Published in final edited form as:
J Oral Rehabil. 2010 January ; 37(1): 11–20. doi:10.1111/j.1365-2842.2009.02017.x.

Psychometric properties of the modified Symptom Severity Index (SSI)

Donald R. Nixdorf[1,2], **Mike T. John**[1], **Melanie M. Wall**[3], **James R. Fricton**[1,4,5], and **Eric L. Schiffman**[1]

[1]Division of TMD & Orofacial Pain, School of Dentistry, University of Minnesota

[2]Department of Neurology, Medical School, University of Minnesota

[3]Division of Biostatistics, School of Public Health, University of Minnesota

[4]Department of Physical Medicine Rehabilitation, Medical School, University of Minnesota

[5]Division of Health Informatics, Medical School, University of Minnesota

Abstract

The psychometric properties of the modified Symptom Severity Index were investigated to assess the relationships among dimensions of pain in temporomandibular disorders. The 15-item instrument is composed of ordinal scales assessing five pain dimensions (intensity, frequency, duration, unpleasantness, and difficulty to endure) as experienced in three locations (temple, temporomandibular joint, masseter). In 108 closed-lock subjects, Cronbach's alpha was used to measure internal consistency resulting in 31 of the 105 pair-wise comparisons ≥0.71. Multilevel exploratory factor analysis was used to assess dimensionality between items. Two factors emerged, termed temple pain and jaw pain. The jaw pain factor comprised the temporomandibular joint and masseter locations, indicating that subjects did not differentiate between these two locations. With further analysis, the jaw pain factor could be separated into temporal aspects of pain (frequency, duration) and affective dimensions (intensity, unpleasantness, endurability). Temple pain could not be further reduced; this may have been influenced by concurrent orofacial pains such as headache. Internal consistency was high, with alphas ≥0.92 for scales associated with all factors. Excellent test-retest reliability was found for repeat testing at 2–48 hours in 55 subjects (ICC=0.97, 95%CI 0.96–0.99). In conclusion, the modified Symptom Severity Index has excellent psychometric properties for use as an instrument to measure pain in subjects with temporomandibular disorders. The most important characteristic of this pain is location, while the temporal dimensions are important for jaw pain. Further research is needed to confirm these findings and assess relationships between dimensions of pain as experienced in other chronic pain disorders.

Keywords

Pain dimensions; pain assessment; reliability; factor analysis; temporomandibular disorders (TMD); orofacial pain

---

Corresponding author: Donald Nixdorf, DDS MS, 6-320 Moos Tower, University of Minnesota, 515 Delaware Street S.E., Minneapolis, MN 55455, (612) 626-5407 phone, (612) 626-0138 fax, nixdorf@umn.edu.

Presented in abstract at the American Association for Dental Research (AADR) meeting in Dallas, TX on April 5th 2008

INTRODUCTION

For clinical trials of treatment interventions for chronic pain conditions, such as the pain that can accompany temporomandibular disorders (TMD), a complete and accurate representation of the pain experience is essential. The outcome variable primarily used in such research is pain intensity (1). Intensity is easily conceptualized, reliably reported by study subjects (2), and considered a robust measure (3). For reasons such as these, pain intensity was recommended as a core outcome variable for assessing pain in clinical trials (IMMPACT Statement) (4).

In opposition to this, it has been suggested that pain intensity alone does not adequately capture the experience of pain (5). Some researchers have assessed other dimensions of pain such as its sensory, affective, and evaluative components (6), as well as pain location (1). This type of measurement approach assumes that the different locations and dimensions of the pain experience can vary individually, and that treatments might have a dimension-specific influence on patients' pain experience. This assumption is the reasoning behind recommendations to capture information about these other dimensions of pain (1,6,7) and to include them as secondary outcome measures in clinical trials (4).

To date, research focusing on the temporal aspects of pain, i.e., frequency and duration, has been minimal (8). Consequently, it is unknown how these temporal aspects are related to other aspects of pain – such as intensity, affective dimensions, or even location – in people with TMD. Furthermore, it is unknown how these temporal dimensions, either alone or in combination, change with time or are affected by treatment. This may be of particular importance when assessing outcomes of therapeutic trials with TMD pain patients, since it is widely known that this pain can and often does range from being brief and episodic to being daily and continuous (9). Increasingly, clinical trials of TMD treatments have evaluated multiple dimensions of pain as an outcome measure, for example, pain duration (10), pain frequency (11,12), and pain unpleasantness (13,14). Therefore, there is a need to systematically assess the relationship between these multiple dimensions of pain, so as to better interpret the outcomes of clinical trials.

To address this need, we investigated the psychometric properties of the modified Symptom Severity Index (SSI), a self-report pain measure consisting of five subscales that measure pain frequency, duration, intensity, unpleasantness, and difficulty to endure (see Fricton et al, 1990 (15)) (Appendix 1). We used existing data to determine the instruments' reliability, item correlation, and the factor loading of the various pain dimensions, with the overall aim of better understanding the benefits and limitations of using such an instrument to measure TMD-related pain. The approach of using a secondary analysis is an efficient preliminary method for assessing these relationships and a key initial step towards refining this instrument for the measurement of core pain-related outcome variables, as performed by others (2).

METHODS

This study was reviewed, approved and conducted in accordance with the University of Minnesota Institutional Review Board regulations.

Participants

We analyzed data that had been previously collected from participants in a four-arm randomized effectiveness study of therapeutic strategies for temporomandibular joint (TMJ) closed-lock (16). The study enrolled patients who were experiencing pain with an anteriorly displaced TMJ disc that did not reduce, a restricted ability to open their mouths, and concomitant myofascial pain of the muscle of masticatory that conformed to diagnoses Ia or Ib in the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD). Patients were excluded if they had systemic rheumatoid arthritis, generalized joint pain, or major psychiatric disease; were pregnant; or were taking analgesic medications. The mean age of the 108 participants at baseline was 32 years, and 92% were female. Outcome measurements, which included the modified SSI, were taken at baseline, 3, 6, 9, 12, 18, 24, 36, 48 and 60 months after randomization to treatment arm (see Schiffman et al, 2007 (16)). For our exploratory study of the modified SSI's construct validity and internal consistency, we used all baseline and follow-up assessments in the analyses, resulting in a sample of 889 observations for the 108 participants.

For test-retest reliability analyses, we recruited a second separate sample of 55 participants, as part of the "Research Diagnostic Criteria [for TMD] – Reliability and Validity" study. This sample of participants had to have an RDC/TMD diagnosis of TMJ arthralgia and mysofascial pain, but did not necessarily have to have TMJ closed-lock (for details see Schiffman et al, 2009 (17)). This convenience sample was drawn from the same clinic population as the original sample.

Modified Symptom Severity Index

The original Symptom Severity Index (SSI) was first reported on in 1990 (15). The modified SSI used in this study asks respondents to rate their pain experience by completing five items (assessing pain frequency, duration, intensity, unpleasantness, or difficulty to endure) – for three different locations (temple, temporomandibular joint, masseter). The locations are diagrammed on the instrument with a lateral view of the head, face, and jaw. For each of the 15 items, respondents mark one of 28 bubbles on an optical scanning form, as an ordinal scale, to represent their pain experience. Potential responses for pain intensity, unpleasantness, and difficulty to endure range from "zero" or "no difficulty" on the far left to "worst imaginable" on the far right. Responses for pain frequency and duration range from "never" on the far left to "constant" on the far right. Within the pain frequency and duration scales, descriptors are periodically posted between the extremes to spread the responses over the full scale. Frequency is defined as the number of episodes of pain that occurred during the past month. Duration is defined as the typical length of time each episode of pain was present during the same time frame, during the past month. For scoring, each scale has a numeric range from 0 with the bubble on the far left to 1 with the bubble on the far right. Bubbles marked in between these extremes are rated linearly between the bounds of 0 and 1, those producing a fraction. All 5 scales are then averaged to produce the overall summary score.

Analyses

Data were collected from 108 participants with up to 9 follow-up visits after baseline. Eighteen percent of the 1080 potential subject-visit evaluations were missing and therefore not included in the analysis. Analyses were performed with the statistical software package STATA (Stata Statistical Software: Release 9. College Station, TX: StataCorp LP) or Mplus version 4.2 when the multilevel data structure was taken into account. Results were considered statistically significant if $P<0.05$.

Inspection of correlation among variables—The central characteristic of a construct is that the indicators of each construct co-vary with each other. For this reason, it is recommended that one assess the correlation matrix that represents all possible relationships between individual responses from the instrument (18). In our study, the data used to assess these relationships were not independent observations; rather, they came from repeated measures of the same individuals over a 5-year period. Consequently, we used multilevel analyses (with Mplus version 4.2) to calculate the appropriate between-item correlation matrix while controlling for the repeated measures. Resulting correlation coefficients of 0.71 or greater (meaning that the items share 50% or more of their variance) were used as a threshold to define large correlations between individual responses. Identifying large correlations in the data allows for highly correlated variables to be clustered or grouped, which is useful for assessing the data's tendency to separate into different dimensions.

Item analysis—To increase our understanding of the modified SSI and to possibly eliminate items from further analyses, we performed an item analysis according to recommendations (19). Means and total standard deviations were computed to compare items in terms of their ability to characterize respondents along the item continuum (item '*difficulty*'). Floor and ceiling effects, such as items with extreme means or near zero variances where no differentiation can take place, were also identified. Item '*discrimination*' – the extent to which an item differentiates between subjects with different levels of the trait – was evaluated by using the corrected item-total correlations, computed as the correlation between each item and the rest of the scale with the item absent. Taking into account the 10 repeated observations measured by the instrument over time, we computed intraclass correlations coefficients for each item as a ratio of variability due to subjects on an item divided by the total variability of that item. In this case, the intraclass correlations indicate the amount of stability in the items over the follow-up period.

Dimensionality—Multilevel exploratory factor analysis was performed by using Mplus 4.2 (20,21) to determine the nature and most appropriate number of factors that could adequately explain correlations among the items. The multilevel factor analysis focused on the estimated between-item correlation matrix that is representative of a random person at a random time point (i.e., within-level correlation matrix), while controlling for person-to-person differences (i.e., between-level correlation matrix). While this type of analysis allows for individuals to have different values on factors over time, it does assume that the overall factor structure itself (i.e., the number of factors and loadings) does not change over time. The number of eigenvalues of the correlation matrix greater than one and the percent of variability explained by each additional dimension were identified and used to indicate the potential number of factors. In addition, to explore whether this factorial structure would change substantially if more factors were extracted, we identified all four factor solutions with eigenvalues > 0.7. Promax rotation of factor loadings were examined. Values larger than 0.50 were considered as indicative of a relationship between the item and the associated underlying factor.

Reliability analysis—Cronbach's alpha (22) was computed as a measure of the scores' internal consistency. Test-retest reliability was assessed by using data from the in the convenience sample of 55 TMD pain subjects drawn from the same clinic population. These subjects had similar age and gender distributions as those in the clinical trial and completed the modified SSI twice, 2 to 48 hours apart. This relatively short interval was chosen because pain is known to vary from day to day (23). Data from other oral health self-report instruments suggest that the subject's memory does not substantially influence their reporting over short periods of repeated assessment (24). The resultant data were used to investigate the temporal reliability of the instrument. Intraclass correlation coefficients (ICC) were calculated using a one-way repeated measures ANOVA, treating the subjects as a random factor. Reliability was assessed for both the instrument's total score and the subscales found with the factor analysis. Furthermore, the method of Bland and Altman (25) was used to compute the standard deviation of the differences between the first and second time points. "Limits of agreement" around the mean difference were calculated as 1.96 times the standard deviation of the differences. Hence, this statistic represents the test-retest differences expected for 95% of the individuals in the sample. If the confidence interval for the mean of the differences excluded zero, it indicated a statistically significant difference between the measures.

RESULTS

Inspection of correlations among the instrument's variables

The correlation matrix scores for the modified SSI items were substantial, but variable (Table 1). Overall, all 15 items correlated highly. Of the 105 possible pair-wise correlations, 31 were larger than or equal to 0.71, indicating that almost one third of the items shared at least 50% of their information with each other. These results are evidence that items in the modified SSI share common underlying factors.

Item analysis

Item analysis results are in Table 2. Item means and standard deviations were calculated based on the 0–27 raw score from the ordinal scale. For this sample of individuals experiencing pain with TMJ closed-lock, TMJ was the location of the greatest pain intensity (mean item scores ranged from 7.1 to 13.5), followed by the masseter (range = 6.3 to 11.7) and then the temple (range =5.8 to 9.1). Among the pain characteristics, *frequency* of pain received the highest ratings for the pain experienced in the masseter and TMJ, whereas *duration* of pain was scored highest for the temple. None of these items showed considerable floor or ceiling effects as evidenced by the fact that their ratings included much of the scale, with few values at the extreme ends of the scale. Considerable between-subject variance was demonstrated by a wide variability in participants' responses. Item-total correlations were between 0.79 and 0.94, indicating a substantial correlation between each item and the construct as a whole. The intraclass correlation coefficients indicate that the temple items were more stable over time than the masseter and TMJ items.

Dimensionality

Several factor structures were explored (Table 3). When the criterion for an eigenvalue greater than 1 was applied, two factors were retained for further analysis. Together, the first two factors explained 72% of the variance, with the first eigenvalue of 9.6 representing 64% of the variance and the second eigenvalue of 1.8 explaining an additional 12%. When only item loadings with correlations greater than 0.5 on the factors were considered (26), a clear and simple structure emerged: All masseter and TMJ location items loaded on the first factor, and only the temple location items loaded on the second factor.

We also explored factor solutions with factors having eigenvalues less than 1. This changed our findings only slightly. When three factors were extracted from the data, the third factor explained an additional 6% of the variance before rotation and had an eigenvalue of 0.95. Interestingly, all temple location items still loaded strongly on the second factor. However, the first factor from the previous two-factor analysis was split. Specifically, masseter and TMJ pain *intensity, unpleasantness*, and *endurability* clustered together, while masseter and TMJ pain *frequency* and *duration* clustered together to form the third factor.

When four factors were extracted from the data, the fourth factor explained an additional 5% of the variance before rotation and had an eigenvalue of 0.7. The previously observed pattern of loading changed only slightly. All five temple pain items still clustered together (first factor), and masseter and TMJ pain *duration* and *frequency* still loaded together (second factor). But the previously observed single factor for *intensity-unpleasantness-endurability* at the masseter and TMJ locations was split into two different factors by pain location: intensity-unpleasantness-endurability in the masseter (third factor) and intensity-unpleasantness-endurability in the TMJ (fourth factor).

In summary, the factors were interpreted and named as follows:

1. **temple pain**, which comprises all 5 dimension of pain – *duration*, *frequency*, *intensity, unpleasantness, and endurability* – for the temple location;
2. **temporality of jaw pain**, which comprises the *duration* and *frequency* of pain for both masseter and TMJ locations; and
3. **extent of jaw pain**, which comprises the *intensity*, *unpleasantness*, and *endurability* of pain at both master and TMJ locations. If further differentiation is warranted, the *extent* factor could be differentiated into two subcomponents according to the location of the pain problem: **extent of masseter pain**, and **extent of TMJ pain**.

Because the factor rotation method is an oblique technique, the factors can be correlated with each other. Correlations among the factors of the 3 solutions ranged from 0.46 to 0.63.

When an orthogonal rotation method (varimax) was used instead of an oblique method, the results were as follows: The 2-factor solution was identical; neither the item assignment to the factor nor the pattern of loadings at 0.50 changed. A slight change was observed in the 3-factor solution; the *TMJ pain frequency* item loaded on both the first and third factor at greater than 0.50, but loaded substantially higher on the third factor than it did in the oblique rotation method. The more complex 4-factor orthogonal solution was similar to the 4-factor oblique solution in that a clear *temple pain* factor still remained, but had difficulty separating the other factors as there were additional loadings at 0.50 or greater on multiple other factors. However, these loadings were in the range of 0.50–0.58 and were always smaller than the other loadings used previously for factor assignment. Although some other substantial loadings indicated a slightly more complex relationship between items and factors, item assignment to factor and the number of factors did not change. Hence, overall the oblique and orthogonal rotation methods yielded a similar factor structure of the modified SSI.

Reliability

Internal consistency was high, with alphas ≥ 0.92 for scales associated with all factors. The alpha for the total instrument was 0.96. No substantial differences among the factors' homogeneity were found. Test-retest reliability was excellent for the instrument's summary score and all factor scores (Table 4). The lower limits of ICCs exceeded the 0.75 value, which is considered excellent reliability (27). Changes in scores, characterized by the mean of the differences between the first time point and the second time point, were small and not statistically significant. Individual score differences from baseline to follow-up were moderate, as shown by the limits of agreement when compared to the variable range.

DISCUSSION

Our study found that multiple dimensions of pain arising from TMJ closed-lock – specifically, pain location, intensity, duration, frequency, unpleasantness, and difficulty to endure – as assessed by the modified SSI are strongly correlated, with a high degree of commonality between items. These results suggest that the instrument is capable of globally assessing pain with excellent reliability and precision. By considering all reliability findings and applying the eigenvalue > 1 criterion for retaining a factor, the most appropriate factor structure of the modified SSI questionnaire is a two-factor solution that includes a jaw pain factor and a temple pain factor.

TMD is currently viewed as a heterogeneous group of disorders affecting the temporomandibular joints and muscles of mastication (28,29), as opposed to the original idea that it is one disorder centred around the TMJ (30,31). The concept of pain being ascribed to a specific anatomic location is how the RDC/TMD classifies this group of disorders (32). Our results partly support the concept that pain arising from TMJ closed-lock is one disorder, since a high degree of information is shared between all 15 items of this questionnaire. Our results also partly support the concept that location is an important dimension, since the two-factor solution suggests that pain in the temple region is distinct from pain in the jaw region. This division is in contrast to that proposed within the RDC/TMD, which classifies pain on the basis of the tissue of origin – pain arising from masticatory muscles verses from the temporomandibular joints (32).

Support for a delineation of pain within TMD by anatomic region comes from a previous report in which subjects perceived pain in the temple region as being separate from pain in the jaw region (33). Conversely, this separation was not observed when aspects of pain quality, some that were not ascertained by the SSI (i.e. frequency, unpleasantness, difficultly to endure) were included in a factor analysis (34); however, this prior study assessed the relationship of pain components for various pain conditions presented in the orofacial region, and not specifically for TMD. The inability for subjects to separate these anatomical regions when pain is present may explain why descriptive studies have repeatedly shown that the majority of people with TMD present with complaints of pain in both masseter and TMJs (35). Furthermore, emerging research indicates that the RDC diagnosis of TMJ pain, otherwise known as arthralgia, has a sensitivity between 0.52 to 0.55 (36). This suggests that the existing literature has underestimated the prevalence of TMJ pain, and therefore has underreported the co-morbid presentation of TMJ and masseter muscle pain complaints.

Pain perceived in the temple location can be associated not only with TMD, but also other disorders such as migraine headaches (37) and referred pain from the neck (38,39). The latter two conditions are known to be commonly co-morbid with TMD (40–42). Therefore, subjects assessing pain perceived in the temple location might be more likely to have these other pain disorders, which may be a reason why temple pain remained a separate factor.

Most research investigating the various dimensions of clinical pain states has used subjects who are experiencing daily continuous pain, rather than episodic pain (2,23,43). This is understandable, since these individuals are often the worst afflicted. Nevertheless, chronically reoccurring pains can also have a negative impact on the individual (44). We found no other published reports that assessed the relationship between pain frequency, defined as the number of episodes of pain during a particular period of time, and pain duration, defined as the typical length of time each episode of pain is present, for episodic pain disorders. In our factor analysis, an assessment of the temporal dimensions of pain in TMD subjects added a small amount of location-specific information: the temporality of jaw pain could be separated from the extent of jaw pain. However, the individual temporal components of pain frequency and duration were highly correlated, regardless of location, and could not be further separated using factor analysis. This suggests that, at least in subjects with TMJ closed-locked pain, both temporal components of pain are interdependent. The finding of such a high correlation between the temporal dimensions is contrary to our clinical impression and that of others (45). We originally assumed that greater amounts of variation between these dimensions would be seen, because they are thought to represent two separate dimensions of the pain experience. However, our findings are consistent with clinical observations that individuals with TMD pain experience simultaneous improvement of both pain frequency and duration with treatment over time. Such a strong correlation between pain frequency and duration remains to be assessed for other pain disorders. The relationship of temporal dimensions may be different in other disorders.

We also found that the affective dimensions of unpleasantness and difficulty to endure, as measured by the modified SSI, are highly correlated. In our analyses, some of the highest correlations were observed between pain intensity and unpleasantness. In addition, when assessed with factor analysis, pain intensity remained clustered with both affective items, providing more evidence that these constructs are very similar and may be collapsed together that can be addressed in the future following a confirmatory analysis.

Limitations to our study must be considered. First, data were derived from subjects experiencing pain associated with TMJ closed-lock. Further research is needed to confirm our findings in different types of TMD conditions and pain conditions. Second, in previous psychophysical research of the taste sensation, it was found that subjects will 'dump' or impart attributes of other characteristics into their responses when they do not have the ability to specifically report those characteristics (46). A similar type of loading has been observed within an experimental pain paradigm (47), thus it is possible that this effect is present in our data. We suspect, however, that this is unlikely to be occurring to a significant extent. 'Dumping' occurs more often when the number of responses is restricted, such as when only asking about pain intensity, but our assessment included several pain characteristics, thereby minimizing this potential source of error. Third, although the modified SSI includes a diagram for location, it does not provide respondents with an operationalized definition for each item or establish a reference pain level, both of which are thought to improve the accuracy of self-reported pain ratings. We did find that subjects consistently scored the questionnaire in the same manner, as evidenced by good to excellent test-retest reliability over a short-term interval of 2 hours to 2 days (ICC = 0.90–0.96) and a medium-term interval of 2 weeks (ICC = 0.89–0.97) (15). This suggests that subjects consistently use the same information to complete questionnaire items.

In conclusion, the modified Symptom Severity Index (SSI) has excellent reliability. The 5 items per location measure closely related constructs, producing an instrument with high levels of precision, which is a prerequisite for use in small clinical trials. Factor analysis revealed two constructs, temple pain and jaw pain, suggesting that location are important dimensions of pain associated with TMJ closed-lock. Jaw pain is a less homogenous construct than temple pain, breaking into a temporal component and an intensity-affective component of the pain experience with a three-factor analysis. Further research is needed to confirm these findings in different types of TMD conditions and other pain conditions, and to explore options for reducing the number of items to yield a shorter instrument.

Acknowledgments

The authors acknowledge, and are grateful for, assistance from Amanda Jackson and Mary Haugan in data collection and management. We also acknowledge Lisa McGuire, who assisted in literature searches, Pat Lenton and Patt Carlson, who assisted in research subject management, and Anne Marie Weber-Main, who edited manuscript drafts.

Support

This research was supported by NIH grants R29-DE08668 (ELS), U01-DE013331 (ELS), and K12-RR23247 (DRN).

APPENDIX 1

Example of SSI form

REFERENCES

1. Jensen, MP.; Karoly, P. Self-report scales and procedures for assessing pain in adults. In: Turk, DC.; Melzack, R., editors. Handbook of pain assessment. 2nd Edition. New York, NY: Guilford Press; 2001. p. 15-34.
2. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical rating scale. Pain 2001;94:149–158. [PubMed: 11690728]
3. Jensen, MP. Validity of self-report and observation measures. In: Jensen, TS.; Turner, JA.; Wiesenfeld-Hallin, Z., editors. Proceedings from the 8th World Congress on Pain; Seattle, WA: IASP Press; 1997. p. 637-661.
4. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9–19. [PubMed: 15621359]
5. de C Williams AC, Oakley Davis HT, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. Pain 2000;85:457–463. [PubMed: 10781919]
6. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain 1975;1:277–299. [PubMed: 1235985]
7. Price DD, Harkins SW, Baker C. Sensory-affective relationships among different types of clinical and experimental pain. Pain 1987;28:297–307. [PubMed: 2952934]
8. Jensen MP. The validity and reliability of pain measures in adults with cancer. J Pain 2003;4:2–21. [PubMed: 14622723]
9. Ohrbach R, Dworkin SF. Five-year outcomes in TMD: relationship of changes in pain to changes in physical and psychological variables. Pain 1998;74:315–326. [PubMed: 9520246]
10. Truelove E, Huggins KH, Mancl L, Dworkin SF. The efficacy of traditional, low-cost and nonsplint therapies for temporomandibular disorder: a randomized controlled trial. J Am Dent Assoc 2006;137:1099–1107. [PubMed: 16873325]
11. Wahlund K, List T, Larsson B. Treatment of temporomandibular disorders among adolescents: a comparison between occlusal appliance, relaxation training, and brief information. Acta Odontol Scand 2003;61:203–211. [PubMed: 14582587]
12. Ekberg E, Vallon D, Nilner M. Occlusal appliance therapy in patients with temporomandibular disorders: A double-blind controlled study in a short-term perspective. Acta Odontol Scand 1998;56:122–128. [PubMed: 9669465]
13. Dao TTT, Lavigne GJ, Charbonneau A, Feine JS, Lund JP. The efficacy of oral splints in the treatment of myofascial pain of the jaws: A controlled clinical trial. Pain 1994;56:85–94. [PubMed: 8159444]
14. Nixdorf DR, Heo G, Major PD. Controlled trial of botulinum toxin A for chronic myogenous orofacial pain. Pain 2002;99:465–473. [PubMed: 12406522]
15. Fricton JR. Musculoskeletal measures of orofacial pain. Anesth Prog 1990;37:136–143. [PubMed: 2085192]
16. Schiffman EL, Look JO, Hodges JS, Swift JQ, Decker KL, Hathaway KM, et al. Randomized effectiveness study of four therapeutic strategies for TMJ closed lock. J Dent Res 2007;86:58–63. [PubMed: 17189464]
17. Schiffman E, Truelove E, Ohrbach R, Anderson G, John M, List T, et al. Assessment of the validity of the research diagnostic criteria for temporomandibular disorders: Overview and methodology. J Orofac Pain. 2009 In press.
18. Pett, MA.; Lackey, NR.; Sullivan, JJ. Making sense of factor analysis - The use of factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage Publications; 2003.
19. Schultz, KS.; Witney, DJ. Measurement theory in action: Case studies and exercises. Thousand Oaks, CA: Sage Publications; 2005.
20. Reise SP, Ventura J, Nuechterlein KH, Kim KH. An illustration of multilevel factor analysis. J Pers Assess 2005;84:126–136. [PubMed: 15799887]
21. Muthen BO. Multilevel covariance structure analysis. Sociol Methods Res 1994;22:376–398.

22. Cronbach LJ. Coefficient alpha and the internal reliability of tests. Psychometrika 1951;16:297–334.
23. Jensen MP, Turner LR, Turner JA, Romano JM. The use of multiple-item scales for pain intensity measurement in chronic pain patients. Pain 1996;67:35–40. [PubMed: 8895229]
24. John MT, Reissmann DR, Schierz O, Allen F. No significant retest effects in oral health-related quality of life assessment using the Oral Health Impact Profile. Acta Odontol Scand 2008;66:135–138. [PubMed: 18568471]
25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310. [PubMed: 2868172]
26. Pett, MA.; Lackey, NR.; Sullivan, JJ. Extracting the initial factors: selecting the number of factors to retain. In: Pett, MA.; Lackey, NR.; Sullivan, JJ., editors. Making sense of the factor analysis: the use of the factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage; 2003. p. 115-116.
27. Fleiss, JL.; Levin, B.; Paik, MC. The measure of interrater agreement. In: Fleiss, JL.; Levin, B.; Paik, MC., editors. Statistical Methods for Rates and Proportions. 3rd Edition. Hoboken, NJ: Wiley & Sons; 2003. p. 598-626.
28. DeBoever, JA.; Carlsson, GE. Etiology and differential diagnosis. In: Zarb, GA.; Carlsson, GE.; Sessle, BJ.; Mohl, ND., editors. Temporomandibular joint and masticatory muscle disorders. 2nd Edition. Copenhagen: Munksgaard; 1994. p. 171-187.
29. Management of Temporomandibular Disorders; NIH Technology Assessment Statement. 1996. p. 1-32.
30. Costen JB. A syndrome of ear and sinus symptoms dependent upon disturbed function of the temporomandibular joint. Ann Otol Rhinol Laryngol 1934;43:1–15.
31. Costen JB. Neuralgias and ear symptoms associated with disturbed function of the temporomandibular joint. JAMA 1936;107:252.
32. Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: Review criteria, examinations and specifications, critique. J Craniomandib Disord 1992;6:301–355. [PubMed: 1298767]
33. Schiffman E, Haley D, Baker C, Lindgren B. Diagnostic criteria for screening headache patients for temporomandibular disorders. Headache 1995;35:121–124. [PubMed: 7721570]
34. Woda A, Turbert-Jeannin S, Bouhassira D, Attal N, Fleiter B, Goulet J, et al. Towards a new taxonomy of idiopathic orofacial pain. Pain 2005;116:396–406. [PubMed: 15979796]
35. Schiffman EL, Fricton JR, Haley DP, Shapiro BL. The prevalence and treatment needs of subjects with temporomandibular disorders. J Am Dent Assoc 1990;120:295–303. [PubMed: 2312948]
36. Look J, John M, Tai F, Huggins K, Lenton P, Truelove E, et al. Research diagnostic criteria for temporomandibular disorders: Reliability of axis I diagnoses and selected clinical measures. J Orofac Pain. 2009 In press.
37. Silberstein, SD.; Lipton, RB.; Goadsby, PJ. Classification and diagnosis of headache. In: Silberstein, SD.; Lipton, RB.; Goadsby, PJ., editors. Headache in clinical practice. 2nd Edition. New York, NY: Martin Dunitz; 2002. p. 11-20.
38. Bogduk N. The neck and headaches. Neurol Clin N Am 2004;22:151–171.
39. Schmidt-Hansen PT, Svensson P, Bendtsen L, Graven-Nielsen T, Bach FW. Increased muscle pain sensitivity in patients with tension-type headache. Pain 2007;129:113–121. [PubMed: 17161538]
40. Ciancaglini R, Radaelli G. The relationship between headache and symptoms of temporomandibular disorder in the general population. J Dent 2001;29:93–98. [PubMed: 11239582]
41. Wright EF. Referred craniofacial pain patterns in patients with temporomandibular disorders. JADA 2000;131:1307–1315. [PubMed: 10986831]
42. Mitrirattanakul S, Merrill RL. Headache impact in patients with orofacial pain. J Am Dent Assoc 2006;137:1267–1274. [PubMed: 16946432]
43. Jensen MP, McFarland CA. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. Pain 1993;55:195–203. [PubMed: 8309709]

44. Lipton RB, Bigal ME, Diamond M, Freitag F, Reed ML, Stewart WF. Migraine prevalence, disease burden, and the need for preventative therapy. Neurology 2007;68:343–349. [PubMed: 17261680]
45. Goss, AN. Toward an international consensus on temporomandibular joint surgery; Report of the Second International Consensus Meeting, April 1992; Buenos Aires, Argentina: Int J Oral Maxillofac Surg; 1993. p. 78-81.
46. Lawless, HT.; Heymann, H. Context effects and biases in sensory judgment. In: Lawless, HT.; Heymann, H., editors. Sensory evaluation of food: Principles and practices. Gaithersburg, MD: Aspen Publishing; 1999. p. 301-340.
47. Levine FM, De Simone LL. The effects of experimenter gender on pain report in male and female subjects. Pain 1991;44:69–72. [PubMed: 2038491]

Table 1

Correlation matrix for the 15 SSI items

		Masseter pain					TMJ pain					Temple pain			
		Freq[1]	Dur[2]	Int[3]	Unpl[4]	Endur[5]	Freq[1]	Dur[2]	Int[3]	Unpl[4]	Endur[5]	Freq[1]	Dur[2]	Int[3]	Unpl[4]
Masseter pain	Dur	**0.78**													
	Int	**0.75**	0.70												
	Unpl	**0.76**	**0.73**	**0.92**											
	Endur	**0.72**	0.70	**0.88**	**0.91**										
TMJ pain	Freq	**0.77**	0.59	0.61	0.63	0.60									
	Dur	0.57	0.70	0.57	0.58	0.57	**0.72**								
	Int	0.60	0.55	**0.79**	**0.74**	**0.73**	**0.71**	0.67							
	Unpl	0.63	0.59	**0.74**	**0.81**	**0.75**	**0.75**	0.69	**0.91**						
	Endur	0.63	0.60	**0.75**	**0.78**	**0.87**	0.68	0.64	**0.85**	**0.88**					
Temple pain	Freq	0.56	0.51	0.47	0.47	0.48	0.54	0.47	0.45	0.47	0.47				
	Dur	0.44	0.55	0.39	0.40	0.41	0.42	0.55	0.37	0.39	0.39	**0.76**			
	Int	0.44	0.45	0.58	0.53	0.53	0.44	0.44	0.58	0.54	0.54	0.70	**0.71**		
	Unpl	0.43	0.44	0.52	0.55	0.53	0.44	0.45	0.52	0.57	0.54	**0.71**	**0.71**	**0.89**	
	Endur	0.46	0.47	0.55	0.57	0.64	0.45	0.44	0.54	0.56	0.65	0.65	0.64	**0.83**	**0.87**

[1] **Frequency,**

[2] **Duration,**

[3] **Intensity,**

[4] **Unpleasantness,**

[5] **Difficulty to Endure**

[*] **correlations ≥ 0.71 are presented in bolded font**

Table 2

Descriptive statistics for samples used in the different steps of instrument evaluation

Pain location	Characteristic	*Item analysis (109 subjects, 889 observations)*			*Test-retest reliability (55 subjects, 110 observations)*
		Mean (SD*)	**Item–total correlation**	**ICC+**	**Mean (SD*)**
Masseter pain	Frequency	11.7 (9.2)	0.79	0.34	7.3 (8.5)
	Duration	11.4 (9.2)	0.79	0.39	6.9 (8.2)
	Intensity	7.9 (6.8)	0.85	0.32	5.3 (6.6)
	Unpleasantness	7.8 (7.1)	0.86	0.31	5.1 (7.1)
	Endurability	6.3 (6.5)	0.83	0.30	4.3 (6.7)
TMJ pain	Frequency	13.5 (9.0)	0.75	0.34	8.8 (9.2)
	Duration	12.7 (9.1)	0.74	0.38	8.5 (9.0)
	Intensity	9.2 (7.1)	0.84	0.29	6.8 (7.5)
	Unpleasantness	9.0 (7.4)	0.86	0.29	6.3 (7.5)
	Endurability	7.1 (6.8)	0.85	0.28	5.0 (7.0)
Temple pain	Frequency	8.7 (8.0)	0.75	0.51	5.4 (7.0)
	Duration	9.1 (8.3)	0.72	0.49	5.6 (6.7)
	Intensity	7.0 (6.8)	0.76	0.51	4.9 (6.4)
	Unpleasantness	7.0 (7.0)	0.76	0.52	4.9 (6.5)
	Endurability	5.8 (6.3)	0.76	0.51	3.0 (4.9)
Total SSI score		124.4 (94.2)			86.6 (83.2)

Possible range is from 0 to 27 for each scale

*SD – standard deviation

+ICC - Intraclass correlation coefficient

Note: means and SD calculated across all 889 observations

Table 3

Factor loadings from the rotated factor structure matrix for the modified SSI: Principal components analysis with promax (oblique) rotation

		Two factors		Three factors			Four factors			
		1	**2**	**1**	**2**	**3**	**1**	**2**	**3**	**4**
Jaw muscle pain	Freq	**0.79**	0.04	0.37	−0.09	**0.66**	0.49	−0.07	**0.58**	0.01
	Dur	**0.69**	0.12	0.26	−0.03	**0.70**	0.45	−0.01	**0.63**	−0.08
	Int	**0.89**	0.01	**0.75**	0.03	0.20	**0.74**	0.05	0.11	0.15
	Unpl	**0.92**	−0.01	**0.77**	0.01	0.21	**0.78**	0.03	0.12	0.14
	Endur	**0.88**	0.03	**0.78**	0.06	0.14	**0.77**	0.08	0.06	0.16
Jaw joint pain	Freq	**0.75**	0.05	0.43	−0.03	**0.50**	−0.01	−0.06	**0.53**	**0.52**
	Dur	**0.65**	0.13	0.32	0.04	**0.52**	−0.07	0.01	**0.56**	0.45
	Int	**0.85**	0.03	**0.83**	0.08	0.03	0.15	0.06	0.03	**0.79**
	Unpl	**0.88**	0.02	**0.83**	0.07	0.08	0.14	0.04	0.08	**0.81**
	Endur	**0.86**	0.05	**0.86**	0.10	0.01	0.34	0.10	0.01	**0.61**
Temple pain	Freq	0.10	**0.74**	−0.10	**0.61**	0.42	−0.03	**0.62**	0.39	−0.05
	Dur	−0.03	**0.83**	−0.29	**0.69**	**0.50**	−0.14	**0.70**	0.47	−0.14
	Int	0.03	**0.90**	0.17	**0.87**	−0.06	0.06	**0.88**	−0.05	0.09
	Unpl	−0.01	**0.94**	0.15	**0.92**	−0.08	0.03	**0.92**	−0.06	0.09
	Endur	0.13	**0.78**	0.29	**0.79**	−0.11	0.18	**0.80**	−0.11	0.10

− Loadings >|0.5| are presented as bolded font − Shaded cells represent the clustering of each factor

Table 4

Test-retest reliability of instrument and the factor summary scores

Factor (no of items; variable range)			ICC (95% CI+)	Mean of the differences (95% CI+)	Limits of agreement
Temple Pain (5; 0–135)			0.96 (0.94 to 0.98)	1.87 (−0.39 to 4.13)	−13.74 to 17.49
Jaw Pain (10; 0–270)			0.95 (0.93 to 0.98)	0.96 (−4.87 to 6.80)	−39.38 to 41.30
	Temporality of Jaw Pain (4; 0–108)		0.94 (0.91 to 0.97)	−2.11 (−5.27 to 1.05)	−23.99 to 19.77
	Extent of Jaw Pain (6; 0–162)		0.94 (0.91 to 0.97)	3.07 (−0.51 to 6.65)	−21.69 to 27.83
		Extent of Jaw Muscle Pain (3; 0–81)	0.95 (0.92 to 0.97)	1.64 (−0.12 to 3.39)	−10.48 to 13.76
		Extent of Jaw Joint Pain (3; 0–81)	0.90 (0.84 to 0.95)	1.44 (−1.26 to 4.13)	−17.19 to 20.06
SSI total score (15; 0–405)			**0.97 (0.96 to 0.99)**	**2.84 (−2.84 to 8.51)**	**−36.40 to 42.07**

+ Confidence interval