Published in final edited form as:
J Pain Symptom Manage. 2013 September ; 46(3): 413–421. doi:10.1016/j.jpainsymman.2012.08.006.

Validity of an Observation Method for Assessing Pain Behavior in Individuals With Multiple Sclerosis

Karon F. Cook, PhD,
Northwestern University, Chicago, Illinois

Toni S. Roddey, PT, PhD,
Texas Women's University, Denton, Texas

Alyssa M. Bamer, MPH,
University of Washington, Seattle, Washington

Dagmar Amtmann, PhD, and
University of Washington, Seattle, Washington

Francis J Keefe, PhD
Duke University, Durham, North Carolina, USA

Abstract

Context—Pain is a common and complex experience for individuals who live with multiple sclerosis (MS) that interferes with physical, psychological and social function. A valid and reliable tool for quantifying observed pain behaviors in MS is critical to understanding how pain behaviors contribute to pain-related disability in this clinical population.

Objectives—To evaluate the reliability and validity of a pain behavioral observation protocol in individuals who have MS.

Methods—Community-dwelling volunteers with multiple sclerosis (*N*=30), back pain (*N*=5), or arthritis (*N*=8) were recruited based on clinician referrals, advertisements, fliers, web postings, and participation in previous research. Participants completed measures of pain severity, pain interference, and self-reported pain behaviors and were videotaped doing typical activities (e.g., walking, sitting). Two coders independently recorded frequencies of pain behaviors by category (e.g., guarding, bracing) and inter-rater reliability statistics were calculated. Naïve observers reviewed videotapes of individuals with MS and rated their pain. Spearman correlations were calculated between pain behavior frequencies and self-reported pain and pain ratings by naïve observers.

Results—Inter-rater reliability estimates indicated the reliability of pain codes in the MS sample. Kappa coefficients ranged from moderate agreement (sighing = 0.40) to substantial agreement (guarding = 0.83). These values were comparable to those obtained in the combined back pain and

Address correspondence to: Karon F. Cook, PhD, Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, 710 N. Lake Shore Drive, Abbott Hall, Suite 729, Chicago, IL 60611, USA, karon.cook@northwestern.edu.

Disclosures
The authors have no financial or personal relationships to disclose that could be perceived to bias the work conducted as part of this project.

arthritis sample. Concurrent validity was supported by correlations with self-reported pain (0.46-0.53) and with self-reports of pain behaviors (0.58). Construct validity was supported by finding of 0.87 correlation between total pain behaviors observed by coders and mean pain ratings by naïve observers.

Conclusion—Results support use of the pain behavior observation protocol for assessing pain behaviors of individuals with MS. Valid assessments of pain behaviors of individuals with MS in could lead to creative interventions in the management of chronic pain in this population.

Keywords

Multiple sclerosis; pain behaviors; observation protocol; validity

Introduction

In the past several decades, there have been substantial shifts in how pain is conceptualized. Once viewed strictly from a biomedical model, pain was defined as a response to disease or tissue damage. The biomedical model, however, has proved inadequate for explaining the perception and impact of pain. Evidence for psychosocial determinants of pain has accumulated in recent years. For example, Keefe and colleagues found coping strategies to be more predictive of self-reported knee pain than radiographic evidence of disease (1). Fordyce was among early, behaviorally-oriented pain researchers who explored the psychosocial role of pain behaviors and their contribution to subsequent disability (2). Pain behaviors include resting, guarding, facial expressions, asking for help, taking medication, and other observable displays. After an injury or pain exacerbation, such behaviors are protective and/or may be effective in eliciting support and assistance (3). When pain behaviors are maintained beyond the acute stage, however, they may contribute to subsequent psychosocial and physical disability (4). Research suggests pain behaviors are useful targets for interventions such as counter-conditioning and that such interventions can reduce subsequent pain-related impairment as well as reports of pain (2).

Historically, pain behavior assessment approaches sought to assess pain behavior using diaries, behavioral interviews, or controlled clinical protocols in which health care professionals are trained to observe and record pain behavior. Keefe and Block developed a standardized approach for sampling and measuring pain behaviors in individuals with low back pain, hereafter referred to as the “observation protocol” (5). In this approach, participants are videotaped performing sequences of standardized activities including walking, standing, reclining, and sitting. A trained observer subsequently codes pain behavior by type (e.g., guarding, bracing, grimacing). The reliability and validity of this approach has been confirmed in individuals with lower back pain (6), arthritis (1, 7), and cancer (8).

Among individuals who live with multiple sclerosis (MS), pain is a common and complex experience that interferes with physical, psychological and social function (9-11). A valid and reliable tool for quantifying observed pain behaviors in MS is critical to understanding how pain behaviors contribute to pain-related disability in this clinical population. To date, however, no study has tested the Keefe and Block observation protocol for measuring pain behaviors in patients with MS. The purpose of the current study was to evaluate the reliability and validity of this protocol in a sample of individuals with MS. The methods and results for this work are reported here as three separate studies. Discussion of findings is combined for all studies. Approval for all studies was obtained from the institutional review boards of the University of Washington, Seattle and Texas Woman’s University, Houston.

Study 1

Study 1 evaluated inter-observer reliability and concurrent validity of the Keefe and Block observation protocol.

Methods

Participants—Participants were recruited from a survey study of self-reported pain behavior as well as from clinician referrals, advertisements, fliers, and web postings. An invitation letter was sent to potential participants along with a contact number. Interested individuals were scheduled for in-person sessions at the study office if they met eligibility criteria: 1) self-reported diagnosis of MS, 2) 18 years of age or older, 3) able to read or understand English, 4) self-reported average pain score for the past seven days of >4 on a 0-10 scale where "0" indicates "no pain" and "10: indicates "worst possible pain," 5) pain that has persisted for six months or longer, and 6) able to walk (with or without assistive device).

Measures—Immediately prior and immediately after the videotaping sessions, participants reported their "current pain" on a 0-10 scale where "0" indicated "no pain" and "10" indicated "pain as bad as you could imagine." Either immediately before or after the videotaping sessions, participants completed a survey that included demographic questions and several pain-related self-report instruments. A body pain map was completed to indicate locations of participants' pain. The pain map is a graphic showing two outlines of a body, one facing forward and the other facing backward (12). Participants also completed self-reported pain behavior and pain interference measures developed by the National Institutes of Health Patient Reported Outcome Information System (NIH-PROMIS; http://www.nihpromis.org/measures/availableinstruments) initiative. The pain behavior measure comprised 12 items selected from the NIH-PROMIS Pain Behavior bank of items (13). The pain interference measure was the PROMIS six-item short form (Version 6b) (14). Additionally, participants reported their "average pain" scores for the past seven days on a 0-10 scale where "0" indicates "no pain" and "10" indicates "worst possible pain."

Procedures—Participants were videotaped doing a standard set of typical everyday activities that included: sitting (one- and two-minute intervals), standing (one- and two-minute intervals), reclining (two one-minute intervals), and walking (two one-minute intervals). The order was randomized to minimize order effects. Each videotaped session took approximately 10 minutes. Participants were informed that the person videotaping would not be interacting with them during the activities.

Analyses

Coding—A physical therapist (Roddey) was trained by one of the developers of the protocol (Keefe) in a series of sessions in which both coded videotapes archived from another pain behavior study. The trainer and trainee watched videotapes together, simultaneously coded the observed behaviors, and then compared codes. This process was continued until agreement of 85% or greater was routinely reached. Videotapes from the current study were divided into coding and recording intervals using dubbed verbal instructions to "observe" (for 20 seconds) and then "code" (10 seconds). When instructed to "observe," the coder began watching the video until the audio voice instructed the coder to "record." Coders recorded the participant's position, movements, and pain behaviors observed during the 20-second "observe" period. The position codes included *sitting, standing*, and *reclining*. The movement codes included *pacing* (taking at least two steps within a three-second interval) and *shifting* (vertical change in position, i.e., moving upward or downward). The pain behavior codes included *guarding, bracing, grimacing, rubbing* and

sighing. Coding of each behavior was dichotomous—it was coded as present or absent. Full descriptions of pain behaviors have been previously published (6, 7). Briefly, *guarding* can occur during sitting, standing, or reclining and includes stiffness, rigid movements, and use of canes or walkers during walking intervals; *bracing* can occur during sitting, standing, or reclining and is defined as at least three consecutive seconds in which a limb is extended to support and maintain an abnormal distribution of weight; *grimacing* includes obvious facial expressions such as clenched teeth or furrowed brow; *sighing* is air exhalation that is obvious or exaggerated.

Descriptive Statistics and Reliability Estimates—Means, ranges, and standard deviations (SDs) were calculated for all position, movement, and pain behavior codes. Inter-rater reliability was estimated using percentage agreement, percentage effective agreement, and kappa coefficients. Inclusion of all three of these statistics was chosen to maximize comparison of results with previous studies. Percentage agreement was defined as the total number of intervals that both coders agreed on the occurrence or non-occurrence of a behavior, divided by the total number of intervals (usually = 20), multiplied by 100. Because some behaviors such as sighing were observed infrequently, percentage effective agreement also was calculated. In this statistic, the numerator from percentage agreement was divided by the number of coding intervals in which agreements plus disagreements occurred (i.e., when at least one coder recorded observing the behavior). To obtain estimates of agreement that were corrected for chance agreements, we calculated kappa coefficients. These were calculated by treating each coding interval for each person as a unique case and estimating the chance-corrected association between raters for these "cases."

Concurrent Validity Analyses—Spearman rho correlations were calculated between number of coded pain behaviors (separately and totaled) and scores on measures of self-reported pain behavior, pain interference, average pain intensity (7 days), current pain intensity before observation protocol (0-10), and current pain intensity after observation protocol (0-10).

Results

Descriptive Statistics and Reliability Estimates—Videotapes of 30 participants with MS were coded by both coders. Table 1 reports demographic characteristics of the sample. The MS participants were predominately female (73.3%), non-Hispanic White (93.3%), and had experienced pain for a mean of more than 11 years. Table 2 reports the descriptive statistics and reliability estimates for positions, movements, and pain behaviors. There was substantial variability in the frequency with which different pain behaviors were reported both across and within participants. This was particularly true for guarding, which was coded 11 times more frequently than sighing and had an SD of 4.13. The finding is similar to results from a test/re-test study of individuals with non-chronic back pain (15). Guarding and rubbing were observed with nearly equal frequency in this study of 61 test participants (guarding mean = 4.82; rubbing mean = 4.21) and 24 re-test participants (guarding mean = 3.68; rubbing mean = 3.72) (15).

All reliability estimates for positions and movements were quite high (percent agreement ranged from 92.6% to 99.0%). Percent agreement for pain behaviors was also quite high, ranging from 88.6% (rubbing) to 97.2% (sighing). Because kappa coefficients take into account agreement resulting from chance, percentage *effective agreement* estimates were higher than kappa estimates. This was especially true for grimacing (effective agreement = 54.5%; kappa = 0.40) and bracing (effective agreement = 56.2%; kappa = 0.61). However, all kappa coefficients indicated at least "moderate agreement" (0.40-0.60) and others indicated "substantial agreement" (0.61-0.80) and higher (16). Kappa coefficients are

attenuated by low prevalence, so the lowest kappa coefficients for grimacing (0.40) and sighing (0.54) may have been an artifact of the infrequency of these behaviors.

Concurrent Validity Analyses—Table 3 reports results from the evaluations of concurrent validity. Total pain behaviors were correlated substantially with measures of pain intensity. Correlations with average pain, and current pain before and after the observation protocol, were 0.46, 0.46, and 0.53, respectively.

Total pain behaviors also correlated substantially with self-reports of pain behaviors (0.58). This is to be expected given that the self-report measure asked about similar pain behaviors as those coded for in the observation protocol (e.g., non-verbal and motor behaviors). Pain behavior also correlated significantly with self-reported pain interference (0.60). Self-reported pain interference is an indicator of the disabling impact of pain. The strong correlation indicates the relationship between pain-related disability and pain behavior.

Study 2

Study 2 compared the inter-rater reliability of the MS protocol to the inter-rater reliability in other populations. The current study is the first to evaluate the observation protocol in a sample of persons with MS. For this reason, any differences in reliability compared with previous studies could be attributed to the consistency of the particular coders for the current study or to differences in pain behavior presentation in MS. To compare the inter-rater reliability of the behavior observation protocol in MS to other pain-related conditions, the two coders for Study 1 both coded observation protocols for a combined sample of persons with back pain and arthritis. The purpose was to hold coders constant while varying the pain condition (MS versus other painful conditions).

Methods

Potential participants were recruited in the same manner as described for Study 1. The single exception was that, in addition to recruiting additional participants with MS, we also recruited a second sample of patients with a self-reported diagnosis of back pain or arthritis. All study participants were videotaped and coded using the observation protocol described above. Because the protocol has already been validated in back pain (6) and in arthritis (7), only a subsample was coded by both raters. Percent agreement, percent effective agreement and kappa coefficients were calculated for each of the five pain behavior categories.

Results

The videotapes of eight individuals with back pain and five with arthritis were coded by both coders. Inter-rater reliability for this combined group was similar to that obtained for those with MS. Whereas in the MS sample, percent agreement ranged from 88.6% to 97.2%, for the arthritis and back pain sample, the range was 89.0% to 98.8%. Percentage effective agreement was slightly higher in MS (54.5%-81.0%) than in the arthritis and back pain sample (40.7-67.4%). The range of kappa coefficients was wider in the MS sample (0.40-0.83) than in the arthritis and back pain sample (0.55-0.77).

Study 3

Study 3 was conducted to examine the construct validity of the observation protocol used with individuals who have MS. Naïve observers viewed videotapes and made global estimates of participants' pain levels. These estimates were then correlated with the total number of behaviors observed by one of the recorders (Keefe). Pain behaviors are hypothesized to communicate pain behaviors to others. The validity of this hypothesis (and the observation protocol) would be supported if a strong correlation was found between

naïve observers' estimates of pain and the standardized pain behavior counts generated in the observation protocol.

Methods

Participants—Four research associates who had not participated previously in pain behavior research studies were recruited to view and rate the pain of individuals with MS who participated in Study 1.

Procedures—The video segments were prepared by fast forwarding to the first time there was a walking segment and then including five additional minutes of video. This approach was used because walking generated the most observed pain behaviors. Each naïve rater recorded pain ratings independently and did not communicate with other raters. Instructions to raters explained that the videos "show people who may have pain." Raters watched each sample video and then immediately rated participants' current pain on a one to five scale corresponding, respectively, to "no pain," "mild," "moderate," "severe" and "very severe."

Analysis—Associations between mean pain behavior frequencies and naïve observers' mean ratings of current and average pain intensity were estimated as Spearman correlation coefficients. Standard deviations were calculated across observers for each individual within each pain scale (current pain and average pain intensity). The mean of these 30 SDs served as an estimate of intra-coder variability across participants who had MS.

Results

The associations between total pain behaviors observed by coders and mean pain ratings by naïve raters were quite high (Table 4). This was equally true for raters' perceptions of participants' average and current pain (0.86 in both cases). This finding is particularly important in establishing the construct validity of the observation protocol.

The highest correlations between naïve observer pain ratings and individual pain behaviors were with guarding (0.83 and 0.76 for average and current pain, respectively). The least highly correlated pain behavior was sighing (0.02 and 0.10). Sighing, however, also was the least likely behavior to be observed and this may have attenuated the calculated correlation. Both grimacing and bracing, however, although observed somewhat infrequently, were relatively strongly associated with naïve observers' ratings of pain.

Discussion

This is the first evaluation of the Keefe and Block pain observation protocol (6) in a sample of persons with MS. The results from the three studies supported the use of the observation protocol with individuals with MS.

The results of Study 1 and Study 2 indicated that the observation protocol was reliable when used with individuals who have MS. Because MS-related pain is often neurological in nature (17), we expected that there might be differences in the reliability of the observation protocol in individuals who had MS compared with those with other conditions and disease. The inter-rater reliability estimates obtained in Study 1 were somewhat lower than those obtained in published studies of persons with diagnoses other than MS. In our MS sample, the range of percentage effective agreement across the five categories of behavior was 55%-81%. This compares to ranges of 91%-100% in rheumatoid arthritis (7), 80%-96% in chronic low back pain (6), and 68%-87% in non-chronic back pain (15). Similarly, the range of kappa coefficients was divergent compared with published results in other populations. For our MS sample, pain behavior kappa coefficients ranged from 0.40 to 0.83. This

compares with ranges for other samples of 0.93-1.00 in rheumatoid arthritis (7), 0.53-0.79 in juvenile rheumatoid arthritis (18), and 0.68-0.93 in non-chronic back pain (19). These results taken alone would suggest that the protocol is less reliable in individuals who have MS. However, in Study 2, when the same investigators coded frequency of pain behaviors in a sample of individuals with arthritis and back pain, the results were more similar to those we obtained in the MS sample than those published in the literature for other clinical populations. Given the range of results reported in other conditions and the comparison values for our sample of individuals with arthritis and back pain, we concluded that the reliability of the protocol was comparable to that obtained in other clinical populations.

Our results confirmed previous findings that the five pain behavior categories vary substantially in frequency within and across diagnoses, suggesting that different pain behaviors may be characteristic of different clinical conditions. In a study of non-chronic back pain, guarding and rubbing were observed with nearly equal frequency in 61 participants at baseline (guarding mean = 4.82; rubbing mean = 4.21) and in 24 re-test participants (guarding mean = 3.68; rubbing mean = 3.72) (15). In studies of pain behavior in persons with low back pain (6) and rheumatoid arthritis (7, 20), rubbing was the most frequently occurring behavior (but with similar variability across patients). In a study of pain behaviors among individuals with lung cancer, bracing was the most frequently observed pain behavior (8). Few replication studies within a given condition have been conducted. Such studies are needed to distinguish between random variation in behavior frequency counts and diagnosis-related variation.

The high associations we obtained between total pain behavior counts and naïve observers' ratings of participants' pain levels (0.86) in Study 3 support the validity of the observation protocol in this clinical population. The correlation coefficients obtained in this study were higher than obtained in some other studies. Keefe and Block estimated a correlation of 0.69 between naïve observers' 0-10 pain rating and total observed pain behaviors (6). McDaniel and colleagues obtained a correlation of 0.55 when naïve observers used a visual analogue scale to rate participants' pain severity (7). The high correspondence between ratings of naïve observers and observed pain behaviors in the current study is important because pain behaviors are hypothesized to signal to others that an individual is experiencing pain. The results indicate that pain behaviors of persons who have MS do, in fact, communicate their subjective experiences of pain to others.

The current study had a number of limitations. Our sample was largely non-Hispanic White (93.3%) and well educated (53.5% with at least some college), reflecting the demographics of the Seattle area and the demographics of MS. The study completed here should be repeated in a more heterogeneous sample to gauge the generalizability of our results. In addition, although the pain observation protocol promotes reliability by creating a highly controlled environment for observing and coding pain behaviors, its applicability to more naturalistic settings should be evaluated. Some evidence for its validity in clinical settings was obtained in a study by Keefe et al. in which observations of pain behaviors were made in a clinical setting during physical exams.

Taken together, the results of the studies reported here support the use of the observation protocol for assessing pain behaviors of individuals with MS. As we have noted, future research should examine the validity of pain behavior observations in less controlled settings. The positive findings by Keefe et al. hint at the potential for such applications. The ability to validly assess the pain behaviors of individuals with MS in clinical settings could lead to creative interventions in the management of chronic pain in this population.

To clarify the role of pain behaviors in MS, future studies should sample pain behaviors of persons with MS in different contexts. Compared with orthopedic pain, neurologic pain is not as likely to be associated with particular triggers. Future research should examine differences in pain behaviors of individuals with MS who are experiencing orthopedic versus neuropathic pain. Such research also could indicate whether additional pain behavior codes are relevant for use in individuals with MS. As we have pointed out, critical to the validity of pain observation is the degree to which those pain behaviors communicate subjective experiences to others. Future research could address this question further by asking caregivers of and health professionals who treat individuals with MS to rate participants' pain levels after viewing videos of the observation protocol. This and other studies recommended here could clarify the nature of pain behaviors in MS and eventually lead to more effective treatments of the pain they experience.

Acknowledgments

The project described was supported by Award Number RC1NR011804 from the National Institute of Nursing Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Nursing Research or the National Institutes of Health.

References

1. Keefe FJ, Caldwell DS, Queen K, et al. Osteoarthritic knee pain: a behavioral analysis. Pain. 1987; 28:309–321. [PubMed: 2952935]
2. Fordyce WE, Brockway JA, Bergman JA, Spengler D. Acute back pain: a control-group comparison of behavioral vs traditional management methods. J Behav Med. 1986; 9:127–140. [PubMed: 2940370]
3. Martel MO, Thibault P, Sullivan MJ. Judgments about pain intensity and pain genuineness: the role of pain behavior and judgmental heuristics. J Pain. 2011; 12:468–475. [PubMed: 21296030]
4. Prkachin KM, Schultz IZ, Hughes E. Pain behavior and the development of pain-related disability: the importance of guarding. Clin J Pain. 2007; 23:270–277. [PubMed: 17314588]
5. Keefe FJ, Bradley LA. Behavioral and psychological approaches to the assessment and treatment of chronic pain. Gen Hosp Psychiatry. 1984; 6:49–54. [PubMed: 6693027]
6. Keefe FJ, Block AR. Development of an observation method for assessing pain behavior in chronic low back pain patients. Behavior Therapy. 1982; 13:363–375.
7. McDaniel LK, Anderson KO, Bradley LA, et al. Development of an observation method for assessing pain behavior in rheumatoid arthritis patients. Pain. 1986; 24:165–184. [PubMed: 3960569]
8. Wilkie DJ, Keefe FJ, Dodd MJ, Copp LA. Behavior of patients with lung cancer: description and associations with oncologic and pain variables. Pain. 1992; 51:231–240. [PubMed: 1484719]
9. Kerns RD, Kassirer M, Otis J. Pain in multiple sclerosis: a biopsychosocial perspective. J Rehabil Res Dev. 2002; 39:225–232. [PubMed: 12051466]
10. Michalski D, Liebig S, Thomae E, Hinz A, Bergh FT. Pain in patients with multiple sclerosis: a complex assessment including quantitative and qualitative measurements provides for a disease-related biopsychosocial pain model. J Pain Res. 2011; 4:219–225. [PubMed: 21887119]
11. Douglas C, Windsor C, Wollin J. Understanding chronic pain complicating disability: finding meaning through focus group methodology. J Neurosci Nurs. 2008; 40:158–168. [PubMed: 18578274]
12. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore. 1994; 23:129–138. [PubMed: 8080219]
13. Revicki DA, Chen WH, Harnam N, et al. Development and psychometric analysis of the PROMIS pain behavior item bank. Pain. 2009; 146:158–169. [PubMed: 19683873]
14. Amtmann D, Cook KF, Jensen MP, et al. Development of a PROMIS item bank to measure pain interference. Pain. 2010; 150:173–182. [PubMed: 20554116]

15. Jensen IB, Bradley LA, Linton SJ. Validation of an observation method of pain assessment in non-chronic back pain. Pain. 1989; 39:267–274. [PubMed: 2533339]

16. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

17. Grau-Lopez L, Sierra S, Martinez-Caceres E, Ramo-Tello C. Analysis of the pain in multiple sclerosis patients. Neurologia. 2011; 26:208–213. [PubMed: 21163193]

18. Jaworski TM. Juvenile rheumatoid arthritis: pain-related and psychosocial aspects and their relevance for assessment and treatment. Arthritis Care Res. 1993; 6:187–196. [PubMed: 7918714]

19. Jensen I. [The lawless elderly]. [in Danish]. Sygeplejersken. 1989; 89:13.

20. Anderson KO, Bradley LA, McDaniel LK, et al. The assessment of pain in rheumatoid arthritis. Validity of a behavioral observation method. Arthritis Rheum. 1987; 30:36–43. [PubMed: 3814196]

Table 1

Study 1 and Study 2 Demographic and Clinical Characteristics of Study 1 Sample

	Study 1: Multiple Sclerosis (N=30) *n* (%)	**Study 2: Arthritis (N=8) *n* (%)**	**Study 2: Back Pain (N=5) *n* (%)**
Age (yrs), mean±SD	53.3 ± 13.3	61.5 ± 13.3	50.8 ± 4.9
Pain duration (yrs), mean± SD	11.2 ± 8.3	5.4 ± 3.4	5.6 ± 3.0
Female Sex	22 (73.3)	8 (100.0)	3 (60.0)
Race			
Non-Hispanic White	28 (93.3)	5 (62.5)	3 (60.0)
Other or Mixed Race	1 (3.3)	3 (37.5)	2 (50.0)
Hispanic	1 (3.3)	-	-
Married/Living with Partner	23 (76.7)	0 (0.0)	1 (20.0)
Education			
≤ High School	3 (10.0)	2 (25.0)	3 (60.0)
Some College/AA	13 (43.3)	3 (37.5)	1 (20.0)
College Degree (BA/BS)	9 (30.0)	2 (25.0)	-
Advanced Degree	5 (16.7)	1 (12.5)	1 (20.0)
Employment			
Part-time	3 (10.0)	1 (12.5)	-
Full-time	9 (30.0)	1 (12.5)	2 (40.0)

MS = Multiple Sclerosis; Arth = Arthritis; BP = Back Pain

Table 2

Frequency Means, Measures of Variability, and Inter-Observer Reliability Estimates of Observed Pain Behaviors Demonstrated by Participants with Multiple Sclerosis

Behavior	Mean	SD	Minimum Frequency	Maximum Frequency	% Agreement	% Effective Agreement	Kappa
Position							
Sitting	6.21	0.86	4	8	97.7	93.5	0.95
Standing	6.31	1.00	4	8	98.3	95.0	0.96
Reclining	3.76	0.91	1	6	99.0	94.7	0.97
Movement							
Position change	6.93	1.51	4	9	97.4	93.4	0.94
Shifting	5.34	1.95	1	9	97.9	92.6	0.95
Pain Behaviors							
Guarding	5.62	4.13	0	12	93.6	81.0	0.83
Bracing	1.52	2.47	0	7	94.8	56.2	0.61
Rubbing	3.82	3.61	0	10	88.6	75.6	0.67
Grimacing	1.45	1.78	0	7	90.9	54.5	0.40
Sighing	0.52	0.99	0	4	97.2	71.2	0.54

Table 3

Spearman Rho Correlations Between Self-Reported Pain and Observed Pain Behaviors *(N=29)*

Behavior	Average Pain (0-10)		Pain Before Observation (0-10)		Pain After Observation (0-10)		Pain Interference (PROMIS-SF)		Pain Behaviors Self-Report (PROMIS-SF)	
	rho	*P*	rho	*P*	rho	*P*	rho	*P*	rho	*P*
Guarding	0.32	*0.09*	0.21	*0.28*	0.28	*0.15*	0.32	*0.09*	0.45	*0.01*
Bracing	0.38	*0.04*	0.35	*0.06*	0.57	*<0.01*	0.43	*0.02*	0.31	*0.10*
Rubbing	0.20	*0.30*	0.29	*0.13*	0.28	*0.15*	0.37	*0.05*	0.46	*0.01*
Grimacing	0.27	*0.16*	0.33	*0.08*	0.39	*0.04*	0.34	*0.07*	0.32	*0.09*
Sighing	0.46	*0.01*	0.54	*<0.01*	0.52	*0.01*	0.56	*<0.01*	0.33	*0.08*
Total	0.46	*0.01*	0.46	*0.02*	0.53	*<0.01*	0.60	*<0.01*	0.58	*<0.01*

PROMIS = Patient-Reported Outcome Measurement Information System; SF = Short Form.

Table 4

Spearman Correlations Between Naive Observers Pain Severity Estimates and Observed Pain Behaviors (VV=29)

Behavior	Average Pain (0-10)		Current Pain Intensity (1-5)	
	rho	*P*	rho	*P*
Guarding	0.83	*<0.01*	0.76	*<0.01*
Bracing	0.65	*<0.01*	0.64	*<0.01*
Rubbing	0.29	*0.13*	0.33	*0.08*
Grimacing	0.61	*<0.01*	0.65	*<0.01*
Sighing	0.02	*0.92*	0.10	*0.62*
Total	0.86	*<0.01*	0.86	*<0.01*