OPEN ACCESS
OPUS AUGSBURG

Universität Augsburg
Universitätsbibliothek

Faces of clinical pain: inter individual facial activity patterns in shoulder pain patients

Miriam Kunz, Kenneth Prkachin, Patricia E. Solomon, Stefan Lautenbacher

Angaben zur Veröffentlichung / Publication details:

Kunz, Miriam, Kenneth Prkachin, Patricia E. Solomon, and Stefan Lautenbacher. 2021. "Faces of clinical pain: inter individual facial activity patterns in shoulder pain patients." *European Journal of Pain* 25 (3): 529–40. https://doi.org/10.1002/ejp.1691.

Nutzungsbedingungen / Terms of use:

CC BY-NC 4.0

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
CC-BY-NC 4.0: Creative Commons: Namensnennung - Nicht kommerziell
Weitere Informationen finden Sie unter: / For more information see:
https://creativecommons.org/licenses/by-nc/4.0/deed.de

Check for updates

DOI: 10.1002/ejp.1691

ORIGINAL ARTICLE

EJP European Journal of Pain

Faces of clinical pain: Inter-individual facial activity patterns in shoulder pain patients

Miriam Kunz[1,2] | Kenneth Prkachin[3] | Patricia E. Solomon[4] | Stefan Lautenbacher[1]

[1]Physiological Psychology, Otto-Friedrich University, Bamberg, Germany

[2]Medical Psychology and Sociology, University of Augsburg, Augsburg, Germany

[3]Department of Psychology, University of Northern British Columbia, British Columbia, Canada

[4]School of Rehabilitation Science, McMaster University, Hamilton, Ontario, Canada

Correspondence
Miriam Kunz, Department of Medical Psychology and Sociology, University of Augsburg, Universitätsstr. 2, Germany – Augsburg.
Email: miriam.kunz@med.uni-augsburg.de

Funding information
The study was supported by research grants from the Deutsche Forschungsgemeinschaft (Ku2294/4) and by Grant No. MOP 53301 from the Canadian Institutes of Health Research.

Abstract

Background: Facial activity during pain is composed of varying combinations of a few elementary facial responses (so-called Action Units). A previous study of experimental pain showed that these varying combinations can be clustered into distinct facial activity patterns of pain. In the present study, we examined whether comparable facial activity patterns can also be identified among people suffering from clinical pain; namely, shoulder pain.

Methods: Facial expressions of patients suffering from shoulder pain ($N = 126$) were recorded while twice undergoing a battery of passive range-of-motion tests to their affected limbs (UNBC-McMaster Shoulder Pain Expression Archive Database), which elicited peaks of acute pain. Facial expressions were analysed using the Facial Action Coding System to extract facial Action Units (AUs). Hierarchical cluster analyses were used to look for characteristic combinations of these AUs.

Results: Cluster analyses revealed four distinct activity patterns during painful movements. Each cluster was composed of different combinations of pain-indicative AUs, with one AU common to all clusters, namely, "narrowed eyes". Besides these four clusters, there was a "stoic" pattern, characterized by no discernible facial action. The identified clusters were relatively stable across time and comparable to the facial activity patterns found previously for experimental heat pain.

Conclusions: These findings corroborate the hypothesis that facial expressions of acute pain are not uniform. Instead, they are composed of different combinations of pain-indicative facial responses, with one omnipresent response, namely, "narrowed eyes". Raising awareness about these inter-individually different "faces of pain" could improve the recognition and, thereby, its diagnostic training for professionals, like nurses and physicians.

Significance: Similar to experimental pain, facial activity during evoked pain episodes in shoulder pain patients could be clustered into distinct faces of pain. Each cluster was composed of different combinations of single facial responses, namely: narrowed eyes, which is displayed either alone or in combination with opened mouth or wrinkled nose, or furrowed brows and closed eyes. These distinct faces of pain

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2020 The Authors. European Journal of Pain published by John Wiley & Sons Ltd on behalf of European Pain Federation - EFIC ®

may inform the training of professionals and computers how to best recognize pain based on facial expressions.

1 | INTRODUCTION

It has been repeatedly shown that the experience of pain is accompanied by a small subset of elementary facial activities (so-called Action Units, AUs; Ekman & Friesen, 1987), namely, *narrowed eyes, furrowed brows, raising the upper lip/nose wrinkling, opening of the mouth and eye closure* (Kunz et al., 2019; Prkachin & Solomon, 2008; Prkachin, 1992). These elementary facial activities are rarely presented all at once during pain but appear mostly partially in different combinations (Craig et al., 2011; Kunz & Lautenbacher, 2014). Given these varying combinations of AUs, it becomes apparent that there is not only one uniform facial expression of pain that can be observed at all times and in each individual. Using cluster analyses, Kunz and Lautenbacher (2014) showed that AUs during experimental heat pain clustered into four distinct active patterns besides a fifth stoic expression (with no discernible facial actions). The four active patterns were combinations of the AU "narrowed eyes" with either (I) "raising the upper lip/nose wrinkling" and "furrowed brows", (II) "furrowed brows" or (III) "opening of the mouth" and, in addition, a pattern (IV) solely composed of "raised eyebrows". These inter-individually different patterns were unrelated to participants` demographic characteristics or their pain sensitivity and may be behavioural synonyms for the internal state "pain". A follow-up study showed that a brief training focusing on these different facial activity patterns improved recognition of pain significantly (observers were better able to differentiate pain from facial expressions of disgust and from neutral expressions) (Kunz & Lautenbacher, 2015). Thus, a better understanding of the "different faces of pain" may be important for clinical practice because it improves the fidelity with which pain is decoded.

However, in order to apply the cluster findings to clinical practice, it has to be first established that these patterns can be replicated for clinical pain. This was the aim of the present study. In general, elementary facial responses occurring during experimental acute pain are very comparable to facial responses occurring in clinical acute pain, with the only exception that eye closure is displayed more intensely in clinical pain (Kunz et al., 2019). Given the similarity of single facial responses to experimental and clinical pain, we hypothesized that facial responses occurring during clinical pain can also be clustered into inter-individually different facial activity patterns similar to those for experimental pain. To test this, facial responses to acute pain in shoulder pain patients undergoing a range of motion tests (videos were taken from the UNBC-McMaster Shoulder Pain Expression Archive Database (Lucey et al., 2011; Prkachin & Solomon, 2008)) were analysed. Given that patients underwent all motion tests twice, we could test (i) replicability of cluster solutions and (ii) stability of "cluster-membership" within individuals across time. Some of the established fundamental influences on pain processing were selected as potential individual determinants of cluster membership, i.e. age, sex, intensity of acute pain, duration of shoulder pain, pain medication and accompanying disability.

2 | MATERIALS AND METHODS

The video recordings of facial expressions were taken from the *UNBC-McMaster Shoulder Pain Expression Archive Database*, data from which have been presented in Lucey et al., 2011; Prkachin & Solomin, 2008.

2.1 | Participants (UNBC-McMaster Shoulder Pain Expression Archive Database)

Sixty-three women and 63 men (mean age = 42.23; *SD* = 14.48) who self-identified as suffering from shoulder pain participated. They were recruited from two active physiotherapy clinics and by advertisements on the university campus. Slightly under 25% were students; 12.4% were unemployed and 16.2% retired. Other identified occupations included healthcare (7.8%), civil service (7.8%), service (7.8%), trades (7.8%), management (7%), manufacturing (2.3%), and homemaker (2.3%), the cultural sector (2%), and primary industry (1.8%). Seventy per cent of the participants were able to provide a pain diagnosis identified by an attending physician or physiotherapist. Diagnoses included arthritis, bursitis, tendonitis, fibromyalgia, subluxation, rotator cuff injury, impingement syndrome, bone spur, capsulitis and dislocation. Thus, although the underlying physical conditions were not uniform, the sample was broadly representative of the types of problems commonly seen in shoulder pain clinics. Other demographic features of the sample are presented in Prkachin and Solomon (2008). In general, the sample was relatively well-educated, with approximately 74% having at least some postsecondary education, but broadly representative in terms of ethnic background of urban regions in Canada. The vast majority (79.8%) reported having experienced their injury more than 6 months prior to testing and consequently met one of the common criteria for chronic pain. Roughly half (50.4%) of the participants reported using medication for their pain. All participants provided written informed consent. The study protocol was approved by the Research

Ethics Boards of McMaster University and the University of Northern British Columbia.

2.2 | Procedure (UNBC-McMaster Shoulder Pain Expression Archive Database)

Participants were tested in a laboratory room, which had a bed for performing passive range-of-motion tests. Four range-of-motion tests were performed: abduction (lifting the arm sideways and up in the frontal plane), flexion (lifting the arm forward and up in the sagittal plane), internal rotation (bending the arm 90 degrees at the elbow, abducted 90 degrees and turning internally) and external rotation of the arm (bending the arm 90 degrees at the elbow, abducted 90 degrees and turning externally). All four tests were performed under active and passive conditions on the affected and the unaffected limb. Active tests were performed before passive tests because that is the order in which they are usually administered clinically. The four tests within active and passive conditions were performed in a random order. Tests were performed by one of two female physiotherapist research assistants who followed a standardized clinical protocol. Active tests were performed while the participant was standing, and passive tests were performed with the participant resting in a supine position on a bed. The physiotherapist moved the limb until the maximum range was achieved or the participant told her to stop. After each test, participants rated the maximum pain it had produced using a visual analogue scale (VAS, 10 cm line, scored by measuring the response in cm), anchored at the ends with the words, ‘‘No pain” and ‘‘Pain as bad as could be” respectively.

After a 20-min break, during which participants completed psychometric tests, the same range-of-motion tests were repeated. Each test was recorded from a frontal view on digital videotape using a Sony digital video camera focused on the face.

In addition to performing the range-of-motion tests, participants were asked to provide demographic information and to complete questionnaires, including the Coping Strategies Questionnaire (CSQ, Abbott, 2010; Rosenstiel & Keefe, 1983) and the Shoulder Pain and Disability Index (SPADI, Roach et al., 1991; Williams et al., 1995).

The CSQ is a 44-item questionnaire that was developed to measure coping in pain patients. As originally constructed, it comprises two general questions and items selected on a rational basis to reflect seven dimensions of coping: Diverting Attention, Reinterpreting Pain Sensations, Coping Self-Statements, Ignoring Sensations, Praying/Hoping, Catastrophizing and Increasing Behavioural Activity. Items are rated on a 7-point scale, with endpoints “never” and “always”. Initial studies established moderate-to-good internal consistency of most of the identified dimensions (Robinson et al., 1997). Reported test–retest reliability values for the various dimensions over timeframes from 1 day to 6 weeks range from 0.48 to 0.91. There has been disagreement about its underlying factor structure; however, several of the originally proposed dimensions (Diverting Attention, Ignoring Pain Sensations, Catastrophizing and Coping Self-Statements) have been identified repeatedly in factor analyses, show moderate-to-good internal consistency and construct validity (Robinson et al., 1997). In previous research, only a few components of the CSQ – in particular, Catastrophizing – have accounted for variability in pain-related expressions (Prkachin & Mercer, 1989). For this reason, we only focussed on the subscale Catastrophizing for further analyses.

The SPADI was developed as a specific measure of pain and disability associated with shoulder pathology. Its 13 items are rated on 11-point scales, grouped into 5 on pain intensity (with the endpoints “no pain” to “worst pain imaginable”) and 8 on disability (“no difficulty” to “so difficult required help”). Studies of the internal consistency of the scales have showed high internal consistency and moderate test–retest reliability (Williams et al., 1995) and a systematic review (Bot et al., 2004) found it to have good evidence for construct validity and sensitivity. A previous report on the present sample (MacDemid et al., 2006) showed internal consistency of the subscales to be high ($\alpha > 0.92$), and evidence of convergent and discriminant validity.

2.3 | Facial expression analyses (conducted specifically for the current study)

For facial expression analyses, we only focussed on passive tests performed on the affected limb, since these elicited more facial expressions of pain than active movements or movements of the unaffected side (thus, limiting the number of stoic expressions).

Facial expressions were coded from the video recordings using FACS, which is based on anatomical analysis of facial movements and distinguishes 44 different Action Units (AUs) produced by single muscles or combinations of muscles. Although the video recordings had been FACS coded previously (Prkachin & Solomon, 2008), we re-analyzed the videos with FACS given that previously only 11 of the 44 AUs had been coded. Thus, the videos were FACS re-coded by two FACS coders who identified the frequency (how often was an AU displayed) and intensity (5-point scale (1–5)) of all 44 AUs. For each movement test, there was a separate video clip; resulting in $126 \times 8 = 1{,}008$ video clips altogether. Five per cent of the video clips ($N = 50$ randomly selected video clips) was coded by both coders to calculate inter-rater reliability (using the Ekman–Friesen formula; Ekman & Friesen, 1987). Inter-rater reliability was 0.88, which is comparable to previous studies (e.g. Kunz et al., 2011; Kunz, Faltermeier, et al., 2012; Kunz & Lautenbacher, 2014). Software designed for the analysis of observational data (Observer Video-Pro; Noldus Information Technology) was used to segment the videos and to enter the FACS codes into a time-related database. All facial responses occurring within the duration of each motion test (each test lasting between 5 and 15 s) were coded. Given that there were four motion tests (passive movement of the affected limb) that were repeated twice (T1 and T2), eight potentially painful motion tests were FACS coded per participant.

For further analyses, we joined frequency and intensity values for each AU by multiplying the two (product term). Furthermore, those AUs that represent similar facial movements were combined (AU1 & AU2; AU6 & AU7; AU9 & AU10; AU25 & AU26 & AU27), as has been done in most preceding studies on facial responses to pain (Kunz et al., 2004, 2007; 2008; Prkachin, 1992). To select AUs to be entered into the cluster analyses, we calculated how often each occurred during the motion tests at each time point (T1 & T2). Only those AUs that occurred in at least 5% of the motion tests both at T1 as well as at T2 were considered for further analyses (see Table 1, these AU being selected are printed in bold and are highlighted in grey).

We were not interested in identifying clusters that would merely reflect different inter-individual degrees of expressiveness, but rather in the extraction of distinct facial activity patterns, which may be constant across a group of individuals regardless of whether faintly or strongly expressed. In order to eliminate differences in the degree of facial expressiveness between participants, and in accordance with our previous cluster analysis of experimental pain (Kunz & Lautenbacher, 2014), AUs (product term) were z-transformed within each participant before entering the selected AUs (see shaded AUs in Table 1) into the cluster analyses.

2.4 | Data analyses

If participants were facially completely nonexpressive in response to the passive motion tests on the affected limb, we excluded them from the cluster analyses because of inapplicable data for this type of analysis (given that a zero expression ("0-values") cannot be statistically clustered into a specific pattern of facial 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 result section ("stoic cluster") to stress the fact that some individuals do not show any pain-related facial activity in situations evoking pain and that a "stoic-face" can presumably also accompany the experience of pain.

2.4.1 | Cluster analyses

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

(i) 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 two time points (T1 and T2), resulting in two 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 in the total within-group sum of squares. If a point is reached where clusters are combined

TABLE 1 Facial Action Units (AUs) with a critical frequency of occurrence of more than 5% during passive range of motion tests of the affected limb. Values are given separately for the two time points (T1, T2)

Action Unit	Description	Frequency of occurrence (in %) T1	T2	Wilcoxen Test: T1 versus. T2 (p-values)
[a]**AU 1/2**	**raised eyebrows**	**14.1**	**14.9**	0.854
[a]**AU4**	**furrowed brows**	**33.3**	**30.8**	0.827
[a]**AU6/7**	**narrowed eyes**	**120.6**	**114.7**	0.634
[a]**AU9/10**	**wrinkled nose**	**40.3**	**37.7**	0.788
AU 12	smiling	9.9	—	0.112
[a]**AU 14**	**dimpler**	**10.3**	**13.3**	0.237
[a]**AU 17**	**chin raiser**	**11.5**	**10.7**	0.911
AU 18	lip pucker	10.5	—	0.132
AU 24	lip tightener	—	5.6	0.499
[a]**AU25/26/27**	**opened mouth**	**56.4**	**52.5**	0.510
[a]**AU43**	**closed eyes > 0.5s**	**33.5**	**38.3**	0.244

Note: Percentage numbers > 100% indicate than an AU or a combination of AUs (e.g. AU6/7) were on average displayed more than once within a painful episode.

[a]AUs selected for the cluster analyses.

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 in the agglomeration coefficient is considered the "natural grouping scheme" (Hair & Black, 2000). To best determine the correct number of clusters, we inspected the rescaled distances as displayed in the hierarchical cluster dendogram, examined the change in the agglomeration coefficient and applied the Mojena stopping rule as a quantitative criterion to define a "significant jump" in the agglomeration coefficient (Milligan & Cooper, 1985; Mojena, 1977).

(ii) In a second step, the cluster means (centroids) that emerged from each hierarchical cluster analysis were used as initial seed points in nonhierarchical k-means cluster analyses as a method to verify the initial cluster solutions (Hair & Black, 2000). This was again done separately for each of the two time points (T1 and T2).

2.4.2 | Cluster replicability and stability across time points (T1 and T2)

In order to ensure that the resultant clusters do not only represent chance findings but instead stable facial encoding patterns of pain, we assessed the replicability and stability of the identified facial expression clusters across the different time points.

Cluster replicability: To assess replicability of facial encoding clusters across time points, McIntyre and Blashfield´s nearest-centroid cross-validation technique was used (McIntyre & Blashfield, 1980). This cross-validation procedure uses the cluster solutions from one time point (e.g. T1) to classify each person at another time point (e.g. T2) based on minimal distances (nearest centroid). Thus, for each individual at each time point, there are now two clustering solutions available (the original clusters and the nearest-centroid assigned clusters). The replication accuracy between cluster solutions was quantified using the kappa coefficient.

Stability of "cluster-membership" : We also calculated cross-tabulations between cluster solutions of T1 and T2 to investigate whether belonging to a certain facial expression cluster in one situation predicts membership in a comparable cluster in another according to kappa statistics. This procedure allows determining the stability of "cluster-membership" of each participant across time.

2.4.3 | Differences in demographic and pain characteristics in the facial expression clusters

As a last step, we investigated whether the identified facial expression clusters might differ with regard to demographic or pain characteristics by means of χ^2 tests (gender, pain medication) or univariate analyses of variance (age, VAS rating of shoulder pain, years of shoulder pain, SPADI and Catastrophizing-CSQ score). Findings were considered to be statistically significant at $\alpha < 0.05$.

SPSS-26 was used for all analyses.

3 | RESULTS

The SPADI yields three scores; one for pain, one for disability and total summary score. For this sample, means and standard deviations were as follows: pain – 53.27 (23.18), disability – 32.36 (22.85) and total – 42.83 (21.80). By comparison, mean scores for the largest population-based study of people complaining of shoulder pain (Hill et al., 2011) were 34.5, 21.7 and 26.6 respectively (standard deviations were not given). In a prospective cohort study of patients scheduled for shoulder arthroplasty (Angst et al., 2008), the respective means and standard deviations for a German version of the SPADI were 33.1 (20.3), 40.1 (19.9) and 36.6 (18.6). Taken together, the findings suggest that the sample recruited were indeed suffering from clinically significant problems with shoulder pain. For detailed analyses in the following, only the SPADI total score was considered.

On average, participants rated the passive range-of-motion tests on their affected side as mildly painful (VAS-T1: 2.8 ± 2.2; VAS-T2: 2.9 ± 2.2), with VAS pain intensity ratings not differing between the first (T1) and second (T2) execution ($t(128)=0.14$, $p=.890$). None of the participants reported "no pain" across the motion tests at T1 and T2. Moreover, facial responses elicited during the motion tests (composite score of all AUs with an occurrence > 5%) also did not differ between T1 (mean FACS score 5.7 ± 6.2) and T2 (mean FACS score 5.3 ± 5.5) ($t(128)=0.98$, $p=.331$). Consequently, the passive range-of-motion tests elicited comparable subjective and facial pain responses across the two time points.

3.1 | Cluster analyses

In both hierarchical cluster analyses (T1 and T2), the dendogram, the agglomeration coefficient and the Mojena stopping rule showed that a four-cluster solution was the most appropriate solution to cluster facial activity during the painful motion tests. The corresponding cluster means of the different AUs are shown in Table 2. We also included the group of nonexpressive individuals in Table 2 as an additional "cluster" and labelled it "stoic cluster". 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 nonhierarchical, k-means cluster analyses and found the results to be very similar to those of the hierarchical analyses (≥91% of the participants obtained the same group membership).

TABLE 2 Facial expression cluster profiles for T1 and T2 across the selected Action Units (z-scores)

Time point	Cluster	AU 1_2 raised eyebrows	AU4 furrowed brows	AU 6_7 narrowed eyes	AU9_10 wrinkled nose	AU 14 dimpler	AU 17 chin raiser	AU25_26_27 opened mouth	AU43 Closed eyes > 0.5s	Occurrence frequency of each cluster N	Time point (%)
T1	I	−0.36	−0.15	**1.91**	−0.25	−0.49	−0.41	0.05	−0.30	**50**	**39**
	II	−0.32	−0.52	**1.71**	−0.26	−0.32	−0.56	**0.68**	−0.42	**32**	**25**
	III	−0.60	0.05	**1.33**	**0.40**	−0.62	−0.36	0.05	−0.59	**23**	**18**
	IV	0.09	**0.58**	**0.77**	−0.42	−0.60	−0.62	**0.35**	**0.27**	**12**	**9**
	Stoic cluster	0	0	0	0	0	0	0	0	**12**	**9**
T2	I	−0.27	−0.17	**2.25**	−0.19	−0.47	−0.54	−0.32	−0.32	**54**	**42**
	II	−0.39	−0.18	**0.31**	−0.49	−0.47	−0.28	**1.94**	−0.45	**17**	**13**
	III	−0.69	0.01	**1.63**	**0.43**	−0.76	−0.59	0.06	−0.62	**31**	**24**
	IV	−0.53	**0.44**	**0.68**	−0.37	0.02	−0.47	−0.35	**0.39**	**13**	**10**
	Stoic cluster	0	0	0	0	0	0	0	0	**14**	**11**

Note: AUs being indicative for the separate clusters are printed in bold and marked in grey.

3.1.1 | Clusters of facial responses to pain

When visually inspecting cluster solutions for T1 and T2, it becomes evident that there were very similar clusters of facial responses to pain across time points (see Table 2). Examples of the found clusters are given in Figure 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) and was labelled as "*narrowed eyes*" (see also Figure 1, left column). This facial activity pattern was displayed by approximately 40% (rounded average) of the participants at both time points (for exact numbers see Table 2, right column) and proved to be the most frequently occurring facial activity pattern during pain.

Cluster II: As can be seen in Table 2, the second cluster scored highest on AU 6_7 (contraction of the muscles surrounding the eyes) and on AU25_26_27 (mouth opening) and was labelled as "*narrowed eyes with opened mouth*" (see also Figure 1, second panel). This facial encoding pattern was displayed by approximately 20% (rounded average) of participants across both time points.

Cluster III: As can be seen in Table 2, the third cluster scored highest on AU 6_7 (contraction of the muscles surrounding the eyes) with a co-activation of AU9_10 (raising the upper lip with wrinkled nose) and was labelled as "*narrowed eyes with wrinkled nose*" (see also Figure 1, third panel). This facial encoding pattern was displayed (similar to cluster II) by approximately 20% of participants across both time points.

Cluster IV: As can be seen in Table 2, the fourth cluster scored highest on AU 6_7 (contraction the muscles surrounding the eyes) with a co-activation of AU4 (contraction of the eyebrows) and AU43 (closing the eyes for longer than > 0.5 s). AU25_26_27 was also apparent in this cluster but only at T1. Thus, this cluster was labelled as "*narrowed eyes with furrowed brows and closed eyes*" (see Figure 1, fourth panel). This pattern was displayed by only 10% of participants at both time points.

"Stoic Cluster": As can be seen in Table 2, approximately 10% did not show any facial responses during painful motion tests.

3.2 | Cluster Replicability and "Cluster-Membership" Stability

3.2.1 | Replicability of cluster solutions across time points (T1 and T2)

Given that participants underwent each motion test twice, we could test cluster replicability across time points T1 and T2. Comparing the original cluster assignments for *T1* to their nearest-centroid classification using the cluster centres of T2 resulted in a good replicability of cluster assignments

FIGURE 1 Examples of facial actions occurring during passive motion tests of the affected limb in shoulder pain patients. The examples illustrate one of the four facial activity clusters: Custer I: please notice the narrowing of the eye aperture (AU6/7) from “Baseline” to “Motion-test”; Cluster II: please notice the narrowing of the eye aperture (AU6/7) and the opening of the mouth (AU25/26/27) from “Baseline” to “Motion-test”; Cluster III: please notice the narrowing of the eye aperture with the eyes not being fully closed (AU6/7) and the deepening of the nasolabial furrow (AU9/10) from “Baseline” to “Motion-test”; Cluster IV: please notice the tense closing of the eyes (AU6/7 with AU43) and the furrowed (AU4) from “Baseline” to “Motion-test”

of 75% and a substantial kappa value of $k = 0.61$ (t = 13.5, $p \leq .001$). Similarly, comparing the original cluster assignments for T2 to their nearest-centroid classification using the cluster centres of T1 again resulted in a good overall replicability of 71% and a substantial kappa value of $k = 0.61$ (t = 12.6, $p \leq .001$). Thus, the cross-validation procedures revealed good replicability of cluster solutions across time points.

3.2.2 | Cluster-membership stability across time points (T1 and T2)

In addition to comparing the replicability of cluster solutions, we also wanted to assess whether participants remained in the same facial activity cluster across time points or in other words how stable an individual displays a certain facial activity pattern across T1 and T2 (“cluster-membership” stability). Cross-tabulation (kappa-statistics) revealed significant “cluster-membership” stability across situations ($k = 0.52$, t = 10.620, $p \leq .001$). However, given that the k –value was around 0.5, the agreement can only be interpreted as moderate. Although the majority of participants did show the same type of facial activity patterns across situations, 39% changed “cluster-membership” between situations. These changes in “cluster-membership” between situations were, however, unsystematic. In other words, no obvious pattern of change (e.g. members of cluster I change most often to cluster II) could be detected when inspecting the cross-tabulation.

3.2.3 | Demographic and pain characteristics in the different facial expression clusters

As can be seen in Figure 2, the facial expression cluster groups (T1) did not differ with regard to age or gender (all p-values > .05). Moreover, VAS ratings of acute pain due to the motion tests, years of shoulder pain, pain medication, intensity and disability of the shoulder pain (SPADI total score) and pain catastrophizing (Catastrophizing-CSQ) did also not differ between cluster groups (all p-values > .05). The same nonsignificant findings were found for the facial expression

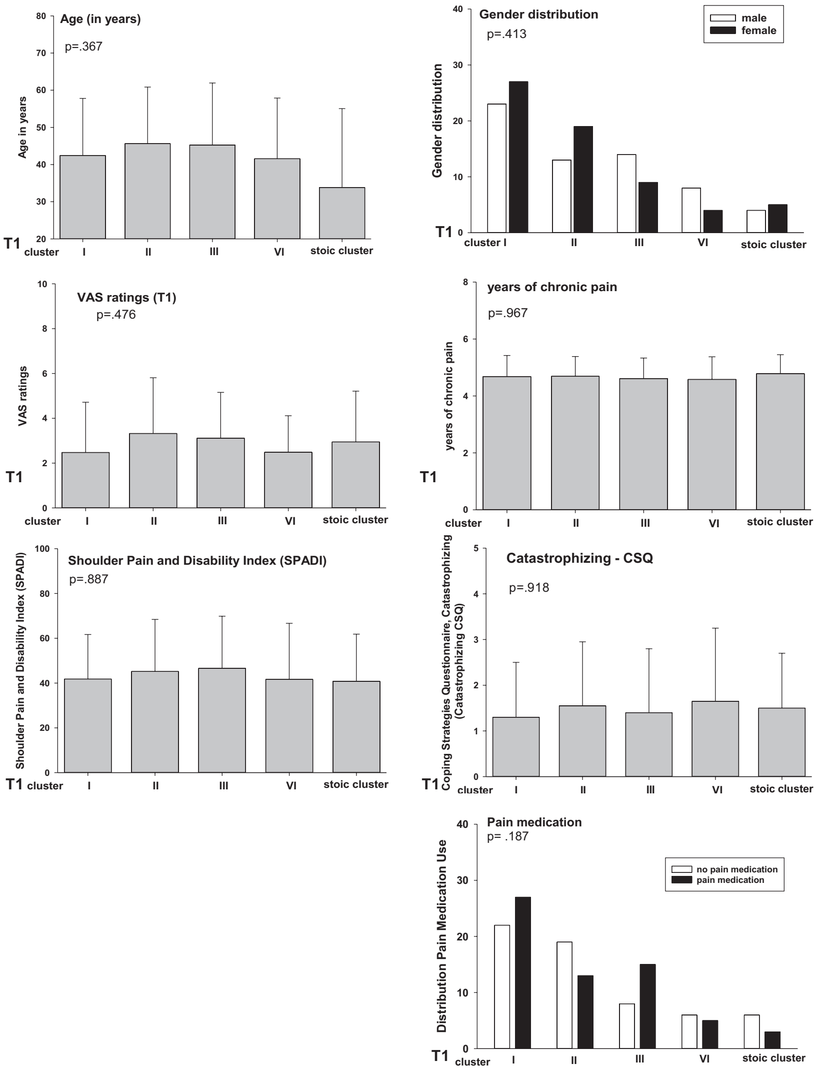

FIGURE 2 Demographic (gender, age) and pain (VAS ratings, years of shoulder pain, SPADI and Catastrophizing-CSQ scores) characteristics for each facial cluster at T1. P-values indicate analyses of variance or Chi-square test outcomes testing for differences between clusters

cluster groups at T2 (gender: p=.852; age: p=.263; VAS pain ratings: p=.172; years of shoulder pain: p=.956; pain medication: p=.210, SPADI total score: p=.261; Catastrophizing-CSQ: p=.296).

4 | DISCUSSION

The aim of the present study was to characterize inter-individual variants of facial expressions of clinical acute pain. This approach complements a previous attempt, aiming at clustering variations in facial expressions during experimental pain (Kunz & Lautenbacher, 2014). Similar to experimental pain, a multiple-cluster solution best described facial activity during motion-induced pain episodes in shoulder pain patients. Each cluster was composed of different combinations of single AUs, namely: *narrowed eyes*, which is displayed either alone or in combination with *opened mouth* or *wrinkled nose*, or *furrowed brows and closed eyes* (for longer than 0.5s). In addition, a small number of individuals did not show any type of identifiable facial responses, forming a “stoic cluster”.

4.1 | Description of the four active facial activity patterns during pain

Narrowed eyes (AU6_7) is omnipresent in all active clusters (except the stoic cluster) and the most frequent facial response compared to all others. This is in line with a recently published review (Kunz et al., 2019), showing that AU6_7 is the most frequent feature of the pain face. However, AU 6_7 is not specific to pain, given that it also occurs in anger, happiness and disgust (Kunz et al., 2013; Simon et al., 2008).

The variants in facial expressions of pain result from combinations of AU6_7 with other AUs, namely, with *wrinkled nose*, *opened mouth* or *furrowed brows* with *closing eyes*. These additions are not negligible extras, but are stable components of facial expressions of pain across clinical and experimental pain studies (Kunz et al., 2019). It was found (Blais et al., 2019; Kunz et al., 2019; Kunz, Lautenbacher, et al., 2012) that *wrinkled nose* and *furrowed brows* are indicative of the affective component, whereas *narrowed eyes* reflects the sensory component of pain. Considering the present data, this may suggest that the sensory component is reliably broadcasted in almost any case, whereas information about the affective component might be variably tied to *wrinkled nose* or *furrowed brows* or may be completely missing.

4.2 | Comparing patterns of facial activity during clinical and experimental pain

The question arises whether the present cluster solutions found for clinical acute pain correspond well to those previously found for experimental pain (Kunz & Lautenbacher, 2014). As affirmative answer, we observed substantial similarity between facial activity patterns produced by these two types of pain. The cluster *narrowed eyes* combined with *opened mouth* was identically present in this study of clinical pain (in around 20%) and in the earlier one with experimental pain (in 20%–29%). Similarly, the cluster *narrowed eyes* combined with *furrowed brows* was very similar in both types of pain (clinical: 10%, experimental: 13%–21%). A slight difference occurs with respect to the cluster *narrowed eyes* and *wrinkled nose* (clinical: 20%, experimental: 37%–53%), in which *furrowed brows* is added in the case of experimental pain. Also, the number of subjects showing no facial responses (stoic cluster) is very similar when comparing clinical (10%) and experimental pain (15%–23%). Thus, facial stoicism is not a feature of experimental pain alone.

However, there are two differences between experimental and clinical cluster solutions. In clinical pain, AU6_7 on its own formed a discrete cluster (approximately 40%), which we did not find in experimental pain. It is difficult to explain, why other AUs did not co-occur more reliably in the presence of *narrowed eyes*. It might be that clinical patients` long experience with pain resulted in an exclusive expression of the sensory dimension of pain and a fading of the affective dimension (Blais et al., 2019; Kunz, Lautenbacher, et al., 2012). Such a hypothesis of a divergence of sensory and affective components over time has to be tested in future longitudinal studies. The second difference is that the pattern *lifting eyebrows* (AU1_2) only occurred during experimental pain. AU1_2 is one of the principal facial movements in Ekman and Friesen's prototype of surprise, which may have been evoked by the sudden onset and varying intensity of our experimental heat pain.

In summary, AU6_7 appears to be the stable uniform backbone of the facial expression of pain, which combines with other key facial responses (AU9/10, AU4, AU25/26/27 and AU43) in both clinical and experimental pain. In accord with earlier studies, the basic frequency of all these other key facial responses was rather low, which does not allow for classifying them as pain indicative when observed alone, although their frequencies are still reliably higher than those of completely pain-irrelevant AUs (Kunz et al., 2019).

4.3 | Potential determinants of facial activity cluster membership

The different patterns of facial expressions of clinical pain were stable across time and relatively stable within individuals. However, we did not find among our demographic and pain variables (age, gender, intensity of acute pain, pain medication, duration, intensity and disability of shoulder pain or pain catastrophizing) any association with our facial activity patterns. This does not imply that it is generally impossible to identify in future studies variables of influence on the facial expression of pain, which help to predict cluster membership.

The stability of clusters across time and within individuals suggests that a given individual mainly expresses pain consistently in the same manner.

4.4 | Implications for the clinical and computational training of search algorithms

It is both a clinical and computational task to address how to best train a professional (Kunz & Lautenbacher, 2015; Rash et al., 2019; Solomon & Prkachin, 1997) or a computer (Asgarian et al., 2019; Ashraf et al., 2009; Hammal & Cohn, 2018; Kunz et al., 2017; Littlewort et al., 2009; Werner et al., 2016) to recognize pain. Our results advocate a hierarchical approach with only two steps, which aims at best possible diagnostic precision combined with high diagnostic economy: First, look for *narrowed eyes* (AU6/7). Second, if AU6/7 is present, look for further inter-individually varying facial signs of pain like *furrowed brows* (AU4), *wrinkled nose* (AU9/10), *opened mouth* (AU25/26/27) or *closing of the eyes* (AU43). If such further signs are present, the likelihood of the existence of pain has increased, which establishes pain as very good hypothesis; if not, further tests or additional information are favourable. However, if AU6/7 remains the only facial response, pain can still be a good hypothesis. If at the starting point, no or other AUs are present but pain is nevertheless very likely, further tests are advisable.

The clinical and computational training protocols developed so far have rarely considered inter-individual variants of facial expression of pain (e.g. Hassan et al., 2019). The four active combinations of AUs found both for experimental and clinical pain appeared to be stable enough to serve as training target for learning inter-individually different facial expressions of pain in future trials. Even if individuals sometimes change their habitual expression over several pain episodes, they appear to switch between the identified clusters and not between any variations, keeping such trainings useful.

4.5 | Limitations

Participants were volunteers who self-identified as having a significant problem with shoulder pain. Although the majority reported specific diagnoses, it is possible that diagnoses were not all accurate. Moreover, given the self-identification of pain, this might not be a standardized patient group which limits application to other patient groups. However, the sample was ethnically diverse and broadly representative of the urban regions in Canada; diagnoses and treatments were quite typical for physiotherapy clinics. Furthermore, pain scores suggest that the pain and disability of our sample were clinically significant.

Although it is already a Herculean effort to assess 129 pain patients twice during clinical provocation conditions and to FACS code their video-taped facial expressions, the number of cases is limited for cluster analyses, potentially causing instability. Thus, more subjects suffering from the same type of clinical pain would have been favourable but have remained out of reach.

We z-transformed the facial responses before entering them into the cluster analyses to highlight the relative contribution of certain AUs to the individual expression. The consequence of this was that in expressive subjects clearly visible AUs remained underemphasized, whereas in non-expressive subjects barely visible AUs became overemphasized. However, omitting the z-transformation would have resulted in clusters merely reflecting different inter-individual degrees of facial expressiveness.

5 | CONCLUSION

We aimed at identifying inter-individual variants in the facial expression of clinical acute pain as we did in an earlier study on experimental acute pain (Kunz & Lautenbacher, 2014). For note, we studied patients with persisting shoulder pain because of their vulnerability to arm and shoulder movements, which allows us precisely triggering motion-induced pain. Thus, we studied rather acute than chronic pain in our understanding.

The facial responses during motion-induced pain could be clustered into distinct expression patterns, namely, *narrowed eyes w*hich are either displayed alone or in combination with *opened mouth*, or with *wrinkled nose,* or with *furrowed brows* and *closed eyes > 0.5s*, These clusters are similar to those derived earlier from experimental data. These insights about the inter-individual variants may inform the training of professionals and computers how to best recognize pain from facial videos. Large datasets of such videos, based on different pain models, are urgently needed to study the inter-individual variants and to replicate the cluster solutions found in our present and previous studies (Kunz & Lautenbacher, 2014).

ACKNOWLEDGEMENTS

We thank Christopher Lischka and Sonja Paehl for the support in data collection and FACS-coding. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

There are no conflicts of interest.

AUTHOR CONTRIBUTIONS

K. P. and P.S. conducted the study in shoulder pain patients. M.K. FACS coded the data. M.K. and S.L designed the cluster analyses procedure. M. K. conducted the analyses. All authors discussed the results and wrote the article together and are responsible for the integrity of the work as a whole.

REFERENCES

Abbott, A. (2010). The Coping Strategy Questionnaire. *Journal of Physiotherapy*, *56*, 63. https://doi.org/10.1016/S1836-9553(10)70061-8

Angst, F., Goldhahn, J., Drerup, S., Aeschlimann, A., Schwyzer, H.-K., & Simmen, B. R. (2008). Responsiveness of six outcome assessment instruments in total shoulder arthroplasty. *Arthritis and Rheumatism*, *59*, 391–398. https://doi.org/10.1002/art.23318

Asgarian, A., Zhao, S., Ashraf, A. B., Browne, M. E., Prkachin, K. M., Mihailidis, A., & Taati, B. (2019). Limitations and Biases in Facial Landmark Detection–An Empirical Study on Older Adults with Dementia. CVPR Workshops, pp. 28–36. arXiv:1905.07446.

Ashraf, A. B., Lucey, S., Cohn, J. F., Chen, T., Ambadar, Z., Prkachin, K. M., & Solomon, P. E. (2009). The painful face—pain expression recognition using active appearance models. *Image and Vision Computing*, *27*, 1788–1796. https://doi.org/10.1016/j.imavis.2009.05.007

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

Blashfield, R. K., & Aldenderfer, M. S. (1988). The methods and problems of cluster analysis. In J. R. Nesselroade & R. B. Cattell (Eds.). *Handbook of multivariate experimental psychology* (pp. 447–474). Plenum Press.

Bot, S. D., Terwee, C. B., van der Windt, D. A., Bouter, L. M., Dekker, J., & de Vet, H. C. (2004). Clinimetric evaluation of shoulder disability questionnaires: A systematic review of the literature. *Annals of the Rheumatic Diseases*, *63*, 335–341. https://doi.org/10.1136/ard.2003.007724

Craig, K. D., Prkachin, K. M., & Grunau, R. V. E. (2011). The facial expression of pain. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment*. (3rd ed.) (pp. 117–133). : Guilford.

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

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

Hammal, Z., & Cohn, J. F. (2018). Automatic, objective, and efficient measurement of pain using automated face analysis. *Social and Interpersonal Dynamics in Pain* (pp. 121–146). Springer.

Hill, C. L., Lester, S., Taylor, A. W., Shanahan, M. E., & Gill, T. K. (2011). Factor structure and validity of the shoulder pain and disability index in a population-based study of people with shoulder symptoms. *BMC Musculoskeletal Disorders*, *12*, 8. https://doi.org/10.1186/1471-2474-12-8

Kunz, M., Chatelle, C., Lautenbacher, S., & Rainville, P. (2008). The relation between catastrophizing and facial responsiveness to pain. *Pain*, *140*, 127–134. https://doi.org/10.1016/j.pain.2008.07.019

Kunz, M., Faltermeier, N., & Lautenbacher, S. (2012). Impact of visual learning on facial expressions of physical distress: A study on voluntary and evoked expressions of pain in congenitally blind and sighted individuals. *Biological Psychology*, *89*, 467–476. https://doi.org/10.1016/j.biopsycho.2011.12.016

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

Kunz, M., & Lautenbacher, S. (2015). Improving recognition of pain by calling attention to its various faces. *European Journal of Pain*, *9*, 1350–1361. https://doi.org/10.1002/ejp.666

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. https://doi.org/10.1016/j.pain.2011.10.027

Kunz, M., Meixner, D., & Lautenbacher, S. (2019). Facial muscle movements encoding pain—a systematic review. *Pain*, *160*, 535–549. https://doi.org/10.1097/j.pain.0000000000001424

Kunz, M., Mylius, V., Schepelmann, K., & Lautenbacher, S. (2004). On the relationship between self-report and facial expression of pain. *The Journal of Pain*, *5*, 368–376. https://doi.org/10.1016/j.jpain.2004.06.002

Kunz, M., Peter, J., Huster, S., & Lautenbacher, S. (2013). Pain and disgust: The facial signaling of two aversive bodily experiences. *PLoS One*, *8*(12). https://doi.org/10.1371/journal.pone.0083277

Kunz, M., Rainville, P., & Lautenbacher, S. (2011). Operant conditioning of facial displays of pain. *Psychosomatic Medicine*, *73*, 422–431. https://doi.org/10.1097/PSY.0b013e318218db3e

Kunz, M., Scharmann, S., Hemmeter, U., Schepelmann, K., & Lautenbacher, S. (2007). The facial expression of pain in patients with dementia. *Pain*, *133*, 221–228. https://doi.org/10.1016/j.pain.2007.09.007

Kunz, M., Seuss, D., Hassan, T., Garbas, J. U., Siebers, M., Schmid, U., Schöberl, M., & Lautenbacher, S. (2017). Problems of video-based pain detection in patients with dementia: A road map to an interdisciplinary solution. *BMC Geriatrics*. https://doi.org/10.1186/s12877-017-0427-2

Littlewort, G. C., Bartlett, M. S., & Lee, K. (2009). Automatic coding of facial expressions displayed during posed and genuine pain. *Image and Vision Computing*, *27*, 1797–1803. https://doi.org/10.1016/j.imavis.2008.12.010

Lucey, P., Cohn, J. F., Prkachin, K. M., Solomon, P. E., & Matthews, I.Painful data: The UNBC-McMaster shoulder pain expression archive database. In Automatic Face & Gesture Recognition and Workshops, 2011 IEEE International Conference pp. 57–64).

MacDermid, J. C., Solomon, P., & Prkachin, K. (2006). Prkachin, K The Shoulder Pain and Disability index demonstrates factor, construct and longitudinal validity. *BMC Musculoskeletal Disorders*, *7*, 12. https://doi.org/10.1186/1471-2474-7-12

McIntyre, R. M., & Blashfield, R. K. (1980). A nearest-centroid technique for evaluating the minimum-variance clustering procedur. *Multivariate Behavioral Research*, *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. https://doi.org/10.1007/BF02294245

Mojena, R. (1977). Hierarchical grouping methods and stopping rules: An evaluation. *The Computer Journal*, *20*, 359–363. https://doi.org/10.1093/comjnl/20.4.359

Prkachin, K. M., & Mercer, S. E. R. (1989). Pain expression in patients with shoulder pathology: Validity, properties, and relation to sickness impact. *Pain*, *39*, 257–265.

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. https://doi.org/10.1016/j.pain.2008.04.010

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

Rash, J. A., Prkachin, K. M., Solomon, P. E., & Campbell, T. S. (2019). Assessing the efficacy of a manual-based intervention for improving the detection of facial pain expression. *European Journal of Pain*, *23*, 1006–1019. https://doi.org/10.1002/ejp.1369

Roach, K. E., Budiman-Mak, E., Songsiridej, N., & Lertratanakul, Y. (1991). Development of a shoulder pain and disability index. *Arthritis Care & Research*, *4*, 143–149. https://doi.org/10.1002/art.1790040403

Robinson, M. E., Riley, J. L. I. I. I., Myers, C. D., Sadler, I., Kvaal, S. A., Geisser, M. E., & Keefe, F. J. (1997). The Coping Strategies Questionnaire: A large sample item-level factor analysis. *Clinical Journal of Pain*, *13*, 43–49. https://doi.org/10.1097/00002508-199703000-00007

Rosenstiel, A. K., & Keefe, F. J. (1983). The use of coping strategies in chronic low back pain patients: Relationship to patient characteristics and current adjustment. *Pain*, *17*, 33–44. https://doi.org/10.1016/0304-3959(83)90125-2

Rovniak, L. S., Sallis, J. F., Saelens, B. E., Frank, L. D., Marshall, S. J., Norman, G. J., Conway, T. L., Cain, K. L., & Hovell, M. F. (2010). Adults' physical activity patterns across life domains: Cluster analysis with replication. *Health Psychology*, *29*, 496–505. https://doi.org/10.1037/a0020428

Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2008). Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *PAIN®*, *135*(1–2), 55–64. https://doi.org/10.1016/j.pain.2007.05.008

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

Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association*, *58*, 236–244. https://doi.org/10.1080/01621459.1963.10500845

Werner, P., Al-Hamadi, A., Limbrecht-Ecklundt, K., Walter, S., Gruss, S., & Traue, H. C. (2016). Automatic pain assessment with facial activity descriptors. *IEEE Transactions on Affective Computing*, *8*, 286–299. https://doi.org/10.1109/TAFFC.2016.2537327

Williams, J. W. J., & Holleman, D. R. J. (1995). Simel, DL Measuring shoulder function with the Shoulder Pin and Disability Index. *Journal of Rheumatology*, *22*, 727–732.

How to cite this article: Kunz M, Prkachin K, Solomon PE, Lautenbacher S. Faces of clinical pain: Inter-individual facial activity patterns in shoulder pain patients. *Eur J Pain*. 2021;25:529–540. https://doi.org/10.1002/ejp.1691