Routledge
Taylor & Francis Group

Journal of Sex & Marital Therapy

ISSN: 0092-623X (Print) 1521-0715 (Online) Journal homepage: http://www.tandfonline.com/loi/usmt20

Further Validation of the Female Sexual Function Index: Specificity and Associations With Clinical Interview Data

Kyle R. Stephenson, Nasreen Toorabally, Leah Lyons & Cindy M. Meston

To cite this article: Kyle R. Stephenson, Nasreen Toorabally, Leah Lyons & Cindy M. Meston (2016) Further Validation of the Female Sexual Function Index: Specificity and Associations With Clinical Interview Data, Journal of Sex & Marital Therapy, 42:5, 448-461, DOI: 10.1080/0092623X.2015.1061078

To link to this article: http://dx.doi.org/10.1080/0092623X.2015.1061078

Accepted author version posted online: 22 Jun 2015.
Published online: 22 Jun 2015.

JOURNAL OF SEX & MARITAL THERAPY, *42*(5), 448–461, 2016

ISSN: 0092-623X print / 1521-0715 online
DOI: 10.1080/0092623X.2015.1061078

Routledge
Taylor & Francis Group

Further Validation of the Female Sexual Function Index: Specificity and Associations With Clinical Interview Data

Kyle R. Stephenson
Department of Psychology, Willamette University, Salem, Oregon, USA

Nasreen Toorabally
Department of Psychology, University of Essex, Colchester, UK, and Department of Psychology, California State University Monterey Bay, Seaside, California, USA

Leah Lyons
Department of Psychology, California State University Monterey Bay, Seaside, California, USA

Cindy M. Meston
Department of Psychology, The University of Texas at Austin, Austin, Texas, USA

Female sexual function is a multi-faceted psychophysiological construct. The Female Sexual Function Index (FSFI) is considered a "gold standard" self-report instrument that assesses the various aspects of sexual function. However, researchers have recently proposed potential limitations of the FSFI, highlighting the need for continued validation research. The aims of the current analyses were (a) to assess the correlations between FSFI scores and information regarding specific rates of functional impairment gained via clinical interview; and (b) to assess the specificity of FSFI subscale scores in reflecting corresponding aspects of sexual function (e.g., whether the Sexual Desire subscale reflects sexual desire specifically rather than sexual arousal, orgasm, etc.). The participants were 97 sexually active women who reported impairments in sexual function. Clinical interview data exhibited moderate-to-strong correlations with FSFI scores. Additionally, FSFI subscales generally exhibited adequate specificity in terms of reflecting their corresponding aspects of sexual function more strongly than other aspects. The results generally supported the validity of the FSFI. Implications for the measurement and conceptualization of female sexual function are discussed.

INTRODUCTION

Female sexual function is a complex construct consisting of multiple related aspects including desire, subjective arousal, lubrication, orgasm, and pain (Rosen et al., 2000). Impairments in female sexual function are highly prevalent, with approximately 58% of women reporting impaired

Address correspondence to Kyle R. Stephenson, Assistant Professor, Department of Psychology, Willamette University, 900 State Street, Salem, OR 97301, USA. E-mail: kstephen@willamette.edu

sexual function in at least one of these areas in the past year (Hayes, Dennerstein, Bennett, & Fairley, 2008). Impairments in sexual function are associated with a higher risk of developing depression and anxiety (Laurent & Simons, 2009) and with lower overall quality of life (Stephenson & Meston, 2015a). Thus, it is important for researchers and clinicians to have an accurate understanding of the nature of female sexual function, and to be able to effectively measure it.

Researchers' understanding of the structure of female sexual function is consistently evolving, and multiple theoretical models have been proposed over the past 20 years. Masters and Johnson initially proposed four sequential phases in their sexual response cycle: arousal, plateau, orgasm, and resolution (Masters & Johnson, 1966). This initial model has been expanded and altered multiple times, including Kaplan's addition of sexual desire preceding sexual activity (Kaplan, 1995) and Basson and colleagues circular response model (Basson et al., 2005), which incorporated multiple additional facets of sexual function and described complex bidirectional interrelationships between these facets. In contrast to these recent complex models, the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-IV-TR*; American Psychiatric Association, 2000) generally maintained a clear distinction between four aspects of sexual function, sexual desire, arousal, orgasm, and pain, with each having a corresponding psychiatric diagnosis. However, the *DSM-5* (5th ed., American Psychiatric Association, 2013) has instituted additional changes, combining sexual desire and arousal into a single diagnostic category. These changes highlight the ever-shifting nature of how female sexual function is defined and organized.

In addition to this diversity of conceptual models of sexual function, there is a wide variety of tools for assessing the construct, ranging from psychophysiological measures such as vaginal photoplethysmography (Sintchak & Geer, 1975) to structured clinical interviews, such as the Derogatis Interview for Sexual Functioning (Derogatis, 1997). However, by far the most common method of assessing sexual function is self-report scales. There are numerous scales currently available (Meston & Derogatis, 2002), many of which have received some empirical support to establish their reliability and validity (e.g., Lorenz, Stephenson, & Meston, 2011). The scale that has received the most attention from researchers is the Female Sexual Function Index (FSFI).

The FSFI is a 19-item self-report measure that provides an overall sexual function score on a continuous interval scale, along with individual domain scores for sexual desire, subjective arousal, lubrication, orgasm, sexual satisfaction, and sexual pain. In the initial validation study (Rosen et al., 2000), the FSFI was found to exhibit good internal consistency and test-retest reliability. It has been shown to differentiate between women with Female Sexual Arousal Disorder and age-matched controls, and between women with Female Orgasmic Disorder or Hypoactive Sexual Desire Disorder and age-matched controls (Meston, 2003). Discriminant validity has been established using a measure of marital satisfaction (Rosen et al., 2000).

Since this initial study, the FSFI has been translated into more than 20 languages (Sun, Li, Jin, Fan, & Wang, 2011) and has been validated in more than 30 countries (Nowosielski, Wróbel, Sioma-Markowska, & Poręba, 2013). It has been explicitly validated for use with multiple populations, including women from different age groups, with diverse medical conditions, and with various sexual dysfunctions (Dargis, Trudel, Cadieux, Villenueve, Préville, & Boyer, 2012; Likes, Stegbauer, Hathaway, Brown, & Tillmanns, 2006; Meston, 2003). Cutoff scores have been established that reliably differentiate between women with and without sexual impairment (Wiegel, Meston, & Rosen, 2005), and a sexual desire cutoff score has been established that differentiates between women with and without Hypoactive Sexual Desire Disorder (Gerstenberger et al.,

2010). The FSFI has also been shown to be sensitive to treatment-induced changes in female sexual function (Rosen, Revicki, & Sand, 2014). Lastly, multiple factor analytic studies have provided some support for the six-factor structure of the FSFI (Opperman, Benson, & Milhausen, 2013). Based on this accumulation of evidence, the FSFI seems to be a well-supported and useful clinical assessment tool. Indeed, the FSFI has been referred to as a "gold standard" instrument for assessing female sexual function (Rosen, Revicki, & Sand, 2014).

However, as with any brief self-report scale of a complex psychophysiological construct, the FSFI has notable practical and theoretical limitations. First, the FSFI requires women to self-report on their typical levels of physiological sexual arousal (i.e., lubrication) despite the fact that women are often fairly inaccurate at estimating their degree of arousal (Chivers, Seto, Lalumière, Laan, & Grimbos, 2010). Second, researchers have noted that the FSFI produces biased results for women who have not been sexually active in the past month (Meyer-Bahlburg & Dolezal, 2007). Specifically, 15 of the 19 questions include a response option of "No sexual activity" or "Did not attempt intercourse" which is scored as a zero. This option is problematic because, when scoring the FSFI, lower scores indicate more severe dysfunction. Of course, there are numerous reasons that women may not engage in sexual activity over a four-week period that do not necessarily imply significant sexual dysfunction (e.g., absent partner).

Third, the FSFI uses relatively vague terminology regarding specific frequency of functional impairment. For example, response options to the question "Over the past 4 weeks, when you had sexual stimulation or intercourse, how often did you reach orgasm (climax)?" include "Most times (more than half the time)" and "A few times (less than half the time)." These response options may limit the FSFI's ability to capture significant distinctions regarding frequency of orgasm. For example, a woman who experiences orgasm 40% of the time and a woman who experiences orgasm 10% of the time may both select the option "A few times (less than half the time)" despite clear differences in their sexual experiences, which increases measurement error.

Fourth, there has been considerable debate regarding the degree to which meaningful distinctions can be made between the various domains of sexual function. This debate is most clearly exhibited in the attempt to define and measure sexual desire and subjective sexual arousal. As discussed above, the initial validation study of the FSFI (Rosen et al., 2000) used factor analytic methods to identify potential subscales. In this study, analyses suggested five domains of female sexual function: desire/subjective arousal, lubrication, orgasm, satisfaction, and pain/discomfort. However, the authors made a "clinically-based decision" to separate sexual desire and subjective arousal into distinct domains because they had traditionally been described as related but independent aspects of sexual function (Kaplan, 1995).

Researchers have questioned the decision to retain desire and arousal as separate aspects of sexual function for a number of reasons. Qualitatively, studies have suggested that lay women often have difficulty distinguishing between desire and arousal in the context of their sexual experiences (Brotto, Heiman, & Tolman, 2009), often conflating descriptions of the two. Quantitatively, studies have found mixed support for this distinction. For example, Forbes and colleagues (2014) conducted an exploratory factor analysis and found results suggesting that desire and arousal items could be combined into a single factor. More informative are studies utilizing confirmatory factor analysis to test a-priori factor structures of the FSFI that either do or do not differentiate between sexual desire and subjective arousal. Opperman and colleagues (2013) found that two different models, one where desire and arousal were combined and one where they were not, exhibited adequate fit to their data. However, the model separating desire and arousal was a

significantly better fit. Kalmbach and colleagues (2014) used similar methods and also found that a model separating desire and arousal was the best fit to their data. However, a limitation of all three of these studies was the use of convenience samples that did not specifically include women with impaired sexual function. Carvalho and colleagues (2012) highlighted the importance of this limitation by showing that the best statistical model of sexual function differed depending on whether women report significant impairments in their sexual function. For women with significant sexual difficulties, a model combining desire and arousal most closely fit the data whereas for women without difficulties, retaining separate desire and arousal factors resulted in better fit. These various findings highlight the unanswered nature of two important questions: How appropriate it is to maintain conceptual distinctions between the various aspects of female sexual function, and is the FSFI able to effectively measure these distinctions?

There is an ongoing debate as to whether the various limitations summarized above make the FSFI a "critically flawed" measure (Forbes, Baillie, & Schniering, 2014), or whether it is simply of limited use in certain situations (Rosen, Revicki, & Sand, 2014). These critiques and responses represent an important conversation regarding the validity of the FSFI. While information is invariably lost when measuring a construct as complex as female sexual function using a brief self-report scale, it is important to determine whether FSFI scores are strongly related to more labor-intensive methods of assessment. More empirical evidence is also needed to determine whether the various aspects of sexual function as measured by the FSFI are indeed distinct.

AIM

To further inform this debate, we conducted a secondary data analysis using a sample of women with self-identified impairments in one or more aspects of sexual function. These analyses had two primary goals. The first goal was to compare information regarding specific rates of impairment gathered during face-to-face clinical interviews (e.g., "What percentage of the time have you not reached orgasm during partnered sexual activity in the past month?") with scores on the FSFI. Based on the extensive research supporting the validity of the FSFI, we predicted that the correlations between interview data and FSFI scores would be moderate to strong. Our second goal was to assess whether each subscale score of the FSFI was specifically predictive of its corresponding aspect of sexual function as measured in a clinical interview (as opposed to indiscriminately reflecting all aspects of sexual function). For example, are scores on the arousal subscale of the FSFI more strongly associated with clinical interview data regarding arousal than they are with interview data regarding desire? We predicted that each FSFI subscale would correlate more strongly with interview data regarding the corresponding facet of sexual function than with other, non-corresponding facets of sexual function.

METHOD

Participants and Procedure

The current study utilized data gathered as part of a larger project examining sexual and relationship experiences for women with sexual difficulties (Stephenson & Meston, 2015b). One hundred participants were recruited from a southern U.S. metropolitan area using flyers and online

advertisements. Participants were required to be adult women currently in a monogamous heterosexual romantic relationship who self-identified as having a difficulty in one or more of the following areas in the past month: low sexual desire, low sexual arousal, impaired vaginal lubrication, difficulty reaching orgasm, or pain/discomfort during or following sexual activity. Further, they were required to be willing to attempt to engage in sexual activity with their partner at least five times during the month after being accepted into the study (this requirement was necessary to fulfill a separate study aim unrelated to the current analyses). Interested individuals were screened by phone before attending an in-person appointment at the Sexual Psychophysiology Laboratory on the University of Texas at Austin campus.

In-person intake appointments included a semi-structured clinical interview conducted by a master's-level clinician that assessed *DSM-IV-TR* (APA, 2000) criteria for Hypoactive Sexual Desire Disorder (HSDD), Female Sexual Arousal Disorder (FSAD), Female Orgasmic Disorder (FOD), and Sexual Pain Disorders (dyspareunia and vaginismus). Participants also completed a number of validated self-report measures, including the FSFI. All study protocol was approved by the institutional review board at the University of Texas at Austin.

Ninety-seven of the 100 women who attended the intake assessment met criteria for the study. The final sample was 27.35 years old on average ($SD = 6.6$), and was 80.4% Caucasian, 13.4% Hispanic, 7.2% Asian American, 5.2% African American, and 2.1% other (participants could select multiple ethnicities). All participants were in monogamous relationships with an average length of 47 months ($SD = 63.81$ months); 27.8% were married. Two percent of participants had earned a high school diploma only, 30.9% had completed some college, 37.1% had earned a bachelor's degree, and 29.9% had earned a graduate degree.

Measures

Clinical Interview

A semi-structured diagnostic interview was used to assess recent sexual function in the areas of desire, subjective arousal, lubrication, orgasm, and pain. *DSM-IV-TR* (APA, 2000) criteria for disorders of sexual function were assessed, and additional questions were included to reflect expansion of theoretical models of female sexual function that have been suggested since the publication of the *DSM-IV-TR* (Basson et al., 2005). For example, when assessing for Female Sexual Arousal Disorder, questions were included that noted the important distinction between problems in physiological arousal alone (e.g., decreased vaginal lubrication) vs. problems in subjective arousal (e.g., feeling "turned on" and "into it" during sexual activity). Similarly, the description of sexual desire provided to participants included both spontaneous desire and reactive desire (i.e., sexual desire in response to a relevant stimulus such as receiving a sexual cue from a partner; Basson, 2000).

As part of this interview, women were asked to estimate specific rates of impairment in each domain of their sexual function in the past month. These specific rates were gathered only in cases where women self-identified as having difficulty in a specific aspect of sexual function. For example, women self-identifying as having a problem with subjective sexual arousal in the past month were asked, "Thinking back to the times you engaged in sexual activity with your partner in the past month, what percentage of the time would you say that your subjective arousal was notably low?" Similar questions were asked regarding lubrication, orgasm, and sexual pain

(Rosen et al., 2000). Given the disconnect between sexual desire and sexual activity (sexual desire can be high or low regardless of whether sexual activity takes place), sexual desire was assessed with a different question: "On what percentage of days during the past month would you say you felt little or no spontaneous or responsive desire for sexual activity?" In each case, women were encouraged to provide the most accurate percentage they could and were encouraged to ask clarifying questions (e.g., about the distinction between physiological and subjective arousal). For each aspect of sexual function, responses were coded as a percentage indicating a specific rate of impairment. For example, participants X and Y both report a significant problem with sexual arousal. However, participant X reports significant impairments in sexual arousal during 40% of partnered sexual episodes in the past month, while participant Y reports impaired arousal during 70% of partnered sexual episodes.

Female Sexual Function Index

The Female Sexual Function Index is a 19-item self report questionnaire developed by Rosen and colleagues (2000) to assess female sexual function in heterosexual women. The instrument measures six areas of sexual function related to sexual activity over the past four weeks: desire, arousal, lubrication, orgasm, satisfaction, and pain. All questions use a five-point Likert scale with higher scores indicating higher levels of function. Additionally, a zero option is available for most questions, indicating no sexual activity in the last month. Each subscale is weighted equally and subscale scores are summed to determine an overall score of sexual function. Full scale scores range anywhere from 2.0 to 36.0. A clinical cutoff score has been established, with women scoring 26.55 or lower likely meeting criteria for female sexual dysfunction (Wiegel, Meston, & Rosen, 2005). Responses of 0 (no sexual activity) were coded as missing data per recommendations (Meyer-Bahlburg & Dolezal, 2007), excluding these participants from analyses. The mean FSFI score in the current sample was 22.88 ($SD = 5.42$). The sexual satisfaction subscale was not utilized due to the well-established distinction between sexual satisfaction and sexual function (Shifren, Monz, Russo, Segreti, & Johannes, 2008; Stephenson, Pulverman, & Meston, 2014). Cronbach's alpha in the current sample was .930, .931, .927, .899, and .864 for sexual desire, arousal, lubrication, orgasm, and pain, respectively.

RESULTS

Descriptive statistics were calculated for all study variables (see Table 1). For interview data assessing specific rates of impairment in sexual function, means and standard deviations were computed. For example, 52 participants self-identified as having a problem with subjective sexual arousal in the past month. These women reported specific rates of impairment ranging from 12.5% to 100%. In other words, women with a self-identified problem with sexual arousal reported experiencing notably low arousal during between 12.5% and 100% of partnered sexual activities. On average, they reported low arousal during 62.30% of partnered sexual activities ($SD = 26.64\%$). Ranges of reported rates of impairment were wide for all aspects of sexual function (desire: 28.6%–100%; arousal: 12.5%–100%; lubrication: 15%–100%; orgasm: 0%–100%; pain: 10%–100%).

TABLE 1
Pearson's Correlations, Means, and SDs for Study Variables

						Clinical Interview		*FSFI Scores*	
Clinical Interview	*FSFI Desire*	*FSFI Arousal*	*FSFI Lubrication*	*FSFI Orgasm*	*FSFI Pain*	M	SD	M	SD
Desire ($N = 56$)	–.**622*****	–.385**	–.278*	.113	–.247+	79.21%	18.29%	3.37	1.41
Arousal ($N = 52$)	–.468**	–.**654*****	–.470**	–.249+	–.133	62.30%	26.64%	3.71	1.33
Lubrication ($N = 55$)	–.385**	–.236+	–.**509*****	.078	–.149	72.65%	28.15%	4.31	1.46
Orgasm ($N = 79$)	.001	–.181	–.067	–.**464*****	–.028	79.72%	27.50%	3.18	1.54
Pain ($N = 52$)	–.252+	–.251+	–.417**	–.084	–.**516*****	62.29%	33.55%	4.57	1.44

Notes. FSFI = Female Sexual Function Index; Clinical Interview = the percentage of time that women reported impaired function in the context of a clinical interview; bolded numbers indicate FSFI subscale and clinical interview data measuring the same aspect of sexual function.

$+ p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Relationship Between FSFI Subscale Scores and Clinical Interview Data

To assess the strength of the relationship between FSFI subscale scores and corresponding data from clinical interviews, Pearson product-moment correlations were computed between the FSFI subscales and interview data (see Table 1). As expected, FSFI subscale scores and interview data were negatively correlated (given that higher FSFI scores indicate better sexual function, whereas higher interview scores indicate worse sexual function). A statistically significant relationship ($p < .001$) was observed between FSFI subscale scores and corresponding interview data in each case. For example, there was a strong and significant correlation between FSFI desire scores and corresponding interview data assessing frequency of sexual desire ($r = –.622$, $p < .001$). Alternatively, the FSFI desire subscale was less strongly correlated with interview data measuring all other aspects of sexual function (r ranged from –.468 to .001). The strength of correlations between each FSFI subscale and its corresponding interview data ranged from –.46 to –.65, which are considered moderate to strong effect sizes.

Differential Relationship Between FSFI Scores and Corresponding vs. Noncorresponding Data

While it was visually clear that FSFI subscales correlated most strongly with corresponding interview data and less strongly with noncorresponding interview data, correlation coefficient t-tests were performed to assess whether the strengths of these correlations were statistically significantly different. In each case, the correlation between an FSFI subscale and its corresponding interview data (e.g., FSFI arousal correlated with interview data regarding arousal at $r = –.65$) was compared to the correlation between the same FSFI subscale and noncorresponding interview data (e.g., FSFI arousal correlated with interview data regarding orgasm at $r = –.18$). In this case, the correlation of –.65 was significantly stronger than the correlation of –.18 ($t\,(51) = 3.8, p < .05$). The results of these analyses are summarized in Table 2. As can be seen, the correlation between FSFI subscales and corresponding interview data was significantly stronger than the correlation between these subscales and noncorresponding interview data in a majority

TABLE 2
Correlation *t*-tests of Pearson's Correlation Coefficients

FSFI Subscale *Interview Data*	T-*values*	*Sig.*
FSFI Desire		
Arousal	1.32	
Lubrication	1.95	+
Orgasm	4.46	*
Pain	2.98	*
FSFI Arousal		
Desire	1.93	+
Lubrication	4.70	*
Orgasm	3.80	*
Pain	3.44	*
FSFI Lubrication		
Desire	1.66	
Arousal	0.39	
Orgasm	2.71	*
Pain	0.76	
FSFI Orgasm		
Desire	3.40	*
Arousal	1.50	
Lubrication	3.21	*
Pain	2.51	*
FSFI Pain		
Desire	1.99	+
Arousal	2.88	*
Lubrication	2.91	*
Orgasm	3.35	*

Notes. FSFI = Female Sexual Function Index; *t*-values represent a comparison between two correlations: FSFI subscale and corresponding interview data and FSFI subscale and noncorresponding interview data.
$+ p < .10$; $^{*}p < .05$.

of cases. Notable exceptions to this pattern were the relatively weak distinction between desire and arousal, and the general inability of the FSFI lubrication subscale to differentiate between interview data regarding lubrication and interview data regarding other aspects of sexual function.

Multiple Regression Analyses Assessing Specificity of FSFI Subscales

Multiple linear regression was used as an additional method of assessing the specificity with which FSFI subscales correlated with interview data. Specifically, these analyses were performed to determine whether variance in interview data regarding a specific domain of sexual function (e.g., frequency of sexual desire reported via interview) was uniquely predicted by its corresponding FSFI subscale (e.g, FSFI desire) over and above the effects of all other FSFI subscales (FSFI arousal, FSFI lubrication, etc.). A series of multiple linear regression models were constructed in which interview data regarding each aspect of sexual function in turn were regressed on all FSFI subscales. The resulting five models are summarized in Table 3. All overall models were

TABLE 3
Multiple Linear Regression Models With Interview Data Regressed on FSFI Subscales

Dependent Variable *Predictor*	*β*	B	SE	*Sig.*	F	R^2
Dependent Variable: Percentage of days no desire					7.72***	.426
FSFI Desire	–.54	–10.84	2.69	***		
FSFI Arousal	–.16	–2.76	2.50			
FSFI Lubrication	–.04	–.55	1.86			
FSFI Orgasm	.24	3.32	1.65			
FSFI Pain	.02	.31	1.76			
Dependent Variable: Percentage of time low arousal					6.94***	.477
FSFI Desire	–.22	–4.30	2.47			
FSFI Arousal	–.54	–12.39	3.59	**		
FSFI Lubrication	–.12	–2.19	2.67			
FSFI Orgasm	–.02	–.27	2.47			
FSFI Pain	.14	2.36	2.50			
Dependent Variable: Percentage of time lubrication low					4.21**	.328
FSFI Desire	–.24	–4.69	3.07			
FSFI Arousal	.05	1.00	3.29			
FSFI Lubrication	–.50	–10.49	3.12	**		
FSFI Orgasm	.19	3.38	2.41			
FSFI Pain	.14	2.66	2.87			
Dependent Variable: Percentage of time orgasm not reached during intercourse					6.68***	.311
FSFI Desire	.19	3.88	2.52			
FSFI Arousal	–.21	–4.60	2.94			
FSFI Lubrication	–.01	–.04	2.29			
FSFI Orgasm	–.52	–10.75	2.14	***		
FSFI Pain	–.10	–2.10	2.43			
Dependent Variable: Percentage of time pain during/after intercourse					4.87**	.372
FSFI Desire	–.06	–1.25	3.39			
FSFI Arousal	.09	2.19	4.15			
FSFI Lubrication	–.28	–6.13	3.48			
FSFI Orgasm	–.15	–3.51	3.17			
FSFI Pain	–.46	–11.83	3.62	**		

Notes. FSFI = Female Sexual Function Index; Clinical Interview = the percentage of time that women reported impaired function in the context of a clinical interview; bolded numbers indicate FSFI subscale and clinical interview data measuring the same aspect of sexual function.

*$p < .05$; **$p < .01$; ***$p < .001$.

significant. Additionally, in each case the corresponding subscale of the FSFI was a unique predictor of the outcome, and no other subscales were unique predictors controlling for other factors in the model. For example, the first model regressed interview data regarding sexual desire (percentage of days with little to no desire reported for women with self-identified problems with desire) on all FSFI subscale scores simultaneously. The overall model was significant (F (5, 57) $= 7.72$, $p < .001$, $R^2 = .426$), and FSFI desire scores predicted unique variance in interview responses when all other subscales were controlled for ($\beta = -.54$, $p < .001$). Noncorresponding scales (FSFI arousal, FSFI orgasm, etc.) did not significantly predict interview data regarding desire.

DISCUSSION

The overarching goal of the current analyses was to provide additional empirical data to inform the current debate over the validity of the FSFI in assessing female sexual function (Forbes, Baillie, & Schniering, 2014; Rosen, Revicki, & Sand, 2014). In particular, we tested the strength of association between FSFI scores and data derived from an in-person clinical interview (convergent validity), as well as the specificity of FSFI subscale scores in predicting their corresponding aspects of sexual function via interview (content validity). Given the already strong evidence regarding the validity of the FSFI (Rosen, Revicki, & Sand, 2014), we expected that the current results would similarly support the measure. Indeed, our results indicated moderate to strong correlations between information gathered via interview and scores on the FSFI. These findings suggest that, even though information acquired using brief self-report scales is inherently less detailed than that acquired using clinical interviews, this loss of information does not appear to invalidate the usefulness of the FSFI in assessing female sexual function.

Furthermore, a majority of FSFI subscales exhibited good specificity, with most subscales being more strongly correlated with corresponding information gathered via clinical interview than with noncorresponding aspects of sexual function. For example, FSFI arousal subscale scores were strongly correlated with corresponding interview information regarding specific rates of low arousal, but were not strongly correlated with interview information regarding specific rates of difficulties with orgasm. However, there were some cases in which the results regarding specificity were not as strong. In particular, the distinction between sexual desire and subjective arousal received mixed support—in some analyses, the FSFI significantly differentiated between interview information regarding these two constructs, and in some analyses it did not. These weaker results are consistent with the occasionally mixed findings of factor analytic studies (Carvalho, Vieira, & Nobre, 2012) and qualitative studies regarding women's perception of desire and arousal (Brotto, Heiman, & Tolman, 2009). It is important to note, however, that the overlap between desire and arousal may partly stem from evolving definitions and conceptualizations of sexual desire. In particular, responsive desire is defined as being experienced *in response* to an already present stimulus, as opposed to being experienced spontaneously before presentation of an explicit sexual stimulus (Basson, 2000). This expansion of desire seems to remove one of the key traditional methods of differentiating between desire and arousal: that desire occurs before sexual activity begins, and arousal occurs during sexual activity. It would be interesting to explicitly test whether the inclusion or exclusion of responsive desire

significantly impacts whether desire and arousal can be meaningfully differentiated using the FSFI.

Results also suggested that the lubrication subscale may indiscriminately reflect multiple aspects of sexual function including both subjective and physiological arousal, as well as sexual pain. While we are not aware of other studies producing such results, these findings are consistent with research suggesting that women may have difficulty accurately reporting their level of physiological arousal. For example, Chivers and colleagues (2010) estimated that the correlation between women's perception of their level of genital arousal and their actual level of genital arousal is between .17 and .30. Indeed, studies suggest that women often do not attend specifically to genital sensations when reporting their sexual arousal (Prause, Barela, Roberts, & Graham, 2013), and the goal of some interventions is to increase women's awareness of their own physiological responses (Silverstein, Brown, Roth, & Britton, 2011). As such, it can be expected that scores on a self-report scale of physiological arousal may be more prone to measurement error and, thus, be fairly unspecific indicators of lubrication in particular.

Despite these exceptions, the results of the current study tend to support the FSFI as a valid measure for assessing self-identified sexual problems in women. Given the time savings of providing a brief self-report measure in comparison to conducting a lengthy interview, it seems justifiable to use the FSFI to assess female sexual function when more comprehensive measurements are not possible. Beyond the use of this specific scale, the current results also have important theoretical implications. In particular, they support the decision to maintain separate diagnoses for impairments in sexual desire, arousal, orgasm, and sexual pain given that women were able to differentiate between these constructs, and patterns of self-report responses suggested that they can be separated empirically. This differentiation is partially reflected in the new *DSM-5* (APA, 2013), which retained distinct diagnoses for difficulty reaching orgasm and sexual pain. However, the *DSM-5* has combined problems with sexual desire, subjective arousal, and physiological arousal into a single diagnostic category: Female Sexual Interest/Arousal Disorder.

The question of whether or not to retain separate diagnostic categories for these various symptoms is important because this nosology shapes the formation of theoretical models and guides research on treatments for these problems. Both the current findings and past research suggest that separate theoretical models may be needed to explain the factors that cause and maintain problems in different aspects of female sexual function. For example, while relatively straightforward behavioral interventions have been shown to be efficacious in treating anorgasmia (Heiman, 2007), success rates are more modest for dysfunctions of desire and arousal (Basson, Wierman, van Lankveld, & Brotto, 2010; ter Kuile, Both, & van Lankveld, 2010). Additional treatment outcome research that assesses if there are significant differences in efficacy depending on whether the primary impairment in sexual function is related to desire or arousal would help determine whether problems in these areas are best explained by distinct theoretical models.

The current study had a number of important limitations. First, the sample of women used was relatively small, which may have limited our ability to detect weak statistical effects. Further, participants were required to be sexually active, likely excluding women who are sexually inactive due to severe impairments in sexual function (e.g., women with intense sexual pain that prevents attempts at sexual activity). Additionally, while all participants self-identified as having impairments in sexual function, and the average FSFI score was below the established clinical cut-off (Wiegel, Meston, & Rosen, 2005), the women in the study were not required to meet full criteria regarding female sexual dysfunction, including significant personal distress or

interpersonal difficulties. As such, further research using samples of women formally diagnosed with sexual dysfunction is necessary to determine whether the current results apply to severely distressed clinical samples.

In addition to sampling issues, retrospective recall bias may have affected participants' responses to both the FSFI and clinical interviews. Inaccurate recall of past experiences is a problem in much of psychological research (Levine & Safer, 2002), including sexuality. For example, Graham and colleagues (2003) compared daily diary reports of sexual activity with recall of the same activity one month later. They found that only 29% of their sample showed no discrepancy between daily reports and retrospective reports of frequency of orgasm, with 54% overestimating their frequency of orgasm. Similarly, it is likely that women in the current study were not entirely accurate when recalling specific rates of impairment in sexual function over the past month whether via interview or FSFI. It is likely that a portion of the correlation between FSFI scores and information gathered through clinical interviews can be attributed to the fact that both reflect this retrospective recall bias to some degree. It will be important for future research to use a wider variety of methods when assessing sexual function and to actively compare the results of these methods.

Lastly, slightly different scales of measurement were used to assess different aspects of sexual function in the clinical interview. For example, rates of low sexual desire were measured in terms of days in a typical week that women experienced little to no desire, whereas rates of low arousal were measured in terms of percentage of partnered sexual activities where arousal was impaired. While these differing metrics make sense conceptually (e.g., desire can be felt, or not, in the absence of any sexual activity) and mirror the language of the FSFI, it is important to note that scores in these areas are not strictly analogous. In fact, the results suggesting overlap in measurement of desire and arousal are even more noteworthy in that this pattern was present despite the different scales of measurement.

Despite these limitations, the current study is the first of which we are aware that assessed the association between scores on the FSFI and specific information regarding rates of sexual impairment measured via clinical interview. The FSFI exhibited moderate-to-strong associations with clinical interview data, and its subscales exhibited good specificity in most cases. These results reinforce the validity and usefulness of the FSFI. Further research, including use of the FSFI in treatment outcome studies and comparison between the FSFI and alternative methods of assessment, is needed to provide further support for this widely used measure.

ACKNOWLEDGMENTS

The authors would like to thank Parrish Williams, Margaret Cerutti, Rebecca Neufeld, Heather Bicoy, and Lauren Coates for their help in data collection.

FUNDING

This publication was supported by Grant Number R01 HD51676 from the National Institute for Child Health and Human Development to Cindy M. Meston. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute for Child Health and Human Development.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

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

Basson, R., Wierman, M. E., van Lankveld, J., & Brotto, L. (2010). Summary of the recommendations on sexual dysfunctions in women. *Journal of Sexual Medicine*, *7*, 314–326. doi:10.1111/j.1743-6109.2009.01617.x

Basson, R., Brotto, L. A., Laan, E., Redmond, G., & Utian, W. H. (2005). Assessment and management of women's sexual dysfunctions: Problematic desire and arousal. *Journal of Sexual Medicine*, *2*, 291–300. doi:10.1111/j.1743-6109.2005.20346.x

Basson, R. (2000). The female sexual response: A different model. *Journal of Sex & Marital Therapy*, *26*, 51–65. doi:10.1080/009262300278641

Brotto, L. A., Heiman, J. R., & Tolman, D. L. (2009). Narratives of desire in mid-age women with and without arousal difficulties. *Journal of Sex Research*, *46*, 387–398. doi:10.1080/00224490902792624

Carvalho, J., Vieira, A. L., & Nobre, P. (2012). Latent structures of female sexual functioning. *Archives of Sexual Behavior*, *41*, 907–917. doi:10.1007/s10508-011-9865-7

Chivers, M. L., Seto, M. C., Lalumière, M. L., Laan, E., & Grimbos, T. (2010). Agreement of self-reported and genital measures of sexual arousal in men and women: A meta-analysis. *Archives of Sexual Behavior*, *39*, 5–56. doi:10.1007/s10508-009-9556-9

Dargis, L., Trudel, G., Cadieux, J., Villeneuve, L., Préville, M., & Boyer, R. (2012). Validation of the Female Sexual Function Index (FSFI) and presentation of norms in older women. *Sexologies: European Journal of Sexology and Sexual Health / Revue Européenne De Sexologie Et De Santé Sexuelle*, *21*, 126–131. doi:10.1016/j.sexol.2012.01.002

Derogatis, L. R. (1997). The Derogatis Interview for Sexual Functioning (DISF/DISF-SR): An introductory report. *Journal of Sex & Marital Therapy*, *23*, 291–304. doi:10.1080/00926239708403933

Forbes, M. K., Baillie, A. J., & Schniering, C. A. (2014). Critical flaws in the Female Sexual Function Index and the International Index of Erectile Function. *Journal of Sex Research*, *51*, 485–491. doi:10.1080/00224499.2013.876607

Gerstenberger, E. P., Rosen, R. C., Brewer, J. V., Meston, C. M., Brotto, L. A., Wiegel, M., & Sand, M. (2010). Sexual desire and the Female Sexual Function Index (FSFI): A sexual desire cutpoint for clinical interpretation of the FSFI in women with and without hypoactive sexual desire disorder. *Journal of Sexual Medicine*, *7*, 3096–3103. doi:10.1111/j.1743-6109.2010.01871.x

Graham, C. A., Catania, J. A., Brand, R., Duong, T., & Canchola, J. A. (2003). Recalling sexual behavior: A methodological analysis of memory recall bias via interview using the diary as the gold standard. *Journal of Sex Research*, *40*, 325–332. doi:10.1080/00224490209552198

Hayes, R. D., Dennerstein, L., Bennett, C. M., & Fairley, C. K. (2008). What is the 'true' prevalence of female sexual dysfunctions and does the way we assess these conditions have an impact?. *Journal of Sexual Medicine*, *5*, 777–787. doi:10.1111/j.1743-6109.2007.00768.x

Heiman, J. R. (2007). Orgasmic disorders in women. In S. R. Leiblum (Ed.), *Principles and practice of sex therapy* (4th ed., pp. 84–123). New York, NY: Guilford Press.

Kalmbach, D. A., Ciesla, J. A., Janata, J. W., & Kingsberg, S. A. (2014). The validation of the female sexual function index, male sexual function index, and profile of female sexual function for use in healthy young adults. *Archives of Sexual Behavior*, doi:10.1007/s10508-014-0334-y

Kaplan, H. S. (1995). *The sexual desire disorders: Dysfunctional regulation of sexual motivation*. Philadelphia, PA: Brunner/Mazel.

Laurent, S. M., & Simons, A. D. (2009). Sexual dysfunction in depression and anxiety: Conceptualizing sexual dysfunction as part of an internalizing dimension. *Clinical Psychology Review*, *29*, 573–585. doi:10.1016/j.cpr.2009.06.007

Levine, L. J., & Safer, M. A. (2002). Sources of bias in memory for emotions. *Current Directions in Psychological Science*, *11*, 169–173. doi:10.111/1467–8721.00193

Likes, W. M., Stegbauer, C., Hathaway, D., Brown, C., & Tillmanns, T. (2006). Use of the Female Sexual Function Index in women with vulvar intraepithelial neoplasia. *Journal of Sex & Marital Therapy*, *32*, 255–266. doi:10.1080/00926230600575348

Lorenz, T. K., Stephenson, K. R., & Meston, C. M. (2011). Validated questionnaires in female sexual function assessment. In J. P. Mulhall, L. Incrocci, I. Goldstein, & R. Rosen (Eds.), *Cancer and sexual health* (pp. 317–337). New York, NY: Humana Press.

Masters, W., & Johnson, V. E. (1966). *Human sexual response*. Oxford, England: Little, Brown.

Meston, C. M. (2003). Validation of the female sexual function index (FSFI) in women with female orgasmic disorder and in women with hypoactive sexual desire disorder. *Journal of Sex & Marital Therapy*, *29*, 39–46. doi:10.1080/713847100

Meston, C. M., & Derogatis, L. R. (2002). Validated instruments for assessing female sexual function. *Journal of Sex & Marital Therapy*, *28*, 155–164. doi:10.1080/00926230252851276

Meyer-Bahlburg, H. L., & Dolezal, C. (2007). The Female Sexual Function Index: A methodological critique and suggestions for improvement. *Journal of Sex & Marital Therapy*, *33*, 217–224. doi:10.1080/00926230701267852

Nowosielski, K., Wróbel, B., Sioma-Markowska, U., & Poręba, R. (2013). Development and validation of the Polish version of the Female Sexual Function Index in the Polish population of females. *Journal of Sexual Medicine*, *10*, 386–395. doi:10.1111/jsm.12012

Opperman, E. A., Benson, L. E., & Milhausen, R. R. (2013). Confirmatory factor analysis of the Female Sexual Function Index. *Journal of Sex Research*, *50*, 29–36. doi:10.1080/00224499.2011.628423

Prause, N., Barela, J., Roberts, V., & Graham, C. (2013). Instructions to rate genital vasocongestion increases genital and self-reported sexual arousal but not coherence between genital and self-reported sexual arousal. *Journal of Sexual Medicine*, *10*, 2219–2231. doi:10.1111/jsm.12228

Rosen, R. C., Revicki, D. A., & Sand, M. (2014). Commentary on 'critical flaws in the FSFI and IIEF'. *Journal of Sex Research*, *51*, 492–497. doi:10.1080/00224499.2014.894491

Rosen, R., Brown, C., Heiman, J., Leiblum, S., Meston, C., Shabsigh, R., ... D'Agostino, R. J. (2000). The Female Sexual Function Index (FSFI): A multidimensional self-report instrument for the assessment of female sexual function. *Journal of Sex & Marital Therapy*, *26*, 191–208. doi:10.1080/009262300278597

Shifren, J. L., Monz, B. U., Russo, P.A., Segreti, A., & Johannes, C. B. (2008). Sexual problems and distress in United States women: Prevalence and correlates. *Obstetrics & Gynecology*. doi:10.1097/AOG.0b013e3181898cdb

Silverstein, R. G., Brown, A. H., Roth, H. D., & Britton, W. B. (2011). Effects of mindfulness training on body awareness to sexual stimuli implications for female sexual dysfunction. *Psychosomatic Medicine*, *73*, 817–825. doi:10.1097/PSY.0b013e318234e628

Sintchak, G., & Geer, J. H. (1975). A vaginal plethysmograph system. *Psychophysiology*, *12*, 113–115. doi:10.1111/j.1469-8986.1975.tb03074.x

Stephenson, K. R., & Meston, C. M. (2015a). The conditional importance of sex: Exploring the association between sexual well-being and life satisfaction. *Journal of Sex & Marital Therapy*, *41*, 25–38. doi:10.1080/0092623X.2013.811450

Stephenson, K. R., & Meston, C. M. (2015b). Why is impaired sexual function distressing for women? The primacy of pleasure in female sexual dysfunction. *Journal of Sexual Medicine*, *12*, 728–737.

Stephenson, K. R., Pulverman, C. S., & Meston, C. M. (2014). Assessing the association between childhood sexual abuse and adult sexual experiences in women with sexual difficulties. *Journal of Traumatic Stress*, *27*, 274–282. doi:10.1002/jts.21923

Sun, X., Li, C., Jin, L., Fan, Y., & Wang, D. (2011). Development and validation of Chinese version of Female Sexual Function Index in a Chinese population—A pilot study. *Journal of Sexual Medicine*, *8*, 1101–1111. doi:10.1111/j.1743-6109.2010.02171.x

ter Kuile, M. M., Both, S., & van Lankveld, J. M. (2010). Cognitive behavioral therapy for sexual dysfunctions in women. *Psychiatric Clinics of North America*, *33*, 595–610. doi:10.1016/j.psc.2010.04.010

Wiegel, M., Meston, C., & Rosen, R. (2005). The Female Sexual Function Index (FSFI): Cross-validation and development of clinical cutoff scores. *Journal of Sex & Marital Therapy*, *31*, 1–20. doi:10.1080/00926230590475206