Psychometric Properties of a DSM-5-Based Screening Tool for Women's Perceptions of Premenstrual Symptoms

Meghan A. Richards and Kirsten A. Oinonen

Department of Psychology, Lakehead University

Paper accepted for publication in Psychological Reports:
https://doi.org/10.1177/0033294120979696.

Cite this article as: Richards, M., & Oinonen, K. A. (2021). Psychometric properties of a DSM-5-based screening tool for women's perceptions of premenstrual symptoms. *Psychological Reports,* 1- 32. https://doi.org/10.1177/0033294120979696

Author Note

Meghan A. Richards https://orcid.org/0000-0002-3358-9259)
Kirsten A. Oinonen https://orcid.org/0000-0003-0200-2993)
Meghan A. Richards is now at the Department of Psychology, University of New Brunswick.

We have no conflicts of interest to disclose.

This research was supported in part by a scholarship to the first author from the Canadian Institutes of Health Research. The authors also thank and acknowledge Dr. Michael Wesner, Psychology Department, Lakehead University, as a portion of this data was collected in his laboratory.

Correspondence concerning this article should be addressed to Dr. Kirsten Oinonen, Department of Psychology, Lakehead University, 955 Oliver Road Thunder Bay, Ontario, P7B 5E1 CANADA. Email: koinonen@lakeheadu.ca

Abstract

A premenstrual screening tool is needed when time constraints and attrition limit the feasibility of daily ratings. The present study examines the utility of a novel, 33-item, retrospective, dimensional, DSM-5-based, screening measure developed to explore women's perceptions of premenstrual symptomatology. This is the first measure that examines perception of impairment for each DSM-5 symptom and assesses the frequency criterion. Participants (N = 331) reported symptoms ranging from none to a level consistent with a provisional DSM-5 diagnosis of Premenstrual Dysphoric Disorder (PMDD). Initial psychometric properties indicated a five-factor structure: (1) affective symptoms; (2) fatigue, sleep, and anhedonia; (3) symptom frequency; (4) impairment and severity of appetite change and physical symptoms; and (5) difficulty concentrating. The total symptom scale and the frequency, severity, and impairment subscales demonstrated high internal consistency. Strong correlations between this dimensional measure and other retrospective and prospective premenstrual symptom measures suggest strong convergent, concurrent, and predictive validity. Premenstrual symptom groups created using this screening measure (minimal, mild/moderate, severe) differed on other retrospective and prospective measures of premenstrual symptoms. There was evidence of divergent validity and lack of an acquiescence bias. We also report data describing women's perceptions of the frequency, level of impairment, and level of severity for each DSM-5 PMDD symptom over a 12-month period and discuss implications for future research on premenstrual phenomenology. Initial evidence for the reliability and construct validity of this symptom screening measure suggests potential value for assessing premenstrual symptomatology in research and practice.

Keywords: premenstrual symptoms, Premenstrual Syndrome, PMDD, screening, psychometrics

Psychometric Properties of a DSM-5-Based Screening Tool for Women's Perceptions of Premenstrual Symptoms

Premenstrual syndrome (PMS) is characterized by the cyclical aggregation of physical, cognitive, and emotional symptoms that occur in the late luteal (premenstrual) phase of the menstrual cycle and remit or abate in the week post-menses. When such symptoms occur in most menstrual cycles; are associated with clinically significant distress or interference with work, school, usual social activities or relationships with others; and cannot be ascribed to another disorder, a DSM-5 diagnosis of Premenstrual Dysphoric Disorder (PMDD) is warranted (American Psychiatric Association (APA), 2013). While up to 80% of reproductive-aged women report experiencing some premenstrual symptoms, 30% to 40% of women report PMS symptoms that require treatment, and 3% to 8% of women report symptoms meeting criteria for PMDD (Ryu & Kim, 2015). Across several studies, prevalence rates between 10% and 35% have been reported for severe PMS (e.g., subthreshold PMDD) (Halbreich et al., 2003; Hylan et al.,1999; Johnson, 2004; Wittchen et al., 2002), with 20.3% of women in one study reporting missing at least one work day during the year due the experience of PMS symptoms (Robinson & Swindle, 2000). While valid prospective rating scales for assessing PMS and PMDD have been developed (Eisenlohr-Moul et al., 2017), there are practical constraints that prohibit many women from completing the required two months of prospective daily symptom monitoring. Further, the symptom frequency diagnostic criteria cannot be addressed with a prospective measure. There remains a need for the development of an accessible DSM-5-based retrospective screening tool to measure women’s experience of premenstrual symptoms in both research and practice.

The heterogeneity of tools for the assessment of PMS and PMDD includes measures that are both prospective (i.e., following/assessing current experiences over time) and retrospective

(i.e., assessment of past/current experiences), with a formal diagnosis requiring prospective charting for two consecutive cycles (Steiner et al., 2003). The most prevalent retrospective measures used to assess for the presence of PMS and/or PMDD include: The Menstrual Distress Questionnaire (MDQ) (Moos, 1968), which has been reported as one of the most well-cited questionnaires developed for the assessment of PMS symptoms (Haywood et al., 2002); the Premenstrual Assessment Form (PAF) (Halbreich et al., 1982); and the Premenstrual Symptoms Screening Tool (PSST) (Steiner et al., 2003). Prospective rating scales for the assessment of PMS symptoms include: the Daily Symptom Rating Scale (DSRS) (Taylor, 1979), the Calendar of Premenstrual Events (COPE) (Mortola et al., 1990), the Cyclicity Diagnoser Scale (CD) (Sundström et al., 1999), and the Daily Record of Severity of Problems (DRSP) (Endicott et al., 2006). None of these measures, however, align with the most recent criteria for PMDD as outlined in the DSM-5, potentially limiting the extent to which these criteria and women's experiences with the symptoms described within them can be investigated.

The current "gold-standard" for the measurement of PMS symptoms and PMDD is the Carolina Premenstrual Assessment Scoring System (C-PASS) (Eisenlohr-Moul et al., 2017). The C-PASS is a standardized scoring algorithm for making dimensional and categorical PMDD diagnoses using two or more months of daily DSRP symptom ratings. The scoring algorithm allows for the establishment of a categorical diagnosis (yes/no PMDD) and several dimensional measures of PMDD, including (1) Dimensional Cyclicity, (2) Dimensional Severity, and (3) Dimensional Frequency/Duration. A total score, individual symptom score, and impairment score as it relates to work, relationships, and/or hobbies can also be calculated for each dimensional index. The C-PASS is thus a robust measure that allows for prospective symptom tracking and the assessment of individual symptoms as dimensions.

The requirement of prospective reporting over two consecutive menstrual cycles is considered a diagnostic necessity for evaluating the presence of PMDD given research suggesting a historical lack of agreement between retrospective and prospective assessments of premenstrual symptoms (Rubinow et al., 1984; Eisenlohr-Moul et al., 2017). High levels of both PMS symptoms and PMS-related functional impairments have also been observed in a cross section of women completing retrospective reports compared to those who completed both prospective and retrospective reports (Cohen et al. 2002). Although retrospective and prospective ratings of affective PMS/PMDD symptoms are usually positively correlated, previous studies suggest that many women misdiagnose their chronic psychiatric symptoms as PMS (Rubinow & Roy-Byrne, 1984).

We are in agreement that prospective ratings are essential in establishing a diagnosis of PMDD but recognize that this gold standard is often not practical or feasible. For women in any of the following contexts, the reliable completion of two months of daily prospective ratings may be difficult or minimally possible: (1) women who suffer from symptoms that are significantly impairing, (2) women with extremely severe symptoms, (3) women with cognitive or executive dysfunction, and/or (4) women who are overwhelmingly or disproportionality affected by social or occupational demands that constrain time. Completion of prospective daily ratings is time consuming and requires intentionality, which may prevent certain women from committing to it. For example, Cohen et al. (2002) reported that 47% of a sample of 976 women screened into a study and asked to complete prospective daily ratings did not agree to do so. As well, the C-PASS validation study reported an attrition rate of 25% after women were invited to complete prospective daily ratings (Eisenlohr-Moul et al., 2017). Thus, while two months of prospective daily ratings are essential to maintaining diagnostic fidelity, it is likely that this practice limits

both research and practice. In research, important information relating to the constructs of PMS and PMDD is lost to attrition when women are unable to complete the ratings. In practice, some women may not receive a PMDD diagnosis or treatment due to the difficulty of the diagnostic requirements.

Despite studies citing a lack of agreement between retrospective and prospective premenstrual symptom ratings, we submit that a validated brief screening tool based on DSM-5 criteria is beneficial for three primary reasons. First, such a tool will identify potential women who might qualify for a provisional diagnosis of PMDD and would be good candidates to complete the more labor intensive two-month prospective rating process (with support if needed). Second, a retrospective measure is the only way to examine the DSM-5 criterion requiring that symptoms be present "for most menstrual cycles that occurred in the preceding year" (APA, 2013, p. 172). Third, the screening tool will help further research into the investigation of women's symptoms and their *perceptions* of the effect(s) that these symptoms have on their day-to-day functioning. From a psychological perspective, the explication of women's retrospective perceptions of their premenstrual symptoms and their effects can help us better understand the psychosocial evolution of these beliefs and their internalization by both individuals and, possibly, a wider culture. Clinically, this understanding can help tailor interventions targeting perceived premenstrual symptom impairment or interference in functioning regardless of whether women meet the diagnostic threshold for PMDD or their premenstrual symptoms exacerbate an underlying disorder such as depression (e.g., PME). This is an important consideration for psychological interventions as psychiatric disorders including major depressive disorder, anxiety disorders, and personality disorders have been reported as co-occurring frequently with PMS and PMDD (Hartlage & Gehlert, 2001). An increased

understanding of women's perceptions of premenstrual symptom type, symptom frequency, and symptom impairment can aid in the development of therapeutic tasks for women seeking treatment for premenstrual-related distress.

Here we propose a new short (e.g., under five minutes) retrospective measure of premenstrual symptoms based upon DSM-5 diagnostic criteria. We also examine the construct validity (i.e., convergent, criterion, and discriminant validity) of the proposed dimensional measure. For each DSM-5 PMDD symptom, we report retrospective data for a 12-month period which highlights women's perceptions of symptom frequency, degree of impairment, and severity. To our knowledge, this is the first study that reports women's perceptions of impairment in functioning for each premenstrual/PMDD symptom. With the exception of one item for which the wording was changed for ease of readability (see measure in the Appendix), the remaining items on this premenstrual symptom screening measure replicate the wording in DSM-5 criteria.

The assessment tool was developed to screen for premenstrual symptoms and provisional PMDD in response to the American Psychiatric Association's call for researchers to further investigate alternative items, thresholds, or durations associated with PMDD (APA, 2000). This new measure is in keeping with the dimensional model of mental illness adopted by the DSM-5 and is capable of measuring premenstrual symptoms along a continuum of severity ranging from an absence of symptoms at one end to the provisional identification of women with PMDD at the other. An advantage of taking a dimensional approach to the measurement of PMS symptoms is that it allows for further explication of subclinical symptoms potentially experienced by large numbers of women and could therefore facilitate an increased scientific and psychosocial understanding of a collective phenomenological experience.

The purpose of the present study was fourfold: 1) to present a retrospective dimensional screening tool for measuring premenstrual symptoms, 2) to provide preliminary psychometric data (e.g., reliability and construct validity data) on its utility as a clinical and research tool, 3) to evaluate the capacity of a dimensional screening tool to provide support for the validity of women's reported experience of premenstrual symptoms as defined by the DSM-5, and 4) to present data reflective of women's 12-month retrospective perception of symptom frequency and symptom impairment. The general construct validity of the proposed dimensional measure was assessed using principal components analysis (PCA). Additional analyses explored criterion (concurrent and predictive) validity, convergent validity, and discriminant validity. The following four predictions were made. (1) The new premenstrual measure will show strong positive correlations with the total premenstrual MDQ scale scores (convergent and concurrent criterion validity). (2) Symptom impairment groups created with the new measure will differ in MDQ premenstrual symptom scores (concurrent criterion validity). (3) Scores on the new tool will correlate positively with prospective lab measures of premenstrual pain, water retention, and negative affect (predictive criterion validity). (4) The new measure will show low or negative correlations with the control subscale of the MDQ, a subscale measuring non-cyclical symptoms (discriminant validity). Unlike previous retrospective measures, the proposed dimensional screening tool uses DSM-5 criteria and evaluates level of impairment for each individual DSM-5 PMDD symptom, rather than using one impairment rating for all symptoms. These impairment ratings may provide useful information regarding which symptoms present the greatest challenge for women's functioning in their day-to-day lives.

Method

Participants

Women were recruited from a primarily undergraduate Canadian university as part of a larger study investigating women's health and the influence of estradiol on visual system functioning. This study was conducted in two phases. Phase one involved the completion of a large screening questionnaire on women's health and reproductive functioning and phase two involved selecting a subset of women to participate in two laboratory sessions. Participants received bonus points in undergraduate psychology courses (phase 1 and 2) and a ten-dollar gift card (phase 2). The project was approved by the institutional research ethics board and informed consent was obtained from all participants.

Three-hundred and thirty-one women (mean age = 20.61 years, SD = 4.60 years) completed the screening questionnaire in phase 1. Of these, 14.5% (N = 48) self-reported a history of diagnosis or treatment for depression and 12.4% (N = 41) reported such history for an anxiety disorder. All women were included in the analyses given that the presence of underlying mental health conditions does not preclude a diagnosis of either PMS or PMDD and underlying hormonal changes could predispose women to both premenstrual symptoms and mood disorders (Hartlage & Gehlert, 2001).

A subsample of 29 women who were not pregnant or taking any form of hormonal contraceptives participated in the prospective symptom laboratory sessions (phase 2) of the study (mean age = 22.03, SD = 5.2). Of those women, half (N = 15) had retrospective MDQ scores for the premenstrual phase equal to or below the phase 1 sample median, and half (N = 14) scored above the phase 1 sample median.

Measures

Background Questionnaire (BQ)

The background questionnaire (BQ) was completed in phase 1 and included questions regarding: demographics (age and education), medical and mental health history, premenstrual symptoms, and reproductive history. Reproductive history was assessed with questions regarding age of menarche, parity, typical length of menstrual cycle, length of menses, time of last menstrual period, and cycle regularity. PMS symptoms were assessed using a 33-item dimensional screening measure of PMS symptom severity based on DSM-5 symptoms and the Menstrual Distress Questionnaire (MDQ) (Moos, 1968).

Screening Measure of Premenstrual Symptoms. A 33-item scale using the DSM-5 criteria for PMDD (APA, 2013) in a new format was used to screen for premenstrual symptom severity, frequency, and impairment. See the Appendix for the items and the question format. These items/criteria were established based on expert opinions and their analysis of available data in the area of premenstrual disorders and PMDD (Epperson et al., 2012). For each of the eleven sets of PMDD symptoms listed in the DSM-5, participants are asked three questions assessing: (1) the severity with which the symptoms are experienced; (2) the degree to which the symptoms impair work, school, or interpersonal performance or functioning; and (3) the frequency with which the symptoms were experienced over the past year. Questions assessing impairment and severity were rated using five-point likert scales anchored by 0 (*not at all*) on one end, and 4 (*extremely*) for impairment; and 0 (*not at all*) and 4 (*extremely severe/debilitating*) for severity. Frequency question response options ranged from 0 to 12 months, indicating how many menstrual cycles a particular symptom is experienced over a one-year period. From these scores, four premenstrual symptom scores can be calculated: frequency, severity, impairment, and total score. As described below, premenstrual symptom impairment

and severity groups were created based on these scores such that women were assigned to one of three symptom groups corresponding to: 0 (minimal; scores of 0 to 11), 1 (mild/moderate; scores ranging from 11 to 22), or 2 (severe; scores > 22).

Menstrual Distress Questionnaire (MDQ). The MDQ (Moos, 1968) consists of 47 items, each of which is rated on a five-point scale corresponding to 0 (*no experience of the symptom*), 1 (*present, mild*), 2 (*present, moderate*), 3 (*present, strong*) and 4 (*present, severe*) with respect to either current experiences (e.g., today) (MDQ – Form T) or experiences during three phases (menstrual, intermenstrual, premenstrual) of the most recent menstrual cycle (MDQ – Form C). The MDQ consists of eight scales derived from empirically distinct, although correlated, sets of symptoms: pain, concentration, water retention, behavior change, negative affect, autonomic reactions, arousal, and control. Of these scales, Moos (1968) reports that all except for arousal and control show large differences between premenstrual and intermenstrual phase scores. Examination of the factor structure of the MDQ provides evidence for the construct validity of all eight scales as effectively representing the structure of menstrual cycle symptoms (Ross et al., 2003). As well, exploratory factor analysis and LISREL analyses support the reliability and validity of the MDQ (Boyle, 1992). The MDQ – Form C was used in the Background Questionnaire.

Laboratory Questionnaire (LQ)

In phase 2, participants completed a brief laboratory questionnaire designed to assess their current affective state and level of cycle-related discomfort during two lab sessions. The four scales (noted below) were completed prospectively within the two laboratory sessions that were scheduled in a counterbalanced manner in both the late follicular (LF) and the premenstrual (PM) menstrual cycle phases and were used as prospective measures of premenstrual symptoms.

Positive and Negative Affect Scale (PANAS). The PANAS (Watson et al., 1988) consists of two 10-item scales for positive affect (PA) and negative affect (NA). Scale reliability is unaffected by the time period specified in the instructions and alphas range from .86 to .90 for PA and from .84 to .87 for NA (Watson et al., 1988). Convergent and divergent validity evidence comes from correlations with measures of anxiety, depression, and stress (Crawford & Henry, 2004). In the present study, both the PA and the NA scales were used as part of the prospective measurement of emotional premenstrual symptoms at two different lab sessions. NA is a general dimension reflecting subjective distress and unpleasant mood states, including anger, contempt, and disgust; while PA is a general dimension reflecting energy, concentration, and meaningful engagement. The NA scale includes items such as: *distressed*, *upset, and irritable;* and the PA scale includes items such as *interested, excited and enthusiastic.* Participants were asked to indicate how they felt at the moment they were completing the surveys in each lab session. They were asked to rate each adjective on a five-point response scale corresponding to 1 *(very slightly or not at all),* 2 *(a little),* 3 *(moderately),* 4 *(quite a bit),* and 5 (*extremely)*.

Menstrual Distress Questionnaire – Form T (MDQ-T). To prospectively assess menstrual cycle related pain and/or discomfort, two scales from the MDQ-T (see above) were included in the Laboratory Questionnaire to examine current experiences (e.g., today) (i.e., the pain and water retention scales). The scales were used as prospective measures of premenstrual symptomatology to examine the initial predictive criterion validity of the new PMS symptom screening measure. Markham (1976) reported that the MDQ-T is internally consistent with alphas ranging from .82 to .98.

Procedure

Phase 1

Participants were recruited for phase 1 of a study examining "visual functioning in women", which involved completing the Background Questionnaire online. All retrospective premenstrual symptom measures were collected at the beginning of the Background Questionnaire (i.e., immediately after the demographic questions), with the Screening Measure of Premenstrual Symptoms following the MDQ. Completion of these measures at the beginning of the questionnaire should minimize any carry-over effects from subsequent measures.

Phase 2

Participants were selected to participate in two laboratory sessions corresponding to two phases of the menstrual cycle when reproductive hormones differ; the LF phase when estradiol is increasing just before ovulation and the PM phase, when estradiol is decreasing. These sessions were scheduled as described below. The first task in each lab session involved completing the Laboratory Questionnaire which included four prospective measures of mood and premenstrual symptoms (i.e., PANAS PA and NA scales; and the MDQ-T pain and water retention scales). Completion of these measures at the start of the sessions reduced the possibility of any carry-over effects from the psychophysical tests completed later (for the larger study).

Scheduling of Laboratory Sessions

The mean laboratory testing days were days 14.48 (*SD* = 1.96) for the LF session and 25.03 (*SD* = 1.38) for the PM session (forward count). To ensure that women attended the laboratory during the appropriate menstrual cycle phases (i.e., late follicular and premenstrual phases), an initial meeting was set up prior to data collection in which participants provided their menstrual cycle dates using calendars and self-report information. This information was

supplemented by the provision of ovulation test strips. These strips are sensitive to the pre-ovulatory surge in luteinizing hormone (LH) and change color when the presence of the hormone is detected in urine. They were used to corroborate the timing of laboratory sessions in the LF phase of the menstrual cycle, when levels of estradiol are rising or high. Participants were instructed on the use of these strips and the interpretation of the results.

Sessions for the PM phase were scheduled in the initial orientation meeting based upon women's current position in the menstrual cycle and the estimate of their typical cycle length. If the timing of the orientation meeting corresponded to the latter half of a woman's menstrual cycle, she was scheduled for a first laboratory session in the PM phase based upon her cycle-length estimate. Women were subsequently asked to contact the researchers upon the start of their next menstrual period, at which point they were instructed to begin using the ovulation test strips in the post-menstrual week to schedule a second laboratory session following a positive result. The scheduling of laboratory sessions was counterbalanced. Confirmation of next menstruation allowed cycle day at the time of laboratory testing to be established using the backwards count method. These days corresponded to -10 to -20 for the LF phase and -1 to -8 for the PM phase.

Premenstrual Symptom Groups

Scores on the PMS symptom screening measure were considered in combination with DSM-5 criteria for PMDD (APA, 2013) to categorize the severity of PMS symptoms (minimal, mild/moderate, and severe), PMS impairment level (minimal, mild/moderate, and severe), and to provide a provisional diagnosis of PMDD (yes, no). Women's responses on the PMS symptom screener were coded, as described below, based on whether they satisfied PMDD (provisional) criteria A through D in the DSM-5.

Criterion A (APA, 2013) reflects symptom frequency and was established in three steps. First, each of the 11 PMDD symptom sets were recoded as "1" if a woman experienced them in seven or more menstrual cycles (i.e., more menstrual cycles than not), and "0" if experienced in six or fewer menstrual cycles in a one-year period. Second, these 11 new dichotomous symptom frequency variables were summed (possible range of 0 to 11). Third, women whose scores were five or higher, reflecting the experience of five or more symptoms in more menstrual cycles than not in a one-year period, met the threshold for criterion A.

Meeting criteria B and C for DSM-5 PMDD (APA, 2013) requires that one of four symptom sets be present from cluster B (items 1 to 4), one of seven symptom sets be present from cluster C (items 5 to 11), and that five symptom sets be present across the two clusters (see the Appendix). The requirement that there be five symptoms across the two clusters was already examined within criteria A. To establish the presence of criteria B and C, the 11 dummy variables created to examine criteria A for the 11 PMS symptoms (e.g., symptoms occurring in seven or more menstrual cycles) were used. Women were classified as meeting criteria B and C if they endorsed at least one symptom set from each of these two separate symptom groups and the cumulative total symptom sets across clusters B and C was five or greater.

Criterion D was established in two different ways using women's ratings for both (a) impairment and (b) severity. Thus, for the purpose of the present study, women were said to meet Criterion D if they met cutoffs for *either* impairment or severity. For both impairment and severity, women were assigned to one of three groups to reflect level of premenstrual symptom severity or impairment: 0 (minimal), 1 (mild/moderate), or 2 (severe). The range of scores corresponding to each of these categories was determined by considering the maximal score a woman would receive if she were to endorse all symptoms as being equally problematic. For

example, endorsing each of the 11 symptoms with a rating of 0 or 1, corresponding to a negligible or mild level of impairment or severity, would result in a score between 0 and 11. Scores within this range were therefore classified as minimal and subsumed a group of women who did not meet criteria for PMDD. This group included women who: (a) did not endorse symptoms as leading to impairment or as being severe, and (b) did not meet criteria A, B, or C but who may have had a score for impairment or severity exceeding 11. Women were classified as experiencing mild to moderate PMDD if they scored between 12 and 22, which required rating (at minimum) all symptoms with a value of mild and at least one symptom with a rating of moderate in terms of impairment or severity. Finally, women were classified as experiencing severe PMDD if their scores for either impairment or severity exceeded 23, which would require endorsing at least all 11 symptoms as moderate and one as severe. Thus, women whose scores were greater than 11 were considered to have met criterion D. Using scores for both impairment and severity, premenstrual symptom classification groups were based on the following scores: 1 to 11 (minimal), 12 to 22 (mild/moderate), and 23 to 44 (severe).

Results

Principle Component Analyses

To determine the underlying dimensions of our screening tool for provisional PMDD and PMS symptoms, an exploratory factor analysis (EFA) was first performed on the 33-items of the PMS symptom screening measure. Horn's Parallel Analysis is a procedure recommended to determine the number of factors to extract (Tabachnick & Fidell, 2014), and was used as described by O'Connor (2000). Parallel analysis suggested the retention of five factors. The principal components analysis (PCA) using oblique (oblimin) rotation indicated that the five factors accounted for 66.11% of the variance. Oblique rotation was selected given the high

potential for factors within our model (e.g., assessment tool) to correlate. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis, KMO = .90, and all KMO values for individual items were > .81, which is above the acceptable limit of .5 (Field, 2009). Bartlett's test of sphericity χ^2 (528) = 9851.13, $p < .001$, indicated that correlations between items were sufficiently large for PCA.

Table 1 shows the factor loadings for the primary five components after rotation, including eigenvalues, variance, and reliability (e.g., Cronbach's α). The items that cluster on these components suggest that component one represents affective symptoms and accounts for 41.80% of the variance; component two is reflective of sleep disruption, fatigue, and decreased interest, and accounts for 8.02% of the variance; component three represents frequency for five symptoms and accounts for 6.81% of the variance; component four represents impairment and severity of appetite change and physical symptoms and accounts for 5.08% of the variance; and component five represents difficulty concentrating and accounts for 4.40% of the variance. Cronbach's alpha for the five components ranged from .62 to .85.

Reliability and Internal Consistency of the Overall Scale and Subscales

Three exploratory PCAs with oblique rotation were then performed for each of the 11-item frequency, impairment, and severity scales. A summary of these analyses as well as means, SDs, and internal consistency values are found in Table 2. These analyses indicated that each scale is principally comprised of one component accounting for between 47.98% and 49.74% of the variance, with internal consistencies ranging from .89 to .90. The overall PMS symptom screening measure (all 33 items) also had high internal consistency, with an overall Cronbach alpha coefficient of .92.

Convergent and Criterion Validity (Concurrent and Predictive)

Convergent validity for the retrospective PMS symptom screener was evaluated by correlating the total item score and individual frequency, impairment, and severity scale scores with the retrospective premenstrual phase MDQ total score and individual scale scores. The MDQ total score was calculated using seven of the eight MDQ scales, omitting the arousal scale, as the items within this scale are not consistent with the experience of premenstrual symptoms (e.g., well-being) as outlined in the DSM-5. For example, (Boyle, 1992) reported that the MDQ arousal scale contains items associated with *positive reactions* across the menstrual cycle and that the scale did not correlate significantly with any other factors derived from the MDQ. The results are presented in the top section of Table 3. The correlations are high, supporting the convergent validity for the PMS symptom screening measure. The correlation between the screening measure's total score and the overall MDQ premenstrual phase score was $r = .70$, $p < .001$, providing evidence of convergent validity for the screening tool.

The three individual subscale scores and the total score of the PMS symptom screening tool were all positively correlated with prospective lab measures of pain, water retention, and NA during the premenstrual phase (see bottom portion of Table 3). The correlations were generally above .50, with total scores best predicting premenstrual phase PANAS NA scores, $r = .70$, $p < .001$. These correlations support the predictive criterion validity of the new screening tool. Additionally, with the exception of PA, the individual scales of the PMS symptom screening tool were all positively associated with the prospective lab change scores, which were calculated by subtracting prospective late-follicular phase scores from premenstrual phase scores. These correlations were strongest for NA and suggest that women with more premenstrual NA symptoms relative to symptoms during the late-follicular phase, have higher scores on the PMS

symptom screening measure. These correlations provide support for the predictive criterion validity of the PMS symptom screening measure.

To further examine the concurrent and predictive criterion validity of the dimensional PMS symptom measure, the symptom impairment classification groups were compared on the other retrospective and prospective premenstrual phase symptom scores (see Table 4). As results were similar for the symptom severity classification groups, these are not presented here (see Supplementary Table 1). These groups represent increasing levels of PMS symptomatology: 0 (*minimal*), 1 (*mild/moderate*), and 2 (*severe*).

Two three-group MANOVAs using Pillai's trace examined whether the three impairment groups created with the PMS symptom screening tool differ in PMS symptoms using the following dependent variables: (1) retrospective premenstrual phase MDQ data from the total sample of women, and (2) prospective laboratory data from the subsample of women in the premenstrual phase of their menstrual cycle. Given unequal variance between group means, the Games-Howell statistic was used for post-hoc analyses. The first MANOVA found that women's scores on the seven retrospective MDQ scales during the premenstrual phase differed as a function of the impairment groups, $V = .55$, $F(14, 636) = 17.41$, $p \leq .001$, $\eta^2 = .28$. Follow-up univariate ANOVAs revealed significant differences between the impairment groups on each MDQ scale in the predicted directions (all $p \leq .001$; see top of Table 4). It is noteworthy that the largest effect sizes were for the behavior change, impaired concentration, and negative affect MDQ scales, suggesting that the impairment group categories are strongly associated with differences in behavior, concentration, and mood. The second MANOVA indicated that the impairment groups differed on prospective premenstrual phase symptoms reported in the lab session, $V = .79$, $F(6, 50) = 5.05$, $p \leq .001$, $\eta^2 = .38$. Follow-up univariate ANOVAs revealed

significant differences between the impairment groups on all prospective measures of premenstrual symptoms (all $p \leq .004$) with the means indicating that the severe premenstrual symptom impairment group endorsed the highest level of prospective symptom expression for all variables, and the minimal group endorsed the lowest level of symptom expression for all variables (see bottom portion of Table 4). The largest effect size was found for the PANAS NA scale suggesting that the premenstrual symptom impairment groups are strongly differentiated in terms of negative emotional symptoms.

For both MANOVAs, Bartlett's Test of Sphericity and Box's Test of Equality of Covariances were significant. Thus, the non-parametric Kruskal-Wallis Test was used as a check to compare means for prospective and retrospective DVs as a function of group. With the exception of group means for prospective MDQ pain ($p = .08$), which was a non-significant trend for both impairment and severity groups, Kruskal-Wallis tests comparing means between groups across all prospective and retrospective DVs were all significant (all $p < .001$ for retrospective ratings and $p < .05$ for prospective ratings). These analyses suggest that the PMS symptom screening tool categorizes women into valid groups that differ in premenstrual symptoms as measured by prospective and other retrospective instruments.

To further examine whether impairment and severity grouping scores (0, 1, 2) from the new PMS symptom screening measure could be predicted using retrospective premenstrual symptom ratings (i.e., MDQ scale scores) and prospective premenstrual rating scores (i.e., two MDQ scales and the PANAS NA scale), four multiple regressions were performed. As in the above MANOVAs, separate regressions were completed on retrospective and prospective premenstrual scores for impairment grouping scores and severity grouping scores. A forced entry model was employed in which all retrospective or prospective premenstrual symptom predictors

of grouping scores were entered into the model simultaneously. For each of these analyses, the Durban-Watson statistic was selected as a test of the assumption of independent error as per Field (2009) and was returned within acceptable limits (range 1.83 to 2.12).

A summary of the two regressions predicting impairment grouping scores is found in Table 5. For the impairment groups, the seven *retrospective* MDQ scale scores predicted group membership scores, F (7, 325) = 31.91 p <.001, adjusted R^2 = .40. The four unique predictors of symptom impairment grouping scores were the NA (p < .001), impaired concentration (p = .01), behavior change (p < .001), and control (p = .005) scales. Higher impairment grouping scores were associated with higher scores on the NA (β = .28), impaired concentration (β = .20), and behavior change (β = .32) scales; and lower scores on the MDQ control scale (β = - .18). Given Moos' (1968) addition of the control scale as a means of identifying women who endorse high levels of symptoms that are not typically associated with the menstrual cycle, the inverse relationship between the MDQ control scales and impairment group scores provides evidence of divergent/discriminant validity for the new premenstrual symptom screener and for the impairment groups.

The second regression using the three *prospective* symptom measures collected during the premenstrual phase of the cycle (N = 29) as predictors of symptom impairment level was also significant, F(3, 28) = 16.22 p <.001, adjusted R^2 = .62. These three prospective premenstrual measures explained 62% of the variance in the impairment grouping scores, which was higher than the 40% explained by the seven retrospective premenstrual symptom measures. Two of the prospective measures were unique predictors of PMS symptom impairment: MDQ water retention (p = .01) and PANAS NA (p = .002). Being screened into a higher symptom impairment group was associated with higher water retention and NA scores (see Table 5).

Given that the regression results were very similar for the symptom severity groups, a summary of the two regressions is provided in Supplementary Table 2.

Discriminant Validity

Some preliminary evidence for the discriminant validity of the new PMS symptom screening measure comes from two findings. First, the regression coefficients (e.g., *B*s) were negative and significant for the retrospective MDQ control scale scores in the prediction of group membership for both PMS symptom impairment and PMS symptom severity (see Table 5 and Supplementary Table 2). This suggests that women who score higher on our dimensional PMS symptom screening tool score lower on the MDQ control scale. Second, correlations between the individual scales of our PMS symptom screening dimensional tool and prospective change in PA were not significant, suggesting that scores on our PMS symptom screening tool are not associated with cyclical change in PA (see Table 3).

To further evaluate the discriminant validity of the PMS symptom screening tool, three partial correlations were performed between retrospective MDQ control scores and each of: *1)* overall scores on the PMS symptom screener, *2)* the PMS screener impairment grouping scores, and *3)* the PMS screener severity grouping scores; while controlling for scores on the other six retrospective MDQ scales. There were negative relationships between MDQ control scores and: overall scores on the PMS symptom screening measure (partial $r = -.19$, $p < .001$), impairment grouping scores (partial $r = -.16$, $p = .005$), and severity grouping scores (partial $r = -.21$, $p < .001$). The MDQ control scale contains items that are not typically associated with the experience of PMS or PMDD. The finding of an inverse association after statistically controlling for responses to the theoretically relevant symptom scales can be regarded as a potential check against acquiescence or over-endorsing response styles. This finding suggests that women who

endorse more symptoms on the new PMS symptom screening measure do not over-endorse all symptoms and that women with more PMS symptoms are, in fact, less likely than others to endorse control symptoms.

Frequency, Severity, and Impairment of Premenstrual Symptoms

Given the above preliminary evidence of the reliability and construct validity of this screening tool, we used it to examine women's experience of premenstrual symptom frequency, severity, and impairment. Means and standard deviations for women's perceptions of the frequency, degree of impairment, and level of severity for each of the eleven symptoms of DSM-5 premenstrual dysphoric disorder over a retrospective 12-month period are reported in Table 6. These results suggest that, over a retrospective one-year period, women perceive that they experience the physical symptoms with the highest mean frequency [7.91 months (SD = 4.44)], degree of impairment, and severity. Correlations between severity ratings and impairment ratings for each symptom are high and significant, (see Table 7), suggesting that for all symptoms, measurement of both severity and impairment may be redundant.

Paired *t*-tests were used to compare mean symptom ratings to determine which symptoms women perceived as being highest and lowest in frequency, severity, and impairment, on average, and which symptom means differed ($p < .05$). In terms of symptom frequency (number of months in the past year), women reported the following symptoms as occurring most to least frequently on average: (1) physical symptoms, (2) appetite change, (3) fatigue = sleep changes = affective lability, (4) anger/irritability, (5) depressed mood = anxiety, (6) decreased interest = feeling overwhelmed, and (7) decreased concentration. With respect to symptom impairment, women reported the following symptoms as being most to least impairing on average: (1) fatigue = physical symptoms, (2) physical symptoms = sleep changes, (3) affective lability =

anger/irritability, (4) anger/irritability = decreased interest = depressed mood = anxiety = appetite change, and (5) appetite change = feeling overwhelmed = decreased concentration. Finally, for symptom severity, the symptoms that were reported as being most to least severe on average were: (1) physical symptoms, (2) appetite changes = sleep changes = fatigue = affective lability, (3) affective lability = anger/irritability, (4) depressed mood = anxiety = feeling overwhelmed, (5) feeling overwhelmed = decreased interest, and (6) decreased concentration.

Table 8 presents the percentage of women reporting each level of impairment for each DSM-5 PMDD symptom. The symptoms that women most frequently reported as causing some level of impairment were physical symptoms (67%), fatigue and lack of energy (64%), and sleep problems (59%). However, some impairment related to each of the emotional symptoms (Criteria B) was reported by about half of the women surveyed (46% to 54%). The symptom that was least likely to be reported as causing impairment was concentration difficulties, with 41% of women reporting at least some impairment. The symptoms that appeared to result in the highest levels of impairment were fatigue and lack of energy with the following levels of impairment noted: 18.5% moderate impairment, 10.6% quite a bit of impairment, and 4.6% extreme impairment. Supplementary Table 3 provides the same information for women's ratings of PMS symptom severity.

Discussion

The purpose of the present study was fourfold: (1) to present a retrospective dimensional screening tool for measuring premenstrual symptoms; (2) to provide preliminary psychometric data (e.g. reliability and construct validity data) on its utility as a clinical and research tool for studying premenstrual symptoms and PMDD; (3) to evaluate the capacity of this dimensional screening tool to provide support for the validity of women's reported experience of DSM-5

premenstrual symptoms; and (4) to use this tool to report women's perceptions as they pertain to the frequency, impairment, and severity of the 11 DSM-5 PMDD symptoms over a 12-month retrospective period. This study lends support to the dimensionality of PMS phenomenology as a whole and, to our knowledge, is the first to report women's retrospective perceptions of their experience with each DSM-5 premenstrual symptom, especially as they relate to frequency and impairment.

A 5-Factor Structure with Internal Consistency and Convergent Validity

The retrospective dimensional screening tool for DSM-5 PMS symptoms is presented in the Appendix. Factor analysis of this new dimensional measure indicates that items converge around five primary clusters reflecting: 1) affective symptoms, 2) sleep disruption, fatigue, and anhedonia, 3) symptom frequency, 4) impairment and severity of appetite and physical symptoms, and 5) concentration difficulties. PCAs of the frequency, impairment, and severity subscales indicate that each subscale is comprised of one primary internally consistent factor. Overall, the entire 33-item measure shows excellent internal consistency.

Correlations between total and individual scale scores of the new PMS symptom screening measure and total and individual scale scores of the retrospective MDQ for items rated during the premenstrual phase suggest that the constructs assessed by these measures are strongly related. Individual correlations between each of the three, dimensional scales and the seven MDQ scales were generally large and highly significant ($p \leq .01$) indicating good convergent validity between the two measures and supporting the capacity of dimensional measurement of the DSM-5 symptoms to capture the essence of premenstrual phenomenology.

Evidence of Predictive and Concurrent Criterion Validity

Additional support for predictive criterion validity of the PMS symptom screening measure was determined using two methods. First, correlations between scores on the PMS symptom screening measure and prospective premenstrual MDQ pain, MDQ water retention, and PANAS NA scale scores (and change scores) collected during two sessions in the lab were highly significant and positive. This suggested good agreement between the new PMS symptom screening measure and prospective symptoms during the premenstrual phase of the menstrual cycle when PMS symptoms should be highest and provides initial evidence for predictive validity.

Second, PMS symptom impairment and PMS symptom severity groups created with the new PMS symptom screening measure (i.e., minimal, mild/moderate, and severe) differed with both retrospective PMS symptom reports and prospective PMS symptom reports, providing evidence for criterion (both concurrent and predictive) validity. Women who were grouped according to their degree of PMS symptom impairment and PMS symptom severity on the screening measure differed significantly in (1) all seven retrospective MDQ scales supporting the dimensional measure's concurrent validity and (2) their prospective premenstrual phase scores on two MDQ scales and the PANAS NA scale, supporting the measure's predictive validity. That is, groups of women with higher PMS symptom impairment and higher PMS symptom severity scored higher on both prospective and retrospective PMS measures.

Multiple regression analyses revealed that, together, retrospective MDQ scale scores explained 41.3% of variance in the PMS screener symptom impairment groupings. Retrospective scores on the MDQ NA, impaired concentration, behavior change, and control scales were unique predictors. The strongest unique effects were for behavior change, NA, and impaired

concentration, suggesting that the premenstrual symptom groups created using the new PMS symptom screening measure maximally capture group differences in these areas of functioning. Interestingly, pain was not a unique predictor.

Impairment groups created with the PMS symptom screening measure were even more strongly predicted by prospective measures of symptoms. Together, scores on the three prospective premenstrual phase measures (pain, water retention, and NA scales) predicted 66.1% of the variance in the groupings for PMS symptom impairment. NA and water retention, but not pain, were unique predictors of group membership for PMS symptom impairment, suggesting that the new PMS symptom screening measure explains variance in emotional symptoms and water retention that are distinct from pain. This is not surprising given the focus on emotional symptoms in the DSM-5 criteria (APA, 2013). Support for the construct validity (i.e., concurrent and predictive criterion validity) of both the dimensionality of PMS symptoms and the grouping criteria of the new measure comes from the fact that group membership based upon increasing levels of impairment and severity using the dimensional tool was predicted using prospective and retrospective MDQ and PANAS data .

Evidence of Discriminant Validity

The finding that the retrospective MDQ control scale was negatively and significantly related to prediction of group membership in the regressions and partial correlations indicates that women who endorse more symptoms on the control scale have lower scores on the PMS symptom screening tool. This finding suggests that women endorsing higher levels of PMS symptoms within DSM-5 criteria, are actually less likely to endorse other symptoms that are not commonly reported to be a part of PMS or PMDD. The MDQ Control scale is populated with items that are not consistent with the experience of PMS but rather, are more consistent with the

experience of other, non-descript, neurological symptoms (e.g., ringing in the ears, heart pounding, and blind spots or fuzzy vision). The negative relationship between women's scores on this new PMS symptom screening tool and MDQ control suggests that: (a) women who score high on the new dimensional measure do not do so because they adopt an acquiescent response style (e.g., a bias towards endorsing all items) when asked to report their experience of PMS symptoms; and (b) women who report a high level of PMS symptomatology are not simply women who tend to complain about everything, or are necessarily high in somatization. The finding lends validity to the overall phenomenology of premenstrual symptoms and premenstrual dysphoria, and provides evidence for the discriminant validity of the PMS symptom screening measure as an assessment tool capable of capturing women's experience of PMS symptomatology. The finding also provides discriminant validity evidence for the DSM-5 symptom set as a measure of premenstrual symptoms and PMDD.

Women's Experiences of Symptom Frequency, Severity, and Impairment

Given preliminary evidence supporting the reliability and construct validity of this dimensional measure, data reflecting women's perceptions of the frequency, impairment and severity of each DSM-5 PMDD symptom were examined. Regarding women's perception of impairment, these data indicate that for each of the eleven DSM-5 PMDD symptoms, between 16% and 33% of women report experiencing a moderate degree of impairment or higher during at least one menstrual cycle over a twelve-month retrospective period. Furthermore, the low and high ends of this range are anchored by: (a) subjective difficulty in concentration and (b) lethargy, easy fatigability, or marked lack of energy, respectively. Across symptoms, a high percentage of women also endorse a level of impairment of moderate or greater for physical symptoms (32.1%) and sleep disruption (30.3%). In terms of comparing women's experience of

the frequency, severity, and impairment of the 11 DSM-5 symptoms; physical symptoms and decreased concentration stand out as the symptoms reported as the highest and lowest on average, respectively, in terms of frequency, severity, and impairment. However, separately measuring the various physical symptoms might lead to different findings.

Notwithstanding the importance of diagnosing PMDD, which is necessary in seeking appropriate treatments and studying its underlying biological mechanisms, it is important to consider the manifestation of women's perceptions of premenstrual symptoms to increase our psychosocial understanding of the construct of PMS and its effect on women's phenomenology. Our data suggest that a high number of women perceive that they are impaired by a number of symptoms associated with a regularly occurring cyclical event. The implication of this perceived burden warrants further investigation to improve our understanding of women's experiences and overall health care.

Strengths and Limitations

While the new PMS symptom screening tool yielded favourable psychometric properties, there are a few limitations to the present study. First, to diagnose PMDD, the DSM-5 stipulates in Criterion F that the temporal occurrence and frequency of symptoms (Criterion A) requires confirmation using prospective daily ratings over two or more cycles, which was not done here. Instead, prospective ratings on the PANAS and MDQ during the LF and PM phases of the menstrual cycle were used as metrics to show change across the menstrual cycle and to validate retrospective ratings. One noteworthy strength of these prospective ratings is that they were both performed in the lab under consistent controlled conditions at a specified date and time, which enhances their validity beyond at-home rating conditions. Future studies using this instrument should use two cycles of daily ratings to examine the ability of this measure to predict PMDD

diagnoses and PMS symptoms with that timeframe. Nevertheless, we submit that the preliminary psychometric properties of this new PMS symptom screening tool suggest that it could be useful in research and to clinically screen for women likely to meet the criteria for PMDD. There may be value in adapting the present scale to an alternative format that could also be used prospectively to evaluate the presence of PMDD symptoms in subsequent studies. Considering the study limitations, future research using this measure should examine test-retest reliability and focus on additional validation studies using a greater variety of women followed over a greater number of menstrual cycles. This would help provide additional evidence of predictive validity and discriminant validity.

Second, with the exception of item nine on our measure, which was adapted to read *sleeping too much or too little* rather than "hypersomnia or insomnia" the remainder of the items were derived from the DSM-5 and the face validity or content validity of these items for a population of university aged women is unclear. Future iterations of this instrument should take into account the literacy level and interpretability of item content for use by different populations of women. It is worthwhile noting, however, that the scales for our PMS symptom screening tool correlate well with the MDQ scales (Table 3), which has historically been widely used and validated, suggesting that at minimum, women were able to infer the content of the items as they pertain to the menstrual cycle. Future studies should also counterbalance the presentation order of the PMS symptom screening tool with any other PMS measure, as that was not done here.

The proposed PMS symptom screening tool has several advantages. It is based on current DSM-5 criteria and is capable of distinguishing women who meet screening criteria for DSM-5 Premenstrual Dysphoric Disorder from the majority of women who experience premenstrual symptoms as non-clinical experiences. Given that the tool is designed to take dimensionality into

account, it would be useful in research settings for those investigating premenstrual symptoms alone or their interactions with other variables of interest. Moreover, given their high internal consistency and content validity, individual scales (i.e., symptom frequency, impairment, or severity) could be employed as standalone measures.

Conclusion

The present study provides preliminary evidence that the proposed screening measure for premenstrual symptoms is a psychometrically sound instrument that: (a) assesses the continuum of premenstrual symptoms from none/mild to severe using either impairment or severity ratings and (b) screens for a provisional diagnosis of PMDD using the most recently published DSM criteria. The retrospective data yielded by this measure shows excellent potential for advancing the study of women's premenstrual symptom experience. In particular, this instrument suggests that there is benefit in studying women's perceptions of impairments in functioning associated with premenstrual disorders, how these impairments translate to their day-to-day experiences (e.g., the functional domains in which women experience the most impairment), and how information from women's ratings of individual symptoms could be used to tailor clinical interventions for their unique symptom presentation

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text revision)*. Author. https://doi.org/10.1176/appi.books.9780890420188.dsm-iii-r

Boyle, G. J. (1992) Factor Structure of the menstrual distress questionnaire (MDQ): Exploratory and LISREL analyses. *Personality and Individual Differences, 13*(1), 1-15. https://doi.org/10.1016/0191-8869(92)90211-7

Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology, 43(Pt 3),* 245–265. https://doi.org/10.1348/0144665031752934

Eisenlohr-Moul, T. A., Girdler, S. S., Schmalenberger, K. M., Dawson, D. N., Surana, P., Johnson, J. L., & Rubinow, D. R. (2017). Toward the reliable diagnosis of DSM-5 Premenstrual Dysphoric Disorder: The Carolina premenstrual assessment scoring system (C-PASS). *American Journal of Psychiatry, 174*, 51-59. doi:10.1176/appi.ajp.2016.15121510

Endicott, J., Nee, J., & Harrison, W. (2006). Daily Record of Severity of Problems (DRSP): reliability and validity. *Archives of Women's Mental Health, 9,* 41-49. doi: 10.1007/s00737-005-0103-y

Epperson, C. N., Steiner, M., Hartlage, S. A., Eriksson, E., Schmidt, P. J., Jones, I., & Yonkers, A. (2012) Premenstrual dysphoric disorder: Evidence for a new category for DSM-5. *American Journal of Psychiatry, 169*(5), 465-475. doi: 10.1176/appi.ajp.2012.11081302

Field, A. (2009). *Uncovering Statistics using SPSS (3rd ed.).* SAGE Publications Ltd.

Halbreich, U., Borenstein, J., Pearlstein, T., & Kahn, L. S. (2003). The prevalence, impairment, impact, and burden of premenstrual dysphoric disorder (PMS/PMDD). *Psychoneuroendocrinology, 28*, 1-23. doi:10.1016/S0306-4530(03)00098-2

Halbreich, U., Endicott, J., Schacht, S., & Nee, J. (1982). The diversity of premenstrual changes as reflected in the Premenstrual Assessment Form. *Acta Psychiatrica Scandanavica, 65,* 46-65. doi: 10.1111/j.1600-0447.1982.tb00820.x

Hartlage, S. A. & Gehlert, S. (2001). Differentiating premenstrual dysphoric disorder from premenstrual exacerbations of other disorders: A methods dilemma. *Clinical Psychology: Science and Practice*, *8*(2), 242–253. https://doi.org/10.1093/clipsy.8.2.242

Haywood, A., Slayde, P., & King, H. (2002). Assessing the assessment tools for menstrual cycle symptoms: A guide for researchers and clinicians. *Journal of Psychosomatic Research, 52,* 223-237. doi: 10.1016/s0022-3999(02)00297-0

Hylan, T. R., Sundell, K., & Judge, R. (1999). The impact of premenstrual symptomatology on functioning and treatment-seeking behavior: Experience from the United States, United Kingdom, and France. *Journal of Women's Health and Gender-Based Medicine, 8,* 1043-1052. doi:10.1089/jwh.1.1999.8.1043

Johnson, S. R. (2004). Premenstrual syndrome, premenstrual dysphoric disorder, and beyond: A clinical primer for practitioners. *Obstetrics and Gynecology, 104,* 845-859. doi:10.1097/01.AOG.0000140686.66212.1e

Markum, R.A. (1976). Assessment of the reliability and the effect of neutral instructions on the symptoms ratings on the Moos Menstrual Distress Questionnaire. *Psychosomatic Medicine, 38,* 163-172. doi: 10.1097/00006842-197605000-00002Mortola, J. F., Girton, L., Beck, L., & Yen, S. S. (1990). Diagnosis of premenstrual syndrome by a simple, prospective, and reliable instrument: The calendar of premenstrual experiences. *Obstetrics and Gynecology, 76*, 302-307.

Moos, R. H. (1968). The development of the menstrual distress questionnaire. *Psychosomatic Medicine, 30*, 853-867. doi: 10.1097/00006842-196811000-00006

O'Connor, B. P. (2000) SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers, 32*(3), 396-402. doi:10.3758/bf03200807

Robinson, R. L., Swindle, R. W. (2000). Premenstrual symptom severity: Impact on social functioning and treatment-seeking behaviors. *Journal of Women's Health & Gender Based Medicine, 9*, 757–768. doi: 10.1089/15246090050147736

Romans, S., Clarkson, R., Einstein, G., Petrovic, M., & Stewart, D. (2012). Mood and the menstrual cycle: A review of prospective data studies. *Gender Medicine, 9*, 361-384. doi: 10.1016/j.genm.2012.07.003

Ross, C., Coleman, G., & Stojanovska, C. (2003). Factor structure of the Modified Moos Menstrual Distress Questionnaire: Assessment of prospectively reported follicular, menstrual, and premenstrual symptomatology. *Journal of Psychosomatic Obstetrics and Gynecology. 24*, 163-174. doi: 10.3109/01674820309039670

Rubinow, D. R. & Roy-Byrne, P. (1984). Premenstrual syndromes: Overview from a methodological perspective. *The American Journal of Psychiatry, 14*, 163-172. doi:10.1176/ajp.141.2.163

Rubinow, D.R., Roy-Byrne, P., Hoban, M.C., Gold, P. W. & Post, R. M. (1984) Prospective assessment of menstrual related mood disorders. *American Journal of Psychiatry*, *141*, 684-686. doi: 10.1176/ajp.141.5.684

Ryu, A. & Kim, T. H. (2015). Premenstrual syndrome: A mini-review. *Maturitas, 82,* 436-440. doi: 10.1016/j.maturitas.2015.08.010

Steiner, M., Macdougall, M., & Brown, E. (2003). The premenstrual symptoms screening tool (PSST) for clinicians. *Archives of Women's Mental Health, 6,* 203-209. doi: 10.1007/s00737-003-0018-4

Sundström, I., Nyberg, S., Bixo, M., Hammarbäck, S., & Backström, T. (1999). Treatment of premenstrual syndrome with gonadotropin-releasing hormone agonist in a low-dose regimen. *Acta Obstetricia et Gynecologica Scandinavica, 78*, 891-899. doi: 10.1034/j.1600-0412.1999.781011.x

Taylor, J.W. (1979). The timing of menstruation-related symptoms assessed by a daily symptom rating scale. *Acta Psychiatrica Scandanavica, 60,* 87-105. https://doi.org/10.1111/j.1600-0447.1979.tb00268.x

Tabachnick B. G., & Fidell, L. S. (2014). *Using Multivariate Statistics (6th ed.).* Pearson.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063

Wittichun, H. U., Becker, E., Lieb, R., & Krause, P. (2002). Prevalence, incidence and stability of premenstrual dysphoric disorder in the community. *Psychological Medicine, 32*, 119-132. doi: 10.1017/S0033291701004925

Appendix

Some women experience changes in mood and physical functioning during the week prior to their menstrual period. As best as you can, please indicate the severity, level of impairment, and frequency of the following 11 symptoms during your PREMENSTRUAL PHASE over the past year.

Symptoms*	Please rate the degree of severity of these symptoms:					To what extent do these symptoms cause impairment in work, school, or relationships with others?					Number of menstrual cycles in which the symptom(s) have been experienced over the past 12 months (0 to 12):
	Not at all 0	Mild 1	Moderate 2	Severe 3	Extremely Severe/ Debilitating 4	Not at all 0	Mild 1	Moderately 2	Quite a Bit 3	Extremely 4	
1. Marked affective lability (e.g., mood swings, feeling suddenly sad or tearful, or increased sensitivity to rejection).	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
2. Marked irritability or anger or increased interpersonal conflicts.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
3. Marked depressed mood, feelings of hopelessness or self-deprecating thoughts.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
4. Marked anxiety, tension, and/or feelings of being keyed up or on edge.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
5. Decreased interest in usual activities (e.g. work, school, friends, hobbies).	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
6. Subjective difficulty in concentration.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
7. Lethargy, easy fatigability, or marked lack of energy.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
8. Marked change in appetite; overeating; or specific food cravings.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
9. Hypersomnia or insomnia.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12

10. A sense of being overwhelmed or out of control.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12
11. Physical symptoms such as breast tenderness or swelling, joint or muscle pain, a sensation of "bloating" or weight gain.	0	1	2	3	4	0	1	2	3	4	0 1 2 3 4 5 6 7 8 9 10 11 12

The symptoms in column one are from the criteria for Premenstrual Dysphoric Disorder (PMDD) as they appear in the *Diagnostic and Statistical Manual of Mental Disorders* (5th Edition; DSM-5), (p. 171-172) by the American Psychiatric Association (APA), 2013, Arlington VA; APA. Reprinted with permission from the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, (Copyright ©2013). American Psychiatric Association. All Rights Reserved.

*Note that item nine was adapted for a lower reading level in our assessment tool and reads as "Sleeping too much or too little."