EJP
European Journal of Pain

ORIGINAL ARTICLE

Efficient information for recognizing pain in facial expressions

C. Roy[1], C. Blais[2], D. Fiset[2], P. Rainville[3], F. Gosselin[1]

1 Département de Psychologie, Université de Montréal, Canada
2 Département de psychoéducation et de psychologie, Université du Québec en Outaouais, Gatineau, Canada
3 Département de stomatologie, Centre de recherche de l'institut universitaire de gériatrie de Montréal, and Centre de recherche sur le système nerveux central (GRSNC), Université de Montréal, Canada

Correspondence
Frédéric Gosselin
E-mail: frederic.gosselin@umontreal.ca

Funding sources
This study was funded by NSERC.

Conflicts of interest
None declared.

Accepted for publication
7 January 2015

doi:10.1002/ejp.676

Abstract

Background: The face as a visual stimulus is a reliable source of information for judging the pain experienced by others. Until now, most studies investigating the facial expression of pain have used a descriptive method (i.e. Facial Action Coding System). However, the facial features that are relevant for the observer in the identification of the expression of pain remain largely unknown despite the strong medical impact that misjudging pain can have on patients' well-being.
Methods: Here, we investigated this question by applying the Bubbles method. Fifty healthy volunteers were asked to categorize facial expressions (the six basic emotions, pain and neutrality) displayed in stimuli obtained from a previously validated set and presented for 500 ms each. To determine the critical areas of the face used in this categorization task, the faces were partly masked based on random sampling of regions of the stimuli at different spatial frequency ranges.
Results: Results show that accurate pain discrimination relies mostly on the frown lines and the mouth. Finally, an ideal observer analysis indicated that the use of the frown lines in human observers could not be attributed to the objective 'informativeness' of this area.
Conclusions: Based on a recent study suggesting that this area codes for the affective dimension of pain, we propose that the visual system has evolved to focus primarily on the facial cues that signal the aversiveness of pain, consistent with the social role of facial expressions in the communication of potential threats.

1. Introduction

The communication of pain provides an adaptive advantage by signalling immediate threats to conspecifics and by enabling the solicitation of protection, appropriate care and moral support (Williams, 2002). Vocalizations, posture, self-report and facial expression are cues an observer can rely on to detect one's pain (Hadjistavropoulos and Craig, 2002). When asked to judge someone else's pain, observers rely mainly on non-verbal components and even more so on facial expression (Poole and Craig, 1992; Williams, 2002).

The constituents of the facial expression of pain have been described in studies using the Facial Action Coding System (FACS; Ekman and Friesen, 1975). It includes brow lowering, cheek raising, lid tightening, nose wrinkling, upper lip raising and eye closing (Craig et al., 1992). Four core actions are particularly consistent: brow lowering, orbit tightening, upper lip raising/nose wrinkling and eye closure (Prkachin, 2009). This pattern of core Action Units can be differentiated from the patterns coding for the expression of the six basic emotions (Simon et al., 2008).

Some data suggest that, when judging others' pain based on facial expression, human observers tend to

What's already known about this topic?
- Specific groups of facial muscles are involved in the facial expression of pain.

What does this study add?
- This study investigates what visual information is efficiently used by the human visual system to recognize pain.
- Accurate pain discrimination relies mostly on the frown lines and the mouth.
- This visual strategy cannot be explained only by the amount of information available in these facial areas.

underestimate the intensity of the sufferer's pain when compared with their own reports; practitioners should thus be more sensitive to this bias, since it seems to result in elevated risks of under-treatment (Prkachin et al., 2007; see however Kappesser and Williams, 2010). Moreover, observers are found to be less accurate at identifying pain than other negative emotions and sometimes mistake pain for disgust, fear and anger (Kappesser and Williams, 2002). Despite the aforementioned findings and the established importance of facial expressions in estimating pain in others, little is known about the visual processes implicated in recognizing pain. Descriptive methods (e.g. FACS) allow researchers to pinpoint the facial components that are involved in the recognition of the expression of pain (Simon et al., 2008). However, they involve making indirect inferences about the information effectively used by the decoder to detect emotions.

Classification image techniques such as the *Bubbles* technique (Gosselin and Schyns, 2001) have proven to be valid and powerful research tools (e.g. Adolphs et al., 2005, 2008; Dotsch et al., 2008; Lee et al., 2011; Nielsen et al., 2006) to reveal which parts of a visual stimulus are responsible for the performance of observers in a specific categorization task. The underlying logic of these techniques is the following. If specific visual information is important for the task at hand, depriving the observer of this information (using additive noise or a mask) will strongly impair his or her performance. In contrast, depriving the observer of non-diagnostic information will not substantially alter performance. Here, observers were asked to identify the emotions expressed by professional actors from small random samples of their face. After performing thousands of trials, we correlated categorization accuracy with the available visual information (pixels that were revealed or not) allowing a direct empirical examination of the diagnostic features used by human observers in pain categorization.

2. Method

2.1 Participants

Fifty participants (17 men) took part in the experiment. All had normal or corrected to normal vision and were paid for their participation. All participants provided informed consent and received monetary compensation for their participation. The local ethics committee approved all procedures.

2.2 Material and stimuli

The stimuli were presented on a calibrated high-resolution CRT monitor with a refresh rate of 60 Hz. Experimental programs were written using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for Matlab (version 7.5; Mathworks, Inc., Natick, MA, USA). Viewing distance was such that stimuli spanned $5.72° \times 5.72°$ (256×256 pixels).

The stimuli were created from a validated database (Roy et al., 2007; Simon et al., 2008) composed of photos of 10 actors successively expressing one of seven emotions (i.e. the six basic emotions and pain) at a comparable strong intensity level or displaying a neutral expression. The six basic emotions and the neutral expression were included in the experiment so that the results would represent the information needed to differentiate pain from other emotions. These 80 grey-scaled base photos had normalized global orientation and lighting. They were also spatially aligned so that the eyes and the nose of all faces were as much as possible at the same position without distorting the faces. A grey mask with an elliptic hole was applied to each face to hide the hair and the background.

On each trial, a 'bubblized' stimulus was created as follows (see Fig. 1 for an illustration of the creation of a stimulus at a given trial). First, each image was decomposed into five spatial frequency bands of one octave each (128–64, 64–32, 32–16, 16–8 and 8–4 cycles/image or 42–85, 21–42, 11–21, 5–11, 3–5 cycles/face width; the remaining bandwidth was used as constant background) using the Laplacian pyramid (Burt and Adelson, 1983). The rationale for sampling the stimulus in different

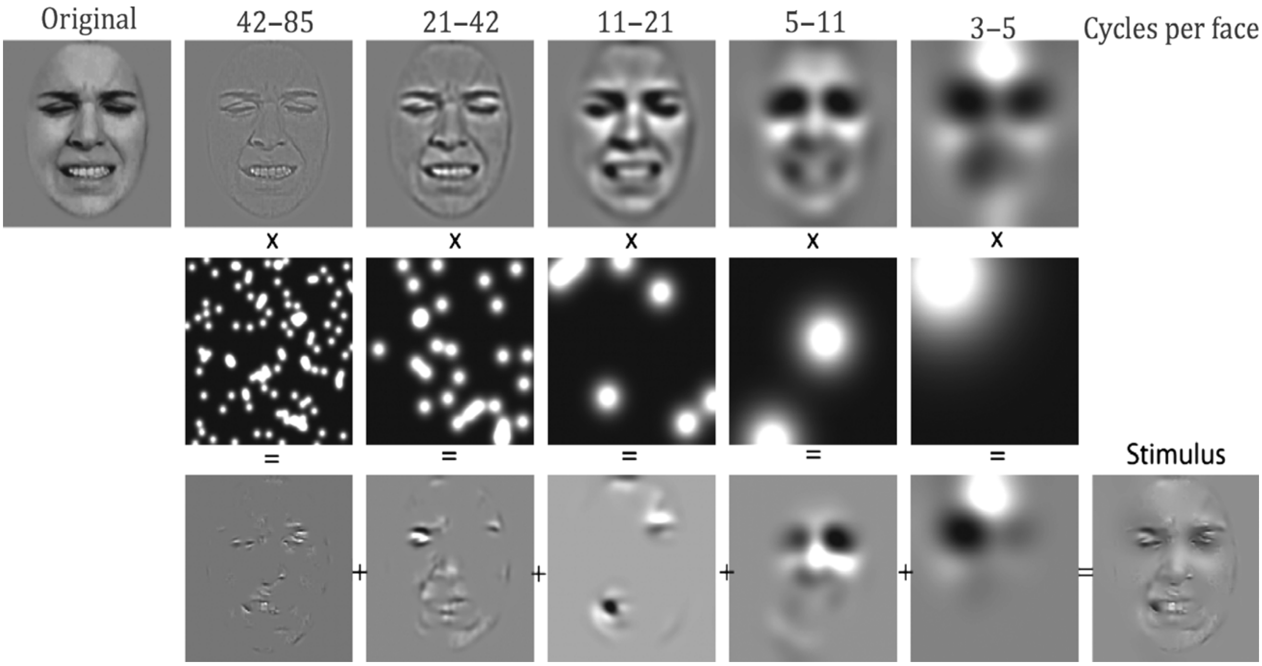

Figure 1 Stimuli creation. The original stimulus is decomposed in five spatial frequency bands (first row). Randomly positioned gaussian apertures are created for each band (second row). The size and number of apertures are adjusted for each band. The first two rows are then multiplied pixel-by-pixel (third row). The five resulting stimuli are then fused to create the final stimulus.

frequency bands is based on a dominant theory in vision. This theory proposes that the human visual system analyses the complex luminance variations that make up a visual stimulation using discrete channels, each tuned to a specific spatial scale (De Valois et al., 1974). The face information at the five scales was then sampled using an opaque mask punctured by randomly located Gaussian holes (henceforth called 'bubbles') to avoid introducing any spatial frequency artefacts. The size of the bubbles was adjusted according to frequency band so that each bubble revealed 1.5 cycles of spatial information (Fig. 1, second row). Because the size of the bubbles increased as the spatial scale became coarser, the number of bubbles differed across scales to keep the size of the sampled area constant across frequency bands. Note that the varying size of the bubbles across spatial frequency bands allowed sampling of features of various sizes ranging from very small (e.g. an eye wrinkle in high spatial frequencies) to very large (e.g. the whole face in low spatial frequencies). Finally, the information revealed by the bubbles was fused across the five frequency bands to produce an experimental stimulus (Fig. 1, third row). This resulted in a stimulus in which different facial features were revealed in different spatial frequency bands, the combinations of which varied across trials. These stimuli were then presented to the participants, and their categorization accuracy was recorded to verify which facial features in which spatial frequency band was correlated with accuracy. The combinations of features and spatial frequency bands that were correlated with accuracy indicated the information used by the participant to discriminate pain facial expressions from other emotions. For more details about the method see Gosselin and Schyns (2001, Experiment 2).

2.3 Procedure

Each trial began with a fixation cross displayed on the centre of the screen for 500 ms. A bubblized stimulus was then presented for 500 ms on the centre of the screen. The chosen stimulus duration is enough to achieve high levels of accuracy in a facial expression categorization task (Calvo and Lundqvist, 2008). Each participant completed a total of 4000 trials (25 blocks of 160 trials). Subjects were asked to identify the perceived emotion by pressing the corresponding key on the computer keyboard. No response time limit was imposed and no feedback was provided. The next trial occurred 500 ms after the participant's response.

Mean accuracy was maintained at 56.25% correct across all emotions (half-way between chance – 12.5% correct – and perfect recognition – 100% correct) by adjusting the number of bubbles (i.e. the amount of visual information revealed) on a trial-by-trial basis using an adaptive procedure (QUEST; Watson and Pelli, 1983).

2.4 Data analysis

We first calculated the unbiased hit rates (see Wagner, 1993) for each emotion, and submitted them to a repeated measure ANOVA with the eight expressions as the within-subject factor. We also calculated the hit and false alarm rates separately. A repeated measure ANOVA was performed on the hit rates with the eight expressions as the within-subject factor. Two chi-square tests were performed on the pain false alarm rates (e.g. respond it's 'pain' when, in fact, the face expresses disgust) and the pain omission rates (e.g. respond 'disgust' when, in fact, the face expresses pain) to evaluate if errors were evenly distributed. These analyses were performed with SPSS 16.0 (SPSS Inc, Chicago, IL, USA).

To pinpoint the visual information used by human observers to discriminate pain from other facial expressions, we computed a 'classification image' for each participant, each expression and each spatial frequency band. A classification image is obtained by calculating a weighted sum of all the bubbles masks presented to the participant, using the accuracy of the participant transformed into z-score values as weights. This procedure amounts essentially to a multiple linear regression on the bubbles masks and on the accuracy. The result of this analysis indicates which facial areas in each spatial frequency band are positively correlated with accurate pain recognition. The classification images were transformed into z-scores using the uninformative area around the face stimulus as a reference noise distribution. A group classification image was then computed for each emotion by summing the individual classification images and by dividing the sum by the square root of 50, i.e. the number of participants. Finally, the *Pixel* test (Chauvin et al., 2005) was applied to the group classification images to determine the critical z-score value for statistical significance ($p < 0.05$ family-wise, one-tailed). The statistical threshold provided by this test corrects for multiple comparisons while taking the spatial correlation inherent to structured images into account.

3. Results

3.1 Identification accuracy

On average, 149.2 bubbles (SD = 127.2) were necessary to maintain accuracy at 56.25%. Average accuracy across all participants was 61.7%, indicating that the algorithm used to adjust performance worked reasonably well. Table 1 displays the hit rates, unbiased hit rates, and proportion of confusion error for each pair of emotions. To summarize, happy expressions were correctly identified on 85.7% of the trials (±9.1%), followed by surprise (69.7 ± 11.5%), sadness (64.1 ± 11.2%), anger (61.5 ± 11.1%), fear (58.7 ± 12.6%), neutral (57.1 ± 10.4%), pain (49.4 ± 14.0%) and disgust (47.4 ± 12.3%). There was a significant effect of emotion on the hit rates [$F(7, 57) = 54.37$, $p < 0.001$]. This was also true for the unbiased hit rates [$F(7, 57) = 136.19$, $p < 0.001$] *Post hoc* analysis using Tukey's HSD criterion indicated that the pain hit rate was significantly lower ($p < 0.05$) than that of all other emotions, except for disgust which did not differ significantly from pain. The same *post hoc* analysis on the unbiased hit rates indicated that pain recognition was significantly lower ($p < 0.05$) than all other emotions, except for disgust and neutral expressions which did not significantly differ from pain. Therefore, in accordance with previous studies (Kappesser and Williams, 2002; Simon et al., 2008), pain exhibited one of the lowest correct recognition rates of all the emotions that we tested.

The confusability matrix (Table 1) shows the distribution of the participant's responses (hit and false alarm rates) for all eight emotions. A chi-square test of independence indicated that pain false alarm rates were not evenly distributed across other emotions ($\chi^2(6, N = 50) = 2.91$, $p < 0.001$). More specifically, *post hoc* paired t-tests with Bonferroni correction for multiple comparisons revealed that sadness and disgust were perceived as pain significantly more often than the other emotions ($p < 0.001$). Other than these two emotions, happiness was also frequently mistaken for pain more often than the other emotions ($p < 0.001$). Inversely, surprise rarely led to pain perception ($p < 0.001$). An uneven distribution of errors was also observed for the pain omission rates ($\chi^2(6, N = 50) = 3.05$, $p < 0.001$). *Post hoc* analysis with paired t-tests indicated that pain expressions were more often mistaken for disgust and for sadness than for any other emotion ($p < 0.001$). Again, pain was less often confounded with surprise than with any other emotions

Table 1 Confusability matrix depicting the proportion of responses (columns) for each target emotion presented (rows) in (A) human observers and (B) an ideal observer.

Emotion perceived	Emotion presented							
	Pain	Disgust	Fear	Happy	Neutral	Anger	Sadness	Surprise
(A)								
Pain	**0.4943 (0.297)**	0.0947	0.0365	0.1502	0.0217	0.0408	0.1533	0.0085
Disgust	0.1581	**0.4744 (0.2987)**	0.0273	0.0273	0.0451	0.1491	0.1119	0.0068
Fear	0.0326	0.0324	**0.5865 (0.3755)**	0.0183	0.0301	0.0427	0.0606	0.1968
Happy	0.0259	0.0103	0.0130	**0.8573 (0.6087)**	0.0595	0.0074	0.0184	0.0082
Neutral	0.0293	0.0365	0.0346	0.0759	**0.5707 (0.3003)**	0.0283	0.1496	0.0751
Anger	0.0348	0.0314	0.0314	0.0418	0.1387	**0.6153 (0.4104)**	0.0256	0.0153
Sadness	0.0773	0.0186	0.0186	0.0225	0.1907	0.0184	**0.6411 (0.3552)**	0.0110
Surprise	0.0077	0.1735	0.1735	0.0156	0.0622	0.0136	0.0208	**0.6972 (0.4798)**
(B)								
Pain	**0.6607 (0.4509)**	0.0557	0.0373	0.0601	0.0460	0.0581	0.0529	0.0293
Disgust	0.0434	**0.6705 (0.4230)**	0.0329	0.0509	0.0507	0.0668	0.0441	0.0407
Fear	0.0339	0.0386	**0.6699 (0.4522)**	0.0422	0.0498	0.0467	0.0498	0.0690
Happy	0.0604	0.0628	0.0371	**0.6226 (0.3900)**	0.0668	0.0559	0.0558	0.0387
Neutral	0.0455	0.0646	0.0527	0.0669	**0.5335 (0.2948)**	0.0740	0.0971	0.0658
Anger	0.0489	0.0787	0.0455	0.0595	0.0683	**0.5839 (0.3413)**	0.0624	0.0527
Sadness	0.0544	0.0488	0.0519	0.0598	0.0973	0.0649	**0.5702 (0.3326)**	0.0527
Surprise	0.0209	0.0435	0.0650	0.0320	0.0531	0.0487	0.0451	**0.6916 (0.4597)**

Hits are presented in the diagonal in bold, unbiased hits between parentheses, while omissions (rows – regular font) and false alarms (columns – regular font) are reported for each emotion in the rest of the matrix.

($p < 0.001$). This pattern of hit and false alarm rates is similar to those previously reported in studies that did not alter the appearance of the facial expressions (Kappesser and Williams, 2002; Smith et al., 2005; Roy et al., 2007; Simon et al., 2008).

3.2 Efficient information for pain identification

The red blobs in Fig. 2A are the areas of the face that were significantly correlated with the accurate discrimination of pain from other facial emotions. To summarize, participants efficiently used the information around the frown lines region at scales from 11 to 85 cycles per face width, the corners of the mouth at scales from 21 to 42 cycles per face width and the entire mouth at 11–21 cycles per face width. No information was used effectively by human observers in spatial frequency bands lower than 11 cycles per face width.

These Bubbles results explain at least qualitatively the pain false alarm rates in the emotion confusion matrix. We can think of the filters that are derived using Bubbles as lenses that the brain uses to look at stimuli. Any facial emotion that looks like pain through those pain filters, i.e. which reveal the frown lines and the corners of the mouth, should lead to pain false alarms. The frown lines of sad faces are remarkably like those of pain faces. The corners of the mouth of happy faces are almost identical to those of pain faces. Both the corners of the mouth and the frown lines of disgust faces resemble those of pain faces. Inversely, any facial emotion that does not look like pain through the pain filters should rarely lead to pain false alarms. Neither the corners of the mouth nor the frown lines of surprise faces resemble those of pain faces.

3.3 Ideal observer

The results obtained with the Bubbles method give information about the visual strategy used by the participants. This strategy may stem from the mental representations of pain the participants have encoded in memory, from the available information in the stimulus or both (Gosselin and Schyns, 2002). Mental representations are not necessarily a perfect copy of the visual world, and may imply, to some extent, an interpretation of the world in which more emphasis is put on some visual information and less on other. It is possible to verify how close the participants' strategy was to one in which they would only rely on the available information by conducting an ideal observer analysis. Because the mental representations given to the ideal observer are a perfect

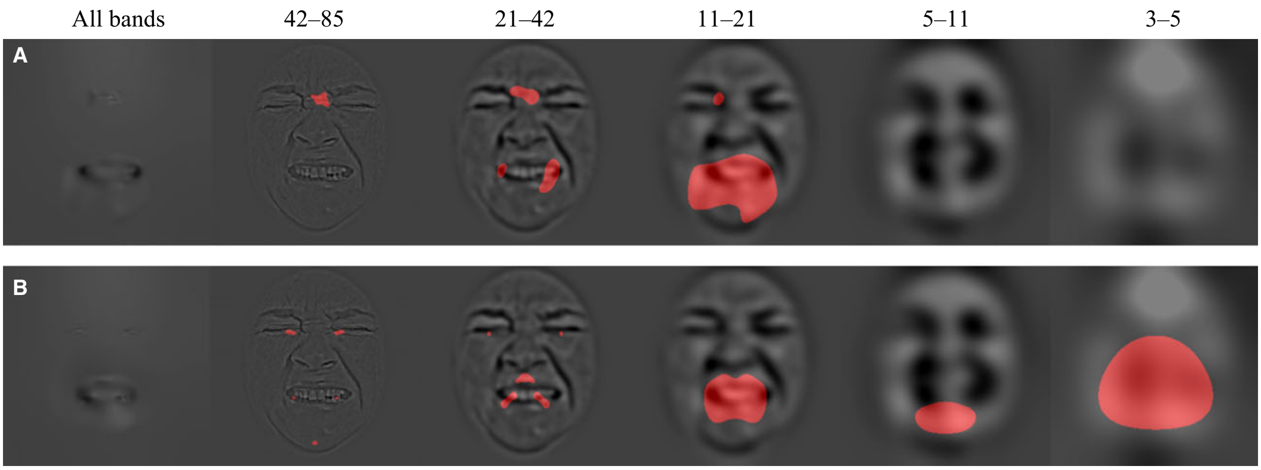

Figure 2 Classification images of the pain expression for the human observers (A) and the ideal observer (B). On the leftward image, the facial areas that were significantly correlated with the correct identification of pain in each spatial frequency band were combined to create the 'optimal' stimulus. The second image on the left up until the rightward image show, in red, the regions significantly correlated with accurate recognition in their corresponding spatial frequency band (indicated in cycle per face).

copy of the visual stimulation, its strategy only reflects the available information in the stimuli. Thus, to uncover the information that was available to our human observers to discriminate the expression of pain from the expression of other emotions, we submitted an ideal observer to the same *Bubbles* experiment as our human observers (e.g. Gosselin and Schyns, 2001). The ideal observer performed the same number of trials per emotion as human participants, and the number of bubbles was set to the average number of bubbles used by the human participants. An adjustable quantity of white Gaussian noise was added to the faces prior to sampling them with Gaussian apertures, to equate human and model performance (i.e. 56.25%). On each trial, the model determined the Pearson's correlation between the sparse input (i.e. the noisy face revealed by bubbles) and each of the 80 base photos as revealed by the same bubble mask. The categorization response was the emotion of the face with the highest correlation with the stimulus. In other words, the ideal observer categorization answer ('pain', 'anger', etc.) corresponded to the emotion expressed by the face in the whole databank that was the most visually similar to the stimulus in the trial when 'viewing' only parts of the expression revealed by the bubbles on this particular trial.

The confusability matrix of the ideal observer is presented in Table 1b. Interestingly, pain has one of the highest hit rates for the ideal observer, whereas it has one of the lowest for human observers. Furthermore, the false alarm and omission rates of the ideal observer are homogenous across emotions (respectively $\chi^2(6, N = 1) = 1.60$, $p = 0.95$ and $\chi^2(6, N = 1) = 2.67$, $p = 0.85$) unlike for human observers.

Next, to reveal directly the information available to discriminate the expression of pain from the expression of other emotions, we computed the classification images of the ideal observer using the same procedure as described in section 2.4. Fig. 2B shows the result of this analysis. Again, red blobs correspond to regions that attained statistical significance. To summarize, the ideal observer used most effectively the information contained in the eye region at high spatial frequencies (scales from 21 to 85 cycles per face width), in parts of the lips at mid-to-high spatial frequencies (scales from 21 to 42 cycles per face width), in the mouth region at lower frequencies (scales from 3 to 21 cycles per face width) and in the nose region at very low spatial frequencies (scales of 3–5 cycles per face width). The differences between the classification images of the human observers and that of the ideal observer imply that human observers are not entirely signal driven, and that they are in part influenced by the mental representation of pain they encoded in memory.

4. Discussion

4.1 Main results and interpretation

The main goal of the present study was to identify the visual information used effectively by healthy human observers to discriminate the facial expression of pain from the facial expression of other emotions. To achieve this aim, we used the *Bubbles* method, which asked observers to discriminate facial emotions from randomly sampled regions of a face. We found that accurate pain discrimination relied mostly on the frown lines region (between 11 and 85 cycles per face width), on the corners of the mouth (between 21 and 42 cycles per face width) and on the entire mouth (between 11 and 21 cycles per face width). No information was used effectively in spatial frequency bands lower than 11 cycles per face width. So, among the facial action units identified in the literature as being present in the expression of pain (nose wrinkling, upper lip raising, brow lowering, cheek raising, lid tightening and eye closing; Prkachin and Craig, 1995), human observers seem to rely only on two: the wrinkles between the eyes and the mouth.

Do human observers use these facial cues because they contain more information for the task at hand than any other facial cue? A rigorous way to characterize the information available to discriminate the expression of pain from the expression of other emotions is to perform an 'ideal observer' analysis. The most informative facial cues for the task are the eyes (from 21 to 85 cycles per face width), parts of the lips (from 21 to 42 cycles per face width), the mouth (from 3 to 21 cycles per face width) and the nose (from 3 to 5 cycles per face width). Thus, both the human observers and the ideal observer use the mouth area, although only the ideal observer uses the mouth area in very low spatial frequencies.

The higher sensitivity of the human visual system to middle spatial frequencies (De Valois et al., 1974) might explain why the human observers did not use the mouth in the lower spatial frequencies. However, the most intriguing finding is the use of frown lines by human observers in the three highest spatial scales despite this facial cue not being used at all by the ideal observer. One potential explanation comes from a study by Kunz et al. (2012). These researchers proposed that the facial expression of pain might encode both the sensory and the affective dimensions of pain. The sensory dimension of pain refers to the quality, intensity and spatio-temporal characteristics of the sensation, whereas its affective dimension refers to the negative valence and aversiveness inherent to the painful sensation (Melzack and Eisenberg, 1968). On the one hand, Kunz and colleagues showed that increasing the affective component of the sufferer's pain experience through suggestions specifically enhances the activation of the *levator labii superioris* muscle (responsible for lip raising) and of the *corrugator* muscle (responsible for the contraction of the eyebrows) in the pain facial expression. On the other hand, they showed that increasing the sensorial aspect of the participants' experience specifically enhances the activation of the *orbicularis oculi muscle* (responsible for the contraction of the muscles surrounding the eyes). In the light of these results, the use of the frown lines by our participants could possibly mean that human observers are tuned to the sufferer's emotional nociceptive experience more than their sensorial experience. This effect contrasts with the pattern observed in the ideal observer where the inferior part of orbicularis did contribute to optimal performance (see Fig. 2). This implies that the sensorial cue of the pain expression was discriminant in our task, but not used by the human observer; and that the affective cue of the expression pain was not discriminant, but was used by human observers. In other words, our results may indicate that more emphasis was put on the affective than on the sensorial visual cue in the participants' mental representations of pain, thus leading them to use more the former than the latter while they attempt to discriminate pain from other expressions. Obviously, this suggestion is speculative and needs some direct empirical support.

The utilization of the mouth area by both the ideal observer and the human observers may appear incongruent with the finding that the appearance of the mouth area is unreliably linked to pain (Prkachin and Solomon, 2009). In the present study, the expressions from a limited set of 10 actors were used, and the finding that the mouth was used by the ideal observer indicates that this area was very informative to discriminate pain from the other emotions expressed by these actors. It is possible that the set of stimuli did not perfectly reflect the individual variations of pain expressions, and that the reliance on the mouth area would have been lower using a set of stimuli comprising more individuals. However, a recent study suggests that modifications in the appearance of the mouth area (e.g. lip raising and mouth opening) during pain expression are quite frequent, and are in fact present in the first two clusters out of four when a cluster analysis is performed on the pain facial expressions across a large sample of individuals (Kunz and Lautenbacher, 2014).

4.2. Methodological implications

The present study proposed an alternative tool to descriptive methods to investigate the visual properties of the facial expression of pain. Description of the facial movements activated in another person's face relies on indirect assumptions and is therefore imprecise when it comes to understanding the observer's visual information processing. Notably, such techniques do not take into account the mental representation of observers and their impact on visual processing. In previous studies, the 'Bubbles' technique allowed significant findings in clinical populations about emotional facial expressions processing (e.g. Adolphs et al., 2008; Lee et al., 2011). More interestingly, literature has shown that identifying visual cues allowing correct emotion recognition in normal population can help normalize patient performances. For example, Adolphs et al. (2005) have described the case of S.M, a patient with bilateral amygdala damage who could not recognize fear in others' face. She was found to use an atypical visual strategy to extract information in facial expressions when compared to peers; she did not spontaneously look at the eyes when asked to categorize facial expressions, although they are critical for identifying fear. Explicitly suggesting to S.M to look at this region when looking at someone's face normalized her performance. Such evidences suggest that investigating visual processing of the facial expression of pain with methods that directly assess information extraction is critical for a complete understanding of correct and incorrect pain identification and judgment.

4.3 Limitations and further perspectives

The stimuli used here have been taken from a databank of professional actors expressing strong pain. Further validation of the present results with spontaneous facial expression of pain varying in their intensity is required. In fact, as a first step to understand the visual processes underlying pain recognition, we chose to use professionally simulated emotions to control the identity of the expressers (the same person expressing all emotions), the emotion intensity (difficult to control in natural settings), purity of emotions (the less mixture possible) and visual properties (luminance, head position, etc.). Of course, this methodological choice limits the generalization of the results. Knowing that the intensity of some facial actions differ in spontaneous and simulated expressions (Littlewort et al., 2009), one could argue that the visual information identified here would have been different with spontaneous pain expressions. This certainly needs further investigation. Similarly, although many studies in the face perception field using the Bubbles method have led to results congruent with studies using a different methodology, it remains possible that the method interacted with the normal visual strategy used to recognize an expression of pain.

Moreover, validating the cues identified here in other types of categorization tasks (e.g. judging the intensity of the pain) would be helpful in determining to what extent the results could be applied to clinical contexts. Knowing that static and dynamic facial expression differ in their treatment (Haxby et al., 2000; Adolphs et al., 2003; Kilts et al., 2003; LaBar et al., 2003), that the rhythm of the expression deployment is of central importance in the recognition process (Kamachi et al., 2001) and that temporal information of pain expression can help discriminate real from faked pain (Hill and Craig, 2002), further studies should use dynamic stimuli.

Finally, one should also consider that the acknowledgement of others' pain is a far more complex phenomenon than the identification of the facial expression and rely on complex considerations about personal and contextual factors in both the sufferer and the observer. Further studies should aim to consider this process in a broader psychosocial context.

Author contributions

All authors were implicated from the inception of the study to the redaction of the manuscript. They all discussed the results and commented on the manuscript.

References

Adolphs, R., Tranel, D., Damasio, A.R. (2003). Dissociable neural systems for recognizing emotions. *Brain Cogn* 52, 61–69.
Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., Damasio, A.R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature* 433, 68–72.
Adolphs, R., Spezio, M.L., Parlier, M., Piven, J. (2008). Distinct face-processing strategies in parents of autistic children. *Curr Biol* 18, 1090–1093.
Brainard, D.H. (1997). The psychophysics toolbox. *Spat Vis* 10, 433–436.
Burt, P., Adelson, E. (1983). The Laplacian pyramid as a compactimage code. *IEEE Trans Commun* 31, 532–540.
Calvo, M.G., Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behav Res Methods* 40, 109–115.
Chauvin, A., Worsley, K.J., Schyns, P.G., Arguin, M., Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *J Vis* 5, 659–667.
Craig, K.D., Prkachin, K.M., Grunau, R.V.E.. (1992). The facial expression of pain. In *Handbook of Pain Assessment*, D. C. Turk & R. Melzack, eds. (New York, NY, USA: Guilford Press) pp. 257–276.
De Valois, R.L., Morgan, H., Snodderly, D.M. (1974). Psychophysical studies of monkey vision. 3. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Res* 14, 75–81.
Dotsch, R., Wigboldus, D.H., Langner, O., van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychol Sci* 19, 978–980.

Ekman, P., Friesen, W.V. (1975). *Unmasking the Face: A Guide to Recognizing Emotions From Facial Clues*. (Oxford, England: Prentice-Hall).

Gosselin, F., Schyns, P.G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Res* 41, 2261–2271.

Gosselin, F., Schyns, P.G. (2002). RAP: A new framework for visual categorization. *Trends Cogn Sci* 6, 70–77.

Hadjistavropoulos, T., Craig, K.D. (2002). A theoretical framework for understanding self-report and observational measures of pain: A communications model. *Behav Res Ther* 40, 551–570.

Haxby, J.V., Hoffman, E.A., Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci* 4, 223–233.

Hill, M.L., Craig, K.D. (2002). Detecting deception in pain expressions: The structure of genuine and deceptive facial displays. *Pain* 98, 135–144.

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception* 30, 875–887.

Kappesser, J., Williams, A.C. (2002). Pain and negative emotions in the face: Judgements by health care professionals. *Pain* 99, 197–206.

Kappesser, J., Williams, A.C. (2010). Pain estimation: Asking the right questions. *Pain* 148, 184–187.

Kilts, C.D., Egan, G., Gideon, D.A., Ely, T.D., Hoffman, J.M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage* 18, 156–168.

Kunz, M., Lautenbacher, S. (2014). The face of pain: A cluster analysis of individual differences in facial activity patterns of pain. *Eur J Pain* 18, 813–823.

Kunz, M., Lautenbacher, S., Leblanc, N., Rainville, P. (2012). Are both the sensory and the affective dimensions of pain encoded in the face? *Pain* 153, 350–358.

LaBar, K.S., Crupain, M.J., Voyvodic, J.T., McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cereb Cortex* 13, 1023–1033.

Lee, J., Gosselin, F., Wynn, J.K., Green, M.F. (2011). How do schizophrenia patients use visual information to decode facial emotion? *Schizophr Bull* 37, 1001–1008.

Littlewort, G.C., Bartlett, M.S., Lee, K. (2009). Automatic coding of facial expressions displayed during posed and genuine pain. *Image Vis Comput* 27, 1741–1844.

Melzack, R., Eisenberg, H. (1968). Skin sensory afterglows. *Science* 159, 445–447.

Nielsen, K.J., Logothetis, N.K., Rainer, G. (2006). Discrimination strategies of humans and rhesus monkeys for complex visual displays. *Curr Biol* 16, 814–820.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis* 10, 437–442.

Poole, G.D., Craig, K.D. (1992). Judgments of genuine, suppressed, and faked facial expressions of pain. *J Pers Soc Psychol* 63, 797–805.

Prkachin, K.M. (2009). Assessing pain by facial expression: Facial expression as nexus. *Pain Res Manag* 14, 53–58.

Prkachin, K., Craig, K. (1995). Expressing pain: The communication and interpretation of facial pain signals. *J Nonverbal Behav* 19, 191–205.

Prkachin, K.M., Solomon, P.E. (2009). The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain* 139, 267–274.

Prkachin, K.M., Solomon, P.E., Ross, J. (2007). Underestimation of pain by health-care providers: Towards a model of the process of inferring pain in others. *Can J Nurs Res* 39, 88–106.

Roy, S., Roy, C., Fortin, I., Ethier-Majcher, C., Belin, P., Gosselin, F. (2007). A dynamic facial expression database. *J Vis* 7, 944.

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, 55–64.

Smith, M.L., Cottrell, G.W., Gosselin, F., Schyns, P.G. (2005). Transmitting and decoding facial expressions. *Psychol Sci* 16, 184–189.

Wagner, H.L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *J Nonverbal Behav* 17, 3–28.

Watson, A.B., Pelli, D.G. (1983). QUEST: A Bayesian adaptive psychometric method. *Percept Psychophys* 33, 113–120.

Williams, A.C.. (2002). Facial expression of pain: An evolutionary account. *Behav Brain Sci* 25,439–455; discussion 455–488.

