Received: 4 December 2018 | Revised: 6 January 2020 | Accepted: 16 January 2020

DOI: 10.1002/ejp.1536

ORIGINAL ARTICLE

EJP European Journal of Pain

Pain assessment for cognitively impaired older adults: Do items of available observer tools reflect pain-specific responses?

Judith Kappesser[1] | **Stefanie Voit**[1] | **Stefan Lautenbacher**[2] | **Christiane Hermann**[1]

[1]Department of Clinical Psychology, Justus-Liebig-University Giessen, Giessen, Germany

[2]Department of Physiological Psychology, University of Bamberg, Bamberg, Germany

Correspondence
Judith Kappesser, Department of Clinical Psychology, Institute of Psychology, Otto-Behaghel-Str. 10 F, 35394 Giessen, Germany.
Email: judith.kappesser@psychol.uni-giessen.de

Funding information
There has been no funding for conducting the study, preparation of data or the manuscript.

Abstract

Background: A number of observational tools are available to assess pain in cognitively impaired older adults, however, none of them can yet be regarded as a "gold standard". An international research initiative has created a meta-tool compiling the facial, vocalization and body movement items of the majority of available tools. Objective of this study was to investigate the pain specificity and the validity of these items.

Method: $N = 34$ older adults with or without cognitive impairment were videotaped in three different conditions (one reference, two painful conditions) in their nursing homes. They were further asked to self-report their pain in each condition. The occurrence of non-verbal behaviours was coded as present or absent using the items of the meta-tool.

Results: The majority of non-verbal behaviours was not pain sensitive as they occurred less than three times across participants and conditions. Of the remaining items, two facial items ("pained expression" and "raising upper lip"), one vocalization item ("using pain-related words") and one body movement item ("guarding") were found to be pain specific and valid. One additional item, the vocalization item "gasping", was pain specific, but not associated with pain self-report, and three additional items, the facial items "frowning" and "narrowing eyes" and the vocalization item "mumbling" were correlated with pain self-report but did not help to separate pain from non-pain conditions.

Conclusions: Systematic evaluation of items of existing observational pain assessment tools under naturalistic conditions seems a promising approach in the process of further investigating and improving tools.

Significance: Only few items stemming from observational pain assessment tools were found to be pain sensitive and specific as well as valid in this study. The investigation of existing tools not only on tool but additionally on item-level can provide helpful insights and thereby can help to improve the original tools and establish a gold standard for nonverbal pain assessment in older adults with cognitive impairments.

1 | INTRODUCTION

Pain in older people with dementia is common (Scherder & Abdulla, 2009). It is estimated, for instance, that up to 80% of people with dementia residing in care homes experience pain on a regular basis (Maxwell et al., 2008). Pain detection and treatment, however, continues to be a challenge for health care professionals (Bullock et al., 2019; Scherder et al., 2009; Schofield & Abdulla, 2018). Results indicate that pain in people with dementia is under-detected as well as under-treated (Benedetti et al., 2006; Herr & Titler, 2009; Kaasalainen et al., 2017; Scherder et al., 2009). Since consequences of undetected and untreated pain are severe (Hadjistavropoulos et al., 2014), a thorough assessment of pain is an essential prerequisite for an adequate and successful treatment (Corbett et al., 2014).

Although patients' self-report is usually most effective to detect the presence of pain, people with dementia often lose their ability to verbally communicate pain, particularly at later stages. For such populations, in which it is difficult to obtain a self-report, observational pain assessment tools are recommended as a substitute for self-report or as complementary information (Hadjistavropoulos et al., 2014; Herr, Coyne, McCaffery, Manworren, & Merkel, 2011).

For older people suffering from dementia, a number of observational pain assessment scales are available (Hadjistavropoulos et al., 2014; Herr, Coyne, McCaffery, Manworren, & Merkel, 2011). These scales mostly rely on non-verbal pain behaviours such as facial expression, vocalization/verbalization and body movement (AGS Panel on Persistent Pain in Older Persons, 2002). A number of systematic reviews as well as a review of these reviews have examined these tools for their psychometric quality (Lichtner et al., 2014) and concluded that some of them seem promising, but many of them lack sufficient evidence for reliability, validity and clinical utility. Therefore, authors conclude consistently that no single scale can be recommended as most appropriate at this point, and that further research is needed (Lichtner et al., 2014).

When carefully evaluating the tools' items, further problems become obvious. Some tools fail to provide item descriptions consistent with empirical evidence. For example the item description for "grimassing" of various scales (Abbey et al., 2004; Husebo, Strand, Moe-Nilssen, Husebo, & Ljunggren, 2010) does not fully match empirically identified pain-specific facial muscle movements. Other tools include items for which it is debatable whether and how they describe behaviours which can be observed (e.g. 'looking sad', 'looking blank'; [Fuchs-Lacelle & Hadjistavropoulos, 2004; Pautex, Herrmann, Michon, Giannakopoulos, & Gold, 2007]). An additional difficulty is that some items are ill-defined or not defined at all. These ambiguities, inconsistencies and inaccuracies suggest that some items may lack validity, hence jeopardizing psychometric quality and clinical utility.

An international initiative (EU-COST action "Pain in impaired cognition, especially dementia"), whose aim it is to develop and establish a meta-tool based on existing instruments (Corbett et al., 2014), has collected and compiled items of all existing scales. The resulting item pool (Pain Assessment in Impaired Cognition (PAIC) [Corbett et al., 2014; de Waal et al., 2019]) enables researchers to systematically evaluate observational pain assessment tools not only at the level of tools, but also at the level of items.

Findings regarding the impact of cognitive impairment on pain behaviours are inconsistent. Yet, the majority of results do not suggest that pain responsivity is attenuated, thus stressing the need to improve pain assessment (e.g. Binnekade et al., 2018; Rajkumar et al., 2017).

The primary aims of this study were to investigate, (a) the difference in the occurrence of non-verbal pain behaviours when pain is induced and in a non-painful situation, and (b) the relationship between non-verbal pain behaviours and pain self-report. To determine the effect of cognitive functioning, we also tested whether cognitive impairment impacts on the ability to self-report pain and/or the likelihood to show non-verbal pain behaviours.

2 | METHODS

2.1 | Participants

The study protocol was approved by the local ethic committee of the Faculty of Psychology and Sports Science (# 2013–0022).

To recruit older adults with and without dementia, local nursing homes and home health care facilities were contacted. They were informed about the study. If interested in collaborating, staff members identified and informed potential participants and/or their guardians about the study. If potential participants or guardians had agreed, our study team contacted them and explained the purpose of the study in more detail. All participants or their guardians gave informed consent for participation. Participants did not receive financial reward for participation.

Some exclusion/inclusion criteria were applied: Specifically, participants had to be 60 years or older. And clearly, participants had to have at least some cognitive and physical capacities to be able to take part in the study, for example. follow instructions or raising their arms. These criteria are likely to have had an effect on the composition of the final sample in which there were no bedridden participants, in which many participants were able to self-report their pain, and in which few were severely demented and of those some were still able to self-report their pain.

$N = 41$ potential participants gave informed consent. However, five of them could not be videotaped: one potential participant had died prior to the assessment, two had longer hospital stays at the time of data acquisition and two declined to take part on the day of video recording. Due to their severe dementia two participants were unable to understand the instructions during video recording and, therefore, had to be excluded from further data analyses. Therefore, the final sample consisted of $n = 34$ older adults described in more detail in Table 1.

No significant differences between participants with no, mild or severe dementia emerged for gender, age, diagnoses potentially causing pain, pain medication or level of depressive symptoms.

TABLE 1 Description of the total sample as well as of subgroups according to dementia status

	Total sample	No dementia	Mild dementia	Severe dementia	Group main effect (ANOVA)
Number	34 (100%)	19 (55.9%)	6 (17.6%)	9 (26.5%)	
MMSE (*M*, *SD*)[a,b]	21.9 (7.2)	27.4 (2.1)	20.2 (1.7)	11.7 (3.6)	**$p < .001$**
Sex [*N* (%) female]	27 (79.4%)	15 (78.9)	5 (83.3)	7 (77.8)	$p = .964$
Current pain diagnosis [*N* (%)]	29 (85.3%)	16 (84.2)	6 (100)	7 (77.8)	$p = .482$
Taking pain medication [*N* (%)]	14 (41.2%)	9 (47.4)	4 (66.7)	1 (11.1)	$p = .072$
Age (*M, SD*)	82.4 (9.5)	81.1 (10.2)	86.7 (2.6)	82.4 (10.8)	$p = .466$
Level of depressive symptoms (*M, SD*)[b]	1.2 (1.6)	1.1 (1.4)	2.2 (2.6)	1.0 (1.2)	$p = .312$

[a]Assessed by the Mini-Mental Status Examination (MMSE; German version [18]); grading of cognitive impairment according to [22]: MMSE 24–30 no dementia, MMSE 18–23 mild dementia, MMSE 0–17 severe dementia.

[b]Assessed by the Geriatric Depression Scale—short version (GDS-8; [37]), ranging from 0–8.

2.2 | Assessment instruments

2.2.1 | Cognitive impairment

Participants were allocated to one of three cognitive impairments groups based upon their score in the Mini-Mental State Examination (MMSE). The MMSE is one of the most widely used screening tools for cognitive performance (Folstein, Folstein, & McHugh, 1975; German version by Kessler, Markowitsch, & Denzler, 2000) and has good psychometric properties (e.g. Beyermann, Trippe, Bähr, & Püllen, 2013; Foreman, 1987; Fountoulakis, Tsolaki, Chantzi, & Kazis, 2000). The interview takes about 10 min and consists of 30 everyday questions, which are easy to be answered correctly by not cognitively impaired people. The total score can range from 0 to 30. We followed the recommendations of the authors of the German MMSE version (Kessler et al., 2000). Hence, to categorize the severity of cognitive dysfunction the grades as defined by Tombaugh and McIntyre (1992) were used: "severe dementia" (0–17), "mild dementia" (18–23) and "no dementia" (24–30).

2.3 | Pain assessment protocol

Participants were videotaped in three different conditions: (a) during a not painful reference condition in which the pain self-report scales were explained (see "Self-report of pain"), (b) while performing a movement presumably eliciting pain, i.e. they were asked to raise their arms up as high above the head as possible (mobilization condition) and (c) while ischaemic pain was experimentally induced using the submaximal effort tourniquet technique (Smith, Lowenstein, Hubbard, & Beecher, 1968). In order to induce ischaemic pain, a blood pressure cuff wrapped around the upper arm was inflated twice. First, participants' actual blood pressure was measured. Then, the blood pressure cuff was inflated up to a pressure 30 mmHg above the previously measured systolic blood pressure. In addition, participants held a weight (sand bag, weighing 250 g) in their hand (arm with blood pressure cuff) and were asked to bend and stretch the forearm. They were asked to continue these movements until the pain became unbearable.

All participants started with the reference condition, continued with the ischaemic pain condition and, lastly, underwent the mobilization condition. The order of the conditions was kept constant across participants, since the explanation of the pain self-report scale always came first and mobilization was expected to affect the blood pressure, therefore mobilization-induced pain was assessed last.

2.3.1 | Self-report of pain

The Faces of Pain Scale-Revised (FPS-R; Hicks, Baeyer, Spafford, Korlaar, & Goodenough, 2001)] was used to assess participants' self-reported pain whenever possible. It consists of six faces showing increasing pain from the left ("no pain" coded as 0) to the right end ("extremely painful" coded as 5). The FPS-R was originally developed to assess subjective report of pain in younger children, but is also successfully used in older adults (Pautex et al., 2005; Ware, Epps, Herr, & Packard, 2006).

The FPS-R was introduced and explained to all participants using the instruction recommended by Hicks et al. (2001). When participants conveyed that they had understood the scale, comprehension was ascertained by asking participants to indicate on the scale what face they would choose in case of "no pain", "extreme pain", "pain of medium intensity". If necessary, these descriptions were paraphrased. Only if participants were not able to point to the faces equivalent to the descriptions after explanation and paraphrase, they were excluded from the study. If they showed comprehension of the scale, they were also asked to rate their current level of pain.

2.3.2 | PAIC items[1]

One of the aims of the EU-COST Action "Pain in impaired cognition, especially dementia" was to identify key items from existing pain assessment tools for older adults with dementia. This was achieved by compiling the items of available tools, by reviewing the psychometric data, by obtaining consensus of international experts on the validity of the items, on the frequency of item occurrence, and by determining to what degree the items complied with the American Geriatric Society's guidelines (AGS Panel on Persistent Pain in Older Persons, 2002; Corbett et al., 2014).

The final PAIC collection in its research version (for the clinical version see Kunz, et al., 2020) is based on items of twelve eligible pain assessment tools which were identified in a comprehensive literature search (Corbett et al., 2014). Each item had to represent one of three main domains of pain (facial expression, vocalization, body movements) as recommended by the American Geriatric Society (AGS Panel on Persistent Pain in Older Persons, 2002). To be further considered, each item had to be part of at least two tools or to be derived from a tool with robust evidence of its psychometric qualities. In a next step, duplicates were removed. Table 2 displays the final research version of PAIC item pool ($N = 36$ items).

TABLE 2 Overview of PAIC items grouped according to American Geriatric Society domains 1–3 (see Corbett et al., 2014)

Facial expression	Vocalization	Body movements
1. Pained expression	1. Using offensive words	1. Freezing
2. Frowning	2. Using pain-related words	2. Curling up
3. Narrowing eyes	3. Repeating words	3. Clenching hands
4. Closing eyes	4. Complaining	4. Resisting care
5. Raising upper lip	5. Shouting	5. Pushing
6. Opened mouth	6. Mumbling	6. Guarding
7. Tightened lips	7. Screaming	7. Rubbing
8. Clenched teeth	8. Groaning	8. Limping
9. Empty gaze	9. Crying	9. Restlessness
10. Seeming disinterested	10. Gasping	10. Pacing
11. Pale face	11. Sighing	
12. Teary eyed		
13. Looking tense		
14. Looking sad		
15. Looking frightened		

2.4 | Videos and coding

2.4.1 | Recording of videos

Videos of participants were recorded by two research assistants using a Canon Legria HF M46 camera each. One research assistant recorded participants' faces, the other participants' bodies. The recordings were cut and edited using Adobe Premiere Pro CS6 (Adobe Systems Software, Ireland). After editing, the clips were exported into commonly used video formats. The final length of videos ranged from 5 to 108 s. The video recordings of the face were used to code facial expression and vocalization items, and the recordings of the body to code body movements.

2.4.2 | Coders and coding

One graduate student (SV) was trained by one of the authors (JK). Training included getting familiar with the definition and explanation of the PAIC items as well as getting practice in coding when using video sequences. For facial expression and body movement JK, and for vocalization SV watched all videos in random order at least once per item and coded the occurrence of behaviours described in the PAIC. In the original PAIC, items are rated on a 0–3 Likert scale with regard to the intensity of their occurrence (0: not at all, 1: slight degree, 2: moderate degree, 3: great degree). In this study, we only coded the presence or absence of each behaviour as described by each item in each video since we were primarily interested in evaluating whether or not the pain behaviours were occurring at all. To assess interrater reliability, SV coded a random sample of 20% of the videos for determining facial expression and body movement items in addition to JK, and JK an additional random sample of 20% of vocalization items.

2.5 | Data analyses

Data were analysed using SPSS IBM Statistics 22 (IBM Germany GmbH, Ehningen, Germany).

Interrater reliability was calculated for each AGS domain (facial expression, vocalization, body movements) using Cohen's kappa adjusted according to Brennan and Prediger (1981). To examine the effect of cognitive impairment on self-reported pain, univariate ANOVAs, and to examine the effect of cognitive impairment on the occurrence of PAIC items, chi-square tests were applied. The occurrence of PAIC items in the different conditions was summed up for the total sample, and a McNemar Test for dependent samples was used to examine potential differences between each of the painful conditions and the reference condition, regardless of dementia status.

Spearman's rho correlation coefficients were used to investigate the relationship between PAIC items and pain self-report.

3 | RESULTS

3.1 | Interrater reliability for PAIC items

Cohen's kappa adjusted according to Brennan–Prediger was very high for all three AGS domains (facial expression: 0.91, vocalization items: 0.93, body movements: 0.92; aggregated kappa across domains: 0.92) and can be considered to be (almost) perfect (Fleiss et al., 2003; Landis & Koch, 1977).

3.1.1 | Effect of cognitive impairment on the ability to self-report pain

Five of the $n = 34$ participants were not able to self-report their pain experience. Four of these five participants suffered from severe (MMSE of 7–11), and one from mild dementia (MMSE = 18). For $n = 29$ participants, self-report was available ($n = 19$ suffering from no dementia, $n = 5$ from mild and $n = 5$ from severe dementia). Univariate ANOVAs (Table 3) revealed no differences for the mean pain self-report between participants with no, mild or severe dementia for any of the three conditions (reference, ischaemic pain and mobilization).

Pain was reported in the three conditions as follows: reference condition—$M = 0.93$ ($SD = 0.22$, Median = 0, Min = 0, Max = 4), mobilization condition—$M = 0.79$ ($SD = 0.19$, Median = 1, Min = 0, Max = 4), ischaemic pain condition—$M = 1.93$ ($SD = 0.20$, Median = 2, Min = 0, Max = 4). Pain self-reports differed significantly between conditions [$F(2;28) = 13.53$, $p = .001$, partial $\eta^2 = 0.33$]. Post hoc comparisons revealed that the ischaemic pain condition elicited significantly higher self-reported pain than the reference condition ($p = .003$) as well as the mobilization condition ($p < .001$).

3.1.2 | Effect of cognitive impairment on occurrence of PAIC items

$N = 108$ chi-square tests (36 items multiplied by the three conditions) were calculated to examine differences in the occurrence of PAIC items depending on cognitive impairment. Of those, only the group comparison for the vocalization item ‘using pain-related words’ in the mobilization pain condition was significant ($\chi^2(2) = 6.34$, $p = .04$) and three had p-values lower than 0.1 (vocalization item ‘using pain-related words’ in the reference condition, $\chi^2(2) = 4.81$, $p = .09$; vocalization item ‘complaining’ in the reference condition, $\chi^2(2) = 4.81$, $p = .09$; vocalization items ‘shouting’ in the reference condition, $\chi^2(2) = 4.81$, $p = .09$). Given the great number of tests that revealed no differences in occurrence of PAIC items between MMSE groups, the occurrence of PAIC items was further investigated based on the total sample, thereby disregarding cognitive impairment.

3.2 | Occurrence of PAIC items and their relation to pain self-report

A considerable number of PAIC items occurred only very rarely (in two or less of $n = 34$ participants per condition, i.e. ≤5.88%). For the domain facial expression, these items were as follows: clenched teeth, empty gaze, seeming disinterested, pale face, teary eyed, looking tense, looking sad and looking frightened. For the domain vocalization, these items were using offensive words, complaining, shouting, screaming, groaning, crying and sighing. And for the domain body movements, these items were freezing, curling up, resisting care, pushing, rubbing, limping and pacing. Due to these low occurrence rates, these items were excluded from further analyses.

3.2.1 | Domain “facial expression”

The occurrence rates of the remaining facial expression items are displayed in Figure 1, correlations between the facial expression items and participants’ pain self-report are illustrated in Figure 2.

Of the remaining seven facial expression items, only two differentiated between the experimentally induced pain and the reference condition: pained expression and raising upper lip. Positive correlations with pain self-report (Figure 2) support the particular relevance of the behaviour “pained expression” as its occurrence correlates positively

TABLE 3 Mean (standard deviation) of pain self-report for each of the three conditions depending on dementia status

Condition of pain self-report	Dementia status: No	Mild	Severe	*df*	*F*	*p*
Reference situation	0.95 (1.27)	1.40 (1.14)	0.40 (0.89)	2;28	0.88	0.43
Ischaemic pain	2.00 (1.15)	1.40 (0.89)	2.20 (0.84)	2;28	0.81	0.46
Movement pain	0.84 (1.12)	1.00 (0.71)	0.40 (0.89)	2;28	0.48	0.62

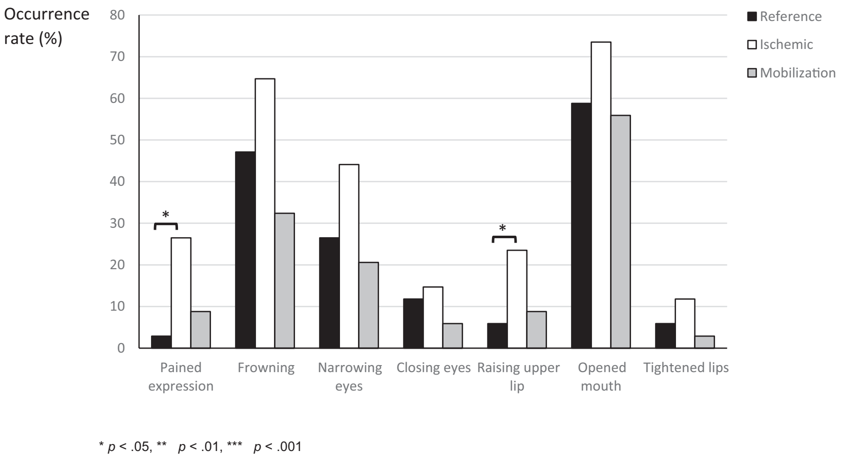

FIGURE 1 Occurrence rate (%) for PAIC facial expression items appearing more than twice in at least one condition as well as significant differences between reference and painful conditions

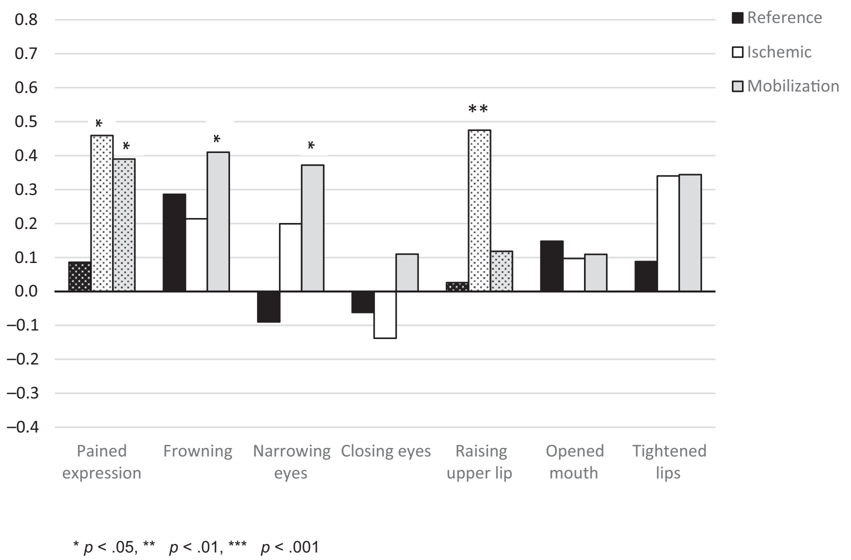

FIGURE 2 Spearman rho correlation coefficients between PAIC facial expression items and pain self-report by conditions. Significant correlations are indicated by stars. Items that had discriminated between pain and reference condition (see Figure 1) are indicated by dotted bars

with participants' pain self-report in both induced pain conditions. Moreover, "raising upper lip" appears relevant as it correlates positively with self-report in the ischaemic pain condition. Interestingly, two additional items, "frowning" and "narrowing eyes", also correlate positively with pain self-report in the mobilization pain condition. These items occurred even more frequently in the ischaemic pain condition than "pained expression" and "raising upper lip". However, they did not differentiate between the ischaemic pain and reference condition because they occurred also frequently in the reference condition.

3.2.2 | Domain "vocalization"

The occurrence rates of the remaining vocalization items are displayed in Figure 3, correlations between vocalization items and participants' pain self-report are shown in Figure 4.

Three of the remaining items, "pain-related words", "repeating words" and "gasping", differentiated significantly between pain conditions and reference condition. "Repeating words" occurred more frequently in the reference than in the mobilization condition and, therefore, is neither pain sensitive nor pain specific. "Pain-related words" and "gasping" occurred more frequently in the ischaemic pain than in the reference condition. The occurrence of "gasping", however, correlated negatively with participants' pain self-report. The positive correlation between "pain-related words" and pain self-report reached significance in the mobilization, but not in the ischaemic condition ($r = 0.30$; $p = .115$). A significant correlation coefficient was also found for "mumbling" in the mobilization condition. However, "mumbling" did not differentiate between reference and pain conditions and, there was a trend for it to occur more frequently in the reference than in the pain conditions. Similar to the item "repeating words", it is, therefore, neither pain sensitive nor pain specific.

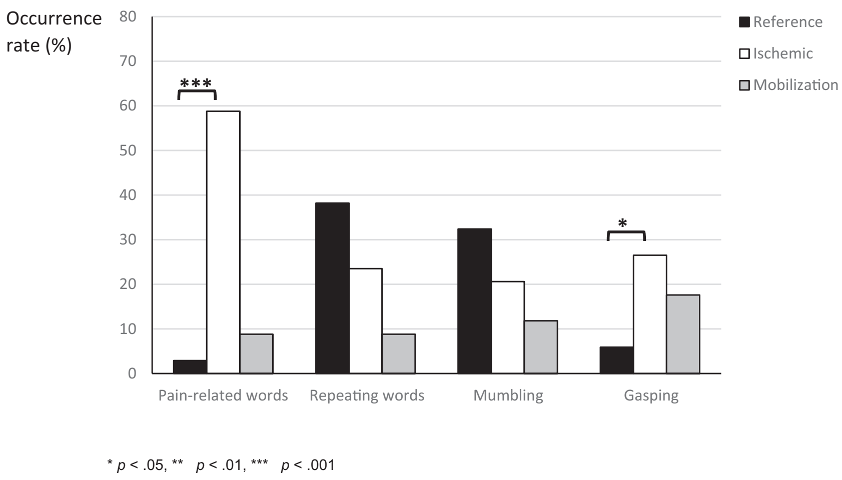

FIGURE 3 Occurrence rate (%) for those PAIC vocalization items appearing more than twice in at least one condition as well as significant differences between reference and painful conditions

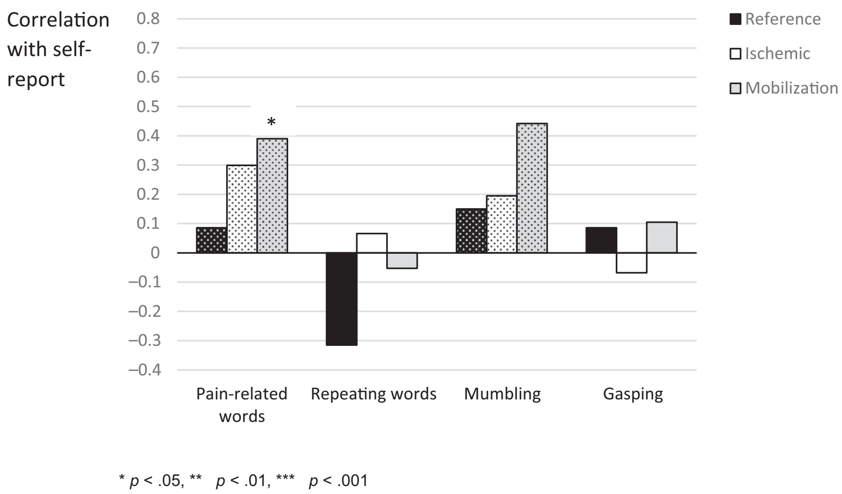

FIGURE 4 Spearman rho correlation coefficients PAIC vocalization items and pain self-report by conditions. Significant correlations are indicated by stars. Items that had discriminated between pain and reference condition (see Figure 3) are indicated by dotted bars

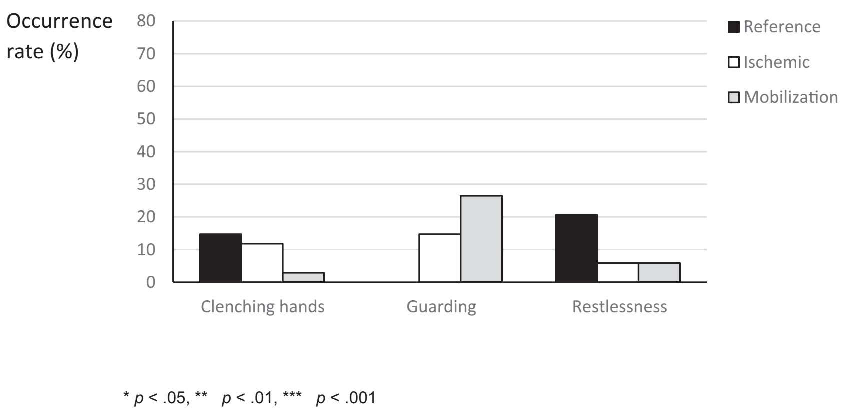

FIGURE 5 Occurrence rate (%) of those body movement items appearing more than twice in at least one condition as well as significant differences between reference and painful conditions

3.2.3 | Domain "body movement"

The occurrence rates of the remaining body movement items are displayed in Figure 5, the correlations between vocalization items and participants' pain self-report are illustrated in Figure 6.

Of the three remaining items only "guarding" differed between mobilization and reference condition. This is interesting as it is contrary to the results for facial expression and vocalization items with the latter differentiating between the ischaemic pain and the reference condition. Furthermore, a significant positive correlation coefficient emerged between "guarding" and participants' pain self-report for the mobilization condition. The positive correlation between "guarding" and pain self-report for the ischaemic pain condition did not reach significance ($r = 0.31$; $p = .102$). Comparable to the facial expression and vocalization items, no significant

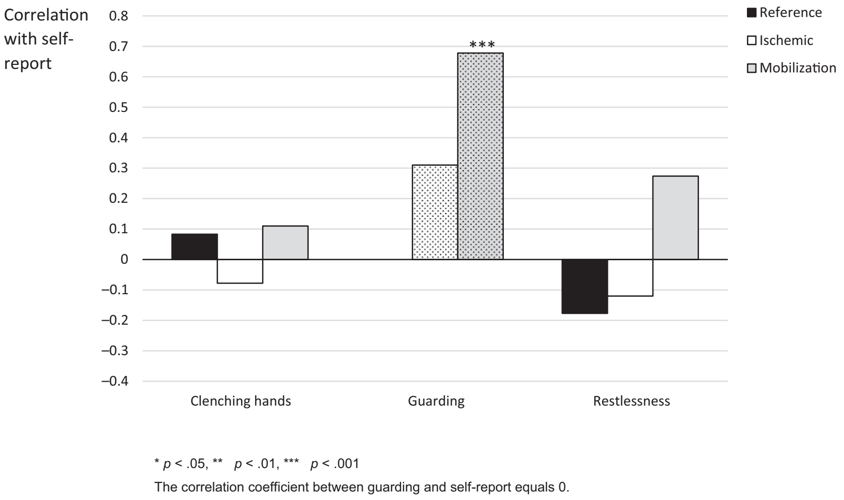

FIGURE 6 Spearman rho correlation coefficients between PAIC body movement items and pain self-report by conditions. Significant correlations are indicated by stars. Items that had discriminated between pain and reference condition (see Figure 5) are indicated by dotted bars

correlation between body movement items and pain self-report emerged for the reference condition.

4 | DISCUSSION AND CONCLUSIONS

Aim of this study was to investigate the occurrence of pain-related behaviours and their relationship with self-reported pain in (non)painful conditions. The majority of non-verbal behaviours was not pain sensitive as they occurred less than three times across participants and conditions. Two facial items ("pained expression", "raising upper lip"), one vocalization item ("using pain-related words") and one body movement item ("guarding") were found to be pain specific and valid. The vocalization item "gasping" was pain specific, but not valid. The facial items "frowning" and "narrowing eyes" and the vocalization item "mumbling" were valid but not pain specific.

4.1 | Ability to self-report pain

Five of the 15 participants with dementia (MMSE 7–18) were not able to use the FPS-R. This finding implies that two thirds of participants with dementia (MMSE 10–23) were able to self-report their pain. The occurrence of pain self-report does not necessarily mean that it is valid. However, our results are in line with studies finding that self-report pain scales can be reliably used in people with dementia (Pautex et al., 2005).

4.2 | Reliability and objectivity when applying behavioural coding

Interrater reliability is not an inherent feature of observational tools, but depends on observers' training and items' definition. In this study, interrater reliability was very high for all domains. This is presumably a result of the intensive training, the clarification of item meaning, and of coding only the presence or absence of pain-indicative behaviours and not their intensity. The item definitions were taken from the original item descriptions some of which are very vague. The item "pained expression", for example, is originally defined as "facial expression of pain", thus being subject to interpretation. We decided to consider a "pained expression" as present whenever at least two empirically derived pain-specific facial movements occurred (Kunz & Lautenbacher, 2014; Kunz, Meixner, & Lautenbacher, 2019; Lautenbacher, Walz, & Kunz, 2017; Prkachin, 1992; Williams, 2002). Due to this operationalization the item 'pained expression' turned out to be particularly relevant in our further analyses. However, operationalized differently the pain specificity of this item may not hold up.

Considerable difficulties emerged when coding several vocalization and body movement items. The distinction between vocalization, verbal utterances and spontaneous self-report is quite difficult due to considerable overlap (take "ouch" as example). Rubbing a painful area requires that observers already know that the patient is in pain and where this pain is located.

4.3 | Differences between pain-related behaviours depending on cognitive impairment

Studies investigating the impact of cognitive impairment on non-verbal pain behaviours have yielded contradictory findings. In this study cognitive impairment had no effect on the occurrence of the pain behaviours. This finding suggests that cognitive impairment is not systematically associated with augmented or dampened (observable) pain responses. Our finding corresponds well with other studies also reporting no effect of cognitive impairment on pain expression

(Hadjistvropoulos et al., 1998). This does not imply, however, that people with dementia experience less pain than people without.

4.4 | Observer ratings by use of pain-related behaviour and relation to self-report

Many of the PAIC items were not endorsed at all in this study, others only very rarely. Although all items have seemingly proven their utility in assessing pain in other scales already, 50–70% of items were only present twice or less across all conditions. This great number of apparently useless items supports the approach of the COST initiative (Corbett et al., 2014; de Waal et al. 2019; Kunz et al., 2020) and other research groups (Chang, Versloot, Fashler, McCrystal, & Craig, 2015; Ersek et al., 2018) that it is absolutely crucial to evaluate psychometric quality at the level of single items.

Items occurring three times or more were explored further by applying two additional criteria: (a) items' specificity, that is their ability to differentiate between pain and reference condition and (b) items' convergent validity, that is their association with self-reported pain in the ischaemic pain condition. Five items met the pain specificity criterion: "pained expression", "raising upper lip", "pain-related words", "gasping" and "guarding". A slightly different list of items emerged when considering convergent validity: "pained expression", "raising upper lip", "pain-related words", "guarding", "frowning", "narrowing eyes" and "mumbling". Yet, "frowning", "narrowing eyes" and "mumbling" were not pain specific. "Mumbling" did occur frequently but was most frequently coded in the reference condition. "Frowning" and "narrowing eyes" occurred frequently in the ischaemic pain but also in the reference condition.

It is interesting that different items distinguished between the two pain conditions: "guarding" differentiated between mobilization and reference condition, "pained expression", "raising upper lip" and "pain-related words" differentiated between ischaemic pain and reference condition. The observed pattern of occurrence of the item 'guarding' might be accounted for by differences between both conditions. It may be easier to assess 'guarding' in the ischaemic pain than in the mobilization condition. In addition, these two situations might also differ in how physically demanding they were. Hadjistavropoulos, LaChapelle, MacLeod, Snider, and Craig (2000), for example found that guarding and facial expression was more frequently shown in situations in which older adults had to be more active (transfer, walk) rather than performing relatively passive movements (sit, recline). These findings support the notion that it is necessary to empirically investigate PAIC items in a variety of different pain provoking conditions.

4.5 | Limitations

To test sensitivity and specificity of the items, participants underwent three conditions. The reference condition was not a neutral rest but a social communication situation introducing the pain self-report where participants were cognitively engaged. We made this choice because of the high external validity. The disadvantage is the higher chance of certain distress behaviours occurring, making it more difficult for certain items to be confirmed as pain specific. This may in particular be true for "frowning" and "narrowing eyes", facial movements that empirically were shown to be pain specific (Kunz & Lautenbacher, 2014; Kunz et al., 2019; Prkachin, 1992; Williams, 2002), but also occur if individuals are cognitively engaged.

It cannot be ruled out that some participants may have experienced pain in the reference condition as some had painful medical diagnoses and others had been prescribed pain medication. By choosing to investigate pain behaviour items during daily routine, we imposed minimal exclusion criteria on participants and, thereby, enhanced ecological validity. However, we did so at the cost of internal validity, for example by not including only participants having no underlying painful diagnoses.

Looking at the pain self-reports, it becomes apparent that, on average, pain ratings in both pain conditions were not high (mobilization: 0.79, ischaemic pain: 1.93). Higher experienced pain generally leads to a greater non-verbal expression of pain (e.g. Kunz, Mylius, Schepelmann, & Lautenbacher, 2004), which may facilitate pain assessment. The comparably low pain self-reports suggest that the occurrence of pain-indicative behaviour was assessed in conditions, which make the judgement of pain more difficult. The low pain self-report ratings may also point to floor effects and associated restrictions in variance may have attenuated the correlations between self-report and observer rating.

The pain self-reports did not differ between mobilization and reference condition, thereby questioning the painfulness of the mobilization condition or the validity of the reference condition. The medians reveal that the mobilization condition causes slightly more pain than the pain reported in the reference condition. Furthermore, raising arms per se is not a painful movement. It becomes painful whenever muscles are sore or joints are affected by age and/or diseases.

We decided a priori not to correct for alpha error risking an unduly increase in significant results. Since the main aim was to explore items and to generate rather than to test hypotheses, we decided to be rather overly inclusive.

4.6 | Conclusion and outlook

Our results indicate that four items ("pained expression", "raising upper lip", "pain-related words", "guarding") were pain specific and valid as they differentiated between (non-) painful conditions and were correlated with self-report. The item "gasping" was pain specific, but not valid. Three items ("frowning", "narrowing eyes", "mumbling") were valid but not pain specific. Our findings suggest that, out of an original pool of 36 items, only four or eight items, if more liberal criteria are applied, are promising as valid pain-related behaviours. This is a somewhat surprising finding given that all 36 items had seemingly earlier proven their utility for pain assessment in patients with dementia.

Importantly, the four or eight items identified in this study must not be understood as a final recommendation. On the contrary, we regard our study as a starting point to empirically investigate items, collect research evidence and, thereby, help improving observational pain assessment tools (e.g. van Dalen-Kok et al., 2017). Some items that were not considered for further analyses in this study may be highly pain sensitive and/or specific in other situations or other samples. A clear definition of items in original tools is crucial for improvement of items and tools. Whenever possible, definitions should be based on empirically derived characteristics. Furthermore, more empirical evidence is needed especially for vocalization and body movement items. It is also important to demonstrate that items are not only sensitive and specific to elicited pain or a change in pain, but also able to assess ongoing pain.

To conclude, there is a number of observational pain assessment tools available for people with dementia which have been developed and psychometrically evaluated. The PAIC meta-tool is the attempt to extract the best possible items out of these tools. The present approach of examining psychometric qualities not only on tool, but also on the item-level seems a promising approach to develop a valid and easy to use pain assessment tool by drawing upon the original observation tools.

ACKNOWLEDGEMENTS

The authors thank Christina Griggel, Marie Sophie Brenner, Jonathan Sebök and Anita Kluck for their help with data collection, all partaking nursing homes for their support and all participants and their relatives for taking part in and supporting this study.

CONFLICT OF INTERESTS

The authors have no conflicts of interest to declare.

AUTHOR CONTRIBUTIONS

All authors contributed substantially to the manuscript: Judith Kappesser and Christiane Hermann by developing the study protocol, Stefanie Voit by collecting the data, Judith Kappesser and Stefanie Voit by analysing the data, Judith Kappesser, Stefan Lautenbacher and Christiane Hermann by writing the manuscript.

ENDNOTE

[1] For more information on item definition and training protocols see Corbett et al. (2014), van Dalen-Kok et al. (2017) or contact the authors.

REFERENCES

Abbey, J., Piller, N., de Bellis, A., Esterman, A., Parker, D., Giles, L., & Lowcay, B. (2004). The Abbey pain scale: A 1-minute numerical indicator for people with end-stage dementia. *International Journal of Palliative Nursing*, *10*(1), 6–13. https://doi.org/10.12968/ijpn.2004.10.1.12013

AGS Panel on Persistent Pain in Older Persons. (2002). The management of persistent pain in older persons. *Journal of the American Geriatrics Society*, *50*(6 Suppl), S205–S224.

Benedetti, F., Arduino, C., Costa, S., Vighetti, S., Tarenzi, L., Rainero, I., & Asteggiano, G. (2006). Loss of expectation-related mechanisms in Alzheimer's disease makes analgesic therapies less effective. *Pain*, *121*(1–2), 133–144. https://doi.org/10.1016/j.pain.2005.12.016

Beyermann, S., Trippe, R. H., Bähr, A. A., & Püllen, R. (2013). Mini-Mental-Status-Test im stationären geriatrischen Bereich [Mini-Mental State Examination in geriatrics: An evaluation of diagnostic quality]. *Zeitschrift Für Gerontologie Und Geriatrie*, *46*(8), 740–747. https://doi.org/10.1007/s00391-013-0488-6

Binnekade, T. T., Scherder, E. J. A., Maier, A. B., Lobbezoo, F., Overdorp, E. J., Rhebergen, D., … Oosterman, J. M. (2018). Pain in patients with different dementia subtypes, mild cognitive impairment, and subjective cognitive impairment. *Pain Medicine*, *19*(5), 920–927. https://doi.org/10.1093/pm/pnx162

Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, *41*, 687–699. https://doi.org/10.1177/001316448104100307

Bullock, L., Bedson, J., Jordan, J. L., Bartlam, B., Chew-Graham, C. A., & Campbell, P. (2019). Pain assessment and pain treatment for community-dwelling people with dementia: A systematic review and narrative synthesis. *International Journal of Geriatric Psychiatry*, *34*(6), 807–821. https://doi.org/10.1002/gps.5078

Chang, J., Versloot, J., Fashler, S. R., McCrystal, K. N., & Craig, K. D. (2015). Pain assessment in children: Validity of facial expression items in observational pain scales. *The Clinical Journal of Pain*, *31*(3), 189–197. https://doi.org/10.1097/AJP.0000000000000103

Corbett, A., Achterberg, W., Husebo, B., Lobbezoo, F., de Vet, H., Kunz, M., … Lautenbacher, S. (2014). An international road map to improve pain assessment in people with impaired cognition: The development of the Pain Assessment in Impaired Cognition (PAIC) meta-tool. *BMC Neurology*, *14*, 229. https://doi.org/10.1186/s12883-014-0229-5

de Waal, M. W. M., van Dalen-Kok, A. H., de Vet, H. C. W., Gimenez-Llort, L., Konstantinovic, L., de Tommaso, M., … Achterberg, W. P. (2019). Observational pain assessment in older persons with dementia in four countries: observer agreememt of items and factor structure of the Pain Assessment in Impaired Cognition. *European Journal of Pain*, *24*(2), 279–296.

Ersek, M., Neradilek, M. B., Herr, K., Hilgeman, M. M., Nash, P., Polissar, N., & Nelson, F. X. (2018). Psychometric evaluation of a pain intensity measure for persons with dementia. *Pain Medicine*, *20*(6), 1093–1104. https://doi.org/10.1093/pm/pny166

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). The measurement of interrater agreement. In J. L. Fleiss, B. Levin & M. C. Paik (Eds.), *Statistical methods for rates and proportions* (pp. 598–626). Hoboken: John Wiley & Sons.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Foreman, M. D. (1987). Reliability and validity of mental status questionnaires in elderly hospitalized patients. *Nursing Research*, *36*(4), 216–219. https://doi.org/10.1097/00006199-198707000-00004

Fountoulakis, K. N., Tsolaki, M., Chantzi, H., & Kazis, A. (2000). Mini mental state examination (MMSE): A validation study in Greece. *American Journal of Alzheimer's Disease and Other Dementias*, *15*(6), 342–345. https://doi.org/10.1177/153331750001500604

Fuchs-Lacelle, S., & Hadjistavropoulos, T. (2004). Development and preliminary validation of the pain assessment checklist for seniors with limited ability to communicate (PACSLAC). *Pain Management Nursing: Official Journal of the American Society of Pain Management Nurses*, *5*(1), 37–49. https://doi.org/10.1016/j.pmn.2003.10.001

Hadjistavropoulos, T., Herr, K., Prkachin, K. M., Craig, K. D., Gibson, S. J., Lukas, A., & Smith, J. H. (2014). Pain assessment in elderly adults with dementia. *Lancet Neurology*, *13*, 1216–1227. https://doi.org/10.1016/S1474-4422(14)70103-6

Hadjistavropoulos, T., LaChapelle, D. L., MacLeod, F. K., Snider, B., & Craig, K. D. (2000). Measuring movement-exacerbated pain in cognitively impaired frail elders. *The Clinical Journal of Pain*, *16*(1), 54–63. https://doi.org/10.1097/00002508-200003000-00009

Hadjistvropoulos, T., LaChapelle, D., MacLeod, F., Hale, C., O'Rourke, N., & Craig, K. D. (1998). Cognitive functioning and pain reactions in hospitalized elders. *Pain Research and Management*, *3*(3), 145–151. https://doi.org/10.1155/1998/621580

Herr, K., Coyne, P. J., McCaffery, M., Manworren, R., & Merkel, S. (2011). Pain assessment in the patient unable to self-report: Position statement with clinical practice recommendations. *Pain Management Nursing*, *12*(4), 230–250. https://doi.org/10.1016/j.pmn.2011.10.002

Herr, K., & Titler, M. (2009). Acute pain assessment and pharmacological management practices for the older adult with a hip fracture: Review of ED Trends. *Journal of Emergency Nursing*, *35*(4), 312–320. https://doi.org/10.1016/j.jen.2008.08.006

Hicks, C. L., von Baeyer, C. L., Spafford, P. A., van Korlaar, I., & Goodenough, B. (2001). The Faces Pain Scale – Revised: Toward a common metric in pediatric pain measurement. *Pain*, *93*(2), 173–183. https://doi.org/10.1016/S0304-3959(01)00314-1

Husebo, B. S., Strand, L. I., Moe-Nilssen, R., Husebo, S. B., & Ljunggren, A. E. (2010). Pain in older persons with severe dementia. Psychometric properties of the Mobilization-Observation-Behaviour-Intensity-Dementia (MOBID-2) Pain Scale in a clinical setting. *Scandinavian Journal of Caring Sciences*, *24*(2), 380–391. https://doi.org/10.1111/j.1471-6712.2009.00710.x

Kaasalainen, S., Zacharias, R., Hill, C., Wickson-Griffiths, A., Hadjistavropoulos, T., & Herr, K. (2017). Advancing the pain management in older adults agenda forward through the development of key research and education priorities: A Canadian perspective. *Canadian Journal of Pain*, *1*(1), 171–182. https://doi.org/10.1080/24740527.2017.1383139

Kessler, J., Markowitsch, H. J., & Denzler, P. (2000). *Mini-mental-status-test*. Weinheim: Beltz Test.

Kunz, M., de Waal, M. W. M., Achterberg, W. P., Gimenez-Llort, L., Lobbezoo, F., Sampson, E. L., … Lautenbacher, S. (2020). The pain assessment in impaired Cognition scale (PAIC- 15): A multidisciplinary and international approach to develop and test a meta-tool for pain assessment in impaired cognition, especially dementia. *European Journal of Pain*, *24*, 192–208. https://doi.org/10.1002/ejp.1484

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 (London, England)*, *18*(6), 813–823. https://doi.org/10.1002/j.1532-2149.2013.00421.x

Kunz, M., Meixner, D., & Lautenbacher, S. (2019). Facial muscle movements encoding pain – a systematic review. *Pain*, *160*(3), 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: Official Journal of the American Pain Society*, *5*(7), 368–376. https://doi.org/10.1016/j.jpain.2004.06.002

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Lautenbacher, S., Walz, A. L., & Kunz, M. (2017). Using observational facial descriptors to infer pain in persons with and without dementia. *BMC Geriatrics*, *18*, 88–98. https://doi.org/10.1186/s12877-018-0773-8

Lichtner, V., Dowding, D., Esterhuizen, P., Closs, S. J., Long, A. F., Corbett, A., & Briggs, M. (2014). Pain assessment for people with dementia: A systematic review of systematic reviews of pain assessment tools. *BMC Geriatrics*, *14*, 138. https://doi.org/10.1186/1471-2318-14-138

Maxwell, C. J., Dalby, D. M., Slater, M., Patten, S. B., Hogan, D. B., Eliasziw, M., & Hirdes, J. P. (2008). The prevalence and management of current daily pain among older home care clients. *Pain*, *138*(1), 208–216.

Pautex, S., Herrmann, F., Le Lous, P., Fabjan, M., Michel, J.-P., & Gold, G. (2005). Feasibility and reliability of four pain self-assessment scales and correlation with an observational rating scale in hospitalized elderly demented patients. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *60*(4), 524–529. https://doi.org/10.1093/gerona/60.4.524

Pautex, S., Herrmann, F. R., Michon, A., Giannakopoulos, P., & Gold, G. (2007). Psychometric properties of the Doloplus-2 observational pain assessment scale and comparison to self-assessment in hospitalized elderly. *The Clinical Journal of Pain*, *23*(9), 774–779. https://doi.org/10.1097/AJP.0b013e318154b6e3

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

Rajkumar, A. P., Ballard, C., Fossey, J., Orrell, M., Moniz-Cook, E., Woods, R. T., … Corbett, A. (2017). Epidemiology of pain in people with dementia living in care homes: Longitudinal course, prevalence, and treatment implications. *Journal of the American Medical Directors Association*, *18*(5), 453.e1–453.e6. https://doi.org/10.1016/j.jamda.2017.01.024

Scherder, E., Herr, K., Pickering, G., Gibson, S., Benedetti, F., & Lautenbacher, S. (2009). Pain in dementia. *Pain*, *145*(3), 276–278. https://doi.org/10.1016/j.pain.2009.04.007

Schofield, P., & Abdulla, A. (2018). Pain assessment in the older population: What the literature says. *Age and Ageing*, *47*, 324–327. https://doi.org/10.1093/ageing/afy018

Smith, G. M., Lowenstein, E., Hubbard, J. H., & Beecher, H. K. (1968). Experimental pain produced by the submaximum effort tourniquet technique: Futher evidence of validity. *Journal of Pharmacology and Experimental Therapeutics*, *163*(3), 468–474.

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society*, *40*(9), 922–935. https://doi.org/10.1111/j.1532-5415.1992.tb01992.x

van Dalen-Kok, A. H., Achterberg, W. P., Rijkmans, W. E., Tukker-van Vuuren, S. A., Delwel, S., de Vet, H. C., … de Waal, M. W. (2017). Pain Assessment in Impaired Cognition (PAIC): Content validity of the Dutch version of a new and universal tool to measure pain in dementia. *Clinical Interventions in Aging*, *13*, 25–34.

Ware, L. J., Epps, C. D., Herr, K., & Packard, A. (2006). Evaluation of the revised faces pain scale, verbal descriptor scale, numeric rating scale, and Iowa pain thermometer in older minority adults. *Pain Management Nursing: Official Journal of the American Society of Pain Management Nurses*, *7*(3), 117–125. https://doi.org/10.1016/j.pmn.2006.06.005

Williams, A. C. D. C. (2002). Facial expression of pain: An evolutionary account. *Behavioral and Brain Sciences*, *25*(04), https://doi.org/10.1017/S0140525X02000080

How to cite this article: Kappesser J, Voit S, Lautenbacher S, Hermann C. Pain assessment for cognitively impaired older adults: Do items of available observer tools reflect pain-specific responses? *Eur J Pain*. 2020;00:1–12. https://doi.org/10.1002/ejp.1536