Pain xxx (2007) xxx–xxx

PAIN

www.elsevier.com/locate/pain

Recognition and discrimination of prototypical dynamic expressions of pain and emotions

Daniela Simon [a,e,*], Kenneth D. Craig [b], Frederic Gosselin [c], Pascal Belin [d], Pierre Rainville [e]

[a] *Department of Clinical Psychology, Humboldt University of Berlin, Germany*
[b] *Department of Psychology, University of British Columbia, Vancouver, Canada*
[c] *Departement de Psychologie, Université de Montréal, Montréal, Canada*
[d] *Department of Psychology, University of Glasgow, Glasgow, UK*
[e] *Departement de Stomatologie, Université de Montréal, Montréal, Canada*

Received 19 February 2007; received in revised form 12 April 2007; accepted 2 May 2007

Abstract

Facial expressions of pain and emotions provide powerful social signals, which impart information about a person's state. Unfortunately, research on pain and emotion expression has been conducted largely in parallel with few bridges allowing for direct comparison of the expressive displays and their impact on observers. Moreover, although facial expressions are highly dynamic, previous research has relied mainly on static photographs. Here we directly compare the recognition and discrimination of dynamic facial expressions of pain and basic emotions by naïve observers. One-second film clips were recorded in eight actors displaying neutral facial expressions and expressions of pain and the basic emotions of anger, disgust, fear, happiness, sadness and surprise. Results based on the Facial Action Coding System (FACS) confirmed the distinct (and prototypical) configuration of pain and basic emotion expressions reported in previous studies. Volunteers' evaluations of those dynamic expressions on intensity, arousal and valence demonstrate the high sensitivity and specificity of the observers' judgement. Additional rating data further suggest that, for comparable expression intensity, pain is perceived as more arousing and more unpleasant. This study strongly supports the claim that the facial expression of pain is distinct from the expression of basic emotions. This set of dynamic facial expressions provides unique material to explore the psychological and neurobiological processes underlying the perception of pain expression, its impact on the observer, and its role in the regulation of social behaviour.

Keywords: Facial expression; Pain; Emotion; FACS; Recognition; Valence; Arousal

1. Introduction

At the end of the 19th century, Darwin recognized the critical importance of facial expressions to communicate emotional states (Darwin, 1872). A recent evolutionary perspective further emphasized that the function of pain expression may be to alarm onlookers in situations of direct threat and/or elicit solicitous behaviour (Williams, 2002). This dual function may be at least partly distinctive from those of basic emotions and the facial expression of pain might arguably be more important for the survival of the species. However, much less is known about pain compared to emotion expression.

Several studies using the Facial Action Coding System (FACS) (Ekman et al., 2002) have reliably identified the occurrence of certain combinations of facial muscles, contractions, or facial action units (AUs), across various

* Corresponding author. Address: Department of Clinical Psychology, Humboldt University of Berlin, Rudower Chaussee 18, 12489 Berlin, Germany. Tel.: +49 30 2093 4823; fax: +49 30 2093 4859.
E-mail address: daniela.simon@psychologie.hu-berlin.de (D. Simon).

0304-3959/$32.00
doi:10.1016/j.pain.2007.05.008

Please cite this article in press as: Simon D et al., Recognition and discrimination of prototypical dynamic ..., Pain (2007), doi:10.1016/j.pain.2007.05.008

acute clinical pain conditions (e.g. Prkachin, 1992). Based on systematic comparisons between studies, pain expression has been described to be unique and distinct from the six prototypical facial expressions of basic emotions (Kappesser and Williams, 2002; see Table 1). While the pain face has been investigated in the past especially in the context of pain assessment (Craig et al., 2001), it has recently become an area of interest in functional neuroimaging (e.g. Botvinick et al., 2005; Simon et al., 2006). This expansion of research on pain communication to the neurobiological domain raises new questions on the specificity of the neural systems responsive to pain and emotion expression. However, this innovative research would be incomplete without a clear demonstration that pain expression can be recognized and discriminated from basic emotions in a within-study design using standardized validated stimuli.

Another aspect that has been neglected in many previous studies is the dynamic nature of facial expression. Indeed, most experimental studies on facial expressions – including those on the facial expression of pain – have been conducted using static facial displays (e.g. Ekman and Friesen, 1976). However, facial movements have been shown to contribute to the identification of facial expression (Harwood et al., 1999; O'Toole et al., 2002; Roark et al., 2003), and discrimination is significantly improved when dynamic properties are available (Ehrlich et al., 2000; Ambadar et al., 2005). Thus, to improve the validity of stimuli used in research on pain expression, there is a need for dynamic stimuli. To date no standardized and validated set of such stimuli containing both pain and basic emotions has been made available to the research community. Studies investigating responses to dynamic facial expressions either used computer-based morphs (e.g. LaBar et al., 2003; Sato et al., 2004) or sets of movie clips comprising some but not all basic emotions (Kilts et al., 2003; Wicker et al., 2003). None of those studies included pain expression. The aim of the present study was to produce and validate a standardized set of dynamic clips of facial expressions of pain and the six basic emotions. We hypothesized that prototypical pain expressions can be readily recognized by normal volunteers and that the discrimination with the basic emotions would reveal high sensitivity and specificity.

2. Method

2.1. Participants

Fifteen healthy volunteers (11 males and 4 females, mean age: 24.1 ± 3.4) were recruited on the campus of the University of Montreal to participate in a study on the perception of facial expressions. All participants provided informed consent and received monetary compensation for their participation (25 CA$). All procedures were approved by the local Ethics Committee.

2.2. Stimuli

Drama students of different theatrical schools in Montréal were hired for the production of the stimuli following the procedure described in Fig. 1. Initially, 11 actors took part in the recording but only eight were used to create this set of stimuli (4 males and 4 females; mean age: 24.4 ± 7.5 y.o.). Selected actors were those who produced the most unambiguous facial expressions as described below. All actors provided written informed consent, transferring the copyright of the produced material to the research group. Their participation was compensated with 100 CA$. The recording sessions took place with a seamless chroma-key blue paper screen background and two diffuse tungsten light sources. In order to minimize contamination of the stimuli by special facial cues, selected actors did not have facial piercing, moustache or beard, and did not wear earrings and make-up during the recording. If necessary they were asked to put on a hairnet. The actors sat comfortably about 1.5 m away from the camera lens.

Actors were given instructions describing the procedure and guiding them to express acute pain in four intensities (mild, moderate, strong, extreme) and six basic emotions (happiness, disgust, fear, anger, sadness and surprise) in three intensities (moderate, strong, extreme). Actors also produced a neutral face as a control. Actors were asked to produce each expression in about 1 s starting with a neutral face and ending at the peak of the expression. They were allowed to include vocalisations but were asked to restrict those to the interjection "Ah!". The actor's performance was monitored online by the filming team positioned outside of the recording cabin. The filming team consisted of one graduate student and one research assistant, who were familiar with the FACS manual and the prototypical combination of facial AUs involved in each intended facial expression (note that detailed FACS coding was done later and independently by two trained coders, as described below; see Section 2.4).

Prior to the recordings, the actors were trained using an instruction guide encouraging them to imagine personal situations to evoke pain and each emotion (Fig. 1, Recording session). Descriptions of short scenes were provided as examples to support vivid imagination of the different emotional states they were asked to display. However, the large majority of the clips (about 90%) were produced using mental imagery of personal emotional situations. If necessary, the actors were also shown photographs of prototypical emotional facial expressions. If the filming team still saw discrepancies between depicted and expected facial expressions, discrete muscles were trained as described by Ekman and Friesen (1975) and Ekman et al. (2002). The performance was repeated until the filming team was convinced that the criteria were met for each facial expression of emotion, as described by Ekman and Friesen (1975) and for pain expression, as described by Williams (2002). At least 2–3 good streams were recorded for each level and in each condition. Given the considerable volume of film clips produced with each actor, a thorough FACS analysis could not be performed online and the filming team primarily relied on their immediate recognition of the target emotions. However, their decision to include a clip in the set was also informed by a list of easily detectable AUs (as per the FACS manual: Ekman et al., 2002), for the online identification of 'expected pain and emotion expressions' (Pain: AUs

Please cite this article in press as: Simon D et al., Recognition and discrimination of prototypical dynamic ..., Pain (2007), doi:10.1016/j.pain.2007.05.008

Table 1
Occurrence rate of AUs relative to the prototypical expression (■ and □) of pain and basic emotions

Emotion	AU 1 Inner brow raiser	2 Outer brow raiser	4 Brow lower	5 Upper lid raiser	6 Cheek raiser	7 Lid tighten	9 Nose wrinkler	10 Upper lip raiser	11 Naso. furrow deepen	12 Lip corner puller	15 Lip corner depressor	16 Lower lip depressor	17 Chin raiser	20 Lip stretch	22 Lip funnel	23 Lip tighten	24 Lip pressor	25 Lips part	26 Jaw drop	27 Mouth stretch	# of target AUs[b]	% of target AUs[c]
Pain prototype			■		■	■	■	■		□				□				□	□	□	6.88	92.3%
Observed	2/8	0/8	8/8[a]	0/8	8/8[a]	8/8[a]	3/8	8/8[a]	0/8	3/8	0/8	3/8	1/8	8/8[a]	0/8	0/8	0/8	8/8[a]	0/8	1/8	(0.30)	
Happiness prototype					■					■											2.00	59.2%
Observed	1/8	1/8	0/8	0/8	8/8[a]	3/8	0/8	0/8	0/8	8/8[a]	0/8	0/8	0/8	3/8	0/8	0/8	0/8	4/8[a]	1/8	0/8	(0.00)	
Anger prototype			■	■		■		□					□		■	■	□	■	■		3.38	89.6%
Observed	0/8	0/8	6/8[a]	3/8	1/8	5/8[a]	1/8	1/8	0/8	0/8	0/8	0/8	3/8	0/8	0/8	6/8[a]	3/8	1/8	0/8	0/8	(0.53)	
Disgust prototype							■ and/or	■	□	□	□					□	□				4.38	59.7%
Observed	0/8	0/8	7/8[a]	0/8	4/8[a]	7/8[a]	3/8	7/8[a]	0/8	0/8	8/8[a]	2/8	7/8[a]	5/8[a]	0/8	1/8	0/8	7/8[a]	1/8	0/8	(0.32)	
Fear prototype	■	■	■	■										□			■ and/or	■ and/or	■		4.63	85.4%
Observed	4/8[a]	3/8	5/8[a]	8/8[a]	1/8	2/8	0/8	1/8	0/8	0/8	0/8	4/8[a]	0/8	2/8	0/8	0/8	0/8	8/8[a]	5/8[a]	0/8	(0.42)	
Surprise prototype	■	■		■															□	□	3.50	81.3%
Observed	7/8[a]	8/8[a]	0/8	5/8[a]	0/8	0/8	0/8	0/8	0/8	0/8	0/8	0/8	0/8	0/8	1/8	0/8	0/8	7/8[a]	6/8[a]	1/8	(0.27)	
Sadness prototype	■		■		□				□		■		□					□	□		4.25	83.8%
Observed	6/8[a]	2/8	7/8[a]	0/8	1/8	2/8	0/8	3/8	0/8	0/8	8/8[a]	0/8	7/8[a]	2/8	0/8	0/8	0/8	3/8	2/8	0/8	(0.49)	

Note. ■ AU which occurs with prototype; □ may occur with prototype and/or major variant [according to FACS investigators' guide and Kappesser and Williams (2002)]; /8: number of actors who activated a certain AU with an intensity of ⩾b (b = slight evidence of facial action).

[a] Occurrence rate ⩾4/8.

[b] Mean number (±SEM) of target AUs activated per clip in each condition.

[c] Percentage of the observed AUs that were part of the prototypes or its major variants in each condition.

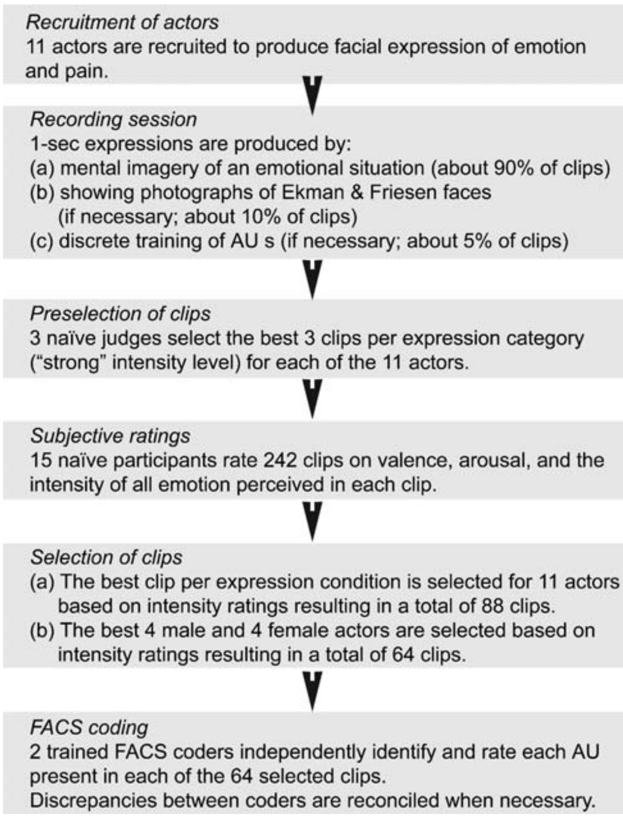

Fig. 1. Flow chart illustrating the method of stimuli production and validation.

4/6/7/10; Happy: AUs 6/12; Anger: AUs 4/7/23; Disgust: AU 10; Fear: AUs 1/4/5/25; Surprise: AUs 1/2/26; and Sadness: AUs 1/4/15). The movie streams were captured in colour by a Canon XL1S video camera and directly transferred and saved to a Pentium computer for off-line editing using Adobe Premiere 6.5. Each recording session lasted about 1.5 h.

In this report we selected stimuli of the "strong" intensity level, as pain and the emotions displayed appeared to be less ambiguous and were rarely judged to be exaggerated at that level. In a pre-rating session three naïve judges (1 male, mean age: 25.3 ± 0.6) independently selected the three best sequences per actor for each expression category. Using Adobe Premiere 6.5 and QuickTime Pro, the clips were examined frame-by-frame by one judge to identify the frame in which facial movement reached its maximum. Clips were then cut backward to assure that the peak expression was always captured within the 1-s clip (image size: 720×480 pixel, frame rate = 29.97 frames per second; mean inter-ocular distance = 100.6 ± 4.5 pixels). Those features were imposed to facilitate their use in experimental studies in which the duration of the stimuli may be critical (e.g. event-related activation in brain imaging experiments). Admittedly, the disadvantage of this choice is that the 1-s duration does not capture the full extent of some dynamic expressions. More specifically, in some clips, the onset of the expression may not match precisely with the onset of the clip (i.e. in some clips the onset to apex slightly exceeded 1 s), and the offset of the expression was excluded from the clips.

2.3. Subjective ratings by naïve observers

All 15 judges participated in one group rating session that took place in a meeting room at the University of Montréal (Fig. 1, Subjective ratings). Participants were trained with the rating scales prior to the session using distinct stimulus material and were asked not to interact with each other during the rating procedure. They then viewed each selected film clip twice and were asked to judge each on three different scales. After the first presentation, participants evaluated valence and arousal on a 9-point Likert scale. Participants were instructed to "*rate how the person in the clip might feel*: *with respect to valence*: *−4 = clearly unpleasant to +4 clearly pleasant*; and *arousal*: *−4 = highly relaxed to +4 = high level of arousal*". Information about valence and arousal was included to allow for the stimuli's use in experimental studies inspired by the dimensional model of emotion (e.g. Lang et al., 1993). Neuroimaging studies have shown that responses of some brain areas are crucially influenced by valence and/or arousal of stimuli (e.g. the amygdala; see Zald, 2003), underlining the importance of controlling for those dimensions in studies investigating emotions. After the second presentation, participants rated each facial expression with respect to the intensity of happiness, disgust, fear, anger, sadness, surprise and pain on a 6-point Likert scale. Participants were instructed to "*rate the intensity of each emotion in the clip from 0 = not at all to 5 = the most intense possible*". Each clip was therefore rated on all emotional categories.

Based on the emotion intensity ratings, the clip with the lowest mean intensity on all non-target emotions was selected for each actor and emotion. The final set comprised 64 one-second clips with each of the 8 actors contributing one clip to each of the eight conditions (Fig. 2).

2.4. Facial action coding system

FACS coding was performed on the 64 selected clips (Fig. 1, FACS coding). This procedure offered the opportunity to compare the results of these facial expressions with the prototypes reported in the literature (Chapter 12, p. 174, Table 1, Ekman et al., 2002; Williams, 2002). All 64 clips were evaluated by two independent accredited coders (1 male and 1 female), who were blind to the target expression in each clip. The FACS ratings comprise information about the occurrence and intensity (a = trace of the action to e = maximum evidence) of each AU. In the first stage of coding, the two coders independently identified the AUs present in each clip and rated their intensity. In the second stage, the differences in coding were reconciled across the two coders to provide a single set of AUs with their intensity for each clip.

2.5. Data analysis

Average ratings were first calculated across subjects for each clip and each rating scale, and the mean and SEM ratings were computed across clips within each expression condition. The frequency of correct classification was calculated for each expression category based on the highest emotion intensity rating to provide an index of sensitivity:

Please cite this article in press as: Simon D et al., Recognition and discrimination of prototypical dynamic ..., Pain (2007), doi:10.1016/j.pain.2007.05.008

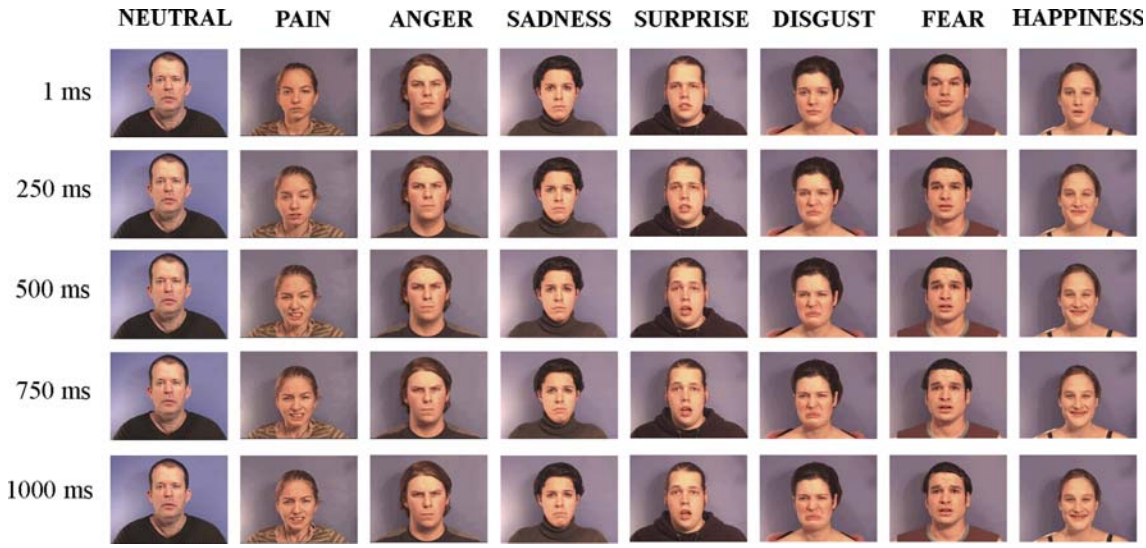

Fig. 2. Examples of four female and four male actors expressing the six basic emotions, pain and neutral taken from the set of 64 selected clips. Five time points in the clip (1, 250, 500, 750, 1000 ms) are displayed.

$$[\text{Hitrate} = \%(\#\text{N of correct classification}) / (\#\ \text{N of correct classification} + \#\text{N of "miss"})].$$

The frequency of correct rejection of stimuli not in the target expression was also determined to provide an index of the specificity of judgements for each expression category:

$$[\text{Specificity} = \%(\#\text{N of correct rejection}) / (\#\text{N of correct rejection} + \#\text{N of "false alarms"})].$$

Additionally, the presence of non-target expression was assessed by calculating a "mixed-emotion index" (mean of the ratings for the non-target expressions) for each condition.

A cluster analysis was computed to insure that each clip was adequately recognized and that clear boundaries existed between the different categories of facial expression (Squared Euclidean distances; Ward-method). The analysis was performed on the 64 clips using the seven intensity ratings (one for each possible expression) averaged across participants. Inter-judge reliability was determined by computation of Cronbach's Alpha. In order to determine whether a portrayed emotion was rated higher on the corresponding scale than on the non-target scales, Fisher's protected least significance difference test was calculated for each face category. Subsequent analyses included repeated measures analysis of variance (ANOVA) on the participants' ratings (intensity, mixed-emotion index, valence and arousal) with factors expression (pain and six basic emotions) and sex of actor (male/female). A Greenhouse-Geisser correction was used for computation of the statistical results ($p \leqslant .05$). Main effects of expression underwent post hoc analysis (multiple comparisons on eight emotional categories) using a Bonferroni correction ($p \leqslant .05/36 = .0013$).

The evaluation of the FACS coder was performed in two successive stages. During the first stage both coders provided independent lists and intensity ratings of the observed AUs. The frequency reliability was determined by 2x the number of agreements divided by the total number of AUs coded by both raters. The intensity reliability was determined by 2× the number of intensity agreements ±1, divided by the total number of AUs on which the coders agreed. During the second stage an agreed set was determined by reconciling the disagreements. The criteria for the reconciliation are provided by the FACS manual (Ekman et al., 2002, p. 88). The reported results refer to this agreed set.

The number of actors showing some involvement of an AU was determined for each expression condition. Additionally, the mean number and % of target AUs observed in each expression category were determined. All AUs that were activated at an intensity level of 'b' (indicating slight evidence of facial action) or more in at least 50% of the actors were considered to be representative of the expression condition (note that level 'a' indicates only a trace). The lists of observed AUs in each category were compared with those previously reported for prototypical emotional facial expression or their major variants (Ekman et al., 1982). While a prototype refers to a configuration of AUs commonly associated with a certain emotion, major variants constitute partial expressions of a prototype (Smith and Scott, 1997). Moreover, in order to evaluate sex differences in facial expression, repeated measurement ANOVAs with within-subjects factor 'AU' and between group-factor 'sex of the actor' were computed for each emotion ($p \leqslant .05/7 = .007$).

3. Results

3.1. Subjective ratings

Analysis of inter-rater reliability revealed a very high agreement of the 15 observers for intensity, valence and arousal ratings (Cronbach's Alpha = .97).

3.1.1. Intensity

The average ratings across participants for each portrayed expression on the corresponding target scales are reported in Table 2. Pairwise comparisons revealed that

Table 2
Intensity rating, sensitivity (Hit rate) and specificity (Correct rejection rate) by expression category

Target expression	Expression Perceived								
	Pain	Happiness	Anger	Disgust	Fear	Surprise	Sadness	Hit rate (%)	Correct rejection rate (%)
Pain	**2.68**[a] **(0.25)**	0.11 (0.06)	0.03 (0.01)	0.62 (0.15)	0.53 (0.21)	0.43 (0.13)	0.27 (0.11)	74.17 (4.95)	96.18 (1.00)
Happiness	0.00 (0.00)	**3.34**[a] **(0.14)**	0.00 (0.00)	0.00 (0.00)	0.02 (0.02)	0.18 (0.07)	0.01 (0.01)	100 (0.00)	96.98 (1.62)
Anger	0.03 (0.02)	0.00 (0.00)	**2.97**[a] **(0.18)**	0.13 (0.07)	0.04 (0.03)	0.08 (0.04)	0.09 (0.06)	97.50 (1.22)	98.84 (0.42)
Disgust	0.43 (0.12)	0.04 (0.02)	0.02 (0.02)	**2.84**[a] **(0.18)**	0.32 (0.09)	0.20 (0.09)	0.13 (0.05)	85.83 (3.66)	95.58 (0.54)
Fear	0.38 (0.15)	0.00 (0.00)	0.08 (0.05)	0.43 (0.10)	**2.48**[a] **(0.16)**	1.08 (0.17)	0.08 (0.04)	74.17 (5.55)	96.69 (0.59)
Surprise	0.00 (0.00)	0.17 (0.08)	0.00 (0.00)	0.00 (0.00)	0.34 (0.19)	**3.27**[a] **(0.14)**	0.01 (0.01)	95.00 (1.67)	94.45 (0.96)
Sadness	0.34 (0.22)	0.00 (0.00)	0.08 (0.04)	0.24 (0.07)	0.16 (0.07)	0.11 (0.05)	**2.40**[a] **(0.14)**	87.50 (6.60)	96.59 (0.71)
Neutral	0.01 (0.01)	0.06 (0.02)	0.06 (0.03)	0.06 (0.03)	0.03 (0.02)	0.02 (0.01)	0.14 (0.05)	71.67 (5.31)	100 (0.00)

Note. Values correspond to the mean (±SEM) intensity rating (ranging from 0 to 5), as well as % (±SEM) correct recognition (Hit rate) and % (±SEM) correct rejection of each expression category (Expression Perceived) for each target expression condition averaged across actors and observers. SEM reflects the variability between actors. Bold values are very important in order to highlight that each perceived emotion corresponds exactly with the emotion intended by the researchers and not with other emotions (diagonal line in bold shows that at first glance).

[a] Mean rating of the target emotion is significantly higher than the mean rating on the other scales (Fisher's protected least significance difference test; all p's < .001).

each emotional facial expression was perceived as more intense than neutral faces (all ps < .001). Moreover, each category of facial expression yielded significantly higher average ratings on the corresponding scale than to the six non-target scales (Fisher's protected LSD; all ps < .001; Table 2). The analysis of the mean intensity ratings of the target emotional category (see diagonal in Table 2) also revealed a main effect of emotional expression (F (7,98) = 58.20; $p < .001$; $\varepsilon = .57$). Facial expressions of happiness and surprise were judged as more intense than sad faces (happiness Vs sadness: $p < .001$; surprise Vs sadness: $p < .001$) while happy faces were perceived as more intense than fear faces ($p < .001$). The remaining emotions did not significantly differ from each other. Intensity ratings were not influenced by the sex of the observer (F (1,13) = 1.24; $p = .285$) or of the actor (F (1,14) = 0.03; $p = .864$; $\varepsilon = 1.00$) (see Table 3).

3.1.2. Discrimination

The cluster analysis identified 8 different subgroups of clips corresponding to each of the 8 expression conditions. According to the observed clustering, all clips were adequately assigned to the target category (i.e. the distance between clips of the same target emotion was always smaller than between clips of different target emotions). The analysis also revealed second-order combinations of clusters for neutral and sadness, as well as pain and fear. Pain and fear also showed some proximity to disgust as reflected by a third-order combination of these clusters.

The sensitivity (Hit rate) and specificity (Correct rejection rate) indices confirmed that participants clearly recognized the target emotion shown in the clips, and that the presence of non-target expression was negligible (Table 2). Sensitivity was higher than 70% for all conditions (mean sensitivity: 86%), well above the 14% chance rate. Specificity was above 94% for all expression (mean specificity: 97%). However, the presence of non-target expression differed between emotions as revealed by a main effect of emotional expression on the mixed-emotion index (F (7,98) = 13.89; $p < .001$; $\varepsilon = .44$). Pairwise comparisons revealed that fear, pain and disgust faces showed slightly more traces of non-target emotions than happy, angry and neutral faces (all ps ⩽ .001). Fear clips additionally showed more traces of non-target emotions than sadness and surprise clips (ps ⩽ .001). While the mixed-emotion index was not influenced by the sex of the observer, a main effect of the sex of the actor was

Table 3
Mean (SEM) ratings of the intensity, valence, and arousal of the expressions displayed by male and female actors

Target expression	Intensity of the target expression		Valence of the target expression		Arousal of the target expression	
	Male actors	Female actors	Male actors	Female actors	Male actors	Female actors
Pain	2.68 (0.23)	2.67 (0.30)	−2.77 (0.19)	−3.00 (0.15)	2.25 (0.28)	2.03 (0.18)
Happiness	3.22 (0.15)	3.47 (0.13)	2.80 (0.16)	3.18 (0.12)	0.18 (0.44)	0.73 (0.43)
Anger	3.00 (0.18)	2.93 (0.21)	−1.57 (0.22)	−1.80 (0.21)	1.20 (0.37)	1.28 (0.30)
Disgust	2.78 (0.20)	2.90 (0.20)	−2.57 (0.17)	−2.65 (0.15)	0.73 (0.31)	1.23 (0.28)
Fear	2.45 (0.20)	2.50 (0.18)	−2.23 (0.12)	−2.20 (0.14)	1.37 (0.29)	1.58 (0.30)
Surprise	3.37 (0.17)	3.17 (0.13)	−0.27 (0.12)	0.25 (0.16)	1.55 (0.35)	1.48 (0.34)
Sadness	2.42 (0.14)	2.38 (0.19)	−2.08 (0.18)	−2.30 (0.18)	−0.83 (0.32)	−0.38 (0.41)
Neutral	0.06 (0.02)	0.05 (0.02)	0.02 (0.03)	−0.18 (0.09)	−1.23 (0.38)	−1.18 (0.36)

Note. Intensity ratings ranged from 0 to 5. Valence and arousal ratings ranged from +4 to −4.

Please cite this article in press as: Simon D et al., Recognition and discrimination of prototypical dynamic ..., Pain (2007), doi:10.1016/j.pain.2007.05.008

detected reflecting slightly higher reports of non-target expression in female actors (F (1,14) = 4.87; p = .044; ε = 1.00). Taken together, these findings indicate that the expressions were easily discriminated and included only some traces of non-target expressions.

3.1.3. Valence

Analysis of the average valence scores clearly indicates that participants perceived happy faces as "pleasant", neutral and surprise faces as "neutral", and the remaining emotional faces (pain, disgust, sadness, fear, anger) as unpleasant (Table 3). Thus, the analysis of these ratings resulted in a main effect of emotional expression (F (7,98) = 216.40; p < .001; ε = .47). As demonstrated by pairwise comparisons, participants judged happy faces as significantly more pleasant than all other face categories (all ps < .001). Furthermore, all facial expressions that had been rated as "unpleasant" were significantly more unpleasant than both neutral and surprised faces (all ps < .001). Additionally, within the group of unpleasant facial displays, pain faces were rated significantly higher on unpleasantness than fear, sad and anger faces (pain Vs fear: p < .001; pain Vs sadness: p = .001; pain Vs anger: p < .001). Furthermore, disgust clips also yielded higher unpleasantness ratings than anger clips (p < .001). While valence ratings were not influenced by the observer's sex, an interaction between emotional facial expression and sex of the actor was found (F (7,98) = 5.55; p < .001; ε = .64). This was due to higher levels of pleasantness to female than male faces expressing surprise (p = .001).

3.1.4. Arousal

The statistical analysis of the average arousal ratings for all emotional facial expressions revealed a main effect of emotional expression (F (7,98) = 24.91; p < .001; ε = .39). Pairwise comparisons showed that participants perceived neutral and sad faces as significantly less arousing than pain, surprise, fear, anger, as well as disgust faces (all ps < .001), while sadness did not differ significantly from neutral. Furthermore, pain clips were rated as the most arousing and differed significantly from fear, disgust and happy faces (ps $\leqslant$ .001). While arousal ratings were not influenced by the observer's sex, a main effect of the actor's sex was observed (F (1,14) = 6.32; p = .025; ε = 1.00). Moreover, a trend for an interaction between the actor's sex and emotion was detected (F (7,98) = 2.31; p = .065; ε = .60). Post hoc comparisons revealed higher levels of arousal to female than male faces expressing happiness (t (14) = 3.21; p = .006) (see Table 3).

3.2. Facial action coding system

The reliability of the FACS coding of the 64 clips across the two coders was excellent (frequency reliability = 82.3%; intensity reliabilities = 90.6%). Table 1 lists all AUs observed in the different condition and shows the overlap with prototypical AUs, as reported in the FACS literature. Considering the FACS investigators guide (Chapter 12, p.174, Table 1, Ekman et al., 2002), the detected activation patterns of AUs for each emotion generally corresponded well with the prototypes or major variants of the intended emotional facial expressions. The selected stimuli always included several of the target AUs in various combinations (Table 1). More specifically, all happiness expressions included the target AUs 6 and 12; anger expressions included 8 out of 10 target AUs; disgust expression included all seven target AUs; fear expressions included 7 out of 8 target AUs; surprise expressions included all 5 target AUs; sadness expressions included 7 out of 8 target AUs; and pain expressions included 9 out of 10 target AUs. Moreover, few non-target AUs were observed. For example, anger stimuli included an overall average of 3.38 target AUs and an only 0.38 non-target AUs (see # of target AUs in Table 1). This meant that among all the AUs observed in the anger clips, 89.6% were target AUs. Happiness and disgust showed the lowest relative rates of target AUs (see Table 1) indicating the presence of additional non-target AUs mainly in those conditions. Pain expression contained the highest mean number (6.88) and proportion (92.3%) of target AUs. There was no main effect of the sex of the actor, and no interaction between the expression condition and the sex of the actor on FACS data, suggesting that both males and females contributed to the observed patterns of facial expression across conditions.

4. Discussion

The aim of the present study was to build a standardized set of dynamic stimuli of prototypical facial expressions of pain and basic emotions and test the distinctiveness of pain expressions. Consistent with the hypothesis, the expressions of pain and emotions were clearly recognized and discriminated by normal volunteers. Those expressions further matched precisely with the previously published prototypical displays based on FACS analysis.

4.1. FACS prototypes and the dynamic expression of emotions

The facial expressions included in this set of stimuli corresponded well with the prototypes and their major variants as reported in the literature (Ekman et al., 2002; Williams, 2002), and shown in Table 1. Moreover, the FACS responses of our emotion stimuli were highly consistent with the results of a reference study on dynamic expression of felt emotions using a similar methodology (Gosselin et al., 1995). This strongly supports the validity of the present stimulus set as representative of the prototypical expression of felt emotions.

4.2. Perceiving dynamic expressions of pain and emotions

Consistent with the instruction given to actors to express strong pain/emotions, clips were rated in the upper half of the intensity scale for the target facial expressions (see Table 2). Additionally, the results of the cluster analysis performed on the observer's ratings of pain and emotions demonstrated that the 64 film clips were correctly classified into eight different groups of emotional facial expressions, in accordance with the intended target emotions (100% accuracy). This reflects the fact that each stimulus received its highest mean intensity rating (across subjects) for the target expression category. Hence, in line with the literature, specificity and distinctness of the selected facial expressions was confirmed by both FACS coding and by the perceptual ratings of pain/emotion intensity obtained in naïve observers (Kappesser and Williams, 2002; Smith and Scott, 1997). Furthermore, global correct recognition rates were well above chance, and in most cases higher than 85%, while correct rejection rates were always above 95%, consistent with the excellent sensitivity and specificity of the stimuli and of recognition processes. The range of correct recognition rates (see Hit rate in Table 2) was comparable to that of the recognition rates reported for the original set of stimuli used by Ekman and Friesen (1976) to develop the FACS coding system. Similar recognition rates are also reported by Kappesser and Williams (2002) in health care professionals identifying prototypical facial expressions of pain and negative emotions from photographs.

The results also suggest that participants perceived a mixture of several emotions in some cases. Misattribution was observed mainly in fear and pain (Table 2). The observed traces of surprise in fear expression have been reported in some cultures (Ekman et al., 1972) as well as by a model observer classifying affective faces (Smith et al., 2005). This is consistent with the similarity of both facial displays as demonstrated by the overlap of target AUs involved in those emotions (Table 1). However, AU4 (brow lower) clearly distinguishes between fear and surprise. In addition, AU25 (lip part) was observed in both fear and surprise although it constitutes a prototypical AU only in fear. This might explain the occasional misattributions observed in those categories. Similarly, pain faces contained traces of disgust, fear and to a lesser extent surprise, and were occasionally misattributed to these non-target categories (Table 2). In the present stimulus set, most of those expressions contained AU25 (lip part), a target AU for pain, disgust, and fear consistent with the prototypes. Additional AUs that were commonly observed in both pain and disgust included AUs 4, 6, 7, 9, 10, and 20. Those AUs are part of the prototypes for pain expression but only AUs 9 and 10 are part of the prototypes for disgust expression. However, in spite of the presence of several non-target AUs, disgust faces were well discriminated. This excellent recognition rate may be explained by the presence of AU15 (lip corner depressor) as this AU is part of the major variants of disgust prototypes and was present in all disgust clips but none of the pain clips. The patterns of AUs associated with more ambiguous (or mixed) expression and the specific facial actions (or groups of AUs) on which recognition relies in those cases should be investigated in future research.

4.3. Valence and arousal

In addition to their classification into discrete categories, facial displays can also be divided into pleasant (happiness), neutral (surprise, neutral) and unpleasant (pain, anger, fear, disgust, sadness) expressions according to their valence (e.g. Feldman Barrett and Russell, 1998). Within the subgroup of negative facial displays, pain was generally perceived as the most unpleasant. This finding is in line with the bio-social significance of pain (Williams, 2002). Only disgust faces were perceived to be almost as unpleasant as pain faces. This may reflect the functional role of pain and disgust expression in signalling immediate noxious events (tissue damage or food poisoning), resulting in an actual threat to the body's integrity (e.g. Wicker et al., 2003). This clearly contrasts with fearful and angry faces that may rather signal impending danger.

With respect to arousal, sad and neutral faces were perceived as least arousing in contrast to all other facial expressions, in accordance with previous factorial analytic findings (e.g. Feldman Barrett and Russell, 1998). Again the pain face yielded the highest arousal ratings, however differences were not significant compared to anger and surprise. This may again reflect the functional role of pain expression in communicating an actual threat.

4.4. Sex differences

While the observer's sex did not show any influence on the ratings, the actor's sex seemed to modulate some of the participant's judgements. This may reflect a perceptual-cognitive effect in the observer rather than differences in expression as the FACS results did not show any significant effect of the actor's sex. More traces of non-target emotions and higher ratings of arousal were given to female expression in general. Moreover, females expressing surprise were perceived as being slightly more pleasant than male actors. However, taking the intensity ratings into account, judgements did not seem to be strongly influenced by gender stereotypes as suggested by previous research (e.g. Algoe et al., 2000; Hess et al., 2004; Plant et al., 2004).

4.5. Some limitations and future directions

As a final note, some limitations have to be mentioned. Although the stimuli corresponded generally well with the FACS prototypes, fear and disgust faces may not be optimal. The imperfect match with FACS prototypes might have affected the recognition of fear faces. However, the facial expression of disgust was well recognized in spite of the slight deviation from the FACS prototypes. Some variability between actors in the FACS results and between the present results and the previously reported prototypes may reflect differences in the specific situations mentally evoked to elicit the emotion or individual differences in expressive style. Those differences also occur in naturalistic contexts (Smith and Scott, 1997) and are consistent with the ecological validity of the present stimulus set.

Future studies may examine in more detail the symmetry and the smoothness of the full temporal unfolding of the expression (including the offset) as those factors may be affected by the genuineness of the expression (e.g. Craig et al., 1991; Gosselin et al., 1995). Those factors were not assessed here as they are typically undetected by observers (Gosselin et al., 1995; Hadjistavropoulos et al., 1996). Therefore, although the stimuli might include some spatio-temporal irregularities, this did not compromise the ability of normal volunteers to recognize the expressions. The development of pattern-recognition algorithms for the analysis of dynamic facial expression (Hammal et al., 2006) as well as random spatio-temporal masking procedures (Gosselin and Schyns, 2001) will likely contribute to future developments in this field.

Finally, one has to emphasize that this stimuli set is relatively small and that spontaneous real-life expressions are of course not restricted to the expressions included here. Validated stimuli displaying more subtle and mixed emotional expressions, including the social emotions (e.g. shame, pride,...), must be developed. Furthermore, since the present stimulus set was based on Western prototypes the generalisation to other cultural contexts might be limited. Recordings of spontaneous facial expression in a more natural setting would of course provide stimuli with better ecological validity but the controlled recording (e.g. lighting, orientation of face...) and the adequate sampling of "relatively pure expressions" (i.e. limiting contamination by mixed emotions), in all pain/emotion conditions and in each model, coming from a variety of cultural backgrounds, is extremely challenging.

5. Conclusions

We present a set of dynamic, prototypical facial expressions of pain, the six basic emotions, and a neutral display, which can be reliably recognized and distinguished by normal individuals. Importantly, the results demonstrate the distinctiveness of pain facial displays from those of basic emotions. For a comparable intensity of expression, the facial display of pain was also perceived as the most unpleasant and arousing, possibly reflecting its higher bio-psychosocial significance. This standardized set of stimuli can be widely used in cognitive, social, clinical, and neuroscience studies on facial expressions and their social communicative functions.

Acknowledgements

This research was supported by grants and salary support from the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Fonds Québécois de la recherche sur la nature et les technologies (FQRNT), the Canadian Fund for innovation (CFI), the Fonds de la recherche en santé du Québec (FRSQ), the Canadian Institutes of Health Research, the Social Sciences and Humanities Research Council of Canada, and a scholarship to DS from the German Academic Research Exchange Service (DAAD). We thank Julie Senécal, David Wong and Annie Hsieh for their help in the realisation of this study.

References

Algoe SB, Buswell BN, DeLamater JD. Gender and job status as contextual cues for the interpretation of facial expression of emotion – statistical data included. Sex Roles 2000;42:183–208.

Ambadar Z, Schooler J, Cohn J. Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. Psychol Sci 2005;16:403–10.

Botvinick M, Jha AP, Bylsma LM, Fabian SA, Solomon PE, Prkachin KM. Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. Neuroimage 2005;25:312–9.

Craig KD, Hyde SA, Patrick CJ. Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. Pain 1991;46:161–71.

Craig KD, Prkachin KM, Grunau RVE. The facial expression of pain. In: Turk DC, Melzack R, editors. Handbook of pain assessment, vol. 2nd ed. New York: Guilford Press; 2001. p. 153–169.

Darwin C. The expression of the emotions in man and animals. London: Albemarle; 1872.

Ehrlich SM, Schiano DJ, Sheridan K. Communicating facial affect: it's not the realism, it's the motion. Proceedings of ACM CHI 2000 conference on human factors in computing systems. New York: ACM; 2000. p. 252–253.

Ekman P, Friesen WV, Ellsworth P. Emotion in the human face: guidelines for research and an integration of findings. Oxford: Pergamon Press; 1972.

Ekman P, Friesen WV. Unmasking the face. Englewoods Cliffs, NJ, USA: Prentice Hall; 1975.

Ekman P, Friesen WV. Pictures of facial affect. Palo Alto (CA): Consulting Psychologist Press; 1976.

Ekman P, Friesen WV, Elsworth P. Research foundations. In: Ekman P, editor. Emotions in the human face. London: Cambridge University Press; 1982.

Ekman P, Friesen WV, Hager JC. Facial action coding system (FACS). Salt Lake City: A Human Face; 2002.

Feldman Barrett LF, Russell JA. Independence and bipolarity in the structure of current affect. J Pers Soc Psychol 1998;74:967–84.

Gosselin P, Kirouac G, Dore FY. Components and recognition of facial expression in the communication of emotion by actors. J Pers Soc Psychol 1995;68:83–96.

Gosselin F, Schyns PG. Bubbles: a technique to reveal the use of information in recognition. Vision Res 2001;41:2261–71.

Hadjistavropoulos HD, Craig KD, Hadjistavropoulos T, Poole GD. Subjective judgments of deception in pain expression: accuracy and errors. Pain 1996;65:251–8.

Hammal Z, Eveno N, Caplier A, Coulon P-Y. Parametric models for facial features segmentation. Signal Process 2006;86:399–413.

Harwood N, Hall L, Shinkfield A. Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. Am J Ment Retard 1999;104:270–8.

Hess U, Adams Jr RB, Kleck RE. Facial appearance, gender, and emotion expression. Emotion 2004;4:378–88.

Kappesser J, Williams AC. Pain and negative emotions in the face: judgements by health care professionals. Pain 2002;99:197–206.

Kilts CD, Egan G, Gideon DA, Ely TD, Hoffman JM. Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. Neuroimage 2003;18:156–68.

LaBar K, Crupain M, Voyvodic J, McCarthy G. Dynamic perception of facial affect and identity in the human brain. Cereb Cortex 2003;13:1023–33.

Lang PJ, Greenwald MK, Bradley MM, Hamm AO. Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology 1993;30:261–73.

O'Toole AJ, Roark DA, Abdi H. Recognizing moving faces: a psychological and neural synthesis. Trends Cogn Sci 2002;6:261–6.

Plant EA, Kling KC, Smith GL. The influence of gender and social role on the interpretation of facial expressions. Sex Roles 2004;51:187–96.

Prkachin KM. The consistency of facial expressions of pain: a comparison across modalities. Pain 1992;51:297–306.

Roark D, Barrett SE, Spence MD, Abdi H, O'Toole AJ. Psychological and neural perspectives on the role of facial motion in face recognition. Behav Cogn Neurosci Rev 2003;2:15–46.

Sato W, Kochiyama T, Yoshikawa S, Naito E, Matsumura M. Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. Cogn Brain Res 2004;20:81–91.

Simon D, Craig KD, Miltner WHR, Rainville P. Brain responses to dynamic facial expressions of pain. Pain 2006;126:309–18.

Smith CA, Scott HS. A Componential approach to the meaning of facial expressions. In: Russell JA, Fernandez-Dols J-M, editors. The psychology of facial expression. New York: Cambridge University Press; 1997. p. 229–54.

Smith ML, Cottrell GW, Gosselin F, Schyns PG. Transmitting and decoding facial expressions. Psychol Sci 2005;16:184–9.

Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G. Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. Neuron 2003;40:655–64.

Williams AC. Facial expression of pain: an evolutionary account. Behav Brain Sci 2002;25:439–88.

Zald DH. The human amygdala and the emotional evaluation of sensory stimuli. Brain Res Rev 2003;41:88–123.

Please cite this article in press as: Simon D et al., Recognition and discrimination of prototypical dynamic ..., Pain (2007), doi:10.1016/j.pain.2007.05.008