



Exploring Contextual Cues for Audio-visual Emotion Recognition

What are specific cues and traits within the sender's context that shape human emotion expression and how do these cues influence and bias emotion perception?

Georges Ataya

Supervisor(s): Bernd Dudzik, Sayak Mukherjee

¹EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 21, 2026

Name of the student: Georges Ataya
Final project course: CSE3000 Research Project
Thesis committee: Bernd Dudzik, Sayak Mukherjee, Stephanie Tan

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Abstract

Computational systems are increasingly integrated in people's daily lives. As this continues, proper emotional detection is needed for human-compatible intelligent systems (Russell, 2021). As the universality theory of emotions, the idea that human emotion expressions are consistent and universal across cultures, starts to fade from scientific view, context is increasingly important for emotion recognition.

Using the SALSA methodology, this systematic literature review explores sender context and attempts to answer the following questions: What specific cues and traits within the sender's context shape emotional displays? Through which mechanisms do these cues and traits influence how emotion is perceived?

We anticipated that the four most impactful context categories that would emerge from the literature would be age, culture, gender and personality.

We found substantial evidence of the impact of age, culture and gender on emotional expression and recognition through four key mechanisms: emotion suppression and display rules, in-group advantage, socially acquired stereotypes and physiological differences. Personality was excluded due to insufficient literature. We discuss the implications of these findings on psychology and development of context-aware emotion recognition (CAER) systems.

1 Introduction

Emotion is fundamental to human social cognition, yet traditional computational models often treat its recognition as a simple pattern recognition problem based on facial expressions of emotions devoid of context. As intelligent systems become more tightly integrated into socially complex environments like healthcare and education, developing human-compatible systems that can accurately infer human experiences and values becomes essential.

The historical assumption in psychology and computing (in relation to emotion recognition) has been that a perceiver can read emotions on a face "like words on a page" (Barrett et al., 2011, p. 1). This assumption is commonly known as the theory of universal emotions, that emotions are universal across culture and context. They further argue that this is an erroneous assumption, demonstrating that facial muscle movements are often ambiguous and that emotion perception is deeply shaped by the context in which those movements occur. Since then the scientific consensus has strongly shifted in favour of this context driven approach. The bigger picture of this context encompasses three primary categories: sender, perceiver and situation context. This paper will dive into sender context.

Research on various sender context cues and traits has significantly developed over the last two decades. There have been studies, experiments and meta-analyses on the broad

categories (e.g. culture Elfenbein et al. (2007)) and their impact on emotion expression and perception. There have also been highly specific studies examining one specific subcategory (e.g. wrinkles and their effect on signal clarity Grandhuis et al. (2021)). However, lacking in the literature is a systematic review of sender context cues across categories, and how specifically they affect emotion expression and recognition, not only how they contradict the theory of universal emotions.

Several fundamental questions in Context-aware Emotion Recognition (CAER) remain unanswered: which contextual cues are most influential across different scenarios? How do these features affect perception in each scenario? And how can they be effectively leveraged in large data-driven CAER systems? It remains particularly unclear to what extent the specific cues of the sender (also known as the expresser) influence the way emotions are expressed and subsequently perceived by others.

The aim of this research is to address these gaps by answering the following research questions:

- **RQ1:** Which contextual cues are most critical/influential to someone's emotional expression?
- **RQ2:** How do these cues influence how this emotional expression is perceived? (Assuming a neutral perceiver)

To answer these questions, this paper utilises a systematic literature review methodology. Following the SALSA (Search, Appraisal, Synthesis, and Analysis) framework, we conduct a structured search and critical appraisal of existing interdisciplinary research to address the gaps found in psychology and cognitive science and their implications for CAER systems. This review will not examine audio data due to limited time (see discussion).

2 Background

The leading theory on emotions by the end of the 19th century was primarily based off of Darwin's findings on expression and emotion. In his 1872 work "The Expression of the Emotions in Man and Animals" he describes emotions as instinctive and evolutionary (Darwin, 1872). In his eyes they are the remains of our adaptation and static across culture. Furthermore he introduces the idea of six basic emotional states: happiness, sadness, fear, anger, surprise, disgust.

Almost a full century later, Silvan Tomkins formalised (and partially revived) Darwin's findings while also introducing Affect Theory in his book *Affect Imagery Consciousness* (1962). He posits that there exist nine major affects defined by intensity labels. These affects were seen as a biologically hard-wired mechanism in the body not subject to change.

These findings were further developed by Paul Ekman, a significant researcher in this field for the latter half of the 20th century and the beginning of the 21st century. Ekman et al. (1969) proposed the theory of six universal emotions in a cross-cultural study examining how people in "both literate and pre-literate cultures" (Ekman et al., 1969, p.86) assigned emotions to various photographs. As a result of his findings he posits the view of "universal facial displays" of emotions, a view that became the predominant one in the next few

decades of psychological affect research. Another important addition to the field was the Facial Action Coding System (FACS), which was not originally developed by Ekman but was refined and published by Ekman and Friesen (1978) (and later updated in 2002 (Ekman et al., 2002)). This facial encoding system for categorising facial expressions quickly became the dominant tool in emotion expression research. At this time, the scientific consensus was predominantly convinced of two axioms:

1. That expression of emotions are universal, regardless of culture, situation or perceiver.
2. That there are discrete categories of emotions that can be used to infer the emotional state of a person.

The second assumption was challenged by Russell (1980), in which he introduced the two dimensional **valence-arousal** model of categorising emotions. Instead of being placed into so-called objective and discrete categories independent of one another, emotions were now seen with more nuance. Wierzbicka (1992) further criticises the discrete categorization of emotions by arguing that they are "cultural artifacts of the English language" (Wierzbicka, 1992, p. 1), essentially arguing that there was heavy western bias in the studies and papers used in these categories and the culture associated with that language. Even in cross-cultural studies, western concepts were brought in non-western contexts.

A pivotal moment in the field was Carroll and Russell (1996), which was one of the first papers that showed a significant impact of context on the expression and decoding of a display of emotions (specifically here, situational context).

Finally, the concrete turning point in redefining the theory of universal expressions was the widely cited (Barrett et al., 2011). After which many more papers and reviews on the topic were made (notably, a significantly more extensive review by Barrett: (Barrett et al., 2019)) as well as many other studies using the context dependent view (Fölster et al., 2014; Hess et al., 2012). The scientific consensus seems to have widely shifted to the more nuanced view that emotions are heavily context dependent and that we're slowly moving away from the foundation that Darwin proposed.

3 Methodology

In conducting this literature review, given the scope of the problem, it is very important to collect sufficient papers to draw proper conclusions. However, it was also important to not cast too wide of a net and end up having imprecise and perhaps unhelpful data, due to the relatively small time frame. Considering this, a structured and systematic data collection approach is very much needed. This section will describe how papers were gathered, analysed and synthesised in a manner that will be reproducible to future peers who aim to build on this research.

The framework used for this paper was the SALSA (Search, Appraisal, Synthesis, Analysis) framework consisting of four main steps which will be described in detail (Grant and Booth, 2009).

3.1 Search

3.1.1 Database Search

As is common practice in literature reviews, academic databases were used during the search phase. Notably the following search engines were used:

- **Scopus:** Due to its broad collection of multidisciplinary papers across a wide range of topics. As well as its advanced search features that allow for a very refined query.
- **PsycInfo:** For more precise search in psychological literature that might not be present in Scopus.
- **Google Scholar:** In case an unpublished source or additional article that could not be found in the previous two databases.

Key concepts were used to derive search terms, the key concepts are as follows: Emotion, Psychology, Context and Human. The specific intersection between these four concepts is enough to encompass the necessary literature to answer the research questions posed above.



Figure 1: Search strategy used for the literature review.

Evidently, these are quite broad concepts which is why they were refined to actually perform a search that's properly scoped. The specific terms that were used in the search can be seen in Table 1.

This query resulted in many papers analysing psychiatric disorders and their impact on emotion, as well as machine learning papers, both of which were out of the scope for this review. Thus, we derived the excluded terms to remove them from the results.

3.1.2 Reference Mining

Furthermore, in conjunction with the database search, reference mining was used to derive key and insightful sources. Namely from papers like Barrett et al. (2019) which pioneered new insights into how important context was for audio-visual emotion recognition. We also used more in-depth literature reviews regarding specific emotional context categories (Fölster et al., 2014). While these papers are primarily literature reviews, two major caveats should be noted

Category	Subcategory
<i>Emotion</i>	emotion recognition
	emotion perception
	emotion expression
<i>Context</i>	cultural context
	culture
	social context
	personality
	gender
	age differences
<i>Psychology</i>	cognitive science
	psychology
	cognition
<i>Human</i>	human
<i>Excluded Terms</i>	disorder
	schizophrenia
	machine learning

Table 1: Search Terms (categories joined by intersection, subcategories joined by union)

as they were important to ensure as little bias as possible while analysing these reviews:

- Reviews will mainly be used for gathering new references and citations. They are never cited directly as primary evidence for empirical claims.
- Reviews will be used sparingly to avoid bias

3.2 Appraisal

3.2.1 Exclusion Criteria

In order to filter through the large amount of papers collected in the search phase, we defined stringent exclusion criteria as defined in Table 2. These criteria were also used to include or exclude papers during the reference mining process.

Exclusion Criteria
1. Non-English literature
2. Papers with focus on facial expressions devoid of context
3. Studies analyzing various mental health disorders
4. Studies unrelated to emotional expression
5. Papers without a focus on sender context
6. Non cognitive science or psychology papers
7. Papers primarily studying non-human subjects

Table 2: Exclusion Criteria

3.2.2 PRISMA

We used the PRISMA reporting guidelines to ensure two important factors: clarity and reproducibility. The following results were obtained from our combined database search and reference mining (see Figure 2).

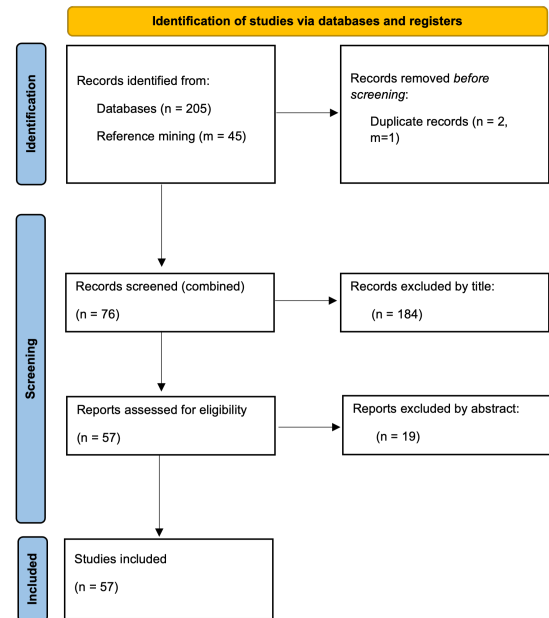


Figure 2: PRISMA reporting diagram

3.3 Synthesis

After the abstract screening phase of this process, the included papers were sorted into six major categories:

- Age - selected 14 papers
- Culture - selected 31 papers
- Gender - selected 14 papers
- Personality - selected 2 papers
- Social class - selected 1 paper
- Background (thematic knowledge that didn't fit into a particular category but had important insights) - selected 17 papers

Given that social class and personality returned an insufficient amount of literature, the categories were not further analysed and all three papers were excluded from the review. The remainder of the synthesis step included further categorising the literature after the final selection was chosen (e.g. the culture papers were further broken down into display rules, in-group advantage and stereotypes).

3.4 Analysis

Once the final selection of papers was collected and categorised, the analysis started on each major category. As described these were further subcategorised and then analysed. The valence-arousal framework (as described in Russell (1980)) was used as opposed to discrete emotion categories. This is in line with current research which suggests that discrete representations of emotions often do not map correctly to a person's actual experience (Barrett et al., 2019, 2011). When discussing a specific emotion label (e.g. anger), that label should be thought of as encompassing the emotions that fall within this emotion category.

The analysis of these subcategories is always in context to how it responds to either of the two primary research questions posed in the introduction. Which means we will be analysing both *which* context cues and traits are most crucial to emotion expression as well as *how* these cues influence perception.

Furthermore, a neutral perceiver was assumed throughout the analysis, with exceptions if the context of the perceiver was relevant to the context of the sender (i.e. for in-group effects).

Finally, the focus of the studies below will primarily be on adult's as senders unless the studying of children is relevant to some cultural phenomena (i.e. cultural upbringing or gender display rules throughout childhood).

4 Results

4.1 Age

The effect of age on emotion has been well researched in the field of cognitive science. The majority of reviews and articles have focused on the perceiver's age, the general consensus on this matter is that as an individual gets older, their capacities for correctly recognising and detecting emotions diminish (Isaacowitz and Stanley, 2011).

When we study the sender's age however, we find strong evidence that emotion expression and recognition isn't static throughout a person's life. Instead, it shifts over time in a consistent manner.

4.1.1 Age Related Effects on Expression

Many recent papers argue that certain displays of emotions on older faces are less accurately recognised than younger faces (Schneider et al., 2022; Grondhuis et al., 2021; Hess et al., 2012; Ebner et al., 2011). A proposed theory of why this happens is that as an individual ages, the manner in which they handle and express emotions changes. This can happen in three major ways.

The first of which is the diminished **physical ability** to express a desired emotion. For instance, Levenson et al. (1991) instructed 20 men and women between the ages of 71-83 to follow muscle-by-muscle instructions to attempt to reproduce various emotions (they were not told which emotion they were imitating). They consistently performed worse than younger adults when it came to replicating the desired emotion. This effect is not only seen through isolated facial movements, in fact people tend to detect smiles more accurately in younger faces than older faces (Murphy et al., 2010). We could infer from this that older adults have a reduced ability to convey certain emotional states due to their age. Muscular control declines with age, including facial muscles which could explain this change in older adults' affective state. In fact with age, "intentional display of facial emotions may become less successful, and displays of unintended blended emotions may become more likely" (Ebner et al., 2011, p. 993). This more regular display of blended emotions could lead to misattributions of various emotional states in these older adults.

Facial artefacts As previously discussed, the faces and emotions of older adults are consistently less accurately decoded than those of younger adults. This could be due to

several facial artefacts making it harder to decode these emotions. An example of one of these artefacts could be, as previously discussed, due to physical ageing of the face, notably wrinkles in older faces. In the experiment conducted by Hess et al. (2012), young adults were asked to rate the emotional expression on the faces of young and old adults. The result of which was a significant difference in accurate emotion identification in the faces of older adults. They found that "wrinkles and folds degraded the clarity of the message conveyed [...] with regard to the specific emotions communicated [...]" (Hess et al., 2012, p. 5).

Emotional suppression There is some evidence that older adults tend to display emotion less and less intensely than younger adults (Magai et al., 2006). Analysing this through the arousal/valence framework of emotion, it seems that the emotion expressions that older adults display consistently are low in arousal and deviate less in valence (from a theoretical neutral state) than younger adults. In fact in Magai et al. (2006) the younger adults had longer bursts of high positive and negative valence emotions than older adults. This could be due to emotional resilience that builds over their lifetime. As an individual lives and experiences highly intense feelings of stress, joy, sadness or other, they may learn to adapt and deal with these situations better. Another theory posited by Magai et al. (2006) is that this reduced display of emotionality could be to "protect social partners from emotional contagion." (Magai et al., 2006, p. 314). Supporting evidence of this was shown in a recent paper that studied over 130 married couples between the ages of 23 and 85, found that older adults use emotional suppression strategies at a much higher frequency than younger adults on a daily basis (Eldesouky and English, 2018).

4.1.2 Own-Age Bias

Throughout much of the literature involving age related displays and recognition of emotion, the idea of "own-age bias" (OAB) in recognition is recurrent (Rhodes and Anastasi, 2012). This theory posits that perceiver's in the same age group as the senders are more likely to accurately decode their emotional displays. This could be because of exposure, you tend to spend time with people in similar age categories as your own, at school, work or a retirement community, therefore learning how to pick up on unique cues. The idea of OAB is also in line with another recurring idea in the study of emotion recognition: in versus out group advantage (Thibault et al., 2006). However OAB has also been criticised, specifically the influence of "perceptual expertise" (Cronin et al., 2019; Civile and Wang, 2025).

4.2 Culture

4.2.1 Cultural variations and display rules

One of the major frameworks prominent in the cultural psychology literature is the impact of **individualistic versus collectivist cultures** on emotion expression. The idea being that individualistic cultures focus on self-expression and independence. Collectivist cultures, on the other hand, place more importance on an individual's role within a group, society and emphasises interdependence.

The relevant research thus focuses on how these differences

between individualistic and collectivist (IC) cultures influence display rules. Matsumoto (1990) first proposed the idea of applying the IC model to emotional expressions and display rules, testing how American (individualist) and Japanese (collectivist) participants would rate the appropriateness of various displays of emotions in different social settings. The findings supported the conjecture but more work was needed. These findings were further developed and argued that the need for "self-regard" is not in fact universal but a strong cultural conception originating from the West (Heine et al., 1999). Markus and Kitayama (1991) argues that different conceptions of the self have strong impact on how someone expresses but also experiences emotion. Further arguing that these displays of emotions are deeply embedded in cultural norms and situation contexts.

Further research conducted attempting to isolate IC as a variable and how it relates to variations in display rules. For instance Matsumoto et al. (1998) studies over 600 participants from four culturally diverse countries (Russia, South Korea, Japan, USA) and rated an "individualism-collectivism" score using a detailed questionnaire measuring over 25 related values. From there they completed another questionnaire evaluating various display rules across four different social relationships for seven basic emotions. This study found that more collectivistic scores were associated with a higher control over socially disruptive emotions (such as anger, contempt, disgust) but *less* control over emotions such as happiness and surprise. Overall a significant effect was found between IC and display rules.

So how does IC affect how and which emotions are displayed? Initial papers suggested that members of collectivist cultures would tend to show more subdued emotions Yuki et al. (2007) in particular negative emotions. Stemming from the idea that since one's display of negative valence emotions are less prioritised than adhering to social norms and not negatively affecting others in the group (Heine et al., 1999). "It is critical for individuals in Japanese culture to act in accordance with the wishes of the group, not on the basis of their own feelings" (Heine et al., 1999, p. 774). Supporting evidence of this theory is Eng (2012) where across five studies studying both Americans and East Asian subjects propensity for emotional suppression, findings were consistent with the idea that collectivist cultures use more suppression more consistently. Another study showed that this is the case even in childhood. Chinese, Japanese and American children were each given "bad" or "disappointing" gifts and their expressions were judged (Ip et al., 2021). Consistently Chinese and Japanese children showed more neutral expressions than the American children. They also found that American children were generally more expressive with both positive and negative displays, demonstrating less emotion suppression or control strategies. However, as of recent, this assumption has been challenged in a recent study (Vishkin et al., 2023). Here the claim is made that individualist cultures are actually *more* likely to adhere to emotional norms. They argue that, as emotional expression is more heavily encouraged in these societies, so is experiencing a specific emotional state (e.g. happiness, excitement). As there is higher social value associated with positive valence emotions, there is higher pressure

to conform. On the other hand, collectivist cultures were argued to place less emphasis on individual emotions, counter-intuitively leading to less need to regulate these emotions in front of others. Nonetheless, this interpretation is novel, with most of the research suggesting the original view on the role of IC on display rules and emotion expression (Friesen, 1972; Matsumoto et al., 1998, 2008; Chentsova-Dutton and Tsai, 2010; Ip et al., 2021).

4.2.2 In group effect

Another critical aspect of emotion recognition is the aforementioned concept of in-group advantage. The idea being that members of a particular cultural group would recognise emotions of members of their own in-group faster and more accurately. "Because the logic of one culture may contradict the logic of another culture, people standing outside a given culture often do not 'get it.'" (Leung and Cohen, 2011, p. 508). One vision of how different cultures vary in expression has been Elfenbein et al. (2007) who presents a view of a "dialect theory" and likens displays of emotions within and between cultures to linguistic dialects and how they might vary. In fact, the literature does suggest that members of a particular in-group have higher accuracy confidence when judging emotions of members belonging to their in-group (Beaupré and Hess, 2006). In-group recognition extends further than just confidence. In an experiment studying Japanese and US participants, both classified and interpreted the expressions shown to them with higher accuracy when shown emotions from people of their own culture (Dailey et al., 2010). In the same study, they then trained two machine learning models, one trained on just Japanese faces and the other on US faces, to classify expressions into six emotion categories. The findings were consistent with the first experiment, supporting the idea of emotion dialects.

Further studies supporting this examined people who were relatively isolated from external (in particular Western) cultural influence. The first conducting studies on two different indigenous societies and seeing how they classified prototypical facial expressions of emotions (Crivelli et al., 2016). The other compared the classification of various emoticons by Tanzanian and Cameroonian participants with Japanese participants (Takahashi et al., 2017). In both cases, expressions of emotions that are traditionally considered universal (Ekman et al., 1969, 2002), were not recognised as such by the people furthest away from Western culture.

Exposure The impact of these findings, however, do not suggest that this in-group advantage is static throughout one's life, in fact they are subject to change with exposure even throughout adulthood. In the highly cited meta analysis, Elfenbein and Ambady (2002), reviewed over 85 articles and found, as expected, that in-group members had a significant advantage in emotion recognition accuracy. However they also found that the effect of the in-group advantage smaller when participants had greater exposure to the other cultures. Whether that be regional proximity or simply media exposure. In the case of minorities who are living in an out-group dominated culture, members of these in-groups recognised more accurately the out-group culture's emotion displays at a better rate than even their own in-group members. This find-

ing was further confirmed in Elfenbein and Ambady (2003), where Tibetans living in China and Africa-Americans living in the United States recognised their host culture's expressions better than their own.

Finally, a recent article suggests that simply labelling of someone as part of an in-group or out-group changes how their emotion is perceived and interpreted (Hess et al., 2022). These findings generally suggest a high in-group advantage between cultures. This advantage, however, is still subject to change as an individual gains exposure to various emotional dialects throughout their lifetime.

4.2.3 Stereotypes & prejudices

Stereotypes and prejudices also shape how emotions are recognised. These stereotypes may be cultural, religious, or racial in nature. Indeed, black men are consistently seen as more threatening than white men (Hutchings and Haddock, 2008; Maner et al., 2005), particularly in white majority countries and regions. Furthermore, racially ambiguous faces portraying anger are consistently more classified as black (Hutchings and Haddock, 2008). This effect is also seen in reverse where the *same* racially ambiguous face is rated higher in intensity and anger if it is labelled as black instead of white.

Cultural specific effects of this have also been documented. In the Netherlands, where anti-Moroccan sentiment is culturally present, Dutch respondents classified Moroccan faces presenting stereotypically consistent emotions (e.g. anger) much faster than other emotions. Dutch faces, on the other hand, saw this effect replicated with quicker sadness attribution (Bijlstra et al., 2014).

4.3 Gender

Throughout the literature there is ample evidence suggesting that emotion expression in men and women are classified differently. Notably, emotions higher in arousal are significantly more quickly and accurately attributed to men (in particular anger) whereas women are attributed lower arousal emotions (e.g. happiness, sadness) (Aguado et al., 2009; García-Gutiérrez et al., 2016; Johnson et al., 2011; Chaplin and Aldao, 2013; Aisha et al., 2026). Below are the major two schools of thought attempting to mechanise this difference in classification.

4.3.1 Morphological differences

Rigorous examination of this effect happened in Becker et al. (2007), where the findings of seven studies on various emotional expressions resulted on angry faces on men and happy faces on women were more accurately and quickly classified. Becker further argued, after examining the results and additional studies, that this could not only be attributed to socially learned differences, but instead to a property of sexual dimorphism. He further posits that anger is an expression of social dominance in men and therefore an evolutionary advantage to naturally convey displays of anger (Becker et al., 2007, p. 189).

In Aguado et al. (2009), three expression classification experiments were conducted. The first one found results consistent with the ones above, that angry faces are more quickly attributed to men, happiness and sadness to women.

Notably, however, this effect was found most strongly for speeded classifications. Aguado suggests that this implies the existence of underlying processes affecting recognition outside of simply gender based stereotypes/social roles.

Finally, Hess et al. (2009) hypothesises that sexually dimorphic facial structures resemble certain emotional displays. They conduct two studies where synthetic digital faces with various emotions were shown and classified. Results aligned with expectations that androgynous angry faces were more likely to be identified as men. This was shown to be linearly proportional with the level of anger that was morphed into the face. Hess also found that the faces of angry women were less quickly classified as women than that of men.

4.3.2 Socially learned biases and roles

More recently however, this morphology theory has increasingly come under scrutiny. For instance, Johnson et al. (2011) strip away all identifiable gender information to purely analyse the perceived effects in four studies. To do so, they analyse and film how a mixed group of 29 participants throw a ball (in a bucket) in four emotional states: angry, happy, neutral and sad. The joints were represented as point light displays. As expected, in the first two studies, angry throws were overwhelmingly assigned to men and sad throws to women. Moreover, in the third study they normalised the velocity of the throws (to further remove any biological cues that could be inferred by the speed/power of the throw). In this the categorisation differences persisted, demonstrating that despite all morphology cues being stripped away, the fundamental gender biases in attribution persist.

Further studies show that anger is socially coded as masculine while happiness (or even smiling) is often coded as feminine (García-Gutiérrez et al., 2016).

A significant meta-analysis studied over 150 studies with around 22,000 participants to see the effect of gender on emotion expression during childhood development (Chaplin and Aldao, 2013). While Chaplin did indeed find differences in the manner in which boys and girls expressed emotion. Boys would express externalising emotions such as anger, girls would express more internalising emotions such as anxiety, sadness or sympathy. The biological fact of these differences were found to be "significant, but very small" (Chaplin and Aldao, 2013, p. 18) for three major reasons. Firstly, this effect was more or less present depending on the social context. For instance, if the child was around their parents, this difference was significantly reduced (especially during earlier years) compared to if they were in the presence of unknown adults. Secondly, depending on the social demand these results also differed. When a girl was given an undesirable present, she tended to display a polite smile more often than boys, and tended to suppress displays of disappointment significantly more often than boys. Finally, the most important finding was developmental stage. The long standing hypothesis in the field was that, if these gender differences were due to socialised expectations, you would see almost no difference in infancy between boys and girls, but a difference would begin and grow along with the children. This theory was indeed confirmed in the literature as many

displays and experiences of emotions were found to have little to no differences in infants (e.g. infant expression of frustration). However as they age, these differences do indeed become more pronounced. "This developmental pattern may reflect a trend for girls to decrease their display of externalising emotions, such as angry outbursts, because they are acquiring an implicit understanding of society's female gender roles" (Chaplin and Aldao, 2013, p. 5). Supporting this idea was a study that examined female and male attorneys emotional delivery across three experiments. Angry emotional displays on male attorneys were perceived as more justified and reasonable than the same displays on female attorneys (Salerno et al., 2018). Moreover, angry male attorneys were perceived as more competent and female attorneys as less effective. This shows how, throughout a boy's life, they are rewarded for displaying these externalising emotions while girls on the other hand are punished for it, leading credence to the theory that women would learn to suppress these displays throughout their lifetime.

Finally, a very recent study examining the effect of pronoun labels for computer recognition software showed an important classification difference. Aisha et al. (2026) looked at how state of the art CAER systems would be biased by a simple change in pronoun labelling and found that only swapping "he, she or they" in sentences led to a significant difference in emotion attribution (up to 7.7% shift for some models).

5 Discussion

Across the three major categories of sender context examined (age, culture, gender), there seem to be recurring themes that systematically impact expression and recognition of emotion. These have direct implications for CAER systems.

5.1 Strong cross-cutting patterns

5.1.1 Emotional suppression

Whether it is gender, age or culture, humans tend to develop varying levels of emotional suppression strategies. This seems to primarily be a pro-social learned behaviour. Collectivist cultures suppress displays of negative emotions to preserve social norms and politeness. Elderly people tend to do so to protect those around them from social contagion. Women have learned that, given the social roles and biases built into society, the way to advance in a given field and be seen as competent is to not display these high-arousal negative emotions. This seems to be the strongest factor influencing emotional expression throughout the various categories. This also means that systems such as FACS should not be used as a universal decoder of emotion, and if used, should be carefully modulated to account for varying levels of emotional suppression.

5.1.2 Stereotypes & prejudices

The role of socially learned stereotypes and prejudices is significant throughout gender, culture, race and even age. As shown in the results, the manner in which society views people with different skin colours, as well as women who show too much anger or externalising emotions, is overwhelmingly negative. This has a significant impact on the various display rules people acquire as they grow up. This effect is the most

challenging one to account for. Systems can increase sensitivity to account for emotional suppression, they can develop culture-specific models to attempt to model various cultural norms and display rules. However ingrained stereotypes and expectations that affect both how emotions are perceived and expressed by those affected need to be addressed and changed at the root cause.

5.1.3 In-group advantage

Another crucial effect seen through multiple of these categories is the in-group advantage. This is especially strong in culture and, to a lesser extent, in age. This is particularly impacted by where someone is born but also what they are exposed to as an adult. Indeed this effect was found to be mitigated when people were simply exposed to a specific culture and their display rules for a sufficient period of time, or if they spent more time around people of different age groups (Macchi Cassia, 2011).

5.1.4 Physiological differences

The final major finding across categories was the physiological differences that interfered with emotional signals. This was found in both age, as people get older and develop wrinkles and their facial musculature decreases, signals become less distinct and are confounded with various emotions. As well as gender, where certain facial structures linked with gender tend to resemble certain emotions. While the effects of this seem weaker than the previous categories, in particular for gender, the differences are worth noting as they do seem to have a non-negligible effect on emotion recognition.

5.2 Implications for CAER systems and psychology

5.2.1 CAER systems

Given the implications of the findings above, CAER systems should be properly calibrated to account for these effects. On one hand, perhaps a longer term or personalised view of a subject should be included with the system. In practice this might mean that a CAER system might need more exposure to a specific person in order to learn about their display rules and level of emotional suppression. This would allow the system to account for varying levels of expressivity that a particular person might regularly display.

On the other hand, these findings might suggest the opposite, that a universal CAER system might not be realistic. In this case, we may want to create purpose-specific systems depending on our needs. For instance, if a nursing home in Japan would like to improve the quality of the lives of its residents with a system that detects negative valence emotions, a specific CAER model could be developed with the particular context of elderly Japanese subjects. This specificity would allow us to further our systems and improve human computer interaction while taking into account the varying contextual differences that impact emotion expression.

Nevertheless, if a system like FACS were to be used in a CAER system, we recommend calibrating the interpretations to account for the various effects described above.

5.2.2 Psychology

The findings above are in accordance with current scientific consensus in this field: that there are no universal expressions of emotions. While this may be increasingly evident with the literature, society has not, as of yet, caught up. Various psychiatric disorders are still being diagnosed in large part by judging the ability of a patient to recognise certain emotions on pictures of faces devoid of context (notably Autism) Barrett et al. (2019).

There are many incentives for ignoring new evidence against universality theory. In particular, scientific research moves faster if we presume that context is nothing more than a nuisance variable. Unfortunately, this is not the case and contextual influences do matter and affect the research.

5.3 Limitations of the research

5.3.1 Available literature

A major limitation while conducting this research was the available literature, throughout the categories most papers focused on **perceiver context** and how it affects emotion recognition. This means there was significantly less data to work with when studying the sender context and its effect on expression and recognition.

Moreover, an important bias in literature, in particular for the cultural section, should be noted. The framing of cultures in two categories individualistic and collectivist is quite helpful when analysing display rules and their potential causes. However it is quite limiting in terms of cultural variation. It serves as a general model for major cultural differences; however, culture is much more complex and has many more dimensions than just IC. An example of this is Leung and Cohen (2011) who distinguish three major categories: honour, face and dignity cultures. As there was not enough literature further using these categories in classifying emotion recognition, we were not able to draw conclusions from them and therefore decided not to report these findings. It is important to note that even this model is an oversimplification of the variation between cultures. For the time being it is sufficient to demonstrate significant cultural differences but should be studied more to analyse the more subtle nuances between cultures and their effects on display rules.

5.3.2 Self-report

Given the move away from the universality theory of emotions, papers have slowly also moved away from an "objective" measure of emotion like FACS. Most recent papers focus instead on self-report data. This means the data used in the studies stems from how the person portraying the emotion describes their emotional state. Evidently, this poses a clear reliability issue when analysing and conducting these studies. Despite this, self-report measures (combined with other context information) remain a more reliable way of researching and reporting emotions than a singular "objective" channel. Further research should therefore also focus on improving and increasing reliability of self-report tools.

5.3.3 Audio

Due to limited time, papers examining audio data were not included in this review. Therefore to confirm and strengthen the

scientific validity of our findings, future research integrating audio data should be conducted.

6 Responsible Research

6.1 Ethics

6.1.1 Studies conducted

A noticeable ethical limitation in the selection of papers was geography. In particular the culture section. Most studies we found analysed four major regions: North America, Europe, China and Japan. With some exceptions for the study of certain indigenous people. This introduces a major gap in the results as most differences and effects were analysed through the lens of relatively affluent and wealthy countries. As discussed above the IC framework is quite limiting but a likely reason of why it was so overwhelmingly present in the literature is the geographical restrictions of the countries listed above. If we expand our search to countries in the Middle East and North Africa (MENA) region we find more nuanced breakdowns of cultural categories (such as the aforementioned honour, face and dignity cultures (Leung and Cohen, 2011)). Therefore we recommend future studies include experiments with subjects from traditionally under-represented regions such as MENA, Latin America, sub-Saharan Africa and south-east Asia.

It is also noteworthy to mention the limitations of relying on exclusively English written literature, particularly in this context. Given that language is such a large part of culture, there are surely many more insights to be gleaned from papers in other languages.

6.1.2 Reproducing prejudices

Many of the papers mentioned in this review examine the effects of racial and gender stereotypes. The handling and presenting of these papers was particularly sensitive and careful as to avoid reproducing or confirming these (possibly harmful) stereotypes.

6.1.3 Author background

It is important to mention that, as a male who grew up in Europe, there might be unconscious biases present in how I analyse and interpret cultural or gender specific results.

6.2 Reproducibility

To ensure reproducibility of both the intermediate and final results we reported the whole process in the Methodology section. Furthermore we outlined each step using the SALSA framework and reported the appraisal step using PRISMA reporting guidelines. Finally, we included strict and clear exclusion criteria that we used throughout our appraisal step to remove papers that were not relevant for our research. The part that somewhat hurts reproducibility of our results is reference mining. Given that reference mining is quite difficult to report with as much precision and transparency as with database searches, the process is less clear. However, a systematic approach was still used while reference mining (which can be seen in the methodology section) as an attempt to counteract the reduced transparency that comes with this method.

6.3 Generative AI usage disclosure

Generative artificial intelligence usage was limited but present throughout the writing of this paper. Tools were used in two primary ways. The first was research, I conducted the research and collection of papers from start to finish using database tools and reference mining (see Methodology). However, during full text screening of certain papers, Claude was used to double-check our findings. An example of a prompt is: "Given the paper attached above, ensure that my interpretation of the paper is consistent with the findings above. My interpretation: [...]"

The second way generative AI was used was for spell and grammar checks to ensure proper grammar and spelling as well as a reasonable academic tone. Example of a prompt: "Can you proof read this section and ensure there are no major grammar or spelling mistakes".

7 Conclusions and Future Work

In this systematic literature review we aimed to answer the following questions: What specific cues and traits within the sender's context shape emotional displays? Through which mechanisms do these cues and traits influence how emotion is perceived?

Throughout the literature we found significant impacts of the following categories: age, culture and gender on emotional expression and perception. We also found four major mechanisms that emerged for *multiple* of these categories: emotion suppression, in-group advantage, socially acquired stereotypes and physiological differences. The fact that these mechanisms are found cross-categorically gives us new key insights into the study of emotions. For instance, an intelligent system that recognises emotional suppression traits for someone might be able to calibrate in order to be more receptive or sensitive to certain facial cues.

Given these results, the careful and thorough creation of intelligent CAER systems is essential. The literature showed significant effects of social stereotypes regarding race and gender on recognising emotion. It is crucial that future development of CAER systems take these biases into account and avoid reproducing them.

Future work An important takeaway from this paper is that, despite this review having analysed emotional expression across many different sender context cues, the scope of this project was quite limited. Future work should include audio as a medium for analysing these differences. Furthermore, as future studies of culture and their impact on emotional expression develop, hopefully so will a more nuanced framework of analysing and categorising various cultures. This includes studying classically under-represented parts of the world and lower income countries.

References

- L. Aguado, A. Garcia-Gutierrez, and I. Serrano-Pedraza. Symmetrical interaction of sex and expression in face classification tasks. *Perception amp; Psychophysics*, 71(1): 9–25, 2009. doi: 10.3758/app.71.1.9.
- Qurat Ul Ain Aisha, Yu-Jin Cho, and Byung Gyu Kim. Pronoun matters: A benchmark for diagnosing gender bias in emotion classification. In Byung-Gyu Kim, Hiroo Sekiya, and Deokwoo Lee, editors, *Multimedia Information Technology and Applications*, pages 98–110, Singapore, 2026. Springer Nature Singapore. ISBN 978-981-95-3141-7.
- L. F. Barrett, B. Mesquita, and M. Gendron. Context in emotion perception. *Current Directions in Psychological Science*, 20(5):286–290, 2011.
- Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M. Martinez, and Seth D. Pollak. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1):1–68, 2019. doi: 10.1177/1529100619832930.
- Martin G. Beaupré and Ursula Hess. An ingroup advantage for confidence in emotion recognition judgments: The moderating effect of familiarity with the expressions of outgroup members. *Personality and Social Psychology Bulletin*, 32(1):16–26, 2006. doi: 10.1177/0146167205277097.
- D. Vaughn Becker, Douglas T. Kenrick, Steven L. Neuberg, K. C. Blackwell, and Dylan M. Smith. The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, 92(2):179–190, 2007. doi: 10.1037/0022-3514.92.2.179.
- Gijsbert Bijlstra, Rob W. Holland, Ron Dotsch, Kurt Hugenberg, and Daniel H. J. Wigboldus. Stereotype associations and emotion recognition. *Personality and Social Psychology Bulletin*, 40(5):567–577, 2014. doi: 10.1177/0146167213520458.
- James M. Carroll and James A. Russell. Do facial expressions signal specific emotions? judging emotion from the face in context. *Journal of Personality and Social Psychology*, 70(2):205–218, 1996. doi: 10.1037/0022-3514.70.2.205.
- Tara M. Chaplin and Amelia Aldao. Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*, 139(4):735–765, 2013. doi: 10.1037/a0030737.
- Yulia E. Chentsova-Dutton and Jeanne L. Tsai. Self-focused attention and emotional reactivity: The role of culture. *Journal of Personality and Social Psychology*, 98(3): 507–519, 2010. doi: 10.1037/a0018534.
- Ciro Civile and Guangtong Wang. Testing the own-age bias in face recognition among younger and older adults via the face inversion effect. *Perception*, 55(6):544–556, 2025. doi: 10.1177/03010066251405714.
- Carlos Crivelli, Sergio Jarillo, James A. Russell, and José-Miguel Fernández-Dols. Reading emotions from faces in two indigenous societies. *Journal of Experimental Psychology: General*, 145(7):830–843, 2016. doi: 10.1037/xge0000172.
- Sophie L. Cronin, Belinda M. Craig, and Ottmar V. Lipp. Emotional expressions reduce the own-age bias. *Emotion*, 19(7):1206–1213, 2019. doi: 10.1037/emo0000517.

- Matthew N. Dailey, Carrie Joyce, Michael J. Lyons, Miyuki Kamachi, Hanae Ishi, Jiro Gyoba, and Garrison W. Cottrell. Evidence and a computational explanation of cultural differences in facial expression recognition. *Emotion*, 10(6):874–893, 2010. doi: 10.1037/a0020019.
- Charles Darwin. *The expression of the emotions in man and animals*. John Murray, London, England, 1872. doi: 10.1037/10001-000. URL <https://doi.org/10.1037/10001-000>. PsycINFO ID: 2004-16316-000.
- Natalie C. Ebner, Yi He, and Marcia K. Johnson. Age and emotion affect how we look at a face: Visual scan patterns differ for own-age versus other-age emotional faces. *Cognition and Emotion*, 25(6):983–997, May 2011. doi: 10.1080/02699931.2010.540817.
- P. Ekman and W.V. Friesen. *Facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press, Palo Alto, CA, 1978.
- P. Ekman, W.V. Friesen, and J.C. Hager. *Facial Action Coding System: Facial action coding system : the manual : on CD-ROM*. Facial Action Coding System. Research Nexus, 2002. ISBN 9780931835018. URL <https://books.google.nl/books?id=wphFzwEACAAJ>.
- Paul Ekman, E. Richard Sorenson, and Wallace V. Friesen. Pan-cultural elements in facial displays of emotion. *Science*, 164(3875):86–88, 1969. doi: 10.1126/science.164.3875.86.
- Lameese Eldesouky and Tammy English. Another year older, another year wiser? emotion regulation strategy selection and flexibility across adulthood. *Psychology and Aging*, 33(4):572–585, 2018. doi: 10.1037/pag0000251.
- Hillary Anger Elfenbein and Nalini Ambady. On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128(2):203–235, 2002. doi: 10.1037/0033-2909.128.2.203.
- Hillary Anger Elfenbein and Nalini Ambady. When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology*, 85(2):276–290, 2003. doi: 10.1037/0022-3514.85.2.276.
- Hillary Anger Elfenbein, Martin Beaupré, Manon Lévesque, and Ursula Hess. Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion*, 7(1):131–146, 2007. doi: 10.1037/1528-3542.7.1.131.
- Joshua S. Eng. Emotion regulation and culture: The effects of cultural models of self on western and east asian differences in suppression and reappraisal. 2012. URL <https://api.semanticscholar.org/CorpusID:46960344>.
- M. Fölster, U. Hess, and K. Werheid. Facial age affects emotional expression decoding. *Frontiers in Psychology*, 5:30, 2014. doi: 10.3389/fpsyg.2014.00030.
- Wallace V. Friesen. *Cultural differences in facial expressions in a social situation: an experimental test of the concept of display rules*. PhD thesis, 1972. URL <https://www.proquest.com/dissertations-theses/cultural-differences-facial-expressions-social/docview/302545694/se-2>.
- Ana García-Gutiérrez, Luis Aguado, Verónica Romero-Ferreiro, and Elisa Pérez-Moreno. Discrimination of face gender and expression under dual-task conditions. *Attention, Perception, and Psychophysics*, 79(2):614–627, 2016. doi: 10.3758/s13414-016-1236-y.
- Maria J. Grant and Andrew Booth. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2):91–108, May 2009. doi: 10.1111/j.1471-1842.2009.00848.x.
- Sabrina N. Grondhuis, Angela Jimmy, Carolina Teague, and Nicolas M. Brunet. Having difficulties reading the facial expression of older individuals? blame it on the facial muscles, not the wrinkles. *Frontiers in Psychology*, 12, 2021. doi: 10.3389/fpsyg.2021.620768.
- Steven J. Heine, Darrin R. Lehman, Hazel Rose Markus, and Shinobu Kitayama. Is there a universal need for positive self-regard? *Psychological Review*, 106(4):766–794, 1999. doi: 10.1037/0033-295x.106.4.766.
- U. Hess, R. B. Adams, K. Grammer, and R. E. Kleck. Face gender and emotion expression: Are angry women more like men? *Journal of Vision*, 9(12):19–19, 2009. doi: 10.1167/9.12.19.
- Ursula Hess, Reginald B. Adams, Annie Simard, Michael T. Stevenson, and Robert E. Kleck. Smiling and sad wrinkles: Age-related changes in the face and the perception of emotions and intentions. *Journal of Experimental Social Psychology*, 48(6):1377–1380, 2012. doi: 10.1016/j.jesp.2012.05.018.
- Ursula Hess, Magdalena Rychlowska, Tobias Storz, Christophe Blaison, Agneta Fischer, and Eva G. Krumhuber. Bridging the empathy gap: or not? reactions to ingroup and outgroup facial expressions. *Journal of Cultural Cognitive Science*, 6(1):77–92, 2022. doi: 10.1007/s41809-022-00097-w.
- Paul B. Hutchings and Geoffrey Haddock. Look black in anger: The role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology*, 44(5):1418–1420, 2008. doi: 10.1016/j.jesp.2008.05.002.
- Ka I Ip, Alison L. Miller, Mayumi Karasawa, Hidemi Hirabayashi, Midori Kazama, Li Wang, Sheryl L. Olson, Daniel Kessler, and Twila Tardif. Emotion expression and regulation in three cultures: Chinese, Japanese, and American preschoolers’ reactions to disappointment. *Journal of Experimental Child Psychology*, 201:104972, 2021. doi: 10.1016/j.jecp.2020.104972.
- D. M. Isaacowitz and J. T. Stanley. Bringing an ecological perspective to the study of aging and recognition of emotional facial expressions: Past, current, and future methods. *Journal of Nonverbal Behavior*, 35(4):261–278, 2011. doi: 10.1007/s10919-011-0113-6.
- Kerri L. Johnson, Lawrie S. McKay, and Frank E. Pollick. He throws like a girl (but only when he’s sad): Emotion

- affects sex-decoding of biological motion displays. *Cognition*, 119(2):265–280, May 2011. doi: 10.1016/j.cognition.2011.01.016.
- Angela K.-Y. Leung and Dov Cohen. Within- and between-culture variation: Individual differences and the cultural logics of honor, face, and dignity cultures. *Journal of Personality and Social Psychology*, 100(3):507–526, 2011. doi: 10.1037/a0022151.
- R. W. Levenson, L. L. Carstensen, W. V. Friesen, and P. Ekman. Emotion, physiology, and expression in old age. *Psychology and Aging*, 6(1):28–35, 1991. doi: 10.1037/0882-7974.6.1.28.
- Viola Macchi Cassia. Age biases in face processing: The effects of experience across development. *British Journal of Psychology*, 102(4):816–829, 2011. doi: 10.1111/j.2044-8295.2011.02046.x.
- Carol Magai, Nathan S. Consedine, Yulia S. Krivoshekova, Elizabeth Kudadjie-Gyamfi, and Renee McPherson. Emotion experience and expression across the adult life span: Insights from a multimodal assessment study. *Psychology and Aging*, 21(2):303–317, 2006. doi: 10.1037/0882-7974.21.2.303.
- Jon K. Maner, Douglas T. Kenrick, D. Vaughn Becker, Theresa E. Robertson, Brian Hofer, Steven L. Neuberg, Andrew W. Delton, Jonathan Butner, and Mark Schaller. Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology*, 88(1):63–78, 2005. doi: 10.1037/0022-3514.88.1.63.
- Hazel R. Markus and Shinobu Kitayama. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2):224–253, 1991. doi: 10.1037/0033-295x.98.2.224.
- David Matsumoto. Cultural similarities and differences in display rules. *Motivation and Emotion*, 14(3):195–214, 1990. doi: 10.1007/bf00995569.
- David Matsumoto, Sachiko Takeuchi, Sari Andayani, Natalia Kouznetsova, and Deborah Krupp. The contribution of individualism vs. collectivism to cross-national differences in display rules. *Asian Journal of Social Psychology*, 1(2):147–165, 1998. doi: 10.1111/1467-839x.00010.
- David Matsumoto, Seung Hee Yoo, and Johnny Fontaine. Mapping expressive differences around the world. *Journal of Cross-Cultural Psychology*, 39(1):55–74, 2008. doi: 10.1177/0022022107311854.
- Nora A. Murphy, Jonathan M. Lehrfeld, and Derek M. Isaacowitz. Recognition of posed and spontaneous dynamic smiles in young and older adults. *Psychology and Aging*, 25(4):811–821, 2010. doi: 10.1037/a0019888.
- Matthew G. Rhodes and Jeffrey S. Anastasi. The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, 138(1):146–174, 2012. doi: 10.1037/a0025750.
- James A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980. doi: 10.1037/h0077714.
- Stuart Russell. Human-compatible artificial intelligence. In Stephen Muggleton, Nicholas Chater, Stephen Muggleton, and Nick Chater, editors, *Human-Like Machine Intelligence*. Oxford University Press, 07 2021. ISBN 9780198862536. doi: 10.1093/oso/9780198862536.003.0001. URL <https://doi.org/10.1093/oso/9780198862536.003.0001>.
- Jessica M. Salerno, Hannah J. Phalen, Rosa N. Reyes, and N. J. Schweitzer. Closing with emotion: The differential impact of male versus female attorneys expressing anger in court. *Law and Human Behavior*, 42(4):385–401, 2018. doi: 10.1037/lhb0000292.
- Jan N. Schneider, Magdalena Matyjek, Anne Weigand, Isabel Dziobek, and Timothy R. Brick. Subjective and objective difficulty of emotional facial expression perception from dynamic stimuli. *PLOS ONE*, 17(6):e0269156, 2022. doi: 10.1371/journal.pone.0269156.
- Kohske Takahashi, Takanori Oishi, and Masaki Shimada. Is smiling? cross-cultural study on recognition of emoticon’s emotion. *Journal of Cross-Cultural Psychology*, 48(10):1578–1586, 2017. doi: 10.1177/0022022117734372.
- Pascal Thibault, Patrick Bourgeois, and Ursula Hess. The effect of group-identification on emotion recognition: The case of cats and basketball players. *Journal of Experimental Social Psychology*, 42(5):676–683, 2006. ISSN 0022-1031. doi: <https://doi.org/10.1016/j.jesp.2005.10.006>.
- Allon Vishkin, Shinobu Kitayama, Martha K. Berg, Ed Diener, Daphna Gross-Manos, Asher Ben-Arieh, and Maya Tamir. Adherence to emotion norms is greater in individualist cultures than in collectivist cultures. *Journal of Personality and Social Psychology*, 124(6):1256–1276, 2023. doi: 10.1037/pspi0000409.
- Anna Wierzbicka. Talking about emotions: Semantics, culture, and cognition. *Cognition and Emotion*, 6(3-4):285–319, May 1992. doi: 10.1080/02699939208411073.
- Masaki Yuki, William W. Maddux, and Takahiko Masuda. Are the windows to the soul the same in the east and west? cultural differences in using the eyes and mouth as cues to recognize emotions in japan and the united states. *Journal of Experimental Social Psychology*, 43(2):303–311, 2007. doi: 10.1016/j.jesp.2006.02.004.