AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

ISSN: 1528-3542

2025, Vol. 25, No. 7, 1750–1763
https://doi.org/10.1037/emo0001516

Is It Pain, Anger, Disgust, or Sadness? Individual Differences in Expectations of Pain Facial Expressions

Arianne Richer[1], Francis Gingras[1, 2], Marie-Pier Plouffe-Demers[1, 2], Daniel Fiset[1], and Caroline Blais[1]
[1] Department of Psychoeducation and Psychology, University of Quebec in Outaouais
[2] Department of Psychology, University of Quebec in Montreal

Humans rely on facial expressions to assess others' affective states. However, pain facial expressions are poorly recognized and are often confused with other negative affective states, such as anger, disgust, sadness, and fear. Previous research has shown that individuals' expectations about the appearance of pain facial expressions are not optimal and do not perfectly reflect the facial features typically observed in individuals expressing pain. In the present study, we verified if expectations about pain facial expressions are also suboptimal by overlapping with other affective states. We relied on two published data sets (data collected between 2017 and 2020) containing images representing the expectations of the appearance of pain facial expressions according to 162 White participants. We then asked an independent group of White participants ($N = 60$, 30 women, $M_{age} = 31.5$) to rate the degree to which they perceived the six basic emotions (anger, disgust, fear, joy, sadness, and surprise) and pain in those images (data collected in 2023). The same pattern of findings was obtained in both data sets. Anger, disgust, and sadness were perceived as highly salient in expectations about pain facial expressions. Most importantly, three clusters of participants with distinct expectations were found. These results support the hypothesis that individual differences exist in how observers expect pain to be expressed. These individual differences might impact the ability of an observer to distinguish an expression of pain from other negative affective states, and raising awareness about them might help reduce mistakes with serious consequences.

Keywords: mental representation, pain facial expression, facial expressions confusion, individual differences, reverse correlation

Supplemental materials: https://doi.org/10.1037/emo0001516.supp

Facial expression is one of the main channels to evaluate pain experienced by others (Hadjistavropoulos et al., 2011; Poole & Craig, 1992; Prkachin, 1986; Williams, 2002). However, facial expressions of pain are often confused with negative emotions, such as disgust (Kappesser & Williams, 2002; Roy et al., 2015; Simon et al., 2008), anger (Dildine et al., 2023; Kappesser & Williams, 2002), fear (Simon et al., 2008), and sadness (Roy et al., 2015). Such confusions can have deplorable consequences, especially when it comes to helping behaviors. For example, if an expressor is trying to communicate pain, but the observer recognizes it as anger, the person in need may not receive help. Situations like these show how important it is to understand mechanisms underlying the confusion between facial expressions of pain and other affective states.

Expectations regarding the appearance of pain facial expressions may contribute to this confusion. Previous research has shown that, on average, individuals hold expectations that do not perfectly match the way pain is most typically expressed (Blais et al., 2019). Although imperfect, these expectations will at least partly reflect facial expressions one has been exposed to in their day-to-day life (Jack et al., 2016). As at least four different facial configurations for pain expressions have been observed (Kunz & Lautenbacher, 2014; Kunz et al., 2021), various degrees of exposure to these configurations may shape expectations differently from one individual to the other.

Individual variations in expectations have already been observed in the overlap between facial expressions of basic emotions

This article was published Online First April 7, 2025.
Hillel Aviezer served as action editor.
Arianne Richer https://orcid.org/0009-0006-4054-8224
Caroline Blais https://orcid.org/0000-0001-6055-8485
All data, analysis code, and research materials are available on the Open Science Framework at https://osf.io/szbm7/?view_only=9964f91627b64fbbb1d926f086b03cd2. The authors have no conflicts of interest to declare. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (Grant RGPIN-2019-06201), as well as by a Canada Research Chair in Cognitive and Social Vision and the Canada Excellence Research Chairs, Government of Canada (Grant 950-232282), both held by Caroline Blais. The authors thank Daphnée Sénécal for her help with the platforms LabMaestro Pack&Go and Prolific and Vicki Ledrou-Paquet for her help with MATLAB programming.

Arianne Richer, Francis Gingras, Marie-Pier Plouffe-Demers, Daniel Fiset, and Caroline Blais conceptualized the study. Arianne Richer, Francis Gingras, and Caroline Blais programmed the experimental and analysis codes. Arianne Richer conducted data collection. Arianne Richer, Francis Gingras, Marie-Pier Plouffe-Demers, Daniel Fiset, and Caroline Blais conducted data analysis. Arianne Richer, Francis Gingras, and Caroline Blais wrote the first article draft. All authors discussed and interpreted the results and commented on the article.

Correspondence concerning this article should be addressed to Caroline Blais, Department of Psychoeducation and Psychology, University of Quebec in Outaouais, C. P. 1250, Succursale Hull, Gatineau, QC J8X3X7, Canada. Email: caroline.blais@uqo.ca

(Brooks & Freeman, 2018). However, few studies have measured expectations of pain facial expressions, and to our knowledge, none have verified how they overlap with the basic emotions. The present study aims to fill this gap by visualizing how expectations of facial expressions of pain overlap with other affective states, while exploring if there are individual differences in patterns of overlap.

Reverse correlation is well suited for measuring expectations about the appearance of pain facial expressions (Blais et al., 2019; Jack et al., 2016). This method has been designed to measure expectations of the appearance of various visual objects (Ahumada, 1996; Ahumada & Lovell, 1971; Mangini & Biederman, 2004). Different versions of the method exist, many of which involve adding visual noise to an image to modulate its overall appearance and asking the participant to perform a decision about the noisy stimulus. This noise varies from one trial to another such that it is possible to infer the noise properties that were associated with the participant's perception. For example, in the context of pain facial expressions, adding visual noise to a face may slightly modify its appearance such that it sometimes looks in pain and sometimes not. The participants' decisions about the similarity between the noisy stimulus and a pain facial expression are recorded across hundreds of trials. Noise associated with the percept of a pain facial expression is then averaged and becomes a proxy of a participant's expectations regarding the appearance of that expression.

In the present study, we used data from two separate studies in which expectations of pain facial expressions had previously been measured (Lévesque-Lacasse et al., 2024; Saumure et al., 2023). We then asked a group of independent participants to rate the degree to which they perceived pain as well as the six basic emotions in those expectations' proxies. We hypothesized that expectations of pain facial expressions would be perceived as overlapping with other negative basic emotions such as anger, disgust, sadness, and fear and that individual differences would be observed in the pattern of overlap.

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study, and the study follows the Journal Article Reporting Standards (Appelbaum et al., 2018). All data, analysis code, and research materials are available at https://osf.io/szbm7/?view_only=9964f91627b64fbbb1d926f086b03cd2 (Richer, 2025). Data were analyzed using IBM SPSS Statistics 29.0.0.0 and MATLAB R2023a (https://www.mathworks.com). Graphs were realized using R, Version 4.3.1, and the package ggplot2, Version 3.4.3. This study's design and its analysis were not preregistered.

Stimuli

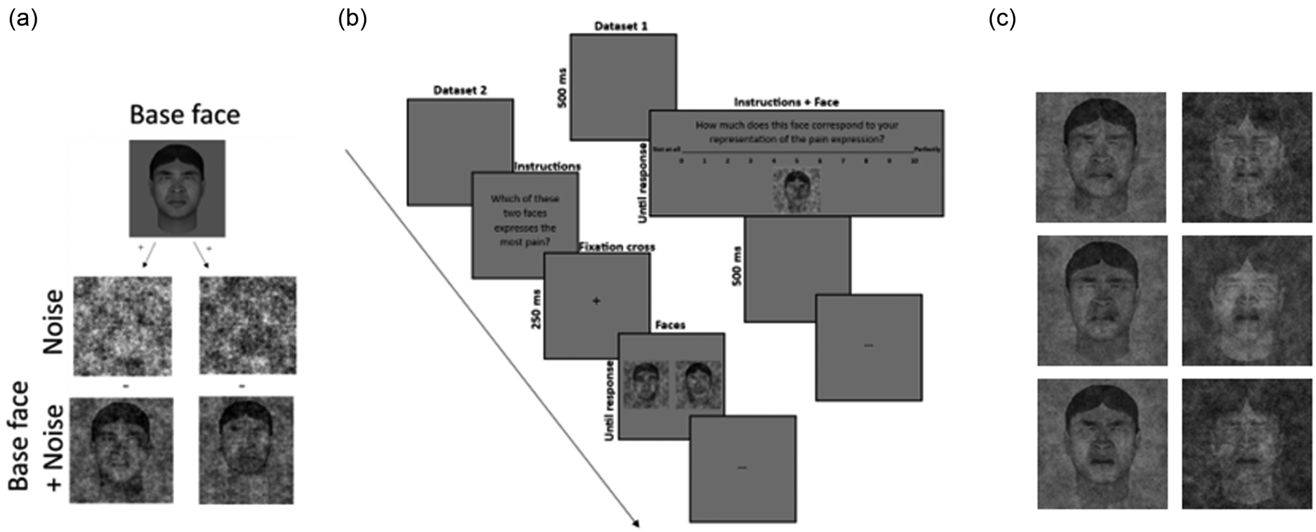
The stimuli utilized in this experiment originate from two previous and distinct studies in which expectations of participants about the appearance of pain facial expressions were extracted using reverse correlation (Ahumada, 1996; Ahumada & Lovell, 1971; Mangini & Biederman, 2004). Figure 1a represents the steps involved in the creation of one stimulus using the version of reverse correlation that was used in those studies (Mangini & Biederman, 2004). More specifically, sinusoidal white noise is added to the picture of a face. The face picture is always the same across trials, but the noise is different on each trial. Because the noise is different on each trial, the appearance of the noisy face varies such that it may become closer to or farther from the participant's expectations of how a pain facial expression looks like. In the study of Lévesque-Lacasse et al. (2024), hereafter called Data Set 1, such noisy faces were presented one at a time to a participant during 500 trials, and the participant's task was to rate, on a scale ranging from 0 to 10, the degree to which the face corresponded to their representation of a pain facial expression. The data for this data set were collected between 2017 and 2020. In the study of Saumure et al. (2023), hereafter called Data Set 2, two noisy faces were presented side-by-side to a participant during 500 trials, and the participant's task was to decide which of the two faces expressed the most pain (see Figure 1b). The data collection for Data Set 2 was conducted between 2017 and 2020. Data Set 1 comprises proxies of expectations of pain facial expressions extracted from 73 White Western participants (37 men). Data Set 2 originally comprised proxies of expectations of pain facial expressions extracted from 30 White Western participants, but data collection was continued afterward to reach a sample of 89 White Western participants (42 men; data not yet published). The reader is invited to read Section A of the Supplemental Material for a more thorough description of the methods used to extract the expectations of pain facial expressions.

The stimulus used in Data Set 1 is a White male avatar generated by FaceGen and FACSGen (Krumhuber et al., 2012; Roesch et al., 2011). This avatar is used as the base face for all trials in Data Set 1. In order to minimize the number of trials, action units (AUs) already linked to facial expression of pain have been slightly activated (Brinkman et al., 2017). Precisely, the AUs activated were brow lowering, cheek raising with eyelid tightening, and nose wrinkling with upper-lip raising (Prkachin, 1992; Prkachin & Craig, 1995; Prkachin & Solomon, 2008). We specifically chose to activate those features in the base face because they have been identified as core features of the facial expression of pain (Prkachin, 1992). For Data Set 2, the base face was created and used in the same way as for Data Set 1, but the avatar was a hybrid composed of the same White male avatar as in Data Set 1 and of an East Asian male face avatar. Adornments, such as hair, ears, and neck, were present in both base faces. While these features may influence participants' evaluation of the stimuli (Jack & Schyns, 2015; see, however, Butler et al., 2010), we chose to retain them because they are inherently part of the interpretative process.

Power

In the present study, an independent set of participants was asked to rate the degree to which they perceived the six basic emotions and pain in a subset of the 162 proxies of expectations included in Data Sets 1 and 2. Importantly, in the analyses, we are interested in the variations *across the proxies of expectations*; we are thus going to verify the systematicity of the ratings across the expectations included in Data Sets 1 and 2. In other words, the power of the study will be determined by the number of proxies, not the number of participants rating the proxies. Two analyses will be conducted. In line with our objective of revealing possible overlap with affective states in expectations of pain facial expressions, we planned a $2 \times 2 \times 7$

Figure 1
Experimental Procedures for the Creation of Pain Expectation Proxies



Note. (a) Steps involved in the creation of two stimuli using the reverse correlation process. Patches of visual noise are added to a face. The face remains constant throughout the experiment, but a new patch of noise is generated on each trial, randomly altering the appearance of the face across trials. (b) Procedure used in the two studies in which Data Sets 1 and 2 were collected. For Data Set 1, participants were asked to rate, on a scale of 0 to 10, the degree to which the face corresponded to their expectation of a pain facial expression. For Data Set 2, participants were asked to decide which of two faces expressed the most pain. (c) Examples of proxies of expectation from participants in Data Set 1 (first column) and Data Set 2 (second column) resulting from the reverse correlation process explained in (a) and (b). These proxies were generated by calculating the weighted sum of all noise patches based on a participant's responses, resulting in a classification image for each participant. This image is then applied to the base face to generate the final proxy. The avatar faces shown here were artificially generated using FaceGen and FACSGen softwares.

mixed analysis of variance (ANOVA) on the between-subject factors of the data set from which the proxies were drawn (Data Set 1 or 2) and of the biological sex of the participants and the within-subject factor of affective states scales. Within this design, the effects that are most pertinent to the objective of the present study, and thus for which a power analysis was conducted, are the main effect of affective states as well as the interaction between participant's sex and affective states. When assuming a medium effect size of Cohen's $f = 0.25$ for the within factor (affective states), an α of .05, and a minimum power of .8, the sample size required needed to be at least 20 proxies. When assuming a medium effect size of Cohen's $f = 0.25$ for the within-between interaction factor (Participant's Sex × Affective States), an α of .05, and a minimum power of .8, the sample size required needed to be at least 28 proxies. In line with our objective of exploring potential individual differences in patterns of overlap, a *k*-means clustering analysis was planned, separately for each data set. It has been shown that samples allowing to reach *N*s of between 20 and 30 participants per cluster lead to sufficient statistical power (Dalmaijer et al., 2022). Thus, we have enough power in both Data Set 1 ($N = 73$) and Data Set 2 ($N = 89$) to include up to three clusters in the analysis.

Participants

Three groups of 20 White participants (a total of 60 participants; 30 men; $M_{\text{age}} = 31.5$, $SD_{\text{age}} = 10.96$) took part in this experiment. Based on the approach proposed by Hehman et al. (2018), we found that 20 observations by proxies were sufficient to attain stable mean ratings. The choice of only recruiting White participants was made in order to avoid the other-race effect (see the Discussion section; Elfenbein & Ambady, 2002; Young et al., 2012). All participants had normal or corrected-to-normal visual acuity. The protocol of this experiment was approved by the Research Ethics Committee of the Université du Québec en Outaouais. All participants were recruited online via the Prolific recruitment platform and received monetary compensation for their participation.

Material

All participants did the task online via the platform LabMaestro Pack&Go from VPixx Technologies (https://vpixx.com/products/labmaestro-packngo/). The experimental program was written in MATLAB, using functions from the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Participants were required to use a desktop computer (i.e., mobile or tablet devices were not allowed) and were instructed to complete the experiment inside a dimly lit room and to set the brightness of their computer screen to 100%.

Procedure

The participants were divided into three groups who each rated 54 of the 162 proxies. This was done in order to make sure that the task did not last more than 30 min. Thus, each proxy was rated by 20 participants (10 men); five out of the 162 proxies only have ratings from 19 participants due to technical issues. This number of participants was chosen in order to obtain stable ratings for each proxy, as the average of the 20 participants was then used as the dependent variable in the aforementioned analyses.

In order to have a stimulus size of approximately 5° of visual angle, each participant started the experiment by completing the credit card test (Li et al., 2020). This test allowed us to approximate the screen resolution of the participant and, thus, the size of the stimulus on screen. The distance needed between the participant and the computer screen for the stimulus to subtend 5° of visual angle was then calculated, and the participant was instructed to respect that distance throughout the experiment.

Each participant completed 64 trials in total. Out of those 64 trials, a car image was presented on 10 trials, and one of the proxies of expectations drawn from Data Set 1 or 2 was presented on the remaining 54 trials. The face and car images were presented in a random order, on the upper part of the screen. Inserting cars among the images served as a measure of attention. On each trial, the participant was first asked to indicate, using different keyboard keys, if a face (proxy of expectations) or a car was presented. Whenever a participant made more than nine mistakes, the experiment was automatically terminated and their data were not included in the study. When a face was presented and correctly categorized, a numerical scale ranging from 1 to 9 appeared underneath it. In a random order, seven questions were then displayed on the screen, asking the participant to indicate the degree to which the displayed face expressed the six basic emotions (anger, disgust, fear, sadness, surprise, and joy) and pain. For each affective state, they needed to provide an answer ranging between 1 (*not at all*) and 9 (*extremely*). This procedure is similar to other affective states evaluation tasks previously used in the field (see Mende-Siedlecki et al., 2020, for an example). The participant's answers were then printed on the screen and remained visible until the end of the trial. After giving a rating for the seven affective states, participants were asked if they wanted to change one of their answers. Participants could change their answers as many times as needed before moving on to the next stimulus. The order in which the proxies and affective states scales were presented varied randomly across participants. At the end of the experiment, the participants were also asked to answer a series of questions to assess the environment and conditions under which they completed the task. These questions are available in Section B of the Supplemental Material.

Analysis and Results

Quality Check: Task Duration, Performance With Catch Trials, and Interrater Agreement

The average completion duration for the online task was approximately 30 min. All the participants had a good performance on the catch trials, with 53 out of the 60 participants reaching an accuracy of 100% and seven participants reaching an accuracy of 98%, suggesting that they paid appropriate attention to the task. Moreover, interobserver agreement was assessed by calculating pairwise correlations between the ratings of the seven scales given by all possible pairs of participants for each proxy. The mean pairwise correlation was then averaged across all proxies and participants, yielding an overall measure of agreement of $r = .435$. This level of agreement is in a similar range to what may be found using pictures of real facial expressions, when data are collected in a lab environment (rather than online). For example, results from an unpublished experiment conducted in our lab showed an average agreement, for negative emotions such as anger, disgust, fear, and sadness, of $r = .601$ for pictures from a database of simulated expressions and of $r = .500$ and $r = .387$ for two databases of spontaneous expressions.

Visualizing How Expectations of Pain Facial Expressions Overlap With Basic Emotions

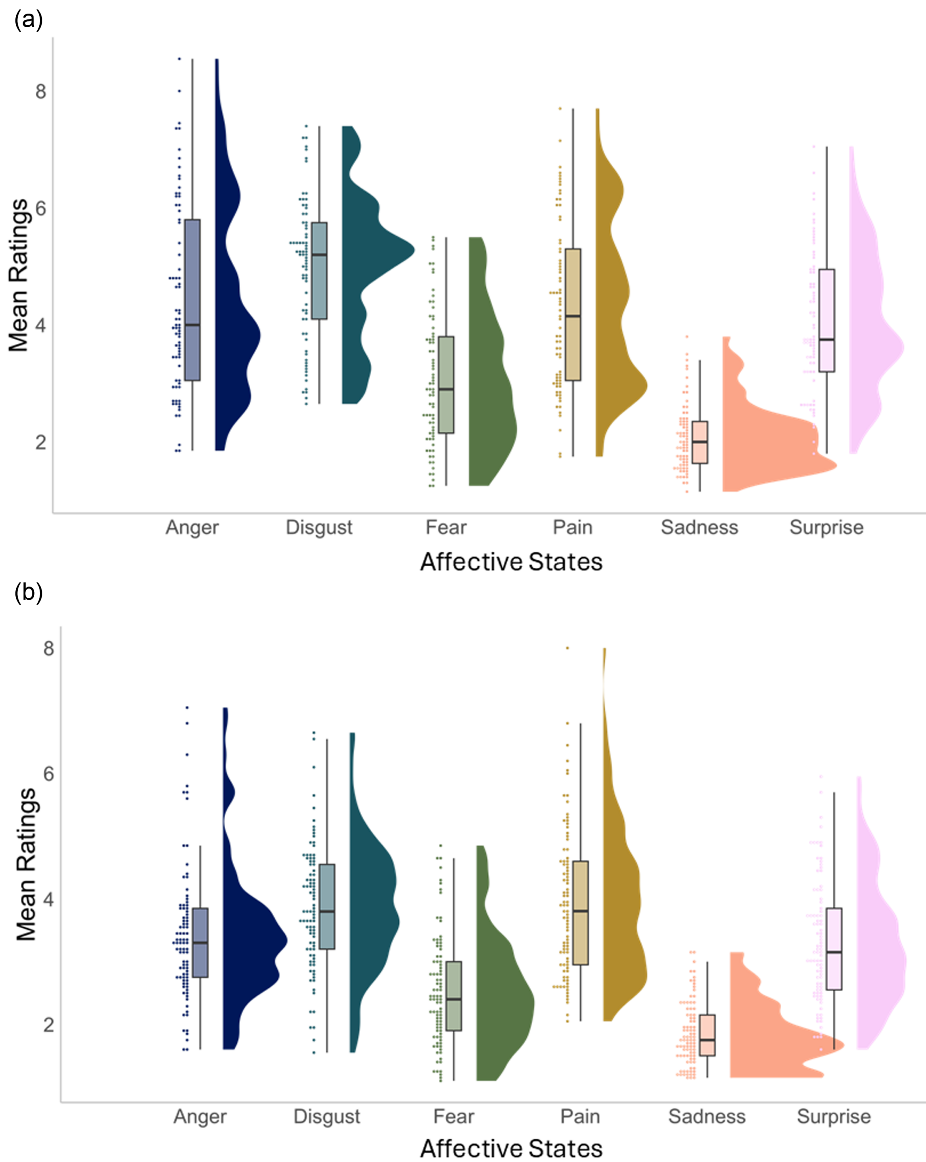
The joy scale was excluded from all subsequent analyses because the variability of ratings was too low, with 50% of ratings contained between 1.05 and 1.15 in Data Set 1 and between 1.05 and 1.2 in Data Set 2. However, running the analyses while keeping joy in the data sets did not change the pattern of results. For each proxy, we calculated mean ratings across all 20 participants separately for each of the six affective states left. Each of the 162 proxies thus ended up with six mean ratings. For all ANOVAs that were conducted in the following section, the Greenhouse–Geisser correction was applied when sphericity was not respected. We first conducted a mixed ANOVA 2 (data sets) × 2 (sexes) × 6 (affective scales) to verify if the data could be pooled across data sets and participants' sex for subsequent analyses. We have opted to include the sex of the participants in this analysis as previous studies have demonstrated differences in performance in recognizing facial expressions of pain between men and women (Hill & Craig, 2004; Plouffe-Demers et al., 2023; Prkachin et al., 2004). This analysis revealed a main effect of the data sets, $F(1, 158) = 43.28$, $p < .001$, $\eta_p^2 = .22$, and a main effect of the affective scales, $F(1.73, 272.63) = 147.81$, $p < .001$, $\eta_p^2 = .48$, but the main effect of sex was not significant, $F(1, 158) = 3.25$, $p = .07$, $\eta_p^2 = .02$. The interactions between affective scales and sex, sex and data sets, or affective scales, sex, and data sets were not significant ($p = .42$ and $\eta_p^2 = .005$, $p = .27$ and $\eta_p^2 = .008$, and $p = .43$ and $\eta_p^2 = .005$, respectively). However, the interaction between affective scales and data sets was significant, $F(1.73, 272.63) = 5.52$, $p < .01$, $\eta_p^2 = .03$.

The main effect of data sets was explained by overall higher ratings in Data Set 1 than in Data Set 2. Since the factor of data sets interacted with the factor of affective scales, the subsequent analyses were conducted separately for each data set. The participants' sex was not included as a factor in the following analyses since it did not interact with any of the variables of interest. Two repeated-measures ANOVAs on the factor of affective scales were conducted. For Data Set 1, the results indicated a main effect of affective scales, $F(1.67, 120.48) = 62.53$, $p < .001$, $\eta^2 = .47$ (see Figure 2a). For Data Set 2, the results also indicated a main effect of affective scales, $F(1.86, 163.53) = 94.96$, $p < .001$, $\eta^2 = .52$ (see Figure 2b).

To identify the affective states with which pain was most confused, paired-samples t tests were conducted. The ratings on each affective scale were compared to the ones given on the pain scale. A Bonferroni correction was applied since five comparisons were conducted, leading to a threshold for significance of $p < .01$. The descriptive statistics and results of the t tests are presented in Table 1. For Data Set 1, results indicated that disgust was perceived to a greater degree than pain, anger and sadness were perceived to a similar degree as pain, and fear and surprise were less perceived than pain. For Data Set 2, results indicated that disgust and sadness were perceived to a greater degree than pain, anger was perceived to a similar degree as pain, and fear and surprise were less perceived than pain.

Figure 2

Raincloud Plots Showing the Distribution of Mean Ratings for Each Affective State Across the Two Data Sets



Note. Panel (a) represents Data Set 1. Panel (b) represents Data Set 2. For each affective state, the figure displays, from left to right, the individual data points, a boxplot, and a density distribution. Within each box, horizontal black lines denote median values; boxes extend from the 25th to the 75th percentile of each group's distribution of values; vertical extending lines denote adjacent values (i.e., the most extreme values within 1.5 interquartile range of the 25th and 75th percentile of each group). The width of the curve in the density distribution represents the density of proxies at each mean rating level presented on the *y*-axis. See the online article for the color version of this figure.

Although facial expression intensity is not necessarily associated with variations in ambiguity, some research has observed a potential link between those two variables, where pain facial expressions of lower intensity are more confused with other affective states (Dildine et al., 2023; Saumure et al., 2018). To rule out the possibility that the relatively low intensity of the proxies might have contributed to the observed confusions, we verified if the perceived intensity of the proxies was associated with the degree to which the proxies were ambiguous, that is, were evaluated as containing a variety of affective signals. To achieve this, we first computed an ambiguity measure by calculating the entropy associated with each proxy. Entropy is a measure quantifying the amount of uncertainty

Table 1
Descriptive Statistics and Results of t Tests Comparing the Mean Ratings for the Six Affective States With the Mean Rating of Pain

Data set	Affective state	*M* (*SD*)	*t*	*p*	Cohen's *d*
1	Anger	4.43 (1.62)	1.35	.18	0.16
	Disgust	4.98 (1.21)	5.17	<.001	0.61
	Fear	3.06 (1.15)	15.55	<.001	1.82
	Sadness	4.29 (1.44)	2.05	.04	0.24
	Surprise	2.07 (0.57)	17.17	<.001	2.01
	Pain	4.07 (1.19)			
	Joy	1.13 (0.2)			
2	Anger	3.41 (1.07)	0.64	.53	0.07
	Disgust	3.87 (1)	4.76	<.001	0.51
	Fear	2.53 (0.88)	19.27	<.001	2.04
	Sadness	3.94 (1.17)	8.86	<.001	0.94
	Surprise	1.85 (0.51)	17.02	<.001	1.80
	Pain	3.31 (0.98)			
	Joy	1.15 (0.14)			

Note. M = mean; SD = standard deviation.

in the potential state of a system—in our case, it reflects the degree to which similar ratings were given across the affective scales. We then conducted an additional experiment in which an independent group of participants was asked to rate the intensity of pain expressed in each proxy. No significant correlation was found between pain intensity ratings and entropy scores, in either data set, confirming that the moderate intensity of the proxies does not explain the observed confusions. A more comprehensive description of this additional experiment as well as of the analyses and results can be found in Section C of the Supplemental Materials.

Exploring Individual Differences in Expectations About Pain Facial Expressions—Variability in the Pattern of Overlap With Other Affective States

To explore individual variability in the degree to which participants expect pain to overlap with other affective states, k-means cluster analyses were carried out separately for the two data sets.

The number of clusters was decided based on two criteria. First, we considered the point of inflection in the graph representing the within-cluster sum of squares as a function of the number of clusters. The graphs are available in Section D of the Supplemental Material. A visual inspection of the graphs suggested that the point of inflection was somewhere around three to four clusters. Second, we considered statistical power, which was sufficient for three clusters but borderline for four clusters. When taking those two criteria into consideration, we opted for three clusters in the main series of analyses. Nonetheless, we provide the results with four clusters as Supplemental Material.

Following the k-means clustering analyses, paired-samples t tests were conducted to further describe which affective states were most salient in each cluster. A Bonferroni correction was applied, leading to a threshold for significance of $p < .003$. The descriptive statistics are available in Table 2, and the results of these t tests are available in Section E of Supplemental Tables S2 and S3.

In both data sets, the three clusters revealed by the k-means clustering algorithm were very similar. The first cluster represented participants with expectations of pain facial expressions that highly overlapped with anger and disgust. In fact, in both data sets, anger and disgust had significantly higher ratings than all other affective states, including pain.

The second cluster represented participants with expectations of pain facial expressions that highly overlapped with sadness. More specifically, in Data Set 1, sadness had the highest mean rating, although it did not significantly differ from the mean ratings of pain and disgust. In Data Set 2, the mean rating for sadness was significantly higher than the mean ratings of all other affective states.

Finally, the third cluster represented participants with expectations of pain facial expressions that equally overlapped with most negative affective states, except for fear. In Data Set 1, the only affective state rated as significantly higher than pain was disgust. While mean ratings for anger and sadness did not significantly differ from those of pain, they also did not differ from disgust, suggesting that all four affective states had overall similar ratings. For Data Set 2, ratings for disgust, anger, and sadness did not differ significantly from one another, although they were all significantly higher than pain. See Figure 3 for all three clusters from Data Set 1 and Figure 4 for clusters from Data Set 2.

Demographics of the participants from Data Sets 1 and 2, including age and sex, are provided in Supplemental Table S4 for Data Set 1 and Supplemental Table S5 for Data Set 2 of Section F of the Supplemental Material. Additional information about the grouping of proxies in each cluster is also presented in the tables. This includes mean reaction time, empathic scores, the number of proxies per cluster, and the number of proxies rated independently by the three groups of participants in the main task of this study separately for the three clusters. All participants from Data Set 1 and Data Set 2 completed an empathy questionnaire as part of the original experiments. Since these data were available to us, we have included them in the tables. Reaction time data are only available for participants from Data Set 2. Importantly, none of those variables seems to drive cluster grouping.

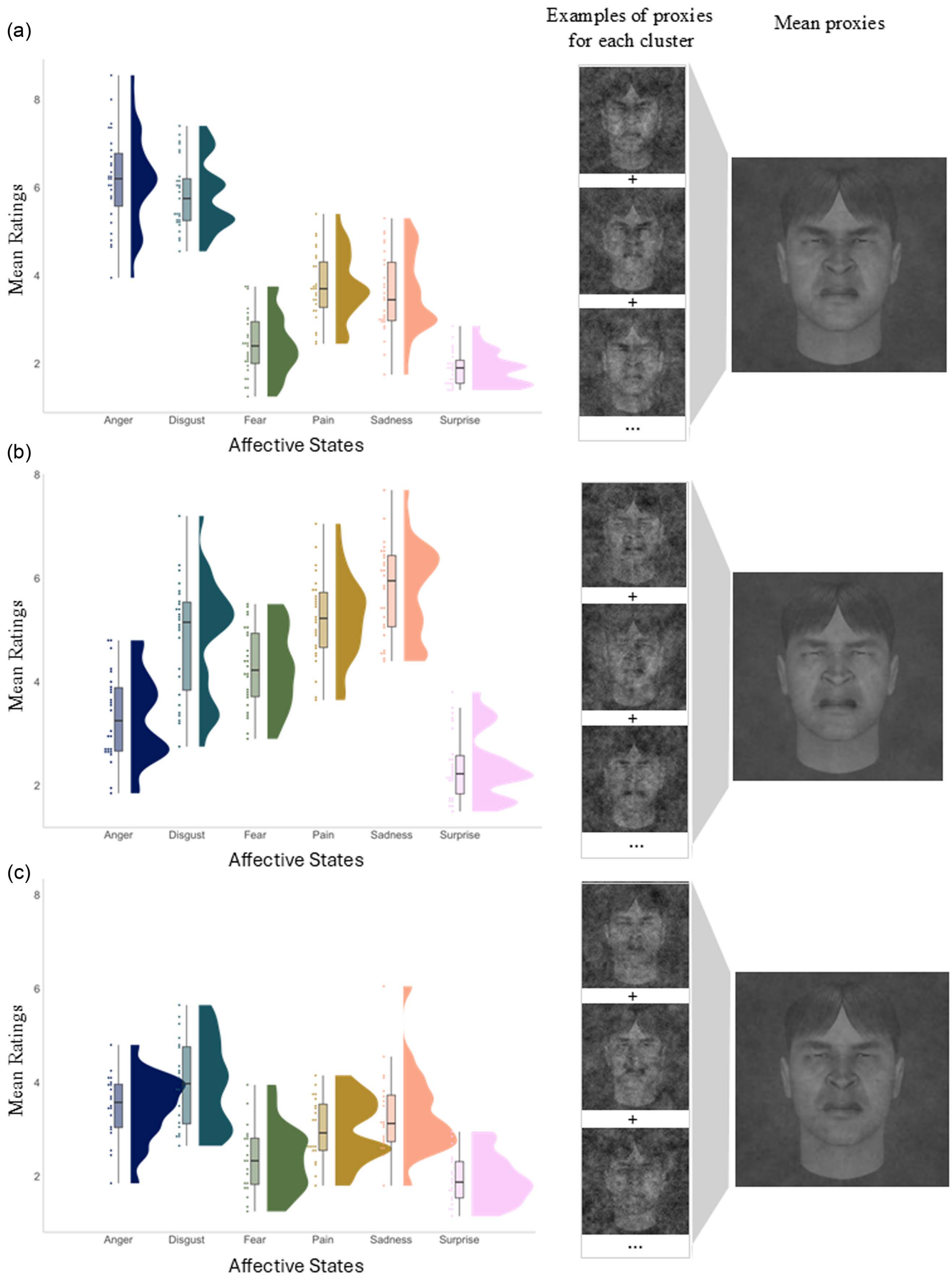
Table 2
Descriptive Statistics for Data Sets 1 and 2

Cluster	Affective state	Data Set 1 *M* (*SD*)	Data Set 2 *M* (*SD*)
1	Anger	6.18 (1.07)	4.41 (1.15)
	Disgust	5.83 (0.80)	4.66 (0.80)
	Fear	2.43 (0.71)	2.38 (0.54)
	Sadness	3.57 (0.92)	3.57 (0.57)
	Surprise	1.88 (0.39)	1.98 (0.46)
	Pain	3.76 (0.77)	3.22 (0.58)
2	Anger	3.32 (0.84)	2.94 (0.56)
	Disgust	4.85 (1.13)	4.02 (0.78)
	Fear	4.26 (0.77)	3.49 (0.69)
	Sadness	5.78 (0.90)	5.26 (0.89)
	Surprise	2.37 (0.65)	2.11 (0.52)
	Pain	5.23 (0.84)	4.40 (0.67)
3	Anger	3.49 (0.70)	2.96 (0.70)
	Disgust	3.98 (0.94)	3.06 (0.66)
	Fear	2.36 (0.72)	1.85 (0.43)
	Sadness	3.32 (0.92)	3.12 (0.72)
	Surprise	1.94 (0.52)	1.51 (0.33)
	Pain	3 (0.68)	2.46 (0.45)

Note. M = mean; SD = standard deviation.

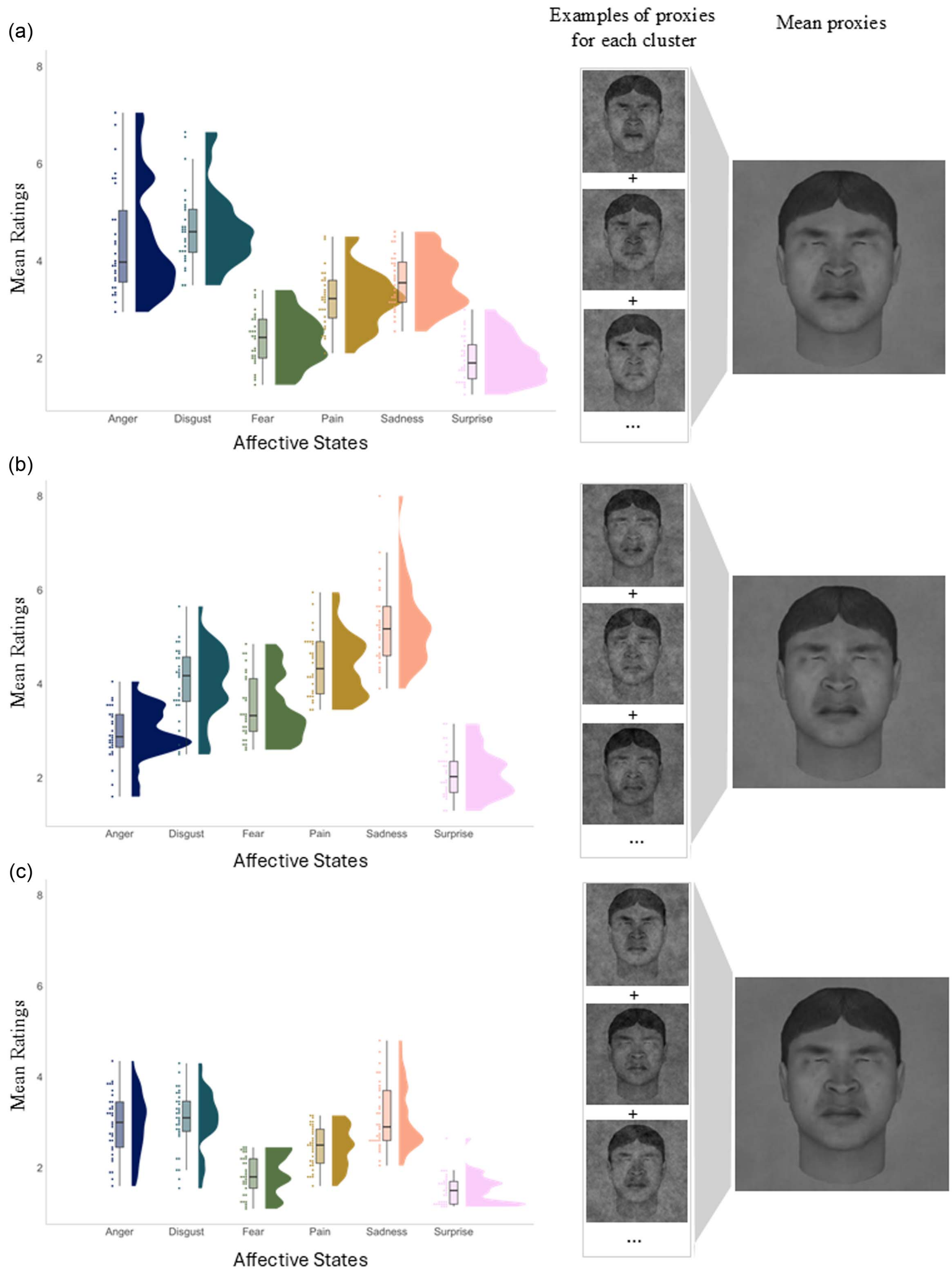
Figure 3

Results of the K-Means Cluster Analysis for Data Set 1



Note. (a) On the left panel, the raincloud plot displays, for each affective state, the individual data points, a boxplot, and a density distribution of ratings (see the figure caption of Figure 2 for a more detailed description). On the right panel, examples of the expectations of some participants from Cluster 1, as well as the average expectation of all participants from that cluster. (b) Same as in (a), but for Cluster 2. (c) Same as in (a), but for Cluster 3. The avatar faces shown here were artificially generated using FaceGen and FACSGen softwares. See the online article for the color version of this figure.

Figure 4
Results of the K-Means Cluster Analysis for Data Set 2



Note. (a) On the left panel, the raincloud plot displays, for each affective state, the individual data points, a boxplot, and a density distribution of ratings (see the figure caption of Figure 2 for a more detailed description). On the right panel, examples of the expectations of some participants from Cluster 1, as well as the average expectation of all participants from that cluster. (b) Same as in (a), but for Cluster 2. (c) Same as in (a), but for Cluster 3. The avatar faces shown here were artificially generated using FaceGen and FACSGen softwares. See the online article for the color version of this figure.

Exploring Objective Differences Across Clusters and Genders

In addition to comparing the way in which the proxies from each of the three clusters are perceived, as was done in the previous analysis, it is also possible to verify if the proxies included in those clusters differ in an objective, pixel-based analysis. In order to do so, the classification image of each participant was first slightly smoothed using a Gaussian kernel with a standard deviation of 12 pixels. Then, we conducted one-sample *t* tests on each pixel of the smoothed classification images to evaluate whether pixel values significantly deviated from zero, indicating areas uniquely associated with each cluster. The statistical threshold for significance was determined using the cluster test from the Stat4CI toolbox (Chauvin et al., 2005), which corrects for multiple comparisons by accounting for spatial dependencies among neighboring pixels. This method leverages random field theory to control the family-wise error rate while preserving sensitivity to localized effects. To run the test, an α value of .008 was used (.025, Bonferroni corrected), and the minimum *t*-score to reach (tC) was set to 3.1. Using those parameters, the minimum cluster size (i.e., number of connected pixels reaching a value of at least 3.1) to reach varied between 180 and 211 pixels, depending on the number of freedom degrees. The significant areas are depicted in white on the middle row of Figure 5a and 5b. The same approach was used for classification images of men and women (α = .013, tC = 3.1; see Figure 5c and 5d). In all groups, the typical areas previously found using a similar approach were found, that is, areas around the mouth and nose regions as well as around the eye and eyebrow regions (Blais et al., 2019; Gingras et al., 2023; Lévesque-Lacasse et al., 2024; Saumure et al., 2023).

In order to statistically compare the proxies across the three clusters revealed in Figures 3 and 4, we conducted three two-sample *t* tests (Cluster 1 vs. Cluster 2, Cluster 1 vs. Cluster 3, and Cluster 2 vs. Cluster 3) for each data set. Again, the statistical threshold for significance was found using the Stat4CI toolbox, using the same parameters (α = .008, tC = 3.1). We also followed the same procedure to statistically compare the proxies captured in men and women. Consistent with prior results in this study, significant differences were identified between clusters, but no significant differences were found between sexes. More specifically, for Data Set 1, areas around the right eye and eyebrow, the bridge of the nose, the upper right cheek, and the left corner of the mouth were significant. For Data Set 2, significant areas were found around the inner corner of the left eye, outer brow corners, and upper part of the left nasolabial fold. Significant areas are highlighted in Figure 5. To help the reader interpret the results, Figure 5 also displays a visualization of the difference between the classification images of each pair of groups. A qualitative inspection of those images supports the findings reported above about the dominant affective states present in the proxies of each cluster. For example, the difference between Clusters 1 and 2 appears angry, which is congruent with the fact that the proxies composing Cluster 1 were on average rated as more similar to anger expressions than the ones in Cluster 2. Reversing the subtraction, that is, subtracting Cluster 1 from Cluster 2, is associated with a residual classification image that appears sad, which is again congruent with the fact that Cluster 2 comprises proxies that were rated as more similar to the expression of sadness.

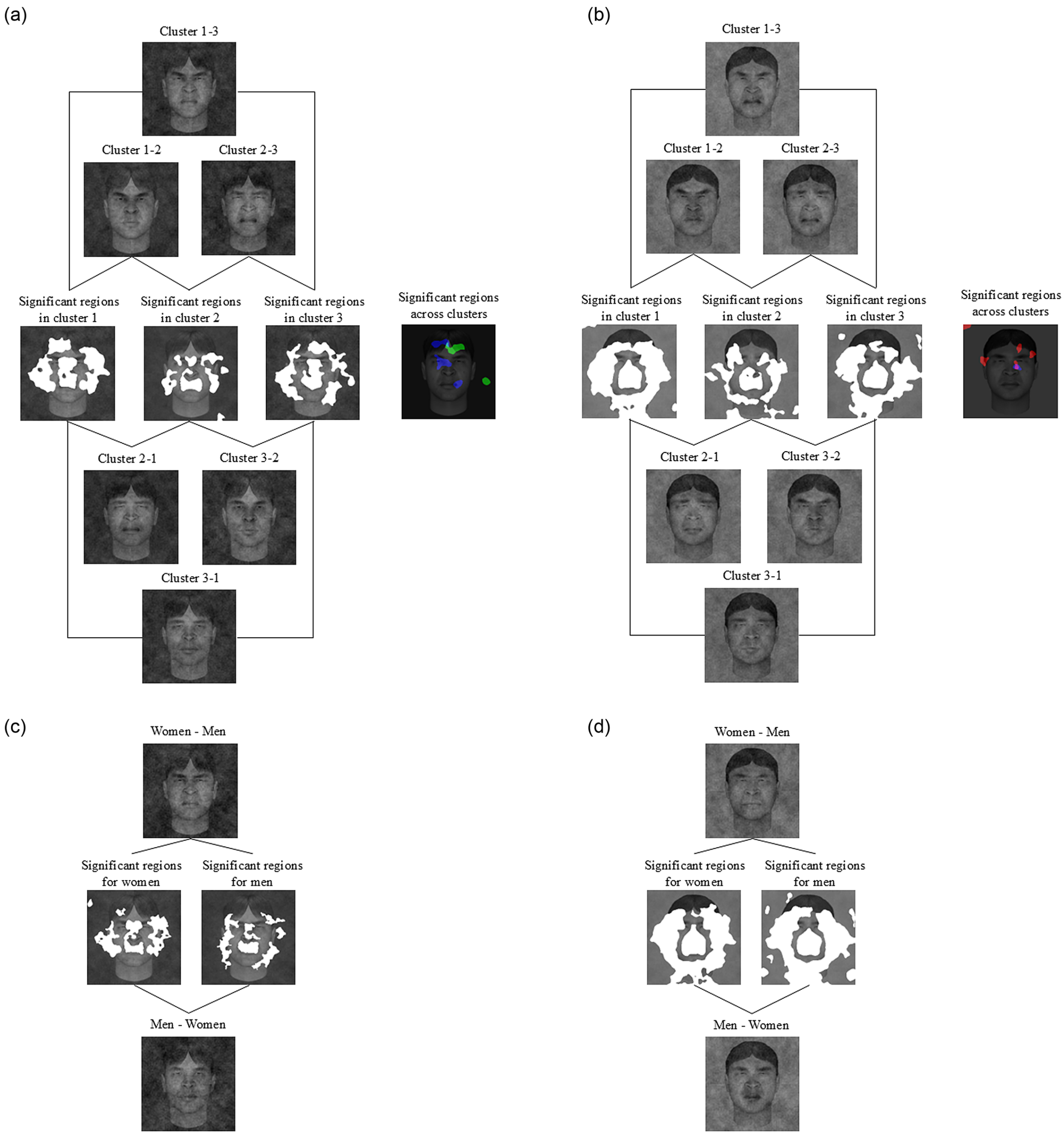
Discussion

Pain facial expressions are often confused with other affective states (Dildine et al., 2023; Kappesser & Williams, 2002; Roy et al., 2015; Simon et al., 2008). Previous research has shown that individuals' expectations about the appearance of pain facial expressions are not optimal and do not perfectly reflect the facial features typically observed in individuals expressing pain (Blais et al., 2019). More specifically, while eye narrowing is the most frequent feature in pain facial expressions, it is not the most salient feature in expectations of pain facial expressions. In the present study, we revealed another way in which expectations of pain facial expressions are not optimal: They highly overlap with other negative affective states. Most importantly, the pattern of overlap is not the same for everyone.

In the first series of analysis, we focused on the average pattern of overlap between pain and other affective states. These analyses showed that overall, disgust, sadness, and anger were perceived to be present at least as much, if not more, than pain in expectations about pain facial expressions. Most importantly, the same pattern was observed in both data sets, with the only difference between the two being that sadness was rated as being higher than pain in Data Set 2, while it did not significantly differ from pain in Data Set 1. These findings are congruent with previous studies focusing on the pattern of confusions obtained during facial expression categorization. In fact, confusions between pain, anger, disgust, and sadness have been repeatedly observed (Dildine et al., 2023; Kappesser & Williams, 2002; Roy et al., 2015; Simon et al., 2008). Some studies have also found confusions with fear, but we did not replicate this finding (Simon et al., 2008). Furthermore, research has shown that adding an emotional context of anger or disgust into a pain scenario does not significantly change the facial configuration of the expression of pain (Tessier et al., 2024). These results highlight the fact that pain and, at least, anger and disgust are associated with highly overlapping expectations regarding facial expressions, corroborating our findings for the first cluster (Tessier et al., 2024).

Interestingly, the pattern of overlap between pain and other affective states did not significantly change as a function of participants' sex. While this finding contrasts with previous studies showing that women are generally better than men at recognizing facial expressions of various affective states, including pain (Hill & Craig, 2004; Prkachin et al., 2004; see, however, Riva et al., 2011), caution is warranted when interpreting this null result. Although our analysis revealed no significant sex difference, several factors suggest that if an effect exists, it is likely small. Specifically, the effect size for the main effect of sex was small, the interaction between sex and affective states was even smaller, no significant regions were found in the objective analysis between proxies created by men and women, and the study had sufficient statistical power to detect a meaningful difference. A possible explanation for this apparent contradiction is that the sex difference in performance at recognizing expressions may not lie in the participant's expectations. In fact, a recent study showed that men and women differ in the efficiency with which they use visual information during the processing of pain facial expressions. More specifically, while both men and women rely on the same facial regions when judging a pain expression, women make better use of these regions (Plouffe-Demers et al., 2023). The advantage for women in the recognition of facial

Figure 5
Results of the Pixel-Based Analysis Across Clusters and Data Sets



Note. (a) Results of the objective (pixel-based) analysis on the three clusters of Data Set 1. The middle lane displays the significant areas, as revealed by the one-sample *t* tests (three leftward images; areas highlighted in white) and the two-sample *t* tests (rightward image; areas highlighted in green and blue). For the two-sample *t* tests, there was no significant difference between Clusters 1 and 2; areas significantly different between Clusters 1 and 3 are highlighted in green, and the ones significantly different between Clusters 2 and 3 are highlighted in blue. The two rows above and below the middle lane display a visualization of the difference between the classification images of each pair of groups. (b) Same as in (a), but for Data Set 2. There was a significant difference between Clusters 1 and 2, which is highlighted in red on the rightward image of the middle lane. There was however no significant difference between Clusters 1 and 3. (c) Results of the objective (pixel-based) analysis on the women and men's classification images of Data Set 1. The middle lane displays the significant areas, as revealed by the one-sample *t* tests (three leftward images; areas highlighted in white). For the two-sample *t* test, there was no significant difference between women and men's CI. The rows above and below the middle lane display a visualization of the difference between the classification images of each pair of groups. (d) Same as in (c), but for Data Set 2. The avatar faces shown here were artificially generated using FaceGen and FACSGen softwares. CI = confidence interval. See the online article for the color version of this figure.

expressions of pain may therefore stem from visual information extraction processes rather than their expectations about the appearance of those expressions. Subsequent studies will be necessary to delve deeper into this question.

We also revealed substantial individual variabilities in the expectations about pain facial expressions. The cluster analysis allowed us to reveal three naturally occurring patterns of overlap between pain and other affective states. More specifically, a first group of participants tended to expect pain facial expressions in which anger and disgust were the most salient affective states. A second group of participants expected pain facial expressions in which sadness was the most salient affective state. Finally, a third group of participants expected pain facial expressions that represented a mixture of negative affective states, with anger, disgust, sadness, and pain being equally salient. Again, the pattern of findings was very similar in Data Sets 1 and 2, supporting the robustness and replicability of our results.

Interestingly, the study from which Data Set 1 was drawn has shown that part of individual variations in the ability to infer the intensity level of pain experienced by another is associated with differences in expectations about pain facial expressions (Lévesque-Lacasse et al., 2024). More specifically, while most individuals have a tendency to underestimate the pain of another (pain estimation bias), those who underestimate less expect pain facial expressions to share similarities with sadness. Moreover, while most individuals have difficulty perceiving subtle changes in the intensity of the pain experienced by another (sensitivity to variations in pain intensity), those who are the most accurate expect pain facial expressions to share similarities with anger and disgust.

This raises the question of how different conceptualizations of pain, as an affective state, impact not only accuracy but also other social and emotional responses. For instance, the clusters in this study suggest systematic differences in how pain is perceived: One group conflates pain primarily with anger and disgust, another with sadness, and yet another with a broader range of negative affective states. Each pattern may hint at distinct underlying frameworks for interpreting and responding to others' pain.

It is possible to hypothesize that these clusters may reflect unique affective structures that predispose individuals to different behaviors when encountering someone in pain. For example, those who link pain with anger and disgust might adopt an avoidant or defensive stance, possibly due to an unconscious alignment of pain with aversive or aggressive cues. Conversely, individuals who see pain as aligned with sadness might adopt a more empathic or supportive approach, interpreting pain through a lens of vulnerability. Those who associate pain with a broad spectrum of negative affective states could exhibit a complex mix of reactions, potentially modulating their responses based on context. However, the same study that identified associations between pain estimation bias, sensitivity to variations in pain intensity, and mental representations did not find a significant relationship between mental representation and empathy (Lévesque-Lacasse et al., 2024). Interestingly, all participants from both data sets in the present study had completed an empathy questionnaire. Supplemental Tables S4 and S5 report the average empathy score per cluster, and a statistical comparison of empathy scores across clusters reveals no significant effect of this variable. These findings suggest that while mental representations of pain expressions shape accuracy in pain perception, they may not directly influence (or be influenced by) empathic responses. Instead, they may indicate distinct pathways by which individuals organize their understanding and responses to others' pain. This hypothesis underscores the importance of further research aiming to clarify how conceptual and perceptual structures may jointly influence social and affective interactions in contexts of pain.

Given that a link has been established before between expectations about facial expressions and the ability to infer the intensity of the pain experienced by another, the present findings raise the question of whether some confusions between pain and other affective states will occur more frequently in some participants than in others. For example, would individuals with expectations similar to our participants in cluster one have a higher tendency to confound pain, anger, and disgust (but not pain and sadness)? Such a pattern of finding would align well with previous findings showing that conceptual knowledge predicts visual perception (Brooks & Freeman, 2018). Moreover, it would be consistent with frameworks suggesting that visual perception depends in part on the visual information available to perform a task and in part on the mental representations, or expectations, an observer has encoded in memory with regard to a category of stimuli (Gosselin & Schyns, 2002). It is also possible to hypothesize that these patterns of confusion, as observed in the present study, may contribute to variations in the recognition of facial expressions of pain across broader social groups, such as sex and gender differences or between ethnicities. Further research is needed to investigate these potential effects.

More generally, the present findings raise the question of how closely the expectations, revealed with an implicit approach such as reverse correlation, may align with expectations that would be measured using a more explicit approach, such as asking participant to manipulate a set of AUs to change the appearance of a face until it fits their expectations of a pain expression. A recent study comparing the conclusions drawn from self-reports versus reverse correlation highlights that these methods do not always align (Axt et al., 2023). This underscores the value of reverse correlation in uncovering implicit information, which may differ from more explicit measures. Most importantly, comparing how well expectations measured using reverse correlation versus a more explicit approach predict behavioral confusions is an important area for further research. Understanding which confusions an individual is most likely to make and the reasons behind them could help improve their performance in pain recognition.

Although our study does not allow us to understand the reason underlying the individual variations we observed in expectations about pain facial expressions, we think that two nonmutually exclusive mechanisms might be at play. Expectations about the appearance of pain facial expressions will likely reflect, at least in part, what a person has been exposed to during their life (Gosselin & Schyns, 2002; Jack et al., 2016). Interestingly, previous studies have shown that people do not all express pain in the same way (Kunz & Lautenbacher, 2014; Kunz et al., 2021). More specifically, they have revealed four clusters of pain facial expression configurations and they showed that the facial configuration typically expressed by an individual is stable over time. One possibility is that some individuals are more exposed to a subset of those facial configurations during their life and build expectations that are more similar to those configurations than to the other ones.

Another possibility is that individuals vary in the way they conceptualize pain. Individual differences in the conceptual structures of basic emotions have been found (Brooks & Freeman, 2018). For example, while anger and disgust are clearly distinct emotions for some individuals, they semantically overlap for others. Most importantly, the conceptual structure of basic emotions has been shown to predict the degree to which expectations about facial expressions of basic emotions overlap with one another (Brooks & Freeman, 2018). Thus, it is possible that some individuals conceptualize pain as closer to anger and disgust, others as closer to sadness, and others as being a mixture of all four affective states. Further research specifically measuring pain conception and expectations about pain facial expressions would allow us to verify that hypothesis.

Limitations of the Present Study (Including Constraints on Generality)

Some limitations must be taken into consideration while interpreting our results. Participants in both data sets as well as those recruited to evaluate the expectations of proxies extracted in Data Sets 1 and 2 were highly homogeneous in terms of race and culture: They were all White and came from Western industrialized countries. This homogeneity was achieved on purpose because of a well-known and robust phenomenon observed in the face perception field, that is, the other-race effect (Elfenbein & Ambady, 2002; Malpass & Kravitz, 1969; Young et al., 2012). The other-race effect is characterized by a marked decrease in the performance of an observer when attempting to recognize the facial expression or the face identity of someone of another race than their own. Most importantly, the other-race effect has been linked with the utilization of different visual, mnemonic, and neuronal mechanisms for own versus other-race faces (Tüttenberg & Wiese, 2023). Moreover, previous work has shown that individuals who have grown up in different cultural environments might have built different expectations about the appearance of facial expressions (Jack, Caldara, & Schyns, 2012; Jack, Garrod et al., 2012; Saumure et al., 2023). Given all this, we took the decision to control the race and cultural background of our participants. However, future work should explore the diversity and individual variations of expectations about pain facial expressions across participants coming from different cultural backgrounds and of different races. Studying the impact of the race of the face stimulus is also of utmost importance, as it may explain a part of the inequalities observed in the health system (Kappesser & Williams, 2010; McCaffery et al., 2000; Seers et al., 2018). Moreover, since the gender of an expressor is known to have an influence on the way pain is perceived by an observer, it would also be important to verify how the present findings generalize to a female face (Göller et al., 2023; Zhang et al., 2021). Another point to acknowledge is the use of an artificial face, rather than a real one, in the reverse correlation experiments from which the proxies in the present study were extracted. This might have limited the ecological validity of the extracted proxies. Yet, some have found minimal variation in recognizing emotions between real and artificial faces (Dyck et al., 2008). Moreover, this artificial base face was overlaid in a large amount of visual noise, which made it difficult to distinguish from a real face. Most importantly, the main pattern of confusion that was found in the present study replicates previous findings in facial expression categorization tasks relying on real face pictures (Dildine et al., 2023; Kappesser & Williams, 2002; Roy et al., 2015; Simon et al., 2008). Considering this, we think that it is unlikely that using an avatar had a substantial impact on the present pattern of findings.

Another limitation to consider is the use of static stimuli in our experiments, which inherently lack the dynamic qualities of real-life facial expressions. Everyday pain expressions unfold in natural, temporal sequences, accompanied by subtle changes in AUs and intensity over time. This dynamic information may provide essential information that observers rely on to perceive and differentiate affective states. To address this limitation, future studies could use a reverse correlation approach involving the temporal dimension, such as the one pioneered by Jack et al. (2016). It is indeed possible that the specific expectations regarding the temporality of the different facial features involved in pain expressions vary across individuals. Nevertheless, the main facial features revealed in the proxies would likely have been the same. In fact, previous studies suggest that the same facial features are relied upon to successfully recognize expressions of pain and other affective states, whether those expressions are static or dynamic (Blais et al., 2017, 2019). Moreover, previous studies have observed confusions between pain and other affective states, even with dynamic expressions (Blais et al., 2017; Simon et al., 2008).

Finally, while the disparities in the methodology used to collect the expectations of participants from Data Set 1 and Data Set 2 may be perceived as a limitation, we think they represent an important strength of the present study. In fact, despite the fact that the race of the face stimulus (White in Data Set 1 and a hybrid between White and East Asian in Data Set 2), the color of the face stimulus background (black in Data Set 1 and gray in Data Set 2), and the task (scale in Data Set 1 and two-alternative forced choices in Data Set 2) differed in both data sets, the pattern of results was almost identical. This is particularly noteworthy considering that the use of cluster analysis may be viewed as exploratory and therefore potentially limiting; in the present case, the cluster analysis conducted on two independent data sets led to highly similar results. We think that the high similarity of the results obtained with both data sets speaks for the robustness of our findings and conclusions.

Conclusion

This study shows that individuals' expectations about pain facial expressions highly overlap with the emotions of disgust, sadness, and anger. Most importantly, the pattern of overlap varies from one individual to the other, with one cluster of individuals expecting pain facial expressions to be closer to disgust and anger, a second cluster expecting them to be closer to sadness, and a third cluster expecting pain expressions to equally overlap with anger, disgust, and sadness. These findings raise important questions about individual variations in the kinds of confusion that are made in a real-life setting when having to recognize a pain facial expression. Further studies investigating the causes and consequences of these individual variations are needed to understand how individuals can more reliably evaluate expressions of pain, and could lead to important progress in helping caregivers to be more efficient at identifying pain expressions.

References

Ahumada, A., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *The Journal of the Acoustical Society of America*, *49*(6B), 1751–1756. https://doi.org/10.1121/1.1912577

Ahumada, A. J., Jr. (1996). Perceptual classification images from vernier acuity masked by noise. *Perception*, *25*(1_suppl), 2. https://doi.org/10.1068/v96l0501

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Axt, J., Siemers, N., Discepola, M.-N., Martinez, P., Xiao, Z., & Wehrli, E. (2023). The mind's "aye"? Investigating overlap in findings produced by reverse correlation versus self-report. *Journal of Experimental Social Psychology*, *107*, Article 104473. https://doi.org/10.1016/j.jesp.2023.104473

Blais, C., Fiset, D., Furumoto-Deshaies, H., Kunz, M., Seuss, D., & Cormier, S. (2019). Facial features underlying the decoding of pain expressions. *The Journal of Pain*, *20*(6), 728–738. https://doi.org/10.1016/j.jpain.2019.01.002

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

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, *28*(1), 333–361. https://doi.org/10.1080/10463283.2017.1381469

Brooks, J. A., & Freeman, J. B. (2018). Conceptual knowledge predicts the representational structure of facial emotion perception. *Nature Human Behaviour*, *2*(8), 581–591. https://doi.org/10.1038/s41562-018-0376-6

Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics*, *72*(6), 1444–1449. https://doi.org/10.3758/APP.72.6.1444

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*(9), Article 1. https://doi.org/10.1167/5.9.1

Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics*, *23*(1), Article 205. https://doi.org/10.1186/s12859-022-04675-1

Dildine, T. C., Amir, C. M., Parsons, J., & Atlas, L. Y. (2023). How pain-related facial expressions are evaluated in relation to gender, race, and emotion. *Affective Science*, *4*(2), 350–369. https://doi.org/10.1007/s42761-023-00181-6

Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R. C., & Mathiak, K. (2008). Recognition profile of emotions in natural and virtual faces. *PLOS ONE*, *3*(11), Article e3628. https://doi.org/10.1371/journal.pone.0003628

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. https://doi.org/10.1037/0033-2909.128.2.203

Gingras, F., Fiset, D., Plouffe-Demers, M. P., Deschênes, A., Cormier, S., Forget, H., & Blais, C. (2023). Pain in the eye of the beholder: Variations in pain visual representations as a function of face ethnicity and culture. *British Journal of Psychology*, *114*(3), 621–637. https://doi.org/10.1111/bjop.12641

Göller, P. J., Reicherts, P., Lautenbacher, S., & Kunz, M. (2023). How gender affects the decoding of facial expressions of pain. *Scandinavian Journal of Pain*, *23*(2), 372–381. https://doi.org/10.1515/sjpain-2022-0063

Gosselin, F., & Schyns, P. G. (2002). RAP: A new framework for visual categorization. *Trends in Cognitive Sciences*, *6*(2), 70–77. https://doi.org/10.1016/S1364-6613(00)01838-6

Hadjistavropoulos, T., Craig, K. D., Duck, S., Cano, A., Goubert, L., Jackson, P. L., Mogil, J. S., Rainville, P., Sullivan, M. J. L., Williams, A. C. C., Vervoort, T., & Fitzgerald, T. D. (2011). A biopsychosocial formulation of pain communication. *Psychological Bulletin*, *137*(6), 910–939. https://doi.org/10.1037/a0023876

Hehman, E., Xie, S. Y., Ofosu, E. K., & Nespoli, G. (2018). *Assessing the point at which averages are stable: A tool illustrated in the context of person perception*. PsyArXiv https://doi.org/10.31234/osf.io/2n6jq

Hill, M. L., & Craig, K. D. (2004). Detecting deception in facial expressions of pain: Accuracy and training. *The Clinical Journal of Pain*, *20*(6), 415–422. https://doi.org/10.1097/00002508-200411000-00006

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General*, *141*(1), 19–25. https://doi.org/10.1037/a0023463

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(19), 7241–7244. https://doi.org/10.1073/pnas.1200155109

Jack, R. E., & Schyns, P. G. (2015). The human face as a dynamic tool for social communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*(6), 708–730. https://doi.org/10.1037/xge0000162

Kappesser, J., & Williams, A. C. d. C. (2002). Pain and negative emotions in the face: Judgements by health care professionals. *Pain*, *99*(1), 197–206. https://doi.org/10.1016/S0304-3959(02)00101-X

Kappesser, J., & Williams, A. C. d. C. (2010). Pain estimation: Asking the right questions. *Pain*, *148*(2), 184–187. https://doi.org/10.1016/j.pain.2009.10.007

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16. https://nyuscholars.nyu.edu/en/publications/whats-new-in-psychtoolbox-3

Krumhuber, E. G., Tamarit, L., Roesch, E. B., & Scherer, K. R. (2012). FACSGen 2.0 animation software: Generating three-dimensional FACS-valid facial expressions for emotion research. *Emotion*, *12*(2), 351–363. https://doi.org/10.1037/a0026632

Kunz, M., & Lautenbacher, S. (2014). The faces of pain: A cluster analysis of individual differences in facial activity patterns of pain. *European Journal of Pain*, *18*(6), 813–823. https://doi.org/10.1002/j.1532-2149.2013.00421.x

Kunz, M., Prkachin, K., Solomon, P. E., & Lautenbacher, S. (2021). Faces of clinical pain: Inter-individual facial activity patterns in shoulder pain patients. *European Journal of Pain*, *25*(3), 529–540. https://doi.org/10.1002/ejp.1691

Lévesque-Lacasse, A., Desjardins, M.-C., Fiset, D., Charbonneau, C., Cormier, S., & Blais, C. (2024). The relationship between the ability to infer another's pain and the expectations regarding the appearance of pain facial expressions: Investigation of the role of visual perception. *The Journal of Pain*, *25*(1), 250–264. https://doi.org/10.1016/j.jpain.2023.08.007

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for participants' viewing distance in large-scale, psychophysical online experiments using a virtual chinrest. *Scientific Reports*, *10*(1), Article 904. https://doi.org/10.1038/s41598-019-57204-1

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*(4), 330–334. https://doi.org/10.1037/h0028434

Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science*, *28*(2), 209–226. https://doi.org/10.1207/s15516709cog2802_4

McCaffery, M., Ferrell, B. R., & Pasero, C. (2000). Nurses' personal opinions about patients' pain and their effect on recorded assessments and titration of opioid doses. *Pain Management Nursing*, *1*(3), 79–87. https://doi.org/10.1053/jpmn.2000.9295

Mende-Siedlecki, P., Qu-Lee, J., Lin, J., Drain, A., & Goharzad, A. (2020). The Delaware Pain Database: A set of painful expressions and corresponding norming data. *Pain Reports*, *5*(6), Article e853. https://doi.org/10.1097/PR9.0000000000000853

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Plouffe-Demers, M.-P., Saumure, C., Fiset, D., Cormier, S., & Blais, C. (2023). Facial expression of pain: Sex differences in the discrimination of varying intensities. *Emotion*, *23*(5), 1254–1266. https://doi.org/10.1037/emo0001156

Poole, G. D., & Craig, K. D. (1992). Judgments of genuine, suppressed, and faked facial expressions of pain. *Journal of Personality and Social Psychology*, *63*(5), 797–805. https://doi.org/10.1037/0022-3514.63.5.797

Prkachin, K. M. (1986). Pain behaviour is not unitary. *Behavioral and Brain Sciences*, *9*(4), 754–755. https://doi.org/10.1017/S0140525X00052316

Prkachin, K. M. (1992). The consistency of facial expressions of pain: A comparison across modalities. *Pain*, *51*(3), 297–306. https://doi.org/10.1016/0304-3959(92)90213-U

Prkachin, K. M., & Craig, K. D. (1995). Expressing pain: The communication and interpretation of facial pain signals. *Journal of Nonverbal Behavior*, *19*(4), 191–205. https://doi.org/10.1007/BF02173080

Prkachin, K. M., Mass, H., & Mercer, S. R. (2004). Effects of exposure on perception of pain expression. *Pain*, *111*(1), 8–12. https://doi.org/10.1016/j.pain.2004.03.027

Prkachin, K. M., & Solomon, P. E. (2008). The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain*, *139*(2), 267–274. https://doi.org/10.1016/j.pain.2008.04.010

Richer, A. (2025, January 7). *Is it pain, anger, disgust or sadness? Individual differences in expectations of pain facial expression*. https://osf.io/szbm7

Riva, P., Sacchi, S., Montali, L., & Frigerio, A. (2011). Gender effects in pain detection: Speed and accuracy in decoding female and male pain expressions. *European Journal of Pain*, *15*(9), 985.e1–985.e11. https://doi.org/10.1016/j.ejpain.2011.02.006

Roesch, E. B., Tamarit, L., Reveret, L., Grandjean, D., Sander, D., & Scherer, K. R. (2011). FACSGen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units. *Journal of Nonverbal Behavior*, *35*(1), 1–16. https://doi.org/10.1007/s10919-010-0095-9

Roy, C., Blais, C., Fiset, D., Rainville, P., & Gosselin, F. (2015). Efficient information for recognizing pain in facial expressions. *European Journal of Pain*, *19*(6), 852–860. https://doi.org/10.1002/ejp.676

Saumure, C., Plouffe-Demers, M.-P., Estéphan, A., Fiset, D., & Blais, C. (2018). The use of visual information in the recognition of posed and spontaneous facial expressions. *Journal of Vision*, *18*(9), Article 21. https://doi.org/10.1167/18.9.21

Saumure, C., Plouffe-Demers, M.-P., Fiset, D., Cormier, S., Zhang, Y., Sun, D., Feng, M., Luo, F., Kunz, M., & Blais, C. (2023). Differences between East Asians and Westerners in the mental representations and visual information extraction involved in the decoding of pain facial expression intensity. *Affective Science*, *4*(2), 332–349. https://doi.org/10.1007/s42761-023-00186-1

Seers, T., Derry, S., Seers, K., & Moore, R. A. (2018). Professionals underestimate patients' pain: A comprehensive review. *Pain*, *159*(5), 811–818. https://doi.org/10.1097/j.pain.0000000000001165

Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2008). Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *Pain*, *135*(1), 55–64. https://doi.org/10.1016/j.pain.2007.05.008

Tessier, M.-H., Mazet, J.-P., Gagner, E., Marcoux, A., & Jackson, P. L. (2024). Facial representations of complex affective states combining pain and a negative emotion. *Scientific Reports*, *14*(1), Article 11686. https://doi.org/10.1038/s41598-024-62423-2

Tüttenberg, S. C., & Wiese, H. (2023). Event-related brain potential correlates of the other-race effect: A review. *British Journal of Psychology*, *114*(S1), 24–44. https://doi.org/10.1111/bjop.12591

Williams, A. C. d. C. (2002). Facial expression of pain: An evolutionary account. *Behavioral and Brain Sciences*, *25*(4), 439–455. https://doi.org/10.1017/S0140525X02000080

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personality and Social Psychology Review*, *16*(2), 116–142. https://doi.org/10.1177/1088868311418987

Zhang, L., Losin, E. A. R., Ashar, Y. K., Koban, L., & Wager, T. D. (2021). Gender biases in estimation of others' pain. *The Journal of Pain*, *22*(9), 1048–1059. https://doi.org/10.1016/j.jpain.2021.03.001

Received June 19, 2024
Revision received December 16, 2024
Accepted January 20, 2025 ■