2021)	USA	84	N/A	N/A
Li & Chignell (2011)	Japan (Studies 1–2) Canada (Studies 3–4)	38	12 female and 26 male Study 1: 2 female/2 male Study 2: 3 female/9 male Study 3: 4 female/6 male Study 4: 3 female/9 male	Overall: 18–65 years Study-specific means: • Study 1: M = 25, SD = 5 • Study 2: M = 26 (SD=N/A) • Study 3: M = 31.5, SD = 11.8 • Study 4: M = 25.1, SD = 2.57
Marmpena et al. (2018)	United Kingdom	20	9 women 11 men	Mean = 23.6 SD = 4.09
Obo and Takizawa (2022)	Japan	16	Male only	20–23
Or and Takanishi (2007)	Japan	40	20 female 20 male	20–34
Penčić et al. (2022)	Serbia	51	29 female 22 male	18–27 Mean = 21.57

Only three of the reviewed papers give precise information concerning nationality or ethnicity. Fernandez De Gorostiza Luengo et al. (2017) reported only Spanish participants, and Or and Takanishi (2007) only Japanese participants, while in the research of Marmpena et al. (2018), 9 participants were British and 11 were from other European countries. Law et al. (2021), Obo and Takizawa (2022) and Penčić et al. (2022) did not provide exact information about the nationality of their participants. Besides, they did not report any information about the participants' socioeconomic status. Further information about the participants in the experiments was scarce. Only Law et al. (2021), Marmpena et al. (2018), Obo and Takizawa (2022) and Penčić et al. (2022) provided further insights, with participants mainly undergraduates and university students.

In conclusion, the composition of the samples is mainly restricted to university students, which makes it hard to generalise the results to broader populations. The fact that the participants' ages were also relatively low, as expected by Fernandez De Gorostiza Luengo et al. (2017), further hinders generalisation.

3.2. *Characteristics of human-robot interaction, collaboration and cooperation*

The type of robot used in the reviewed papers and the type of HRI presented in the experiments are also essential aspects (Fig. 3 and Table 3). First of all, different robots were used in every experiment. Fernandez De Gorostiza Luengo et al. (2017) report that they did not use an actual robot; they presented participants with an artificial sound hypothetically associated with a robot and thus examined the non-verbal sound communication aspect of HRI. Law et al. (2021) report the usage of a Cozmo robot, which presented the visual movement aspect of an HRI combined with some basic speech elements. In Marmpena et al. (2018), a Pepper robot was reported to be used, with non-locomotive body motions (movements of the robot's body that do not involve changing location) and LED lights as the form of communication in the experiment (serving as additional communication channel to enhance or reinforce the emotional content of the gestures, like colour changes and pulsing or blinking patterns). Obo and Takizawa (2022) utilised a Parlo robot and LED lights as the examined form of non-verbal communication. In the research of Or and Takanishi (2007), a specific Wesada Belly Dancer No.2 type of robot was used, but participants did not actually meet with it. Instead, they watched video recordings of the robot presenting body movements as a form of communication. Li and Chignell (2011) similarly relied on video-based presentations of a social robot, examining how simple head and arm movements of a teddy-bear-like robot conveyed emotional content

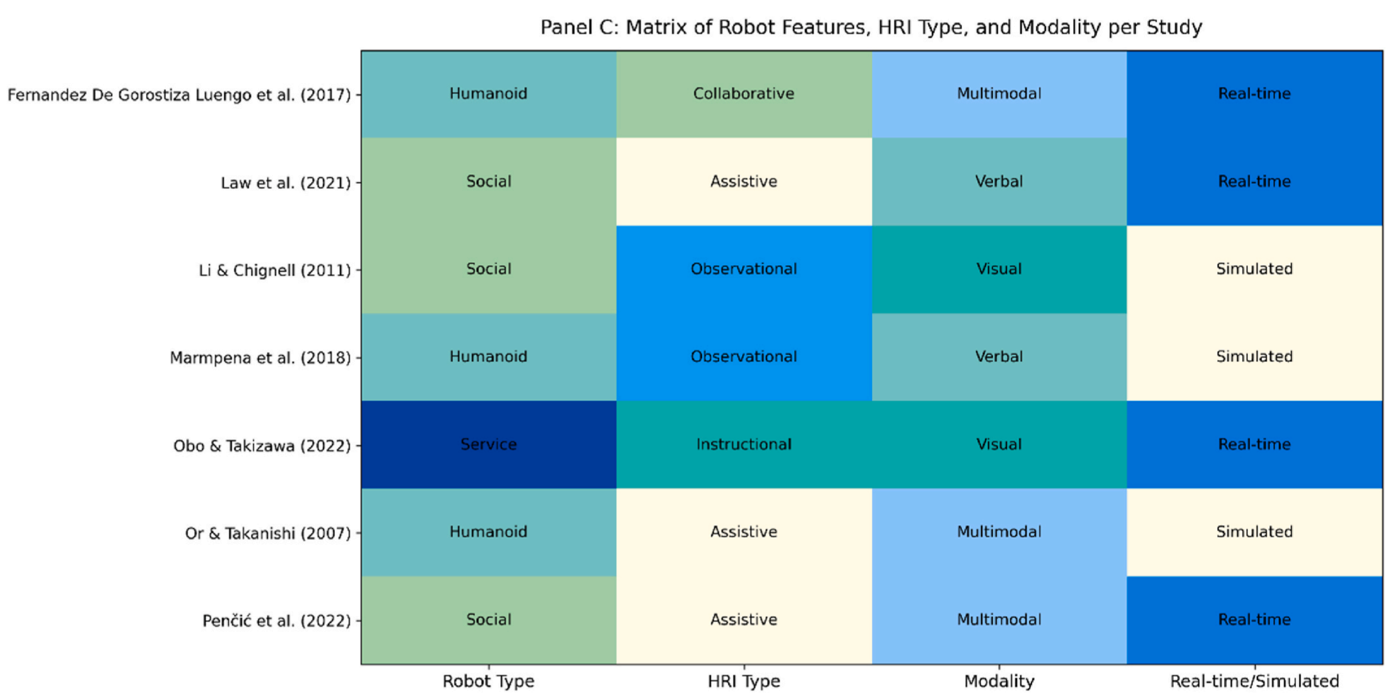

Fig. 3. Summary of robot features, type of HRI and modalities.

Table 3
HRI characteristics in the papers involved in the systematic review and their most important characteristics.

Paper	Robot used in the experiment	Characteristics of HRI	Study design	Applied instruments	Type of communication	Characteristics measured
Fernandez De Gorostiza Luengo et al. (2017)	Only artificial sounds were used	No actual HRI, non-verbal sound communication was simulated	Survey (Online questionnaire)	Generated sound files Experiment specific questionnaire	One-sided, participants received sound stimuli	Expression categories: affirmation, hesitation, denial, question, summon, hush, encouragement, greeting, laughing
Law et al. (2021)	Cozmo	Direct HRI, Cozmo was physically present and used movement and speech	Experiment	Ultimatum game Experiment specific questionnaire about the robot's emotional valence	Two-sided, participants communicated through writing, Cozmo communicated through speech and movement	Movement of the robot Two conditions: positive or negative emotions
Li and Chignell (2011)	RobotPHONE social robot	Indirect interaction, humans create robot gestures by physically manipulating the robot, other humans observe and interpret recorded robot gestures	Experimental, analytical cross-sectional design	Emotion recognition tasks	One-sided The robot expresses emotions via non-verbal gestures (head and arm movements) Human participants do not respond interactively in real time to the robot	Likert scales (1–7) for: Lifelikeness, Likeability, and Confidence in emotion recognition Negative Attitudes toward Robots Scale (NARS) Effect of motion characteristics Effect of author expertise
Marmpena et al. (2018)	Pepper	Direct HRI, Pepper was physically present and used body motion (without locomotion) and LED lights	Experiment (stimuli were presented then the questionnaire answered)	Affect measuring questionnaire	One-sided, participants launched a robot movement, then rated it	Movement of the robot Excited, calm, tired x negative, neutral, positive emotions
Obo and Takizawa (2022)	Parlo	Direct HRI, Parlo was physically present and used speech and LED lights	Experiment	Impression evaluation questionnaire	Two-sided, verbal communication	Visual non-verbal communication (LED lights) Evaluation of the robot's inner (good, friendly, social, altruistic, warm) and outer aspects (kind, cheery, soft, animate, lively, active, prudent, quiet, emotional) and communication (secure, cheerful, stable, altruistic, unfettered, comfortable, pleasant)
Or and Takanishi (2007)	Wesada Belly Dancer No.2	No actual HRI; participants viewed videos of the robot	Survey (Questionnaire - forced-choice)	Experiment specific questionnaire	One-sided, participants viewed videos of a robot	Body movement Expressions: confident, disgust, happy, relieved, patient, angry, sexy
Penčić et al. (2022)	Marko	No actual HRI; participants viewed videos of the robot	Mixed method (Experiment + Questionnaire)	FEI - facial expression identification instrument	One-sided, participants viewed videos of a robot	Facial expressions: anger, happiness, fear, sadness, surprise, disgust

through non-verbal gestures, without direct interaction between participants and the robot. According to Penčić et al. (2022), participants were, again, only presented with video recordings of a robot, in this case, a Marko robot. Similarly to the study of Or and Takanishi (2007), in the research of Penčić et al. (2022), the body movements of the robot were the examined form of communication (see Table 4).

In conclusion, concerning the robots used in the experiments, only three of the reviewed seven papers examined actual HRI, as Or and Takanishi (2007) and Penčić et al. (2022) only presented video recordings to the participants, and Fernandez De Gorostiza Luengo et al. (2017) only sound recordings. Only Law et al. (2021) and Obo and Takizawa (2022) report actual HRI settings in their experiment. In the paper of Marmpena et al. (2018), the robot was also physically present in the experiment; it did not interact with the participant and only presented predefined movements for the participants to evaluate. Similarly, Li and Chignell (2011) relied on simulated HRI, in which participants observed video or animated recordings of a social robot expressing emotions through simple head and arm movements, without any real-time, bidirectional interaction. Overall, in the simulated studies, participants observed pre-recorded robot behaviours and responded based on their perception, rather than interacting in real time. While these studies provide useful insights into human interpretation of robot non-verbal cues, they differ methodologically from studies involving actual embodied interaction, which may engage different cognitive and affective processes. Importantly, findings from Li and Chignell (2011) demonstrate that even robots with minimal expressive morphology can successfully communicate emotional content through carefully designed non-verbal motion. However, recognition accuracy and perceived lifelikeness were strongly influenced by contextual information and gesture complexity. Also, humanoid robots with rich expressive capabilities, such as flexible spines or articulated facial components, consistently outperformed non-anthropomorphic robots in terms of emotional clarity and user interpretation. For example, the Waseda Belly Dancer (WBD-2) conveyed complex affective states more effectively than robots that relied on minimal visual signals. Additionally, studies using integrated multi-modal systems that combine gestures, gaze, and facial motion, tended to report the highest recognition accuracy and user engagement levels.

3.3. Methodology

Regarding methodology, half of the reviewed papers used self-reported questionnaires only (Table 2). In Fernandez De Gorostiza Luengo et al. (2017), Or and Takanishi (2007) and Penčić et al. (2022), participants were presented with the given audio or visual stimuli and then reported their feelings about it in questionnaires constructed explicitly for the given experiment. Law et al. (2021) report an experimental setup in which participants played an Ultimatum game with the robot and then completed a questionnaire assessing the robot's emotional valence. Marmpena et al. (2018) and Obo and Takizawa

Table 4
Main characteristics of the human-robot interactions and the robots applied in the investigations.

Paper	Robot used in the experiment	Robot type	HRI type	Characteristic of the robot
Fernandez De Gorostiza Luengo et al. (2017)	Only artificial sounds were used	non-robotic	simulated	–
Law et al. (2021)	Cozmo	toy	direct	Cosmo is an Educational Toy Robot. It is a miniature robot created by Anki. Cozmo's standard model features a compact design, showcasing a small robot in white and grey with red accents. To replicate human emotions, it employs unique expressions known as the "emotion engine", which can be described as a computational or mechanical system designed to simulate, generate, or interpret emotional expressions in robots (e.g., based on affective computing or behavioral scripts). Subsequent versions were released in red and white, along with another variant in blue. Anki chose to integrate essential machine learning into its character. For instance, it displays curiosity when it requires information about its environment. The development team designed a unique collection of algorithms known as an "emotion engine," enabling Cozmo to effectively replicate human emotions.
Li and Chignell (2011)	RobotPHONE (teddy-bear–like social robot)	social	simulated	Small, non-humanoid, teddy-bear–like robot with limited expressive morphology; emotion conveyed exclusively through simple head and arm movements; no facial expressions, no speech, no sound; gestures presented via video or animation (simulated HRI)
Marmpena et al. (2018)	Pepper	humanoid	direct	Pepper is a humanoid robot developed by SoftBank Robotics that is designed to interact with humans. Pepper is about 120 cm tall and is equipped with various sensors, cameras, and microphones to perceive its environment and communicate effectively. It can recognise human emotions and respond appropriately through its voice, gestures, and screen interface. Pepper's primary applications include customer service, education, and healthcare, where it assists by providing information, facilitating communication, and offering companionship, which provides a sound basis for applying it in sports and sports psychology as well.
Obo and Takizawa (2022)	Parlo	humanoid	direct	PALRO is a compact robot, standing at around 40 cm and weighing roughly 1.8 kg. It utilizes an open architecture with general-purpose components and features functions for communication and autonomous movement, including the ability to recognise human emotions and learn through various methods such as speech recognition, speech synthesis, motion detection, personal recognition, and face recognition. The robot is equipped with a switch that enables users to manage the number of interactions when it detects the presence of individuals nearby. PALRO can conduct lectures on

(continued on next page)

Table 4 (*continued*)

Paper	Robot used in the experiment	Robot type	HRI type	Characteristic of the robot
				physical relaxation exercises and offers quizzes related to professional baseball, prefecture names, the Japanese word game shiritori, horseracing, and sumo wrestling. Additionally, it can dance, sing, or perform internet searches and report the results. It is often used in senior care facilities and homes.
Or and Takanishi (2007)	Wesada Belly Dancer No.2	humanoid	simulated	The robot boasts a total of 28 degrees of freedom: six for each leg, four for each arm, and eight for the spine. In order to enhance the robot's ability to convey emotions through full-body movements, we incorporated an additional two degrees of freedom in the spine. Most notably, we upgraded the lower body of the previous prototype with a more effectively designed version, wherein the motors for roll and pitch axes at the hip and ankle joints have been orthogonalized and aligned at the same levels. This modification reduces the robot's overall height.
Penčić et al. (2022)	Marko	humanoid	simulated	The robot has 4 DOFs eyeballs, 4 DOFs eyelids, and 3 DOFs eyebrows, as well as LEDs for the mouth and ears.

(2022) also report experimental setups, but in both cases, no further methods were used beyond stimuli presented by the robot. Again, participants had to complete a questionnaire about the robot's emotional states after the stimuli were presented. Similarly, Li and Chignell (2011) applied an experimental, laboratory-based design in which participants evaluated pre-recorded robot gestures using a combination of emotion recognition tasks and self-report Likert-scale questionnaires assessing perceived lifelikeness and likeability, without involving behavioural or physiological measures.

3.4. Verbal and nonverbal attributes and patterns

Considering the attributes and patterns of the communication (Table 2), as partially stated in previous chapters, only two of the seven reviewed papers reported actual two-sided interaction, with Obo and Takizawa (2022) utilising verbal communication combined with non-verbal LED lights and Law et al. (2021) utilising verbal communication combined with non-verbal body movement. Fernandez De Gorostiza Luengo et al. (2017) examined only non-verbal auditory communication, focusing on communicating the following emotional states: affirmation, hesitation, denial, question, summon, hush, encouragement, greeting and laughing. Marmpena et al. (2018) focused on the robot communicating three emotional states (excited, calm, tired) in three different styles (negative, neutral, positive). Obo and Takizawa (2022) report using LED lights to convey additional information by the robot during a verbal interaction, where participants were left free to interpret these. In Or and Takanishi (2007), participants were presented with the full-body movements of a robot representing the following states: confident, disgust, happy, relieved, patient, angry and sexy (i.e. mimicking expressive, rhythmic human dance, particularly belly dancing). In Penčić et al. (2022), the robot's facial expressions were in focus, including anger, happiness, fear, sadness, surprise and disgust. Similarly, Li and Chignell (2011) examined one-sided, non-verbal communication, in which participants observed a social robot expressing emotions exclusively through simple head and arm gestures, without speech or facial expressions, and were asked to interpret the conveyed emotional states. The study focused primarily on basic emotions (e.g. happiness, sadness, anger, fear), while also highlighting the role of contextual information in disambiguating gesture-based emotional cues.

In conclusion, most experiments focused on core emotions such as anger, happiness, and disgust. In some cases, specific emotional states or intentions were also examined, such as surprise, hesitation, excitement, or presenting a sexy body movement (referring to fluid, rhythmic full-body movements inspired by dance that are perceived as emotionally expressive or suggestive).

Overall, the cross-study comparison of recognition rates reveals notable variation across different modalities used in non-verbal emotion expression. For instance, gesture-based expression achieved relatively high recognition accuracy, ranging from 83 % to 90 % in studies using full-body humanoid robots (e.g. Or & Takanishi, 2007). Li and Chignell (2011) further demonstrate that even robots with limited expressive morphology can reach above-chance and, in some conditions, high recognition accuracy when gestures are sufficiently distinctive, although performance was strongly influenced by gesture complexity and the availability of contextual cues. Similarly, facial expression-based cues, particularly those involving anthropomorphic robotic eyes, were associated with recognition rates of above 80 % for emotions such as surprise and happiness. However, accuracy declined for more subtle affective states like fear or disgust (Penčić et al., 2022). In contrast, simpler modalities such as LED light signals demonstrated lower recognition rates, approximately 60–70 %, depending on the timing and context of the signal, as observed in Obo and Takizawa (2022) which studied turn-taking behaviour using LED cues.

3.5. Main results found

According to the results reported by the reviewed papers, most emotions and emotional states may be easily interpreted by participants. In the reports of Fernandez De Gorostiza Luengo et al. (2017), which used auditory stimuli, only encouragement and greeting were more often confused with other states. Denial was the easiest to recognise, and so were question, summon, hush, and laughing. The authors also stated that the level of intention favoured higher recognition rates for these states. Both Law et al. (2021) and Or and Takanishi (2007) report that participants accurately recognised the body movements of the robot. Law et al. (2021) confirmed that humans tend to react to non-verbal cues from a robot in the same manner that they would react to similar non-verbal cues from humans, and Or and Takanishi (2007) stated that the movements of the robot dancer were often as recognisable as the movements of the human dancer. Therefore, body movements provided enough information for participants to recognise feelings independently of the robot's type. Similarly, Li and Chignell (2011) demonstrated that emotions expressed through simple robot head and arm gestures could be recognised by participants with accuracy significantly above chance level, despite the robot's limited expressive morphology. Their results further showed that recognition performance varied across emotional categories and was strongly influenced by gesture distinctiveness and the availability of contextual information, with context substantially improving correct interpretation. Marmpena et al. (2018) conclude that valence is harder to guess than arousal regarding robot emotional states. Arousal is more easily perceived based on body expressions than valence, whereas facial expressions represent a more stable visual modality for valence recognition. Also, low confidence usually leads to evaluating the expression as neutral. Obo and Takizawa (2022) concluded that their proposed LED light communication during verbal interaction was most effective when the light stimuli were set to 2 s in duration. Under this condition, the number of nodding and giving responses increased significantly, and participants obtained high evaluation scores for many items in the impression evaluation. The assessment of the robot's communication revealed that individuals had a more favourable impression when there was a time gap before the display of the "diamond pattern," specifically in terms of stability, comfort, and pleasantness. Consequently, the effectiveness of the interaction between humans and the robot may be influenced by the length of time between changes in the LED display pattern. Penčić et al. (2022) examined whether the robot eyes and eyebrows were capable of effectively expressing emotions to human subjects. The authors report that participants correctly perceived anger and sadness, but surprise was only partially correct. In these cases, the level of intensity was also correctly identified. Disgust, happiness, and fear were all poorly identified.

In conclusion, it is interesting to note that although pure non-vocal auditory communication seems to be the least examined aspect of non-verbal Fernandez De Gorostiza Luengo et al. (2017), who utilised this, reported successful interpretation of most of their presented emotional states. Findings from Li and Chignell (2011) further indicate that even minimal, gesture-based non-verbal communication can support reliable emotion recognition, particularly for basic emotions, although performance decreases for more subtle affective distinctions and benefits greatly from contextual support. Combining the results reported by Law et al. (2021), Or and Takanishi (2007) and Penčić et al. (2022), the body expressions, both in terms of movement and facial expressions, can be fairly accurately interpreted, especially in with regards to positive and negative states, but as more specific emotions are presented, people tend to mistake them for other states.

4. Discussion

In this paper, we aimed to review all empirical studies examining how non-verbal communication and emotional expression affect the quality of human-robot interaction and draw attention to this specific field of research. The systematic literature review showed that only a few studies focus on the robot's emotional expression through nonverbal communication during human-robot interaction. The systematic analysis of the manifestation of emotions in human-robot interaction (HRI) has yielded significant insights into how different forms of communication, including verbal and non-verbal, can be effectively interpreted by humans. The studies reviewed highlight several key findings and implications for developing more effective HRI.

One of the most relevant findings is that humans can recognise emotional states through nonverbal cues. Studies confirm that the body movements of robots are often as recognisable as those of humans, suggesting that nonverbal communication plays a crucial role in HRI (Admoni et al., 2016; Saunderson & Nejat, 2019). This finding emphasises the importance of designing robots with expressive body language to facilitate better human-robot interactions.

However, despite the overall positive results, some emotional states are more challenging to identify. Some emotions are more obvious and relatively easy to detect, while others proved challenging to identify accurately. Human robots are intuitively attractive to humans, establish an emotional bond with their human counterparts and have the ability to adapt and learn from their human partners, utilising their input to enhance their interactions (Kolling et al., 2016; Spezialetti et al., 2020; Szabóová et al., 2020). However, this discrepancy highlights a crucial area where robot design needs to advance, underscoring the need for more complex, nuanced expressions that can effectively represent a broader spectrum of emotions.

Some participants' personality characteristics may also have had a significant impact on emotion recognition, as evidenced by differences in confidence and valence. Results indicate that, in addition to being able to convey emotions with clarity, robots should also be built to reinforce these expressions in a way that increases user confidence in their interpretations. Emotional intelligence has a significant impact on how people perceive and comprehend emotions. People with higher levels of empathy and self-awareness can identify better subtle emotional signs in themselves and others (Drigas & Papoutsi, 2018; Filice & Weese, 2024).

Reflecting on our first research question, among the reviewed findings, the most consistently supported result is the recognition of basic emotions, particularly anger, happiness, and sadness, through body language and non-verbal cues such as movement or posture. Participants were generally able to correctly identify these emotional states, even when expressed by robots with limited anthropomorphic features (e.g. Law et al., 2021; Li & Chignell, 2011; Or & Takanishi, 2007; Penčić et al., 2022). This indicates strong empirical support for the centrality of body language in emotion recognition during HRI. In contrast, concepts such as robotic politeness (i.e. social appropriateness, norms, or user comfort), intentionality (the robot's apparent capacity to act with purpose or goals), and timing of LED patterns (e.g., Obo & Takizawa, 2022), while promising and theoretically important, are less frequently studied and supported by fewer data points. Li and Chignell (2011) additionally showed that basic emotions can be effectively conveyed through simple, low-complexity robot gestures, especially when contextual information is available. These areas should be considered emerging or exploratory themes warranting further empirical attention. Making this distinction helps readers see which findings are most useful for designing better robots, like focusing on clear, expressive gestures. It also highlights where there's still room to explore, such as understanding politeness or intentionality in robot behaviour. The duration and timing of non-verbal cues also significantly impact their effectiveness. An immediate response may not be the best solution, because giving time to perceive and understand non-verbal cues could lead to a much more lifelike experience, with greater stability, comfort, and pleasantness. This aligns with the key statements of Organic Human-Robot Interaction (O-HRI) that blurs boundaries in how social robots should be designed by incorporating psychological features into their behaviour (Őrsi et al., 2024). This approach integrates imperfections, mistakes, and intentionality into the robot behaviour, recognizing their importance in fostering trust, engagement, and sociability According to O-HRI, such errors in decision-making or during task execution can portray robots as lifelike and natural social actors, thereby making them more easily accepted by society. Intentionality is also crucial; for example, by clearly communicating behavioural intentions, robots can cue humans about the actions they will carry out, making the interaction more fluid and predictable. Finally, the emphasis on politeness is coupled with the idea that robots should perform a variety of tasks that require social skills grounded in culture and individuals, such as prompting adequate eye contact and appropriate responses to social signals. In conclusion, the O-HRI framework seeks to fulfil the objective of building trust and engagement with robots as their actions come closer to human expectations, producing human-robot interactions that feel organic. This framework acknowledges the importance of embedding social and emotional features into robotics, enabling robots to act as real social partners (Őrsi et al., 2024).

Furthermore, improving the quality of non-verbal cues and emotional expression in human-robot interaction could be a key factor for Industry 5.0, which also aims to put human creativity and well-being at the heart of the industry (Mohamed Hashim et al., 2024). Thus, better perception and understanding on the non-verbal level could also foster creativity during cooperative work and improve how social robots work in supporting mental health and well-being (Ghobakhloo et al., 2024; Trivedi et al., 2024).

While this review primarily focused on how human users interpret robot-expressed non-verbal cues, it is equally important to reflect on how robots interpret human non-verbal signals, and how these processes align or diverge. In HRI, emotion recognition is often asymmetric: humans rely on embodied intuition and social cognition to decode robotic behaviour, while robots depend on programmed algorithms and sensor data to interpret human gestures or facial expressions (Jirak et al., 2022; Spezialetti et al., 2020). This asymmetry has design implications: whereas humans may fill in gaps using empathy and context, robots typically require explicit input channels and training data to perform similar tasks (Park & Whang, 2022; Schömbs et al., 2023). For example, affective computing enables robots to detect human emotions using multimodal inputs such as facial expressions and vocal tone (Picard, 1997). Still, these interpretations often lack the contextual sensitivity that humans bring to emotional understanding. In contrast, humans can interpret robot intention even from limited cues like gaze direction or timing delays (Obo & Takizawa, 2022), suggesting that human perception compensates for robotic constraints. A more reciprocal model of interaction would recognise both interpretative processes as complementary but fundamentally different: humans rely on context and intuition, while robots rely on predefined parameters and probabilistic models. Future research should explore how these differences affect mutual understanding and miscommunication in dynamic settings (Zonca et al., 2021). This comparative approach could contribute to designing systems that not only express but also interpret affective signals in ways that feel natural and intuitive to both parties.

Reflecting on our second research question, we could observe that most studies employed experimental designs with small, convenience samples, often composed of university students with limited diversity in age, gender, or socioeconomic status. Across the research, several methodological limitations were identified, including a small sample size and a need for more diverse demographic representation. The results may have been skewed by the fact that the majority of participants were young adults, who are typically more familiar with modern technologies. Also, the low number of participants may hinder the efficacy of the experiments. Involving young people in research may be easier because younger people are usually more familiar with digital technologies, artificial intelligence, and robots. However, it would be necessary to compare the attitudes, reactions and general and content-specific characteristics of older people as well, not only relying on young adults. It would be crucial to involve people with various sociodemographic and academic backgrounds albeit a key limitation across the included studies is the inconsistent and often absent reporting of participant demographics. Details such as gender, cultural background, and socioeconomic status, which can significantly shape emotional expression and perception, were rarely disclosed. This limits the generalisability of the findings and raises concerns about cultural and contextual biases in HRI research. For example, interpretations of non-verbal cues may vary widely across cultures, but such differences could not be examined due to missing data. We have to emphasize that in many cases, students studying IT and engineering participated in the studies; however, students studying in other fields (having less specific knowledge on this topic) as well as other age groups and different generations should also be represented in such research and experiments. Measuring the socioeconomic status of the participants may also allow the researcher to conduct comparative investigations following the social background, since some previous research has proved the impact of sociodemographic characteristics and socioeconomic background on attitudes toward robots. For instance, the Special Eurobarometer (European Commission, 2022) survey findings indicate that individuals' attitudes towards robots are influenced by gender, age, education level, and socioeconomic status (SES). In general, women, older individuals, those with lower education, and individuals facing financial strain appear to be less inclined towards embracing robots. To guarantee that the results are generalisable, future research should strive to include a broader range of ages and socioeconomic backgrounds.

As reflecting on the third research question, we should state that a notable limitation of the research is that three of the seven included studies applied simulated HRI, relying on video or audio playback rather than real-time, embodied interaction. While such methods are common in early-stage HRI research, offering controlled environments for investigating user perceptions, they do not fully reflect the reciprocal, dynamic, and context-sensitive nature of actual human-robot interactions. This distinction is important, as the construct validity of findings may be affected when user responses to non-verbal cues are assessed outside of interactive contexts. Simulated settings may elicit different emotional or cognitive responses compared to direct engagement with a robot. Therefore, conclusions drawn from these studies should be interpreted with caution, particularly when considering their applicability to real-world HRI scenarios. One should also mention that only a few studies examined how individual user characteristics, such as empathy, confidence, cultural background, or familiarity with technology, influenced the interpretation of robot non-verbal communication. Most studies reported aggregated group-level findings without exploring interpersonal variability. This represents a significant gap in the literature, as previous research suggests that traits like social sensitivity or prior exposure to robots can meaningfully shape emotional interpretation and interaction dynamics (Daruwala, 2025; Faria et al., 2025; Salma et al., 2025; Yang & Xie, 2024). Future research should consider integrating user profiling or psychometric assessments to better understand how personal factors mediate responses to non-verbal cues in HRI.

We should also mention that although our review identifies several important research streams, it also highlights that historically influential modalities such as gaze behaviour, facial expressions, proxemic patterns, and multimodal cue integration are inconsistently represented in current empirical work. Despite decades of foundational research, these areas remain under-explored in modern robot platforms, especially in studies involving real-time, embodied interaction. Incorporating a stronger historical perspective helps situate current findings within a broader developmental trajectory and clarifies why certain modalities (e.g., gaze or proxemics) need renewed empirical attention. Emerging evidence suggests that isolated non-verbal channels may be insufficient to convey complex affective states reliably, whereas coordinated combinations of gaze direction, facial animation, body movement, spatial positioning, and timing may support more robust and interpretable emotional communication (Bečková et al., 2025; Stedtler et al., 2025). As robotic platforms continue to evolve, future research should therefore prioritise multimodal integration and examine how these cues dynamically converge in real-time interaction. Such an approach may enable next-generation social robots to communicate emotions in a manner that is not only more expressive but also more context-sensitive, socially intelligible, and aligned with human expectations.

Additionally, methodological variability across the included studies may have weakened the overall strength of the evidence. While most studies demonstrated sound design and ethical rigour, key limitations were observed, including inconsistent operationalisation of non-verbal cues and limited sample diversity, which could affect the generalisability of the findings and suggest that future research should adopt more rigorous, transparent designs. Furthermore, the absence of standardised measures for emotion recognition in HRI complicates cross-study comparisons. Thus, while the reviewed studies collectively highlight the relevance of non-verbal communication, the conclusions should be interpreted with caution due to these methodological constraints. Furthermore, the research frequently lacked thorough explanations of the control measures and experimental settings, which are crucial for reliability and validity. Researchers should put greater emphasis on introducing the participants in detail and transparently presenting test conditions, demonstrating the control over the most important confounding factors.

5. Limitations

Although this review offered a detailed comparison between the studies included in the literature, the research is not without limitations. First, although the initial systematic search followed PRISMA guidelines, it relied primarily on a single aggregated search engine (EBSCO Discovery Service). While this platform provides access to a wide range of indexed databases, its coverage is not exhaustive. To mitigate this limitation, we conducted supplementary validation searches in OpenAlex using the same refined Boolean logic (including field tags, truncation, and proximity operators). These additional searches did not identify further eligible studies, yet variations in indexing practices across platforms mean that the possibility of missed publications cannot be entirely excluded. Regarding additional sampling procedures, forward snowballing was not applied in the present review. Although forward citation tracking can be useful for identifying emerging research, it often reflects visibility, disciplinary trends, or publication recency rather than the methodological quality of studies. In fast-moving and highly interdisciplinary fields such as human–robot interaction, this approach may place disproportionate emphasis on newer or highly cited papers that have not yet been replicated or thoroughly validated. Consequently, forward snowballing may introduce temporal or citation-based bias and reduce conceptual consistency across included studies. To minimise these risks, we prioritised a structured, database-driven search strategy complemented by backward reference screening in order to ensure methodological reliability and coherence of the reviewed evidence.

Although only seven studies were ultimately included from the initial pool, this low number was expected due to our deliberately narrow inclusion criteria. We chose to focus on a clearly defined conceptual area rather than take a broad approach that might include loosely related topics. Our goal was to provide depth and clarity in exploring a niche but under-researched aspect of human-robot interaction. That said, we recognise that this focused strategy may have excluded broader or influential studies that address similar issues but use different terminology or frameworks. Only experiments and pilot studies could have been detected, which did not allow us to investigate the changing nature of emotion recognition. Moreover, only English papers were included, which may also hinder the detection of relevant papers. Also, the review excluded conference papers and other non–peer-reviewed sources. Although this decision ensured a consistent level of methodological quality, it may have limited the breadth of the evidence base. While we did gather extensive data on factors such as trial quality, participants, and interventions, some unaccounted-for heterogeneity remained across the trials. Given the small number of included studies and their substantial methodological heterogeneity, no statistical pooling (e.g., meta-analysis) was feasible. Therefore, a narrative synthesis approach was adopted, consistent with guidance for reviews with limited and heterogeneous evidence. Taken together, these limitations suggest that the results should be interpreted with caution, and future research would benefit from broader database coverage, inclusion of grey literature, and more standardised methodological approaches.

6. Practical implications

Based on the findings of this review, several actionable recommendations can be summarised for researchers and designers aiming to improve emotional communication in human-robot interaction. First, designers should consider the integration of expressive non-verbal modalities, such as gestures, posture, and facial animation, particularly in humanoid robots, to enhance emotional clarity and user engagement. Future studies should also focus on the topics detected as underrepresented themes to better understand the nature of emotion recognition in HRI. While the current review focused on empirical studies of non-verbal emotional cues in HRI, the observed findings offer interpretive touchpoints for broader, emerging frameworks such as Organic Human-Robot Interaction (O-HRI, Őrsi et al., 2024) and Industry 5.0. For instance, the recognition that timing and duration of LED cues influence comfort and perceived naturalness (Obo & Takizawa, 2022) aligns with O-HRI's emphasis on integrating socially situated imperfections and intentional delays to foster lifelike engagement. Similarly, findings that user confidence and emotional valence shape the perception of robot expressions (e.g. Law et al., 2021; Marmpena et al., 2018) align with Industry 5.0's prioritisation of well-being and trust in human-machine collaboration. It is important to note, however, that these frameworks were not directly assessed in the included studies. The connections proposed here are therefore interpretive extensions rather than evidence-based conclusions, and they highlight promising directions for future research to empirically examine how such paradigms might inform the design of emotionally attuned, socially integrated robotic systems.

Furthermore, longitudinal investigations should be carried out to better understand how this aspect could affect the quality of human-robot interaction and collaborative work in the work environment, and how the quality of understanding non-verbal cues could develop over time and with practice. Researchers are encouraged to expand empirical studies beyond university student samples by including participants from diverse age groups, cultural backgrounds, and socioeconomic statuses, enabling more generalisable insights. Future studies should also move beyond simulated interactions to real-time, bidirectional HRI settings that better grab the dynamic and reciprocal nature of emotional communication. Moreover, the influence of user traits such as empathy, confidence, or familiarity with technology should be systematically examined using psychometric profiling.

CRediT authorship contribution statement

Balázs Őrsi: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Péter Korondi:** Writing – review & editing, Writing – original draft, Funding acquisition, Conceptualization. **Csilla Csukonyi:** Writing – review & editing. **Dávid Papp:** Writing – original draft, Funding acquisition, Conceptualization. **Panna Márkus:** Writing – original draft, Conceptualization. **Bence Tamás Selejó Joó:** Writing – original draft, Conceptualization. **Husam A. Neamah:** Writing – review & editing, Visualization, Supervision, Methodology, Conceptualization. **Karolina Eszter Kovács:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization.

Data availability statement

Not applicable.

Declaration of the use of AI

The authors did not use generative AI and AI-assisted technologies in the writing process.

Funding statement

This research was funded by the HUNGARIAN RESEARCH FUND grant number OTKA K143595 and the EKÖP-24 University Research Scholarship Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Peter Korondi reports financial support was provided by National Research Development and Innovation Office. David Papp reports financial support was provided by National Research Development and Innovation Office. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Supported by the University of Debrecen Program for Scientific Publication.

Appendix

Table A1
List of databases searched by the EBSCO Discovery search engine

Databases searched by the EBSCO Discovery search engine
Accucoms—COVID-19 resources,
ACM Digital Library
Arts & Humanities (Proquest)
Bibliotheca Corviniana Digitalis
Biological Abstracts 2000–2004
Biomedical & Life Sciences Collection
BMJ Journals
Business Source Premier
CAB Abstracts
Cambridge Journals
ChemSpider
CNKI
Cochrane
COMPASS
Congress.gov
De Gruyter Journals
Directory of Open Access Journals (DOAJ)
Ebook (Springer)
Ebook Collection (Ebsco)
EbookCentral (Proquest)
EBSCOHost
Elsevier
Elsevier—SciVal
ELTE Reader
Emerald
EMIS University— Central and Southeast Europe
EndNote
ERIC
European Parliament Legislative Observatory
EUR-Lex
Europeana Collections
EUROSTAT
FSTA (Food Science and Technology Abstracts)
GALE Literary Sources (GLS)
Gale Reference Complete
Global Health and Human Rights Database
Grove Music Online
HUMANUS
HUNGARICANA
IJOTEN
Impact Factor (Journal Citation Reports)
InCites
International Human Rights Network
Internet Archive
JSTOR
MATARKA
MathSciNet
MathSciNet (EBSCOhost)
MEDLINE (EBSCOhost)
MEDLINE (PubMed)
Medscape
Nature Journals
NEJM Group—COVID-19 resources
Nutrition and Food Sciences
Oxford Handbooks Online (OHO)—Criminology and Criminal Justice
Oxford Handbooks Online (OHO)—Law
Oxford Scholarship Online (Law Collection)

(*continued on next page*)

Table A1 (*continued*)

Databases searched by the EBSCO Discovery search engine
Oxford University Press (OUP) Journals
Project Gutenberg
ProQuest—One Academic
PubMed
PubMob
RefWorks
SAGE Journals
Science Direct
Science Magazine
SciFinder
Scifinder-n
SCImago Journal and Country Rank (SJR)
SciTech (Proquest)
SCOPUS
SHERPA/RoMEO
SpringerLink
STADAT
Statista
SzocioWeb
Taylor and Francis Online Library
The Historical Map Portal
United Nations Treaty Collection
UpToDate Advanced
Web of Science
Wiley Online Library
World Biographical Information System
zbMATH
International Directory of Music Resources

References

Abdulghafor, R., Abdelmohsen, A., Turaev, S., Ali, M. A. H., & Wani, S. (2022). An analysis of body language of patients using artificial intelligence. *Healthcare (Basel, Switzerland), 10*(12), 2504. https://doi.org/10.3390/healthcare10122504

Admoni, H., Weng, T., Hayes, B., & Scassellati, B. (2016). Robot nonverbal behavior improves task performance in difficult collaborations. In *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)* (pp. 51–58). https://doi.org/10.1109/HRI.2016.7451733

Ajibo, C. A., Ishi, C. T., & Ishiguro, H. (2025). Negative expressions by social robots and their effects on persuasive behaviors. *Electronics, 14*(13), Article 13. https://doi.org/10.3390/electronics14132667

Al Moubayed, S., Skantze, G., & Beskow, J. (2013). The furhat back-projected humanoid head-lip reading, gaze and multi-party interaction. *International Journal of Humanoid Robotics, 10*(1), Article 1350005.

Argyle, M. (1988). *Bodily communication* (2nd ed.). Routledge. https://doi.org/10.4324/9780203753835

Baddoura, R., & Venture, G. (2015). This robot is sociable: Close-up on the gestures and measured motion of a human responding to a proactive robot. *International Journal of Social Robotics, 7*(4), 489–496. https://doi.org/10.1007/s12369-015-0279-x

Barker, T. H., Habibi, N., Aromataris, E., Stone, J. C., Leonardi-Bee, J., Sears, K., Hasanoff, S., Klugar, M., Tufanaru, C., Moola, S., & Munn, Z. (2024). The revised JBI critical appraisal tool for the assessment of risk of bias for quasi-experimental studies. *JBI Evidence Synthesis, 22*(3), 378–388. https://doi.org/10.11124/JBIES-23-00268

Bečková, I., Pócoš, Š., Belgiovine, G., Matarese, M., Eldardeer, O., Sciutti, A., & Mazzola, C. (2025). A multi-modal explainability approach for human-aware robots in multi-party conversation. In *Computer vision and image understanding, 253*, Article 104304. https://doi.org/10.1016/j.cviu.2025.104304

Belhassein, K., Fernández-Castro, V., Mayima, A., Clodic, A., Pacherie, E., Guidetti, M., Alami, R., & Cochet, H. (2022). Addressing joint action challenges in HRI: Insights from psychology and philosophy. *Acta Psychologica, 222*, Article 103476. https://doi.org/10.1016/j.actpsy.2021.103476

Bethel, C.l., & Murphy, R.r. (2008). Survey of non-facial/non-verbal affective expressions for appearance-constrained robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, 38*(1), 83–92. https://doi.org/10.1109/TSMCC.2007.905845

Bohus, D., & Horvitz, E. (2009). Dialog in the open world: Platform and applications. In *Proceedings of the 2009 international conference on multimodal interfaces* (pp. 31–38). https://doi.org/10.1145/1647314.1647323

Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems, 42*(3–4), 167–175. https://doi.org/10.1016/S0921-8890(02)00373-1

Breazeal, C. (2004). Social interactions in HRI: The robot view. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 34*(2), 181–186. https://doi.org/10.1109/TSMCC.2004.826268. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews).

Chevalier, P., Kompatsiari, K., Ciardo, F., & Wykowska, A. (2020). Examining joint attention with the use of humanoid robots-A new approach to study fundamental mechanisms of social cognition. *Psychonomic Bulletin & Review, 27*(2), 217–236. https://doi.org/10.3758/s13423-019-01689-4

Chi, R., Zhang, J., & Pan, M. (2024). The effect of anthropomorphic competence-warmth congruence of service robots on recommendation intention. *Current Psychology, 43*(7), 6570–6583. https://doi.org/10.1007/s12144-023-04825-5

Corrales-Paredes, A., Sanz, D. O., Terrón-López, M.-J., & Egido-García, V. (2023). User experience design for social robots: A case study in integrating embodiment. *Sensors, 23*(11), Article 11. https://doi.org/10.3390/s23115274

Damiano, L., & Dumouchel, P. (2018). Anthropomorphism in human–robot Co-evolution. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.00468

Daruwala, N. A. (2025). How multimodal narrative and visual representations of human-like service robots shape attitudes and social connection. *Frontiers in Robotics and AI, 12*. https://doi.org/10.3389/frobt.2025.1568146

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 679–704. https://doi.org/10.1098/rstb.2006.2004

de Kervenoael, R., Hasan, R., Schwob, A., & Goh, E. (2020). Leveraging human-robot interaction in hospitality services: Incorporating the role of perceived value, empathy, and information sharing into visitors' intentions to use social robots. *Tourism Management, 78*, Article 104042. https://doi.org/10.1016/j.tourman.2019.104042

Depaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In*The handbook of social psychology* (4th ed., Vols. 1–2, pp. 3–40). McGraw-Hill.

Dourish, P. (2001). *Where the action is: The foundations of embodied interaction*. The MIT Press. https://doi.org/10.7551/mitpress/7221.001.0001

Drigas, A. S., & Papoutsi, C. (2018). A new layered model on emotional intelligence. *Behavioral Sciences, 8*(5), 45. https://doi.org/10.3390/bs8050045

El-Raheb, K., Kougioumtzian, L., Kalampratsidou, V., Theodoropoulos, A., Kyriakoulakos, P., & Vosinakis, S. (2025). Sensing the inside out: An embodied perspective on digital animation through motion capture and wearables. *Sensors, 25*(7), Article 7. https://doi.org/10.3390/s25072314

Ende, T., Haddadin, S., Parusel, S., Wüsthoff, T., Hassenzahl, M., & Albu-Schäffer, A. (2011). A human-centered approach to robot gesture based communication within collaborative working processes. In *2011 IEEE/RSJ international conference on intelligent robots and systems* (pp. 3367–3374). https://doi.org/10.1109/IROS.2011.6094592

European Commission. (2022). Special eurobarometer 525—Sport and physical activity. *Európai Unió*.

Faria, D. R., Godkin, A. L., & da Silva Ayrosa, P. P. (2025). Advancing emotionally aware child–robot interaction with biophysical data and insight-driven affective computing. *Sensors, 25*(4), Article 4. https://doi.org/10.3390/s25041161

Fernandez De Gorostiza Luengo, J., Alonso Martin, F., Castro-Gonzalez, A., & Salichs, M. a. (2017). Sound synthesis for communicating nonverbal expressive cues. *IEEE Access, 5*, 1941–1957. https://doi.org/10.1109/ACCESS.2017.2658726

Filice, L., & Weese, W. J. (2024). Developing emotional intelligence. *Encyclopedia, 4*(1), Article 1. https://doi.org/10.3390/encyclopedia4010037

Fiore, S. M., Wiltshire, T. J., Lobato, E. J. C., Jentsch, F. G., Huang, W. H., & Axelrod, B. (2013). Toward understanding social cues and signals in human–robot interaction: Effects of robot gaze and proxemic behavior. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00859

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems, 42*(3), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego: Academic Press (with Internet Archive) http://archive.org/details/humanfacialexpre0000frid.

Ghobakhloo, M., Iranmanesh, M., Fathi, M., Rejeb, A., Foroughi, B., & Nikbin, D. (2024). Beyond Industry 4.0: A systematic review of Industry 5.0 technologies and implications for social, environmental and economic sustainability. *Asia-Pacific Journal of Business Administration*. https://doi.org/10.1108/APJBA-08-2023-0384. *ahead-of-print*(ahead-of-print).

Haddadi, A., Croft, E. A., Gleeson, B. T., MacLean, K., & Alcazar, J. (2013). Analysis of task-based gestures in human-robot interaction. In *2013 IEEE international conference on robotics and automation* (pp. 2146–2152). https://doi.org/10.1109/ICRA.2013.6630865

Jirak, D., Aoki, M., Yanagi, T., Takamatsu, A., Bouet, S., Yamamura, T., Sandini, G., & Rea, F. (2022). Is it me or the robot? A critical evaluation of human affective state recognition in a cognitive task. *Frontiers in Neurorobotics, 16*. https://doi.org/10.3389/fnbot.2022.882483

Kolling, T., Baisch, S., Schall, A., Selic, S., Rühl, S., Kim, Z., Rossberg, H., Klein, B., Pantel, J., Oswald, F., & Knopf, M. (2016). Chapter 5—What is emotional about emotional robotics? In S. Y. Tettegah, & Y. E. Garcia (Eds.), *Emotions, technology, and health* (pp. 85–103). Academic Press. https://doi.org/10.1016/B978-0-12-801737-1.00005-6.

Kolomaznik, M., Petrik, V., Slama, M., & Jurik, V. (2024). The role of socio-emotional attributes in enhancing human-AI collaboration. *Frontiers in Psychology, 15*. https://doi.org/10.3389/fpsyg.2024.1369957

Kopp, S., & Wachsmuth, I. (2010). Gesture in embodied communication and human-computer interaction. *Lecture Notes in Computer Science, 5934*. https://doi.org/10.1007/978-3-642-12553-9

Law, T., de Leeuw, J., & Long, J.h. (2021). How movements of a non-humanoid robot affect emotional perceptions and trust. *International Journal of Social Robotics, 13*(8), 1967–1978. https://doi.org/10.1007/s12369-020-00711-3

Li, J., & Chignell, M. (2011). Communication of emotion in social robots through simple head and arm movements. *International Journal of Social Robotics, 3*(2), 125–142. https://doi.org/10.1007/s12369-010-0071-x

Li, L., Xu, Q., Wang, G.s., Yu, X., Tan, Y.k., & Li, H. (2015). Visual perception based engagement awareness for multiparty human-robot interaction. *International Journal of Humanoid Robotics, 12*(4). https://doi.org/10.1142/S021984361550019X

Marmpena, M., Dahl, T.s., & Lim, A. (2018). How does the robot feel? Perception of valence and arousal in emotional body language. *Paladyn, 9*(1), 168–182. https://doi.org/10.1515/pjbr-2018-0012

McColl, D., & Nejat, G. (2014). Recognizing emotional body language displayed by a human-like social robot. *International Journal of Social Robotics, 6*(2), 261–280. https://doi.org/10.1007/s12369-013-0226-7

Mehrabian, A. (1972). *Nonverbal communication*. Routledge. https://doi.org/10.4324/9781351308724

Mirnig, N., Weiss, A., Skantze, G., Al Moubayed, S., Gustafson, J., Beskow, J., Granström, B., & Tscheligi, M. (2013). Face-to-face with a robot: What do we actually talk about? *International Journal of Humanoid Robotics, 10*(1), Article 1350011. https://doi.org/10.1142/S0219843613500114

Mohamed Hashim, M. A., Tlemsani, I., Mason-Jones, R., Matthews, R., & Ndrecaj, V. (2024). Higher education via the lens of industry 5.0: Strategy and perspective. *Social Sciences & Humanities Open, 9*, Article 100828. https://doi.org/10.1016/j.ssaho.2024.100828

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*(1), 1. https://doi.org/10.1186/2046-4053-4-1

Moola, S., Munn, Z., Sears, K., Sfetcu, R., Currie, M., Lisy, K., Tufanaru, C., Qureshi, R., Mattis, P., & Mu, P. (2015). Conducting systematic reviews of association (etiology): The Joanna Briggs Institute's approach. *International Journal of Evidence-Based Healthcare, 13*(3), 163–169. https://doi.org/10.1097/XEB.0000000000000064

Moon, A., Parker, C. A. C., Croft, E. A., & Van der Loos, H. F. M. (2013). Design and impact of hesitation gestures during human-robot resource conflicts. *Journal of Human-Robot Interaction, 2*(3), 18–40. https://doi.org/10.5898/JHRI.2.3.Moon

Obo, T., & Takizawa, K. (2022). Analysis of timing and effect of visual cue on turn-taking in human-robot interaction. *Journal of Robotics and Mechatronics, 34*(1), 55–63. https://doi.org/10.20965/jrm.2022.p0055

Or, J., & Takanishi, A. (2007). Effect of a flexible spine emotional belly dancing robot on human perceptions. *International Journal of Humanoid Robotics, 4*(1), 21–48. https://doi.org/10.1142/S0219843607000935

Őrsi, B., Csukonyi, C., & Korondi, P. (2024). Organic human-robot interactions: Psychological aspects to help social robots become sociable. In *2024 IEEE/SICE international symposium on system integration (SII)* (pp. 1038–1044). https://doi.org/10.1109/SII58957.2024.10417356

Park, S., & Whang, M. (2022). Empathy in human–robot interaction: Designing for social robots. *International Journal of Environmental Research and Public Health, 19*(3), 1889. https://doi.org/10.3390/ijerph19031889

Penčić, M., Čavić, M., Oros, D., Vrgović, P., Babković, K., Orošnjak, M., & Čavić, D. (2022). Anthropomorphic robotic eyes: Structural design and non-verbal communication effectiveness. *Sensors, 22*(8). https://doi.org/10.3390/s22083060

Picard, R. W. (1997). *Affective computing* (pp. xii–292). The MIT Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sabelli, A. M., Kanda, T., & Hagita, N. (2011). A conversational robot in an elderly care center: An ethnographic study. *Proceedings of the 6th International Conference on Human-Robot Interaction, 37–44*. https://doi.org/10.1145/1957656.1957669

Salma, Z., Hijón-Neira, R., Pizarro, C., & Abdul Moqeet, A. (2025). Effectiveness of robot-mediated learning in fostering children's social and cognitive development. *Applied Sciences, 15*(7), Article 7. https://doi.org/10.3390/app15073567

Saunderson, S., & Nejat, G. (2019). How Robots Influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics, 11*(4), 575–608. https://doi.org/10.1007/s12369-019-00523-0

Schellen, E., Bossi, F., & Wykowska, A. (2021). Robot gaze behavior affects honesty in human-robot interaction. *Frontiers in Artificial Intelligence, 4*. https://doi.org/10.3389/frai.2021.663190

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). Oxford University Press.

Schömbs, S., Klein, J., & Roesler, E. (2023). Feeling with a robot—The role of anthropomorphism by design and the tendency to anthropomorphize in human-robot interaction. *Frontiers in Robotics and AI*, 10. https://doi.org/10.3389/frobt.2023.1149601

Sheikholeslami, S., Jung Moon, A., & Croft, E.a. (2017). Cooperative gestures for industry: Exploring the efficacy of robot hand configurations in expression of instructional gestures for human-robot interaction. *The International Journal of Robotics Research, 36*(5–7), 699–720. https://doi.org/10.1177/0278364917709941

Skantze, G. (2017). Towards a general, continuous model of turn-taking in spoken dialogue using LSTM recurrent neural networks. In K. Jokinen, M. Stede, D. DeVault, & A. Louis (Eds.), *Proceedings of the 18th annual SIGdial meeting on discourse and dialogue* (pp. 220–230). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-5527.

Spezialetti, M., Placidi, G., & Rossi, S. (2020). Emotion recognition for human-robot interaction: Recent advances and future perspectives. *Frontiers in Robotics and AI, 7*. https://doi.org/10.3389/frobt.2020.532279

Stedtler, S., Fantasia, V., Tjøstheim, T. A., Brinck, I., Johansson, B., & Balkenius, C. (2025). Gaze and movement adaptation in response to delayed robotic movement during turn-taking. *Scientific Reports, 15*, Article 34098. https://doi.org/10.1038/s41598-025-17140-9

Szabóová, M., Sarnovský, M., Maslej Krešňáková, V., & Machová, K. (2020). Emotion analysis in human–robot interaction. *Electronics, 9*(11), Article 11. https://doi.org/10.3390/electronics9111761

Trivedi, C., Bhattacharya, P., Prasad, V. K., Patel, V., Singh, A., Tanwar, S., Sharma, R., Aluvala, S., Pau, G., & Sharma, G. (2024). Explainable AI for industry 5.0: Vision, architecture, and potential directions. *IEEE Open Journal of Industry Applications, 5*, 177–208. https://doi.org/10.1109/OJIA.2024.3399057. IEEE Open Journal of Industry Applications.

Yang, W., & Xie, Y. (2024). Can robots elicit empathy? The effects of social robots' appearance on emotional contagion. *Computers in Human Behavior: Artificial Humans, 2*(1), Article 100049. https://doi.org/10.1016/j.chbah.2024.100049

Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence, 4*, Article 100134. https://doi.org/10.1016/j.caeai.2023.100134

Zonca, J., Folsø, A., & Sciutti, A. (2021). The role of reciprocity in human-robot social influence. *iScience, 24*(12), Article 103424. https://doi.org/10.1016/j.isci.2021.103424