EJP
European Journal of Pain

ORIGINAL ARTICLE

The faces of pain: A cluster analysis of individual differences in facial activity patterns of pain

M. Kunz[1,2], S. Lautenbacher[1]

1 Physiological Psychology, Otto-Friedrich University Bamberg, Germany
2 Biological Psychology, Ludwig Maximilian University, Munich, Germany

Correspondence
Miriam Kunz
E-mail: miriam.kunz@uni-bamberg.de

Funding sources
These studies were supported by a research grant from the Deutsche Forschungsgemeinschaft (Ku2294/4; study 1) and a FNK grant of the University of Bamberg (study 2).

Conflicts of interest
None declared.

Accepted for publication
5 October 2013

doi:10.1002/j.1532-2149.2013.00421.x

Abstract

Background: There is general agreement that facial activity during pain conveys pain-specific information but is nevertheless characterized by substantial inter-individual differences. With the present study we aim to investigate whether these differences represent idiosyncratic variations or whether they can be clustered into distinct facial activity patterns.
Methods: Facial actions during heat pain were assessed in two samples of pain-free individuals ($n = 128$; $n = 112$) and were later analysed using the Facial Action Coding System. Hierarchical cluster analyses were used to look for combinations of single facial actions in episodes of pain. The stability/replicability of facial activity patterns was determined across samples as well as across different basic social situations.
Results: Cluster analyses revealed four distinct activity patterns during pain, which stably occurred across samples and situations: (I) narrowed eyes with furrowed brows and wrinkled nose; (II) opened mouth with narrowed eyes; (III) raised eyebrows; and (IV) furrowed brows with narrowed eyes. In addition, a considerable number of participants were facially completely unresponsive during pain induction (stoic cluster). These activity patterns seem to be reaction stereotypies in the majority of individuals (in nearly two-thirds), whereas a minority displayed varying clusters across situations.
Conclusion: These findings suggest that there is no uniform set of facial actions but instead there are at least four different facial activity patterns occurring during pain that are composed of different configurations of facial actions. Raising awareness about these different 'faces of pain' might hold the potential of improving the detection and, thereby, the communication of pain.

1. Introduction

There is general agreement that facial activity during pain is not unspecific grimacing but conveys pain-specific information (Williams, 2002; Hadjistavropoulos et al., 2011). Evidence for this can be mainly taken from two sources. First, only a small subset of facial activities has been repeatedly observed during pain, namely, tightening of the muscles surrounding the eyes, furrowed brows, raising the upper lip/nose wrinkling and eye closure (Prkachin, 1992; Prkachin and Solomon, 2008). Second, when actors are taught to display this subset of 'pain prototypical' facial expressions, observers can recognize pain among other emotions above chance level (Simon et al., 2008). However, despite the evidence that these key facial activities reliably occur during pain, this does not imply only one uniform facial expression of pain that can be observed at all time and in each individual (Craig et al., 2011). Instead, the frequencies of occurrence of these key activities during pain usually only range from 10% to 60% (e.g., Kunz et al., 2008), and thus the likelihood that all four key facial activities occur simultaneously is very low. Rather, individuals

What's already known about this topic?

- Facial pain expressions are not unspecific grimacing but convey pain-specific information.
- There are a limited number of key facial actions that are displayed reliably in the context of pain.
- Nevertheless, there are also great variations in facial expressions of pain, with individuals often displaying only parts of these key facial actions in different combinations.

What does this study add?

- We could show that there are at least four different facial activity patterns occurring during pain that are composed of different configurations of facial actions (different faces of pain).
- These patterns occur across samples and across situations.
- Raising awareness about the different 'faces of pain' might hold the potential to improve pain assessment.

often display only parts of this subset, sometimes even blending it with a limited range of other facial activities (e.g., smiling; Hale and Hadjistavropoulos, 1997; Kunz et al., 2009).

These inter-individual variations suggest that there might be various 'faces of pain', and this variability might interfere with the accurate communication of pain. Indeed, several studies have demonstrated shortcomings of observers to accurately recognize and estimate pain through facial activity (e.g., Kappesser et al., 2006; Eritz and Hadjistavropoulos, 2011) and compared with almost all six basic emotions, the recognition accuracy for pain seems to be among the lowest (Kappesser and Williams, 2002; Simon et al., 2008). This is especially alarming given that the accurate decoding of facial pain activity informs the interaction and transaction between patient and caring person, and can hereby determine the success of pain management (Hadjistavropoulos et al., 2011). Thus, a better knowledge and awareness of 'different faces of pain' is urgently needed since it holds the potential to improve the accurate communication of pain. Although we are not the first ones acknowledging inter-individual variations, our aim was to increase our understanding of facial expressions by a systematic search for regular and stable patterns within the many variations in facial expressions of pain.

For that purpose, we used hierarchical cluster analyses to investigate whether variations in facial activity during pain can be clustered into distinct facial activity patterns or simply represent idiosyncratic variations. If we indeed find stable cluster solutions, this would represent a promising approach to characterize types of inter-individual differences in facial activities during pain. We intend to derive the best-fitting cluster solution in one sample and verify its appropriateness in a second sample as a test on the *replicability of clusters across samples*. Furthermore, to test whether potential facial activity clusters represent stable reaction stereotypies, we also included different social situations to test the *situational stability of 'cluster membership'*. In sum, the present approach should allow for revealing the existence of facial activity patterns of pain, testing their stability of occurrence within and between individuals, and identifying potential individual determinants of cluster membership (e.g., sex and catastrophizing).

2. Materials and methods

2.1 Participants

Sample 1 consisted of 128 subjects [64 men, 64 women; M = 39.87 years; standard deviation (SD) = 13.59; 64 heterosexual couples being in a relationship >6 months] who were recruited via advertisements in the local newspaper (Bamberg). Couples were selected to test participants in different social situations (alone, in the presence of the partner, in the presence of the experimenter) and, thus, to investigate whether facial expression clusters are replicable across different social situations. Furthermore, this allows for assessing how stable one individual shows the same type of facial encoding pattern across situations (stability of cluster membership). If individuals were facially completely non-expressive in response to the noxious stimulation, we excluded them from the cluster analyses because of insufficient data for this type of analysis [given that a lack of expression ('0 values') cannot be clustered into a specific type of facial expression, besides being a lack of expression]. However, even though these individuals will not be directly entered into the cluster analyses, we will treat this group as an additional cluster in the Results section ('Stoic cluster') to stress the fact that individuals might be experiencing pain although they do not show any pain-related facial activity and that a 'stoic face' can also be an expression of pain. Sample 2 consisted of 112 students (54 men, 58 women; M = 22.93 years; SD = 4.7) who were recruited via advertisements posted in the university buildings of the University of Bamberg. Again, we excluded those participants who were facially not expressive to our pain induction procedure from the cluster analysis but included them in the Results section as a 'stoic cluster'.

Exclusion criteria for both samples were current experience of acute or chronic pain, psychological and physical illnesses. Participants taking psychotropic drugs or analgesics were excluded from participation as well. These criteria led to the exclusion of 8% of the participants who had shown interest in participating. All participants provided informed consent and received either course credit or monetary compensation. The study protocols were approved by the Ethics Committee of the University of Bamberg.

2.2 Procedure

We investigated facial activities elicited during painful heat stimulation in two different samples. We used two different samples in order to be able to compare facial encoding clusters between samples and, thus, test the cluster replicability across samples. The pain stimulation protocol was the same in both samples. The main procedural difference between samples was that sample 1 underwent three blocks of thermal stimulation (with 10 painful stimuli each), whereas sample 2 only underwent one block (of 10 painful stimuli). The three blocks in sample 1 were necessary to compare cluster solutions across different social situations within one sample. Accordingly, we used a within-subject design and tested participants once while they were alone in the room, once with the partner present and once with the experimenter present. Since we had found stable cluster solutions across situations in sample 1 (see Results section of the paper), we decided to test sample 2 in only one situational context (while being with the experimenter). The order of situations in sample 1 was randomized across subjects.

2.2.1 Pain stimulation

Pain stimulation protocols were exactly the same in both samples. Following a previous protocol that has been shown to successfully elicit facial expressions of pain (Kunz et al., 2011, 2012), pain was induced using a Peltier-based, computerized thermal stimulator (Medoc TSA-2001; Medoc Ltd, Ramat Yishai, Israel) with a 3×3 cm^2 contact probe attached to the outer part of the left lower leg (midpoint between ankle and knee). To ensure that temperature intensities were perceived as painful but not too painful in all subjects (in order to prevent floor as well as ceiling effects), temperature intensities were tailored to the individual pain threshold. Thus, heat pain thresholds were determined first, using the method of adjustment. Subjects were asked to adjust a temperature starting from 38 °C, using heating and cooling buttons, until they obtained a level that was barely painful. A constant press of the buttons produced a heating or cooling rate of 0.5 °C/s. Following a familiarization trial, there were four trials and the average of these trials was used to constitute the threshold estimate.

Following the assessment of pain thresholds, phasic heat stimuli [5 s (plateau); rate of change: 4 °C/s; baseline temperature: 38 °C; inter-stimulus intervals of 10–15 s] were applied to the lower leg. Two different stimulus intensities were applied in a randomized order, namely, painful (+3 °C above the pain threshold) as well as non-painful (−1 °C below the pain threshold) intensities. The non-painful intensities were interspersed in order to reduce the chance of habituation or sensitization effects. We refrained from entering facial responses occurring during non-painful stimulation into the cluster analysis because half of the participants did not display any facial response during the non-painful stimulation. In each experimental block, participants received 10 painful and 10 non-painful stimuli in a random order. Given that sample 1 underwent three experimental blocks (namely, while being alone, while being with the experimenter and while being with the partner), we changed the site of stimulation between experimental blocks (moving the thermode 3–5 cm upwards or downwards, respectively) to also reduce the chance of local sensitization effects.

2.2.2 Facial actions during pain

The face of the participants in sample 1 and 2 was videotaped throughout the pain induction procedures. The camera was positioned in front of the subject (>1 m distance) and provided a full-face view (see Fig. 1). A light-emitting diode visible to the camera, but not to the subject, was lighted concurrently with the thermal stimuli to mark the onset of stimulation. To ensure that the face would always be upright and in a frontal view during stimulation, subjects were asked to look at the computer screen or at a picture in front of them throughout the thermal heat stimulation. Subjects were also instructed not to talk during thermal stimulation.

Facial expressions were coded from the video recordings using the Facial Action Coding System (FACS; Ekman and Friesen, 1987), which is based on anatomical analysis of facial movements and distinguishes 44 different action units (AUs) produced by single muscles or combinations of muscles. FACS coders identified the frequency and the intensity (5-point scale) of different AUs. Given the enormous amount of video data and the time-consuming FACS coding procedure, five FACS coders were needed to analyse the video material. Among them was a certified FACS coder (the author M.K.; qualified by passing an examination given by the developers of the system), who additionally FACS coded 5% of the video material from each coder to calculate inter-rater reliability [using the Ekman–Friesen formula (Ekman and Friesen, 1987); inter-rater reliability lay between 0.84 and 0.93]. A software designed for the analysis of observational data (the Observer Video-Pro; Noldus Information Technology, Wageningen, the Netherlands) was used to segment the videos and to enter the FACS codes into a time-related database. Given that each painful stimulus lasted for 5 s, we selected these time segments of 5 s (beginning just after stimulus had reached the target temperature) for scoring the facial response occurring during painful stimulation. In total, 10 painful segments (5 s each) per experimental block were analysed in each subject. As has been performed in most preceding studies on facial responses to pain (Prkachin, 1992; Kunz et al., 2004, 2007, 2008), we combined those AUs that represent similar facial movements (AU 1_2, AU 6_7, AU 9_10 and AU 25_26_27).

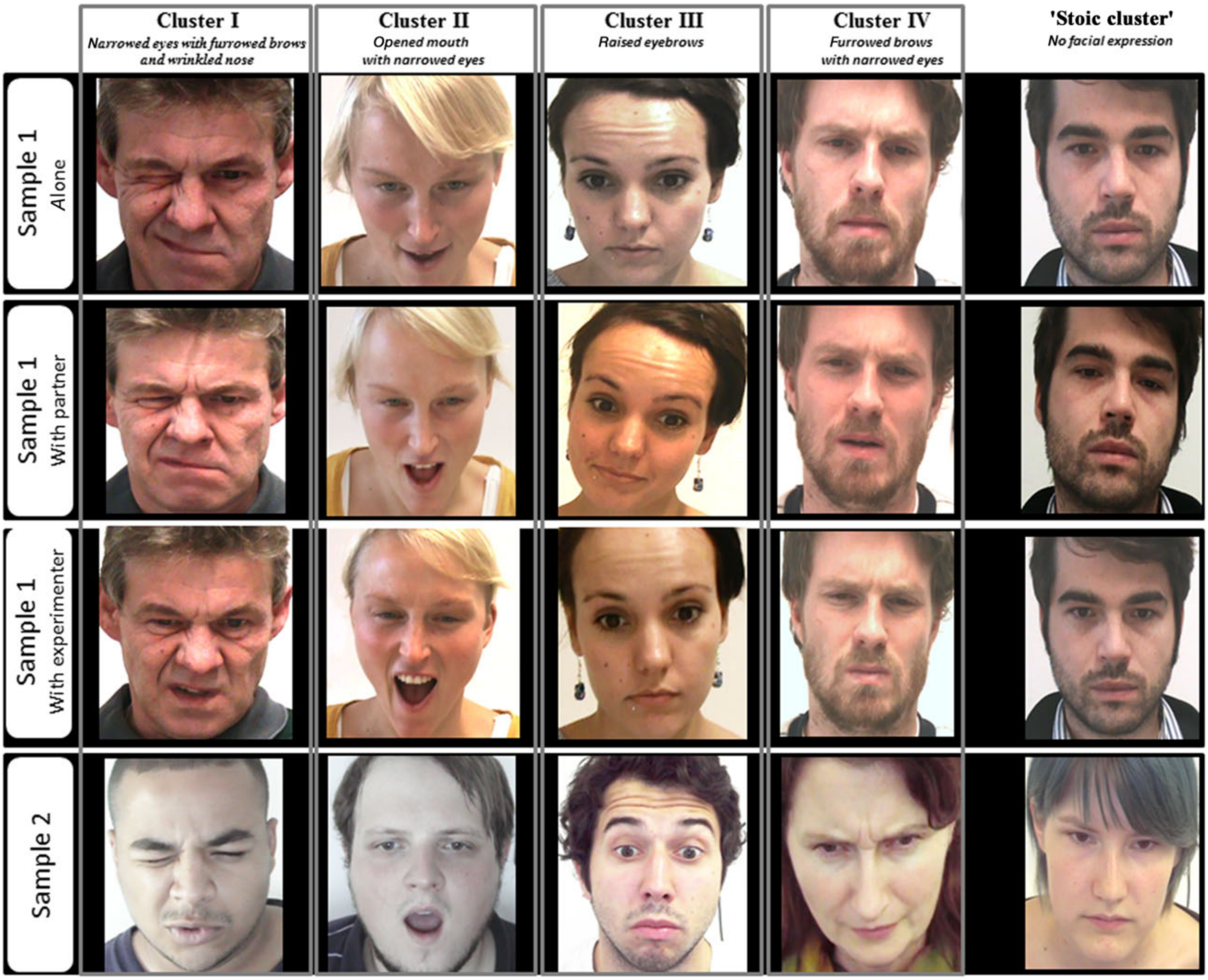

Figure 1 Examples of facial actions occurring during pain that belong to one of the four facial activity clusters as well as to the 'stoic cluster'. Examples are given separately for sample 1 (for all three social situations) and for sample 2.

To select AUs to be entered into the cluster analyses, we calculated how often each AU occurred during painful stimulation in sample 1 (separately for each of the three situations) and in sample 2. Only those AUs that occurred in at least 5% of the painful segments recorded are listed in Table 1 and were considered for further analyses. We decided to only include in the cluster analyses those AUs that occurred in more than 5% in both samples and in all three situations (these were eight Aus; see Table 1). However, before entering the selected eight AUs into the cluster analyses, AU values (product term: frequency × intensity) were z transformed within each participant in order to eliminate differences in the degree of facial expressiveness between participants. This was necessary because otherwise highly expressive individuals would contribute disproportionately to the formation of facial expression clusters. However, we were not interested in clusters merely reflecting different degrees of expressiveness (referring simply to the vigour and number of facial responses), but instead, in the extraction of distinct facial activity patterns, which may be constant across a group of individuals regardless of whether being faintly or more strongly expressed. Individual differences in the degree of expressiveness were thus eliminated using z transformation within each participant.

2.2.3 Self-report ratings

After the application of each thermal pain stimulus, subjects always rated the intensity on a visual analogue scale (VAS). The scale was labelled with the verbal anchor 'faintly painful' in the centre so that all non-painful sensations were rated below 50 and all painful ones at 50 or above (subjects were instructed that the lower end meant no felt temperature and the upper end would indicate extremely strong pain).

2.2.4 Pain catastrophizing

In contrast to one of our previous studies (Kunz et al., 2008), others repeatedly have found significant associations between

Table 1 Facial action units (AUs) with a critical frequency of occurrence of more than 5% in painful segments.

		Frequency of occurrence (%)			
		Sample 1			Sample 2
Action unit	Description	Alone	With the partner	With the experimenter	With the experimenter
[a]AU 1_2	Raised eyebrows	17.7	22.9	15.1	13.2
[a]AU 4	Furrowed brows	37.7	43.5	38.7	29.8
[a]AU 6_7	Narrowed eyes	57.1	69.3	56.4	57.0
[a]AU 9_10	Wrinkled nose	31.0	32.6	25.7	12.0
AU 12	Smiling	9.9	15.9	–	12.8
[a]AU 14	Dimpler	14.4	13.2	18.1	11.0
[a]AU 17	Chin raiser	6.4	9.7	8.9	7.6
[a]AU 18	Lip pucker	10.5	8.9	7.9	5.5
AU 23	Tightened lips	7.3	6.5	7.8	–
[a]AU 25_26_27	Opened mouth	44.8	36.8	34.5	23.0

Values are given separately for the three social situations in sample 1 and differently for samples 1 and 2.
[a]AUs selected for the cluster analyses.

pain catastrophizing and facial pain displays (Sullivan et al., 2006; Vervoort et al., 2011). We were interested to investigate whether pain catasrophizing might not only be related to the quantity of facial responses (or the degree of expressiveness) but also to the more qualitative aspects of facial responses to pain (what types of facial actions are displayed). We therefore included a German translation of the Pain Catastrophizing Scale (PCS; Sullivan et al., 1995; Huber et al., 2010) to analyse whether the trait 'catastrophizing' might help to explain cluster formation in facial encoding of pain. (We decided to use catastrophizing as trait rather than a state variable because we assume that facial encoding patterns of pain are also more trait characteristics.) The PCS was developed as a measure of catastrophic thinking related to pain. Participants are instructed to reflect on thoughts or feelings during the past painful experiences. The scale contains 13 items that are rated on a 5-point scale, with the endpoints 'not at all' and 'all the time'.

2.3 Data analyses

2.3.1 Cluster analyses

Following previous approaches and recommendations (Blashfield and Aldenderfer, 1988; Hair and Black, 2000; Rovniak et al., 2010), a two-step clustering procedure was used.

In the first step, hierarchical cluster analyses were performed using Ward's method (Ward, 1963) and squared Euclidean definition of distances to determine the number of cluster groups within each of the three situations in sample 1 and in sample 2 (resulting in four hierarchical cluster analyses). The clustering process starts with the same number of clusters as there are cases and reduces the number of clusters by step-wise combining those clusters whose combination results in a minimum increase of the total within-group sum of squares. If a point is reached where clusters are combined that are dissimilar, the within-group sum of squares noticeably increases (as can be seen in the agglomeration schedule). The number of clusters prior to this rapid increase of the agglomeration coefficient is considered the natural grouping scheme (Hair and Black, 2000). To best determine the correct number of clusters, we inspected the rescaled distances as displayed in the hierarchical cluster dendrogram, examined the change in agglomeration coefficient and applied the Mojena stopping rule as a quantitative criteria to define a 'significant jump' in the agglomeration coefficient (Mojena, 1977; Milligan and Cooper, 1985).

In the second step, the cluster means (centroids), emerged from each hierarchical cluster analysis, were used as initial seed points in non-hierarchical, k-means cluster analyses as a method to verify the initial cluster solutions (Hair and Black, 2000). This was again carried out separately for each of the three situations in sample 1 and for the cluster analysis in sample 2.

2.3.2 Cluster replicability across samples/situations and cluster membership stability across situations

In order to ensure that the found clusters do not only represent chance findings but instead represent stable facial encoding patterns of pain, we assessed the replicability and stability of the identified facial expression clusters both across different situations (with the experimenter, alone, with the partner; sample 1) and across different samples (samples 1 and 2).

2.3.2.1 Cluster replicability across samples and situations

To assess replicability of facial encoding clusters across samples, we compared cluster solutions of sample 1 with cluster solutions in sample 2 (in sample 1, we selected the situation when subjects were tested in the presence of the experimenter, which was the situation common to both samples). We also compared cluster solutions between situations in sample 1 to test replicability across situations. Replicabitity was tested using McIntyre and Blashfield's (1980) nearest centroid cross-validation technique. This cross-validation procedure uses the cluster solutions from one sample (e.g., sample 1) to classify another sample (e.g., sample 2). This is performed by assigning, for example, each person in sample 2 to one of the cluster centres of the first sample based on minimal distances to these centres, and vice versa (cross-validation). Thus, for each individual in each sample there are now two clustering solutions available (the original clusters and the nearest centroid-assigned clusters). We did the same for cluster solutions in different social situations in sample 1. The replication accuracy between cluster solutions was quantified using the kappa coefficient.

2.3.2.2 Stability of cluster membership between situations

In sample 1, we calculated cross tabulations and kappa statistics between cluster solutions in different situations (alone vs. being with the experimenter, alone vs. being with the partner, being with the partner vs. being with the experimenter) to investigate whether belonging to a certain facial expression cluster in one situation predicts the membership to a comparable cluster in another situation. This procedure allows determining the stability of cluster membership of each participant across situations.

2.3.3 Demographic and pain-processing characteristics in different facial expression clusters

As a last step, we investigated whether the identified facial expression clusters might differ with regard to demographic characteristics (age, sex) or pain sensitivity (pain threshold, VAS ratings, PCS) using χ^2 tests (sex) as well as univariate (age, pain threshold, VAS ratings, PCS score) analyses of variance. Findings were considered to be statistically significant at $\alpha < 0.05$.

Statistical Package for the Social Sciences version 20 (IBM SPSS, Chicago, IL, USA) was used for all analyses.

3. Results

The heat intensities used to elicit pain (these were adjusted to be +3 °C above the individual's pain threshold) were comparable in both samples [sample 1: 49.3 °C (SD 1.2); sample 2: 49.0 °C (SD 1.4); $p > 0.05$ for sample differences]. Moreover, VAS pain intensity ratings (ranging from 0 to 100, with 50 indicating pain threshold) did also not differ between both samples or between the different situations in sample 1 {sample 1: alone [81.4 (SD 14.4)], with the partner [81.0 (SD 13.4)], with the experimenter [80.8 (SD 14.1)]; sample 2: 78.2 [SD 15.3]; all p-values > 0.05 for sample and situation differences}. Consequently, we were able to elicit comparable moderate-to-strong pain experiences in both samples and across social situations.

3.1 Cluster analyses

In all three hierarchical cluster analyses for sample 1 as well as in the hierarchical cluster analysis for sample 2, the dendrogram, the agglomeration coefficient and the Mojena stopping rule showed that a four-cluster solution was always the most appropriate solution to cluster facial activity during pain into different encoding patterns. The corresponding cluster means of different AUs are shown in Table 2 for sample 1 and for sample 2. In a second step, we verified these initial cluster solutions by using the centroid values of the hierarchical cluster analyses as initial seed points in non-hierarchical, k-means cluster analyses and found the results to be very similar to those of the hierarchical analyses (≥95% of the participants obtained same-group membership).

In addition to these four clusters, we also included the group of non-expressive individuals in Table 2 as an additional cluster and labelled it stoic cluster.

3.1.1 Facial expression clusters

When visually inspecting cluster solutions found in sample 1 for different situations as well as for sample 2, it becomes evident that very similar facial expression clusters were found across situations and across samples (see Table 2 and Fig. 1).

3.1.1.1 Cluster I

As can be seen in Table 2, the first cluster scored highest on AU 6_7 (contraction of the muscles surrounding the eyes) with a slighter co-activation of AU 4 (contraction of the eyebrows) and (in sample 1) of AU 9_10 (upper lip raise/nose wrinkle). This cluster was labelled as 'narrowed eyes with furrowed brows and wrinkled nose'. To better illustrate this facial activity pattern, photos are displayed in Fig. 1 (left column). This facial activity pattern was displayed by 46% of subjects in sample 2 and by 37–53% of subjects in sample 1 (depending on the social situation) and proved to be the most frequently occurring facial activity pattern during pain in both samples.

3.1.1.2 Cluster II

As can be seen in Table 2, the second cluster scored highest on AU 25_26_27 (mouth opening) with a

Table 2 Facial expression cluster profiles for sample 1 (separately for the three social situations) and sample 2 across the selected action units (*z*-scores).

	Situation	Cluster	AU 1_2 Raised eyebrows	AU 4 Furrowed brows	AU 6_7 Narrowed eyes	AU 9_10 Wrinkled nose	AU 14 Dimpler	AU 17 Chin raiser	AU 18 Lip pucker	AU 25_26_27 Opened mouth	Occurrence frequency of each cluster *n*	%
Sample 1	Alone	I	−0.54	0.42	1.89	0.20	−0.51	−0.64	−0.56	−0.26	36	28
		II	−0.58	−0.29	0.27	−0.27	−0.21	−0.38	−0.20	1.75	28	22
		III	1.82	−0.20	−0.02	−0.33	−0.22	−0.60	−0.41	−0.04	18	14
		IV	−0.56	1.30	0.20	−0.50	0.68	−0.47	−0.51	−0.09	16	13
	Stoic cluster		0	0	0	0	0	0	0	0	30	23
	With the partner	I	−0.41	0.44	1.71	0.17	−0.56	−0.51	−0.63	−0.11	52	41
		II	−0.59	−0.32	0.30	−0.33	−0.23	−0.51	−0.31	1.99	18	14
		III	1.69	−0.03	−0.23	−0.26	−0.20	−0.70	−0.08	−0.20	15	12
		IV	−0.48	1.04	0.21	−0.38	0.88	−0.18	−0.66	−0.01	13	10
	Stoic cluster		0	0	0	0	0	0	0	0	30	23
	With the experimenter	I	−0.60	0.21	2.02	0.20	−0.30	−0.51	−0.51	−0.24	39	31
		II	−0.61	0.03	0.13	−0.46	−0.18	−0.46	−0.27	2.08	15	12
		III	0.78	−0.10	−0.34	−0.51	0.55	0.01	−0.45	−0.04	23	18
		IV	−0.47	1.9	0.23	0.02	−0.08	−0.49	−0.49	−0.55	21	16
	Stoic cluster		0	0	0	0	0	0	0	0	30	23
Sample 2	With the experimenter	I	−0.32	0.30	2.20	−0.26	−0.37	−0.42	−0.51	−0.24	44	39
		II	−0.46	0.09	0.62	−0.44	0.08	−0.46	−0.66	1.04	19	17
		III	1.82	−0.24	−0.19	−0.58	−0.19	−0.10	−0.49	−0.04	13	12
		IV	−0.50	2.02	0.66	−0.48	−0.41	−0.43	−0.53	−0.32	19	17
	Stoic cluster		0	0	0	0	0	0	0	0	17	15

Table 3 Demographic (sex, age) and pain sensitivity (pain threshold, VAS ratings, PCS score) characteristics for each cluster.

Situation		Cluster	Sex Men : women	*p*	Age Mean (SD)	*p*	Pain threshold Mean (SD)	*p*	VAS pain ratings Mean (SD)	*p*	PCS score Mean (SD)	*p*
Sample 1	Alone	I	16:20	0.395	39.4 (15.0)	0.756	46.5 (1.2)	0.615	80.3 (12.2)	0.320	14.3 (7.8)	0.364
		II	12:16		42.6 (13.3)		46.1 (1.2)		83.8 (11.7)		14.3 (7.1)	
		III	9:9		39.2 (14.4)		46.4 (1.3)		77.4 (13.4)		14.5 (6.8)	
		IV	9:7		41.7 (13.5)		46.3 (1.1)		84.5 (12.7)		16.5 (8.5)	
		'Stoic'	18:12		38.3 (12.0)		46.5 (1.3)		82.5 (11.4)		12.0 (6.8)	
	With the partner	I	24:28	0.486	40.1 (14.9)	0.875	46.5 (1.1)	0.105	82.7 (12.2)	0.386	14.8 (6.9)	0.284
		II	8:10		42.5 (14.6)		45.8 (1.3)		78.5 (11.5)		16.4 (6.9)	
		III	7:8		40.7 (14.2)		46.5 (1.1)		76.0 (18.0)		13.1 (6.5)	
		IV	7:6		40.8 (10.8)		46.0 (1.2)		82.3 (13.4)		13.7 (7.5)	
		'Stoic'	18:12		38.3 (12.0)		46.5 (1.3)		81.0 (12.0)		12.0 (6.8)	
	With the experimenter	I	17:22	0.230	39.2 (14.9)	0.523	46.4 (1.1)	0.851	82.4 (11.3)	0.446	14.6 (8.0)	0.448
		II	6:9		41.2 (13.6)		46.1 (1.1)		84.8 (9.5)		15.6 (6.2)	
		III	10:13		39.2 (14.2)		46.2 (1.3)		78.8 (15.8)		15.0 (7.7)	
		IV	13:8		44.6 (12.4)		46.5 (1.4)		79.2 (15.8)		13.8 (7.3)	
		'Stoic'	18:12		38.3 (12.0)		46.5 (1.3)		79.4 (11.8)		12.0 (6.8)	
Sample 2	With the experimenter	I	22:22	0.697	23.2 (5.3)	0.408	46.1 (1.5)	0.828	77.0 (16.9)	0.186	18.1 (7.3)	0.361
		II	7:12		23.0 (6.2)		46.0 (1.3)		80.4 (10.2)		20.5 (5.5)	
		III	6:7		23.8 (3.9)		45.9 (1.9)		75.4 (15.1)		15.9 (7.7)	
		IV	8:11		21.0 (2.1)		46.0 (1.1)		78.7 (15.0)		18.3 (8.8)	
		'Stoic'	10:7		23.5 (3.3)		45.6 (1.1)		69.9 (8.3)		15.9 (7.7)	

p-values indicate differences between clusters. Values are given separately for sample 1 (for all three situations) and sample 2. PCS, Pain Catastrophizing Scale; SD, standard deviation; VAS, visual analogue scale.

slighter co-activation of AU 6_7 (contraction of the muscles surrounding the eyes). This cluster was labelled as 'opened mouth with narrowed eyes'. To better illustrate this facial activity pattern, photos are displayed in Fig. 1. This facial encoding pattern was displayed by 20% of subjects in sample 2 and by a comparable percentage of 15–29% in sample 1 (depending on the social situation).

3.1.1.3 Cluster III

As can be seen in Table 2, the third cluster scored highest on AU 1_2 (lifting the eyebrows) and was only found in one analysis (sample 1, with the experimenter) accompanied by AU 14 (dimpler). Given that in three out of four analyses, this cluster was solely composed of AU1_2; this cluster was simply labelled as 'raised eyebrows'. To better illustrate this facial activity pattern, photos are displayed in Fig. 1. This facial encoding pattern was displayed by 14% of subjects in sample 2 and by 16–24% of subjects in sample 1 (depending on the social situation).

3.1.1.4 Cluster IV

As can be seen in Table 2, the fourth cluster scored highest on AU 4 (contraction of the eyebrows) with a slighter co-activation of AU 6_7 (contraction of the muscles surrounding the eyes). In two out of four analyses, this cluster also showed a co-activation of AU 14 (dimpler). However, given that this co-activation only occurred in two of the analyses, we refrained from including it in the labelling. Thus, this cluster was labelled as 'furrowed brows with narrowed eyes'. To better illustrate this facial activity pattern, photos are displayed in Fig. 1. This facial encoding pattern was displayed by 20% of subjects in sample 2 and by a comparable percentage of 13–21% in sample 1 (depending on the social situation).

3.1.1.5 Stoic cluster

As can be seen in Table 2, a considerable number of participants (23% in sample 1 and 15% in sample 2) did not show any facial responses during pain stimulation.

3.2 Cluster replicability and cluster membership stability

3.2.1 Replicability across samples

Comparing the original cluster assignments for the participants in sample 1 (while being tested with the experimenter) with their nearest centroid classification using the cluster centres of sample 2 resulted into an impressive overall agreement of 84% and a kappa value of $\kappa = 0.80$ ($t = 13.3$; $p \leq 0.001$). Similarly, comparing the original cluster assignments for the participants in sample 2 with their nearest centroid classification using the cluster centres of sample 1 (while being tested with the experimenter) resulted into an impressive overall agreement of 87% and a kappa value of $\kappa = 0.82$ ($t = 13.4$; $p \leq 0.001$). Given that for the interpretation of the κ-value, a value of 0.8 or higher indicates a very good agreement (Altman, 1991); the cross-validation procedures revealed high agreements and high replicability of facial activity clusters across the two samples.

3.2.2 Replicability across situations

Given that participants in sample 1 were tested in three situations, we could test the cluster replicability also across situations in the same sample. As cluster centres, we used the cluster classifications from the situation when participants were tested with the experimenter (to make it comparable with the replicability analyses across samples). Comparing the original cluster assignments for the *alone* situations with their nearest centroid classification using the cluster centres of the *with the experimenter* situation resulted into an impressive overall agreement of 88% and a kappa value of $\kappa = 0.83$ ($t = 13.8$; $p \leq 0.001$). Similarly, comparing the original cluster assignments for the *partner* situation to their nearest centroid classification using the cluster centres of the with the experimenter situation again resulted into an impressive overall agreement of 83% and a kappa value of $\kappa = 0.80$ ($t = 12.9$; $p \leq 0.001$). Thus, the cross-validation procedures revealed high agreements and high replicability of facial activity clusters also across different social situations.

3.2.3 Cluster membership stability across situations

In addition to comparing the replicability of cluster solutions across situations, we also wanted to assess whether participants of sample 1 remained in the same facial activity cluster across situations or in other words how stable an individual displays a certain facial activity pattern across situations (cluster membership stability). Kappa statistics revealed significant cluster membership stability across situations (*alone vs. with the partner*: $\kappa = 0.433$, $t = 6.73$, $p \leq 0.001$; *alone vs. with the experimenter*: $\kappa = 0.456$, $t = 7.85$, $p \leq 0.001$; *with the experimenter vs. with the partner*: $\kappa = 0.414$, $t = 6.43$,

$p \leq 0.001$). However, given that κ-values were around 0.4, the agreement can only be interpreted as moderate. Likewise, when considering the cross tabulations, it becomes evident that the majority of participants did indeed show the same type of facial activity patterns across situations, but nevertheless, 34–38% of the participants also changed cluster membership between situations. These changes in cluster membership between situations were, however, unsystematic. This means that no obvious pattern of change (e.g., members of cluster I change most often to cluster II) could be detected.

3.3 Demographic and pain-processing characteristics in the different facial expression clusters

In Table 3, the demographic (age, sex) and pain-processing characteristics of each facial expression cluster of sample 1 (for all three situations) and sample 2 are presented. Neither for age nor for sex did we find significant differences between cluster groups (see Table 3). Moreover, pain processing (as measured by pain threshold, VAS pain ratings as well as by the tendency to catastrophize about pain) did also not differ between cluster groups (see Table 3).

4. Discussion

The aim of the present study was to use hierarchical cluster analyses to look for possible systematic combinations of single facial activities in episodes of pain (distinct facial activity patterns of pain). So far, studies on facial responses during pain have mainly followed the rational of research on facial expressions of emotions (e.g., anger, disgust, joy) by trying to identify a uniform set of facial responses that might distinguish pain from other types of affective states. Prkachin (1992) and Prkachin and Solomon (2008) have been quite successful in describing such a set of prototypical facial expressions of pain. However, when considering individual responses during pain, considerable inter-individual variations can be observed (Craig et al., 2011), and only a very small number of individuals actually display the complete set of facial actions of the so-called pain prototypical expression. Although we are not the first to acknowledge these variations (e.g., Craig et al., 2011; Ahola Kohut et al., 2012), we have tried to contribute to the discussion by systematically separating regular from idiosyncratic facial activity patterns during pain. We were able to show that in fact a multiple-cluster solution seems best fitting for categorizing facial activity during pain. Each cluster was composed of different combinations of single facial responses in pain episodes, namely, (I) narrowed eyes with furrowed brows and wrinkled nose; (II) opened mouth with narrowed eyes; (III) raised eyebrows; and (IV) furrowed brows with narrowed eyes (see also Fig. 1 for better illustration). The fact that these four facial activity patterns could be replicated across situations, as well as across two different samples, strongly supports the robustness of our findings. Since the patterns appeared to characterize better the individual than the situation, one might consider them – with caution (given the only moderate replicability) – as individual response stereotypies. Besides these four facial activity patterns of pain, a considerable number of individuals (approximately 20%) did not show any type of facial responses during painful stimulation, forming a kind of stoic cluster. Thus, even a lack of expression can be a face of pain.

4.1 Detailed description of the four facial activity patterns of pain

The most frequent facial activity pattern – in both samples as well as in all situations (alone, with their partner, with the experimenter) – was 'narrowed eyes with furrowed brows and wrinkled nose', displayed by approximately 40–50% of the participants. This facial activity pattern was not only the most frequent but was also impressively replicable, given that the nearest centroid cross-validation technique (McIntyre and Blashfield's, 1980) revealed agreement rates of 93–97% between samples and between situations. In comparison, the average agreement rates (averaged across all four clusters) lay between 80% and 87%. Likewise, cluster IV ('furrowed brows with narrowed eyes'; in contrast to cluster I, the furrowed brow here is the dominant action) also showed impressive replicability with agreement rates of 95–100% between samples and situations. However, this activity pattern was less prevalent (15–20% in both samples) than the first one. It is interesting that these two very stable facial activity patterns of pain show a great overlap with regard to the single facial activities involved, given that narrowed eyes and furrowed brows appeared in both clusters. Nevertheless, the phenomenology is quite different. Whereas the orbicularis oculi activation (narrowed eyes) is the most prominent response in the first cluster, this response is surpassed by the corrugator activation (furrowed brows) in the second cluster (see Fig. 1). Interestingly, the orbicularis oculi activation (narrowed eyes) was also part of a third cluster (frequency = 15–28%), namely, 'opened mouth with narrowed eyes'; however, here it was accompanied by a strong movement of the mouth (see Fig. 1).

Finding orbicularis oculi activation to be part of three distinct facial activity patterns is well in line with previous reports that found narrowed eyes to be the most common facial response to pain (Craig et al., 2011). Moreover, we could recently demonstrate that the orbicularis oculi activation is most closely associated with the sensory dimension of pain (Kunz et al., 2012), and thus three out of four facial activity patterns seem to include facial elements encoding the sensory dimension of pain. There was only one facial activity pattern that lacked orbicularis oculi activation, namely, 'raised eyebrows'. This facial behaviour has been most closely linked to the experience of surprise in earlier studies (Ekman and Friesen, 1987). In the context of pain, raised eyebrows might encode the suddenness/novelty component of the pain experience. Those facial responses that seem to be mostly linked to the affective component of pain, namely, 'furrowed brows' and 'wrinkled nose' (Kunz et al., 2012) occurred in two of the four facial expression patterns (cluster I and cluster IV). Accordingly, the four distinct activity patterns might be considered behavioural synonyms for pain, which encode slightly different aspects of pain.

The behavioural synonyms might be comparable with synonyms in verbal languages, which convey the same but mainly not identical messages, allowing for subtle differences in meaning. When looking at the photos in Fig. 1, some of the activity clusters indeed seem to convey slightly different meanings such as pain blended with surprise in cluster III and pain blended with disgust in cluster I. In future studies, specific self-report measures assessing these other affective states should be included to test whether they might help to explain cluster membership. These potential slight differences in meaning of the facial activity patterns might lead to different interactions and transactions (Hadjistavropoulos et al., 2011) between sender (patient) and receiver (caring person) in pain management.

4.2 A uniform set of facial responses versus different facial activity patterns of pain

When comparing these four different facial activity patterns with the so-called prototypical facial expression of pain (Prkachin, 1992; Craig et al., 2011) it becomes apparent that most of the single facial actions within this – as 'pain prototypically' postulated – set of facial responses can also be found in three out of the four cluster solutions. Accordingly, there is a considerable amount of agreement between our findings and previous reports. However, there is at least one divergent conclusion. With the present findings, we can clearly demonstrate that these single facial actions that often occur in the context of pain do not form a uniform set of facial expression with some unsystematic variants but combine to regular and distinct facial activity patterns of pain. These combinations do never include all single facial actions at once.

The necessity to consider more than one facial activity pattern might explain the weak effects of decoder trainings, making use of the prototypical set of facial activities (Solomon et al., 1997). The improvement in decoding performance after such trainings was only small to moderate. Training observers to recognize the four different facial activity patterns of pain (instead of one set of prototypical single actions) might lead to better results.

Interestingly, however, we did not find any demographic or pain-related psychological variable explaining which individual shows which type of facial activity pattern. Neither sex, age, pain sensitivity (pain threshold, pain ratings) nor pain catastrophizings could explain cluster membership. Whereas pain sensitivity (e.g., Kunz et al., 2004) and pain catastrophizing (Sullivan et al., 2006; Vervoort et al., 2011) have been found to explain differences in the vigour and amount of facial expressions of pain,[1] we have not yet detected the critical characteristics explaining which individual shows which type of facial activity pattern.

5. Limitations

Although the observed facial activity patterns stably occurred across situations as well as across different samples, our findings are nevertheless limited to experimentally induced pain. It is possible that facial activity, for example, during clinical pain might be clustered into different facial activity patterns than the ones we found, which has to be examined in future studies. In addition, by nature of cluster analyses all activity patterns were grouped, which does not exclude the existences of idiosyncratic and atypical responses. However, they were not too prevalent because the goodness of fit of the four-cluster solution would have been lower. Furthermore, we – as most of previous studies – did not consider the temporal sequence of facial actions in an activity pattern, which might form a further source of information in the encoding and decoding processes. Such types of analyses are even

[1]In the present study, however, catastrophizing could not explain differences in vigour and amount of facial expressions of pain (neither in sample 1 nor in sample 2).

more challenging, but should be performed as a next step to fully capture the facial language of pain.

6. Conclusion

With the present study, we were able to show that pain is not accompanied by one invariant and uniform facial activity pattern but instead is encoded in at least four different activity patterns (as well as in a stoic expression) that could be verified in two independent samples of subjects. The only moderate stability of these facial activity patterns within individuals across situations suggests the existence of individual facial reaction stereotypies for pain in the majority of participants but not in all. Whether the four facial activity patterns represent behavioural synonyms for the internal state 'pain' with slightly differing meanings remains to be determined. Raising awareness in observers for these different activity patterns might hold the potential of improving the detection of pain. This might be especially important in clinical practice, particularly for the assessment of pain in non-verbal individuals (e.g., patients with moderate-to-severe dementia).

Acknowledgements

We thank Anna Karmann, Florian Bauer, Nina Kunzelmann, Jessica Peter and Viktoria Zorn for the support in data collection and FACS coding.

Author contributions

M.K. and S.L. designed the study together and are responsible for the integrity of the work as a whole. M.K. conducted the study and did the analyses. Both authors discussed the results and wrote the article together.

References

Ahola Kohut, S., Pillai Riddell, R., Flora, D.B., Oster, H. (2012). A longitudinal analysis of the development of infant facial expressions in response to acute pain: Immediate and regulatory expressions. *Pain* 153, 2458–2465.

Altman, D.G. (1991). *Practical Statistics for Medical Research* (London: Chapman and Hall) p. 611.

Blashfield, R.K., Aldenderfer, M.S. (1988). The methods and problems of cluster analysis. In *Handbook of Multivariate Experimental Psychology*, J.R. Nesselroade, R.B. Cattell, eds. (New York: Plenum Press) pp. 447–474.

Craig, K.D., Prkachin, K.M., Grunau, R.V.E. (2011). The facial expression of pain. In *Handbook of Pain Assessment*, 3rd edn, D.C. Turk, R. Melzack, eds. (New York: Guilford) pp. 117–133.

Ekman, P.E., Friesen, W.V. (1987). *Facial Action Coding System* (Palo Alto, CA: Consulting Psychologists Press).

Eritz, H., Hadjistavropoulos, T. (2011). Do informal caregivers consider nonverbal behavior when they assess pain in people with severe dementia? *J Pain* 12, 331–339.

Hadjistavropoulos, T., Craig, K.D., Duck, S., Cano, A., Goubert, L., Jackson, P.L., Mogil, J.S., Rainville, P., Sullivan, M.J., de C Williams, A.C., Vervoort, T., Fitzgerald, T.D. (2011). A biopsychosocial formulation of pain communication. *Psychol Bull* 137, 910–939.

Hair, J.F., Black, W.C. (2000). Cluster analysis. In *Reading and Understanding More Multivariate Statistics*, L.G. Grimm, P.R. Yarnold, eds. (Washington, DC: American Psychological Association) pp. 147–205.

Hale, C., Hadjistavropoulos, T. (1997). Emotional components of pain. *Pain Res Manag* 2, 217–225.

Huber, C., Kunz, M., Artelt, C., Lautenbacher, S. (2010). Attentional and emotional mechanisms of pain processing and their related factors: A structural equations approach. *Pain Res Manag* 15, 229–237.

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., Prkachin, K.M. (2006). Testing two accounts of pain underestimation. *Pain* 124, 109–116.

Kunz, M., Chatelle, C., Lautenbacher, S., Rainville, P. (2008). The relation between catastrophizing and facial responsiveness to pain. *Pain* 140, 127–134.

Kunz, M., Chen, J.I., Lautenbacher, S., Vachon-Presseau, E., Rainville, P. (2011). Cerebral regulation of facial expressions of pain. *J Neurosci* 31, 8730–8738.

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.

Kunz, M., Mylius, V., Schepelmann, K., Lautenbacher, S. (2004). On the relationship between self-report and facial expression of pain. *J Pain* 5, 368–376.

Kunz, M., Prkachin, K., Lautenbacher, S. (2009). The smile of pain. *Pain* 145, 273–275.

Kunz, M., Scharmann, S., Hemmeter, U., Schepelmann, K., Lautenbacher, S. (2007). The facial expression of pain in patients with dementia. *Pain* 133, 221–228.

McIntyre, R.M., Blashfield, R.K. (1980). A nearest-centroid technique for evaluating the minimum-variance clustering procedure. *Multivariate Behav Res* 15, 225–238.

Milligan, G.W., Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50, 159–179.

Mojena, R. (1977). Hierarchical grouping methods and stopping rules: An evaluation. *Comput J* 20, 359–363.

Prkachin, K.M. (1992). The consistency of facial expressions of pain: A comparison across modalities. *Pain* 51, 297–306.

Prkachin, K.M., Solomon, P.E. (2008). The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain* 139, 267–274.

Rovniak, L.S., Sallis, J.F., Saelens, B.E., Frank, L.D. (2010). Adults' physical activity patterns across life domains: Cluster analysis with replication. *Health Psychol* 29, 496–505.

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.

Solomon, P.E., Prkachin, K.M., Farewell, V. (1997). Enhancing sensitivity to facial expressions of pain. *Pain* 71, 279–284.

Sullivan, M.J., Martel, M.O., Tripp, D., Savard, A., Crombez, G. (2006). The relation between catastrophizing and the communication of pain experience. *Pain* 122, 282–288.

Sullivan, M.J.L., Bishop, S., Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychol Assess* 7, 524–532.

Vervoort, T., Caes, L., Trost, Z., Sullivan, M., Vangronsveld, K., Goubert, L. (2011). Social modulation of facial pain display in high-catastrophizing children: An observational study in schoolchildren and their parents. *Pain* 152, 1591–1599.

Ward, J. (1963). Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58, 236–244.

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