CrossMark

ORIGINAL PAPER

The Validation of the Female Sexual Function Index, Male Sexual Function Index, and Profile of Female Sexual Function for Use in Healthy Young Adults

David A. Kalmbach · Jeffrey A. Ciesla · Jeffrey W. Janata · Sheryl A. Kingsberg

Abstract Past research has typically used clinical samples to evaluate the validity of sexual function measures. As normal variations in sexually healthy individuals are of important research and clinical interest, evaluating the applicability of common sexual function measures to these populations is important. Factor structures of the Female Sexual Function Index (FSFI), Male Sexual Function Index (MSFI) (adapted for this investigation), and Profile of Female Sexual Function (PFSF) were examined in young, healthy men and women. We predicted the factor structures to be consistent with past evaluations. In a cross-sectional study, 1,258 participants (M age = 19.56 years; 59 % women) completed these measures. Confirmatory factor analyses did not initially support the factor structures. However, factor loadings showed marked differences between positively and negatively worded items. As such, each measure's factor structure was tested using multi-trait multi-method confirmatory factor analysis which accounted for variance due to item valence. These models supported the predicted structures of the FSFI and MSFI, whereas the PFSF's Responsiveness scale required modification, resulting in an Avoidance scale for both genders. This study was one of few to validate the FSFI in young, healthy adults and the

first to examine the MSFI and PFSF in these populations. Additionally, this investigation was the first to propose a reconceptualization of the PFSF Responsiveness scale into an Avoidance scale. Lastly, our study highlights the significant impact of item valence on how individuals respond to questions regarding their sexual functioning.

Introduction

Many aspects of daily living, such as physical health, psychological well-being, and partner relationship quality, influence sexual functioning in otherwise sexually healthy individuals (Dunn, Croft, & Hackett, 1999; Kalmbach, Ciesla, Janata, & Kingsberg, 2012; Levine, 2003; Rosen & Bachman, 2008). Presently, evidence for the utility of self-report instruments of sexual functioning in men and women primarily comes from clinical samples (e.g., individuals with sexual dysfunction). As such, it is presently unclear whether existing measures of sexual functioning are sensitive to normative variations. Indeed, prior investigations have typically relied on using measures of sexual functioning which have not been validated for normative, healthier populations who are less likely to be encumbered by physical complaints and psychiatric illness. In order to advance the study of the influences and outcomes of normative sexual functioning, studies must establish the applicability of these measures to healthier populations. Additionally, the field presently suffers from a dearth of self-report instruments of male and female sexual functioning that allow for comparison between men

D. A. Kalmbach · J. A. Ciesla
Department of Psychology, Kent State University, Kent, OH, USA

D. A. Kalmbach (✉)
Sleep and Chronophysiology Laboratory, University of Michigan, Ann Arbor, MI 48109, USA
e-mail: dkalmbac@gmail.com

J. W. Janata · S. A. Kingsberg
Departments of Reproductive Biology and Psychiatry, Case Western Reserve University School of Medicine, Cleveland, OH, USA

and women. Parallel measures of male and female sexual functioning would allow for important comparisons between men and women in both clinical and research settings. As such, the present study sought to examine the utility of existing measures of sexual functioning in non-clinical samples of young adult men and women.

The Female Sexual Function Index (FSFI) was introduced by Rosen et al. (2000) who examined the measure by comparing middle-aged (M age = 40 years) women with female sexual arousal disorder with healthy controls. Using a sample of 259 women, exploratory techniques provided support for a five-factor model. This factor structure was consistent with the a priori model, which includes pain, lubrication, orgasm, satisfaction, and desire/arousal. However, Rosen et al. made a clinically-based decision to split desire and arousal into separate scales, as each construct can be defined independently. Rosen et al. also found the FSFI factors to have acceptable test–retest reliability and high internal consistency. Notably, Wiegel, Meston, and Rosen (2005) replicated these findings when comparing a sample of middle-aged women with various sexual disorders and a healthy control sample.

The FSFI has been validated for use in a variety of populations in which individuals may experience sexual difficulties. Indeed, analyses revealed FSFI factors to have high internal consistencies and have shown expected differences in sexual difficulties between middle aged women who are healthy and the following clinical populations: women who suffer from hypoactive sexual desire disorder (Meston, 2003), female sexual orgasmic disorder (Meston, 2003), vulvar intraepithelial neoplasia (Likes, Stegbauer, Hathaway, Brown, & Tillmanns, 2006), chronic pelvic pain (Verit & Verit, 2007), multiple sclerosis (Borello-France et al., 2008), and vulvodynia (Masheb, Lozano-Blanco, Kohorn, Minkin, & Kerns, 2004).

Carvalho, Vieira, and Nobre (2012) examined the factor structure of the FSFI in both clinical and non-clinical samples of Portuguese women. Unlike much of the prior research, they used confirmatory techniques to compare different models of the FSFI in their samples. Carvalho et al. compared a four-factor and five-factor model for clinical samples: in the four factor model, the desire and arousal were combined, but separated in the five-factor model (sexual satisfaction was excluded from analyses). Despite the five-factor showing better model fit as indicated by the fit indices, Carvalho et al. reasoned that, because the five-factor model violated a .70 measure distinctness threshold between the desire and arousal scale, the four-factor model was the better model for the clinical sample.

Though much empirical support of the FSFI has been found in clinical samples, two recent studies have examined its utility in non-clinical samples. Carvalho et al. (2012) found a five-factor model to best fit the data: desire, arousal,

lubrication, orgasm, and pain (satisfaction was not examined). Opperman, Benson, and Milhausen (2013) used confirmatory techniques to examine the utility of the FSFI in a non-clinical sample of young Canadian women. They found that the original six-factor structure was the best fit for their data.

Similarly, studies have also provided empirical support for the utility of the Profile of Female Sexual Function (PFSF). McHorney et al. (2004) examined the PFSF in a sample of 580 oophorectomized women with hypoactive sexual desire disorder and healthy controls from North America, Europe, and Australia. Using exploratory factor analysis, they found a seven-factor structure: desire, arousal/orgasm, responsiveness, pleasure, self-image, concerns, and disinterest. However, McHorney et al. discarded the disinterest scale as it accounted for a small portion of the total variance. Further, McHorney et al. split the arousal/orgasm scale as arousal and orgasm are clinically observed to be separate constructs. In support of the measure's reliability, McHorney et al. found good internal consistencies and test–retest reliability for each of the scales. The PFSF showed good clinical utility in its ability to discriminate participants experiencing sexual difficulties from the healthy controls. In further support, DeRogatis et al. (2004) found that the PFSF differentiated between clinical samples and healthy controls and provided support for the measure's factor internal consistencies.

Many of the validation studies of the FSFI and PFSF have used samples of middle-aged adults and individuals with various sexual dysfunctions and medical conditions. However, these instruments may also be useful for studying more normative variations in sexual functioning, as would be seen in healthier populations. By validating these measures in non-clinical samples, we can evaluate the sensitivity of these measures to be able to quantify the more subtle differences anticipated within normative levels of functioning. To our knowledge, the utility of the FSFI in non-clinical samples has only been evaluated in two studies (Carvalho et al., 2012; Opperman et al., 2013) whereas the PFSF has not been examined in non-clinical samples.

Both the FSFI and PFSF were developed for the assessment of female sexual functioning. Male sexual functioning has typically been assessed through the use of other instruments, such as the International Index of Erectile Function (IIEF) (Rosen et al., 1997). However, measures of female and male sexual functioning are not parallel to each other. For example, though the FSFI and the IIEF share some similarities (e.g., both measures assess sexual satisfaction), there remain some significant differences (e.g., the IIEF does not assess subjective arousal whereas the FSFI does). The development of maximally similar measures for use in both genders would allow researchers to compare the relative influence of variables on female and male sexual health. For instance, the present study was part of a larger effort interested in the relations between affective (i.e., depression and

anxiety) and sexual problems and how these relations differ between men and women (Kalmbach et al., 2012). The use of non-parallel measures would complicate such research, as any observed gender differences could either be due to true differences between women and men or to differences between the measures themselves.

The FSFI and PFSF were chosen to be validated in men rather than validating male sexual function measures in women. Measures of female sexual function place a lesser emphasis on the physical aspects of sexual function than measures of male sexual function. As many of the gender differences in sexual function are physiological, similar measures would need to place lesser emphasis on these physiological, gender-specific difficulties (e.g., lubrication and erection difficulties) though we also did not want to ignore them. As such, any measure with gender-specific items would have to be modified to appropriateness for men (e.g., replace lubrication difficulty items with erection difficulty items); however, this practice should not be considered an attempt to equate the two constructs. Notably, some measures of male sexual function attend more so to difficulties with premature ejaculation than to delayed ejaculation. However, this study was part of a larger effort in examining the relations between affective and sexual problems. Kennedy, Dickens, Eisfeld, and Bagby (1999) showed that delayed ejaculation has twice the prevalence of premature ejaculation in depressed men. As such, we chose to validate in men the FSFI and PFSF, which assess delayed and difficult orgasm.

The present study sought to examine the utility of the FSFI, Male Sexual Function Index (MSFI) (adapted for use in this study), and PFSF in young adults. As the FSFI has gender-specific items, we modified the FSFI and to create the MSFI. The MSFI was created to be maximally similar to the FSFI to allow for comparisons between men and women, while minimizing measurement confound. Unlike the FSFI, as the PFSF does not contain gender-specific items, it was unaltered for administration to male participants. As past research has outlined the factor structures of these measures, we employed confirmatory factor analysis (CFA). We hypothesized that the previously identified factor structures of the FSFI and PFSF would hold true for our sample of young female adults. Although our review of the literature did not include male samples, we also elected to employ CFA for the examination of the MSFI and PFSF in young male adults. We hypothesized that the structure of the MSFI in the male sample would be parallel to the previously identified factor structure of the FSFI in women. Similarly, we predicted that the previously identified factor structure of the PFSF would also hold true for men.[1]

Method

Participants

The sample consisted of 1,258 undergraduate students (748 women) who were screened to be antidepressant-free for at least 2 weeks prior to participating in the study. A total of 731 participants (409 women) reported sexual activity with a partner (oral, anal, or vaginal sex, sexual foreplay) in the past 30 days. The participant age range was 18–29 years (M = 19.56). The mean age that participants first engaged in sexual intercourse was 16.22 years (SD = 2.16) for female participants and 16.18 years (SD = 2.47) for male participants. As for lifetime history of sexual experiences among the entire sample, 81.2 % reported having engaged in sexual intercourse, 82.9 % had received oral sex, 79.7 % had performed oral sex, and 78.8 % reported having performed self-stimulation. The mean number of lifetime sexual partners (defined as any persons with whom they engaged in any form of sexual activity) was 6.17 (SD = 2.16) for men and 3.81 (SD = 4.54) for women. The mean number of partners within 30 days of participating was .98 (SD = 1.08) for men and .76 (SD = .63) for women. Participants were largely heterosexual (87.3 %), with a smaller proportion being bisexual (10.7 %) or homosexual.

Procedure

Individuals were recruited from introductory psychology classes at a large midwestern university in the US and received course credit for their participation. The present study used a cross-sectional design. All instructions were changed to ask participants to report their experiences over the past month (30 days) to both standardize assessment window across all measures as well as to encompass the average menstrual cycle. Additionally, definitions of sexual terms were provided to individuals prior to participation. The Institutional Review Board approved this study. All participants were required to provide written informed consent prior to participation.

Measures

The FSFI (Rosen et al. 2000) is a 19-item self-report questionnaire of female sexual functioning. The FSFI assesses sexual desire, psychological arousal, lubrication, pain, satisfaction, and orgasm achievement over the previous 30 days. We slightly modified the FSFI's definitions of sexual activity and intercourse to also include anal sex, as we felt that it was a

---

[1] Though we hypothesize that the factor structures will be the same across men and women, we do not mean to imply that sexual functioning is the same for both genders. Rather, factor analyses regard the structure

Footnote 1 continued

of the measure and the way in which items are correlated with one another and we predicted that item correlations would be similar for both men and women.

relevant sexual activity for both our heterosexual and homosexual participants. As some participants in the study were not expected to have sexual partners in the assessment window, items specific to sexual activity with a partner were given an additional response choice of "I have not had a sexual partner in the past 30 days" (e.g., Item 14). For items that were specific to sexual activity with a partner, we treated non-sexually active participants' data to be missing and computed prorated (using mean substitution) scores for the scale (Brotto, 2009; Meyer-Bahlburg & Dolezal, 2007). However, individuals' factor scores were treated as missing if more than 25 % of data in a given factor was missing. Our rationale was that we believed treating a response of "No Sexual Activity" as zero would artificially bias scores into indicating higher dysfunction whereas proration allowed us to estimate the total scale score based on their other responses within the same scale. We decided to only prorate when the response rate was 75 % and above so as to minimize the impact of our estimation on the data.

The MSFI is a 16-item self-report questionnaire of male sexual functioning over the previous 30 days that was created by modifying the FSFI. As previously mentioned, the MSFI was created rather than using an existing measure of male sexual functioning to allow for some comparison of sexual problems between genders. That is, utilizing maximally similar measures in order to decrease confounding between measures and biological sex was important. Similar to the FSFI, we slightly modified the MSFI's definitions of sexual activity and intercourse to also include anal sex, as we felt that it was a relevant sexual activity for both our heterosexual and homosexual participants. In creating the MSFI, we replaced FSFI items assessing lubrication items with items assessing erection difficulties [see Kalmbach et al. (2012) for complete list of new items]. Also, the Pain scale was removed when adapting the MSFI due to the low prevalence of sexual pain in men (see Rosen, 2000). However, the desire, psychological arousal, orgasm, and satisfaction items were unchanged from their FSFI phrasings. Like the FSFI, the MSFI contains items that are specific to sexual activity with a partner. Again, partner sex items were accompanied by a response option indicating an absence of partner sex during the assessment window. These data were considered missing, which allowed for proration of the scale scores.

The PFSF (McHorney et al., 2004) is a 37-item self-report questionnaire assessing sexual function. The PFSF is an assessment tool of sexual difficulties, including desire, psychological arousal, orgasm achievement, pleasure, responsiveness, concerns, and self-image over the past 30 days. As items on the PFSF are not specific to females, the items were unchanged for male participants. Further, because the PFSF contains items that are specific to sexual activity, participants were provided with a response option indicating an absence

of sexual activity during the assessment window. Similarly to the FSFI and MSFI, these data were considered missing, which allowed for proration of the scale scores.

Data Analysis

All analyses were conducted separately for each gender. To evaluate the utility of the FSFI, MSFI, and PFSF in young adults, we examined the structure and reliability of these measures. To examine the factorial structures of the FSFI, MSFI, and PFSF, we conducted CFAs for each measure in its corresponding sample (i.e., FSFI and PFSF in women, MSFI and PFSF in men). A notable benefit of using CFA over exploratory techniques is that it allowed us to test the validity of the FSFI and MSFI two-item desire scales. Exploratory factor analysis requires a minimum of three to five items for a factor and, thus, cannot validate a two-item scale whereas CFA models with multiple factors require that each factor have two or more indicators (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Kline, 2013; Velicer & Fava, 1998). We predicted that the FSFI and PFSF's original structures would be supported in our sample of young, healthy women. Similarly, we hypothesized that the MSFI's structure would correspond to its female counterpart's original structure, and that the PFSF's original structure would be supported in our male sample. To examine the reliability of these measures, we used Cronbach's alpha to investigate the internal consistency of the scales of each questionnaire.

Results

Factorial Structure

FSFI in Women

We tested the six-factor structural model of the FSFI with all latent variable variances set to 1.0 and, based on Hu and Bentler's (1998) recommendations, and found adequate to good model fit, $\chi^2(137) = 683.28$, $p < .001$, CFI = .91, TLI = .88, RMSEA = .07. Examination of the factor loadings revealed notable differences between items that were worded positively versus those that were worded negatively. As such, we suspected that item-valence markedly influenced the observed structure. Items worded in a negative direction tended to cluster together, as did the items worded in a positive direction. Therefore, we employed a multi-trait multi-method (MTMM) CFA (Kline, 2011) to model method variance due to item-valence. This approach allowed for modeling item covariance as a function of (1) the associated sexual functioning construct and (2) whether items were worded in a negative or positive direction.

Fig. 1 Partial model of the Female Sexual Function Index multi-trait multi-method confirmatory factor analysis (error terms not shown in model)
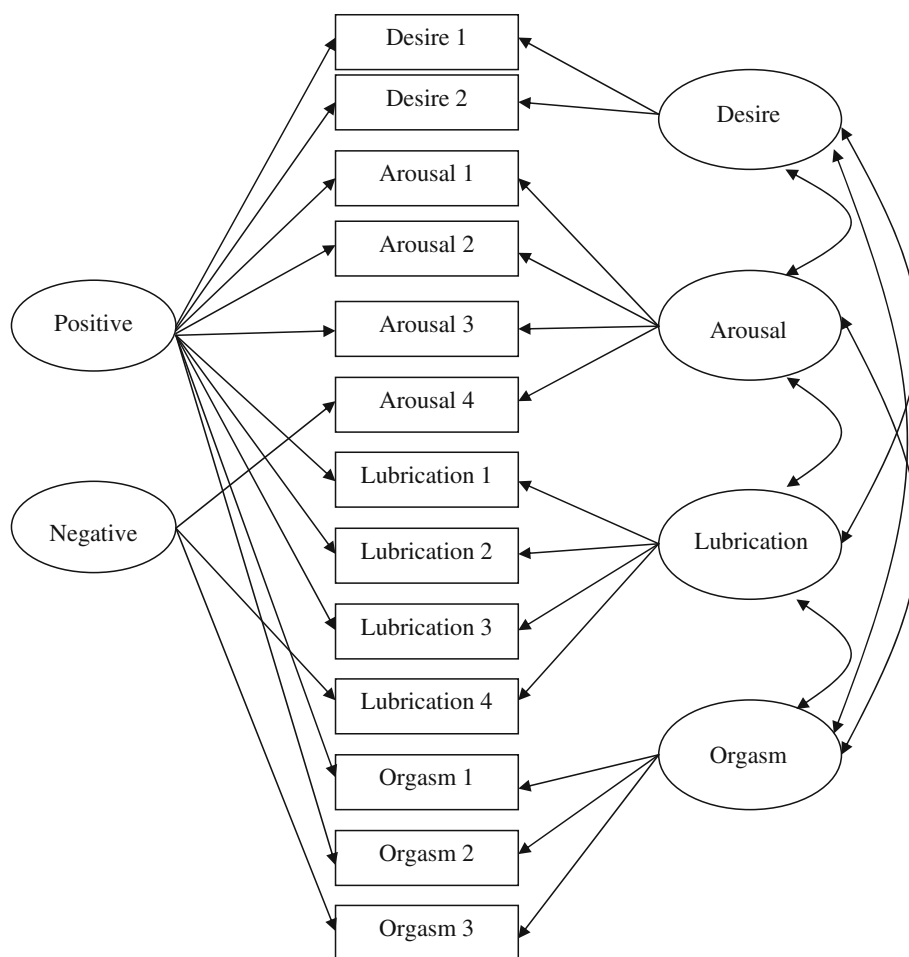
To create our MTMM CFA model, we used two exogenous factors such that all positively-valenced items (e.g., How often did you become lubricated during sexual activity or intercourse?) were loaded onto one factor and all negatively-valenced items (e.g., How difficult was it to become lubricated during sexual activity or intercourse?) were loaded onto a second factor. These loadings were in addition to the already established loadings of the items onto their corresponding sexual functioning factors (see Fig. 1 for a partial model). We then tested this MTMM CFA model, which produced very good model fit, $\chi^2(118) = 303.01$, $p < .001$, CFI = .97, TLI = .95, RMSEA = .04.[2] Additionally, analyses showed that the MTMM CFA was a significant improvement over the initial model, $\Delta\chi^2(19) = 380.27$, $p < .001$. All items significantly loaded onto their respective factors (see Table 1 for factor loadings).

---

[2] Carvalho et al. (2012) showed that desire and arousal were distinct in non-clinical women whereas they were unitary in clinical women. To examine if we were able to replicate their findings, we also ran a separate MTMM CFA with desire and arousal combined into a single construct. When comparing the models, analyses revealed desire and arousal to be distinct constructs, $\Delta\chi^2(5) = 134.77$, $p < .001$. These analyses are consistent with Carvalho et al.'s findings regarding non-clinical women.

PFSF in Women

We tested the seven-factor structural model of the PFSF and found poor model fit, $\chi^2(608) = 4540.59$, $p < .001$, CFI = .86, TLI = .84, RMSEA = .09. Similar to the FSFI, substantial variance in the PFSF items appeared to be due to item-valence. Thus, we tested a MTMM CFA model similar to the previously described, which produced good model fit. However, the Responsiveness factor did not conform to the original structure. Upon examination, we found that the first two items produced high loadings of .57 and .63 whereas the latter five item loadings ranged from .02 to .19. These data suggested that the first two items of the Responsiveness scale were not measuring the same construct as the latter five items. Inspection of the scale items showed that the latter five items appeared to measure sexual avoidance (e.g., "I avoided having sex") whereas the first item appeared to measure sexual initiation and the second item measured responsiveness. As the data suggested that the Responsiveness scale was measuring more than one construct, we removed the first two items of the scale, thus revising the Responsiveness factor into an avoidance factor. After re-running the MTMM CFA, our model yielded good fit, $\chi^2(573) = 2488.53$, $p < .001$,

Table 1 Female Sexual Function Index, multi-trait multi-method confirmatory factor analysis factor loadings, and standardized regression weights

Item	Factor	Estimate
1	Desire	.78
2	Desire	.86
3	Arousal	.57
4	Arousal	.59
5	Arousal	.65
6	Arousal	.63
7	Lubrication	.65
8[a]	Lubrication	.76
9	Lubrication	.67
10[a]	Lubrication	.77
11	Orgasm	.79
12	Orgasm	.85
13[a]	Orgasm	.83
14	Satisfaction	.75
15	Satisfaction	.91
16	Satisfaction	.83
17	Pain	.79
18	Pain	.80
19	Pain	.81

[a] Negatively-valenced item

CFI $= .93$, TLI $= .91$, RMSEA $= .06$. Additionally, analyses showed that the MTMM CFA was a significant improvement over the initial model, $\Delta\chi^2(35) = 2052.26$, $p < .001$. All item loadings were significant (see Table 2 for factor loadings).

MSFI in Men

We tested the five-factor structural model of the MSFI and found poor model fit, $\chi^2(94) = 440.82$, $p < .001$, CFI $= .86$, TLI $= .80$, RMSEA $= .08$. However, similar to the analyses for women, we once again observed substantial variance in the MSFI items due to item-valence. We then tested a MTMM CFA model accounting for item valence, which yielded good fit, $\chi^2(78) = 229.51$, $p < .001$, CFI $= .94$, TLI $= .89$, RMSEA $= .06$[3] and was a significant improvement over the initial model, $\Delta\chi^2(16) = 211.31$, $p < .001$. Item loadings on each factor were significant (see Table 3 for factor loadings).

---

[3] Carvalho et al.'s (2012) examination of the FSFI showed that desire and arousal were distinct in non-clinical women whereas they were unitary in clinical women. To examine if we were able to replicate their findings in our sample of non-clinical men, we also ran a separate MTMM CFA with desire and arousal combined into a single construct. Comparison of the models provided empirical evidence of the distinction between desire and arousal, $\Delta\chi^2(4) = 30.46$, $p < .001$.

Table 2 Profile of Female Sexual Function multi-trait multi-method confirmatory factor analysis factor loadings and standardized regression weights

Item	Factor	Estimate
1	Desire	.68
2	Desire	.72
3	Desire	.70
4	Desire	.66
5[a]	Desire	.30
6	Desire	.54
7[a]	Desire	.29
8	Desire	.49
9	Desire	.52
10[a]	Arousal	.65
11[a]	Arousal	.77
12[a]	Arousal	.71
13	Orgasm	.41
14[a]	Orgasm	.83
15[a]	Orgasm	.95
16[a]	Orgasm	.80
17	Pleasure	.51
18	Pleasure	.55
19	Pleasure	.53
20	Pleasure	.58
21	Pleasure	.53
22	Pleasure	.57
23	Pleasure	.60
24[a]	Concerns	.77
25[a]	Concerns	.71
26[a]	Concerns	.87
27[a]	Avoidance	.66
28[a]	Avoidance	.68
29[a]	Avoidance	.54
30[a]	Avoidance	.73
31[a]	Avoidance	.72
32	Self-image	.71
33	Self-image	.80
34	Self-image	.77
35[a]	Self-image	.39

[a] Negatively-valenced item

PFSF in Men

We tested the seven-factor structural model of the PFSF and found poor model fit, $\chi^2(608) = 3495.86$, $p < .001$, and additional fit indices, CFI $= .83$, TLI $= .80$, RMSEA $= .09$. After again observing substantial variance due to item-valence, we then tested a MTMM CFA model, which produced good model fit. However, the Responsiveness factor once again did not conform to its original structure. Similar to what was found in the female sample, the first two items did

Table 3 Male Sexual Function Index multi-trait multi-method confirmatory factor analysis factor loadings and standardized regression weights

Item number	Factor	Estimate
1	Desire	.68
2	Desire	.80
3	Arousal	.49
4	Arousal	.56
5	Arousal	.62
6	Arousal	.68
7	Erection	.47
8[a]	Erection	.51
9	Erection	.76
10[a]	Erection	.76
11	Orgasm	.65
12	Orgasm	.51
13[a]	Orgasm	.73
14	Satisfaction	.78
15	Satisfaction	.90
16	Satisfaction	.81

[a] Negatively-valenced item

not correlate with the latter five items. After employing the same modification used in the female sample, we re-ran the MTMM CFA and our model produced adequate to good model fit, $\chi^2(573) = 2255.29, p < .001, \text{CFI} = .90, \text{TLI} = .88, \text{RMSEA} = .07$, and was a significant improvement over the initial model, $\Delta\chi^2(35) = 1007.33, p < .001$. All item loadings were significant (see Table 4 for factor loadings).

Reliability

Internal consistency was examined for each measure to evaluate inter-item correlations within each measure's scales. The FSFI produced high internal consistencies in each of its subscales (see Table 5), with Cronbach's alphas ranging from .81 (lubrication) to .89 (desire). Similarly, the internal consistencies in the PFSF for the female sample ranged from high (orgasm; $\alpha = .84$) to very high (pleasure; $\alpha = .99$) (see Table 5). The MSFI yielded internal consistencies ranging from adequate (orgasm; $\alpha = .66$) to high (desire; $\alpha = .85$) (see Table 6). The PFSF for the male sample produced adequate (orgasm; $\alpha = .71$) to very high (pleasure; $\alpha = .97$) internal consistency (see Table 6).

Discussion

The present study sought to examine the validity of commonly used measures of sexual functioning in a sample of

Table 4 Profile of Female Sexual Function multi-trait multi-method confirmatory factor analysis factor loadings and standardized regression weights

Item	Factor	Estimate
1	Desire	.70
2	Desire	.76
3	Desire	.78
4	Desire	.67
5[a]	Desire	.18
6	Desire	.61
7[a]	Desire	.25
8	Desire	.35
9	Desire	.47
10[a]	Arousal	.61
11[a]	Arousal	.74
12[a]	Arousal	.67
13	Orgasm	.28
14[a]	Orgasm	.68
15[a]	Orgasm	.84
16[a]	Orgasm	.64
17	Pleasure	.58
18	Pleasure	.61
19	Pleasure	.55
20	Pleasure	.65
21	Pleasure	.51
22	Pleasure	.64
23	Pleasure	.66
24[a]	Concerns	.74
25[a]	Concerns	.55
26[a]	Concerns	.74
27[a]	Responsiveness	.68
28[a]	Responsiveness	.64
29[a]	Responsiveness	.47
30[a]	Responsiveness	.41
31[a]	Responsiveness	.41
32	Self-image	.75
33	Self-image	.74
34	Self-image	.61
35	Self-image	.44

[a] Negatively-valenced item

healthy, young adults. Findings of our confirmatory techniques supported the original structures of the FSFI and MSFI, the latter of which was adapted for this study to be maximally similar to the FSFI (from which it was based). However, when evaluating the PFSF, the original structure required modification in both the male and female samples. Specifically, we revised the PFSF Responsiveness scale by removing the first two items, thus creating a Sexual Avoidance scale. This modification resulted in a supported structure

Table 5 Means, SDs, internal consistencies, and correlations among sexual functioning measures (women)

Factor	M (SD)	α	FSFI–desire	FSFI–arousal	FSFI–lubrication	FSFI–orgasm	FSFI–satisfaction	FSFI–pain	PFSF–desire	PFSF–arousal	PFSF–orgasm	PFSF–pleasure	PFSF–concerns	PFSF–avoidance
Female Sexual Function Index														
Desire	6.65 (2.02)	.89												
Arousal	16.33 (3.59)	.87	.61*											
Lubrication	17.69 (3.05)	.81	.33*	.55*										
Orgasm	11.17 (3.52)	.88	.20*	.46*	.34*									
Satisfaction	12.28 2.99	.87	.19*	.50*	.36*	.40*								
Pain	12.63 (2.77)	.86	.12	.24*	.48*	.17*	.19*							
Profile of Female Sexual Function														
Desire	34.61 (9.92)	.94	.81*	.60*	.34*	.20*	.19*	.15*						
Arousal	15.24 (3.14)	.92	.27*	.48*	.52*	.27*	.36*	.40*	.33*					
Orgasm	16.86 (5.05)	.84	.14	.36*	.32*	.82*	.36*	.18*	.16*	.40*				
Pleasure	28.64 (11.85)	.99	.56*	.70*	.48*	.47*	.46*	.26*	.63*	.29*	.16*			
Concerns	6.19 (3.54)	.89	-.05	-.37*	-.40*	-.33*	-.61*	-.36*	-.06	-.45*	-.25*	-.23*		
Avoidance	9.11 (4.99)	.88	-.19*	-.41*	-.53*	-.24*	-.39*	-.51*	-.24*	-.45*	-.48*	-.21*	.43*	
Self-image	16.80 (4.86)	.88	.40*	.50*	.36*	.30*	.49*	.22*	.47*	.36*	.14*	.58*	-.43*	-.29*

FSFI Female Sexual Function Index, *PFSF* Profile of Female Sexual Function

* Correlation is significant at the .01 level

Table 6 Means, SD, internal consistencies, and correlations among sexual functioning measures (men)

Factor	M (SD)	α	MSFI– desire	MSFI– arousal	MSFI– erection	MSFI– orgasm	MSFI– satisfaction	PFSF– desire	PFSF– arousal	PFSF– orgasm	PFSF– pleasure	PFSF– concerns	PFSF– avoidance
Male Sexual Function Index													
Desire	7.59 (1.76)	.85											
Arousal	17.33 (2.89)	.82	.46*										
Erection	18.17 (2.84)	.76	.10	.53*									
Orgasm	17.77 (2.84)	.66	.15*	.49*	.52*								
Satisfaction	12.16 (2.64)	.82	.12*	.49*	.30*	.36*							
Profile of Female Sexual Function													
Desire	38.94 (8.23)	.90	.79*	.46*	.19*	.16*	.08						
Arousal	15.80 (3.06)	.93	.16*	.41*	.61*	.46*	.20*	.18*					
Orgasm	19.28 (3.58)	.71	.10	.41*	.41*	.61*	.15*	.10	.55*				
Pleasure	31.40 (10.19)	.97	.40*	.53*	.29*	.39*	.34*	.47*	.14*	.16*			
Concerns	6.79 (3.83)	.89	.02	−.40*	.29*	−.34*	−.45*	−.04	−.43*	−.25*	−.34*		
Avoidance	8.11 (4.61)	.91	−.11	−.28*	−.46*	−.42*	−.21*	−.16*	−.56*	−.48*	−.03	.37*	
Self-image	17.57 (4.60)	.86	.25*	.46*	.34*	.21*	.44*	.35*	.21*	.14*	.57*	−.51*	−.14*

MSFI Male Sexual Function Index, *PFSF* Profile of Female Sexual Function

* Correlation is significant at the .01 level

of the PFSF in both sexes.[4] Further supporting the utility of these measures, reliability analyses revealed adequate to very high internal consistencies among the scales of the three measures.

As reviewed in the Introduction, the validity of the FSFI and PFSF has largely been supported in numerous studies using older female samples with medical and/or sexual difficulties. To our knowledge, only two previous investigations have examined the FSFI in non-clinical women whereas the PFSF had yet to be examined in sexually healthy women. Additionally, with our use of confirmatory techniques, our findings supported the separation of the desire and arousal scales in the FSFI for use in young women. This research is important in that it advances the study of sexual health by providing empirical evidence supporting the appropriateness of administering the FSFI and PFSF to younger, healthier female populations. We believe that the ability to detect normal variations in sexual response is crucial to identifying biological and psychosocial influences and vulnerabilities that correspond to pre-morbid changes in sexual function. Therefore, we believe that our findings supporting the use of

the FSFI and PFSF in non-clinical samples are important to this etiological research.

This study was also the first examination to support the MSFI and PFSF for use in young, healthy men. As previously stated, the MSFI was created for this study to be maximally similar to the FSFI whereas the PFSF, as administered to our male sample, was unaltered from its original form. Empirical support of maximally similar measures for use in men and women will allow researchers to evaluate the relative influences on female and male sexual health. By using these parallel measures, we were elsewhere able to demonstrate that affective problems were differentially predictive of sexual functioning for women versus men (Kalmbach et al., 2012). Due to the numerous biological differences between men and women, we caution against equating scores across gender. However, the development of maximally similar measures better enables researchers to make comparisons across gender, without differences in measures confounding their results. For example, parallel measures will allow researchers to better examine such hypotheses as: Do men and women differ in their normative fluctuations in desire across the lifespan? Does life stress differentially predict sexual dysfunction for men and women? Our findings showed that the FSFI, MSFI, and the modified PFSF can be helpful answering these types of questions for researchers who are interested in male and female sexual functioning.

This study was also the first to propose a reconceptualization of the Responsiveness scale of the PFSF. In both female and male samples, analyses revealed that the PFSF did not conform to its original structure. Specifically, the first two

---

[4] It is important to clarify that our findings do not imply that sexual functioning is the same for both sexes or that there is a singular, unisex model of sexual functioning. In summary, our findings showed that the structure of the measures and the manner in which these items correlated with one another were similar for men and women. For reasons previously outlined, we initially used the female models as a priori models and found that they applied to men as well. If these structural models of the measures did not fit men, then our analyses would not have supported them and the MTMM CFAs would not have worked.

items did not correlate with the latter five items (even after accounting for item valence), results that were replicated in both genders. This finding resulted in removing the first two items of the Responsiveness scale to form an Avoidance scale; the renaming of the scale was based on inspection of the items that revealed a theme of sexual avoidance. This revised Avoidance scale was supported in both men and women. Importantly, our finding not only argues for future studies to administer or score the PFSF in such a way as to account for this modified avoidance scale, but the finding also invites future investigations to create a more coherent responsiveness scale. We contend that both sexual avoidance and responsiveness are important aspects of sexual behavior. As such, both constructs warrant research interest. Because empirical evidence exists in support of both the PFSF's responsiveness scale (McHorney et al., 2004) and avoidance scale (the present study), further psychometric studies of these scale conceptualizations are needed to examine their validity.

Lastly, this investigation supported the factor structure of these instruments, but it is important to note that this was enabled through the use of an MTMM approach to the factor analysis, which attempted to isolate the effects of item valence. Though this statistical approach may appear to simply account for nuisance variance in individual self-report, we believe that the significant impact of item valence across all measures for both genders has important psychometric implications for the study of sexual health. The superiority of MTMM models conceptually demonstrated that the wording of items led to observable differences in response. To illustrate, our findings suggest that asking a man how easily he can produce an erection will likely generate a response that is not simply the inverse of asking the same participant how difficult it is for him to produce an erection. Strongly agreeing with the former was empirically distinct from strongly disagreeing with the latter. To highlight the strength of the effect, even after employing multi-method techniques to account for this variance due to item valence, some factor loadings still showed differences between how positively and negatively worded items loaded onto factors of sexual functioning in which a mixture of both valences appear (see Tables 1, 2, 3 and 4). Based on those differences, the mixture of positive and negative items may have tempered the internal consistencies for these scales. Men appeared to particularly respond differently to positive and negative items on the erection and orgasm scales, which was reflected in their lower internal consistencies as compared to other factors. It remains uncertain whether writing items to detect *dys*function rather than writing items to detect healthy functioning produces more or less valid results for either men or women. Future studies are needed to examine these possibilities.

Though we believe this investigation contributes to our knowledge of normative sexual functioning and the assessment of sexual functioning, some study limitations should be

acknowledged. One limitation regards the omission of a premature ejaculation scale in either measure of male sexual functioning. If using the MSFI and/or PFSF, researchers studying sexual health in young men may want to use additional measures or scales, if interested in premature ejaculation, which is a common sexual complaint among men (Laumann, Paik, & Rosen, 1999). Also, future validation studies that wish to examine the utility of male sexual function measures for use in normal populations would benefit from examining measures that assess premature ejaculation. Furthermore, researchers have recently identified female premature orgasm as an area of interest in sex research (Carvalho et al., 2011). Thus, future investigations would benefit from not only examining premature ejaculation in men, but women as well. Additionally, the present study did not attempt to examine the utility of measuring male sexual pain in our young, healthy sample. Though male sexual pain receives little attention in the field of sexual health research (Davis, Binik, & Carrier, 2009; Luzzi & Law, 2006), our understanding and treatment of male sexual pain would improve with both increased attention to the phenomenon and better methods to measure it. Further, we cannot generalize our findings in our sample of young, healthy men and women to other populations. Notably, though our samples consisted of bisexual and homosexual men and women, these groups were not large enough to allow for examination of sexual orientation as a factor in our study.

Sexual health research would greatly benefit from future studies that build upon the findings of this study. Though much research exists supporting these measures in various female populations, the MSFI and PFSF warrant evaluation in male populations other than young adult men as studied here. Indeed, just as the validity of the FSFI and PFSF warranted examination in younger women, studies determining the applicability of the MSFI and PFSF to older and more sexually dysfunctional men deserve attention. Should these measures prove to be appropriate for these other male populations, the field of research examining determinants and consequences of male sexual difficulties would greatly benefit.

Additionally, to our knowledge, this study was the first to identify item valence and message framing as an important aspect to psychometrics in sex research. We observed that individuals responded to questions about their sexual functioning very differently depending on whether these questions were framed positively or negatively. Additionally, even after employing techniques to best account for variance due to positively and negatively worded items, which significantly improved all four models, our findings still showed some evidence of item valence playing a role in factor loadings (see Tables 1, 2, 3 and 4). Indeed, assessing the sexual response in functional and dysfunctional women and men remains difficult. Patient reported outcomes are becoming the primary endpoints for clinical trials in treatments for sexual

dysfunction since they are better suited to capture the multidimensional and subjective information collected in this research (Kingsberg & Althof, 2011). Therefore, much is unknown about the importance of this phenomenon and it is presently unknown whether or not one valence (i.e., positive or negative) is more valid than the other. Future studies are needed to address this question, comparing self-reported sexual functioning to functioning as assessed through an alternative methodology (e.g., corroborating reports from partners; physiological assessments). Also, it is possible that the item valence phenomenon may differ in other cultures with different attitudes toward sex. These findings not only highlight the importance of message framing in the psychometric validation of sexual functioning measures, but also to other areas of sexual research.

Importantly, the use of confirmatory techniques is what allowed us to partition out the method variance due to item valence. For this reason, we urge future investigations to employ similar techniques, as the ability to account for message framing variance may result in more consistent findings across studies. Additionally, as the FSFI, MSFI, and IIEF contain two-item scales, future studies should avoid using exploratory factor analysis for these measures, given that EFAs require that each scale have a minimum of three to five indicators. As such, we strongly urge future investigations of these measures to utilize confirmatory techniques to allow for validation of two-item scales, which may also yield more reliable findings across studies.

The implications of this study also extend to clinical practice. Indeed, clinicians now have empirical evidence showing that these measures are appropriate for use in these populations. Therefore, these tools can be used to evaluate function and change in function in patients with subclinical sexual concerns as well as diagnosed sexual disorders. Also, not only are maximally similar measures of sexual function between men and women important to research, but it can be particularly useful in couples counseling when the dyad is heterosexual. Having the ability to administer similar measures to both members of the couple allows for easy comparison. For instance, if both partners complete the same questionnaire, it allows the clinician to more easily and efficiently compare levels of, say, sexual desire or arousal between partners.

References

Borello-France, D., Dusi, J., O'Leary, M., Misplay, S., Okonski, J., Leng, W., … Chancellor, M. (2008). Test-retest reliability of the urge-urinary distress inventory and Female Sexual Function Index in women with multiple sclerosis. *Urologic Nursing, 28*, 30–35.

Brotto, L. A. (2009). The Female Sexual Function Index [Letter to the Editor]. *Journal of Sex and Marital Therapy, 35*, 161–163. doi:10.1080/00926230802712327.

Carvalho, S., Moreira, A., Rosado, M., Correia, D., Maia, D., & Pimentel, P. (2011). Female premature orgasm: Does this exist? *Sexologies, 20*, 215–220. doi:10.1016/j.sexol.2011.08.008.

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.

Davis, S. N., Binik, Y. M., & Carrier, S. (2009). Sexual dysfunction and pelvic pain in men: A male sexual pain disorder? *Journal of Sex and Marital Therapy, 35*, 182–205. doi:10.1080/00926230802716310.

DeRogatis, L. R., Rust, J., Golombok, S., Bouchard, C., Nachtigall, L., Rodenberg, C., … McHorney, C. A. (2004). Validation of the Profile of Female Sexual Function (PFSF) in surgically and naturally menopausal women. *Journal of Sex and Marital Therapy, 30*, 25–36. doi:10.1080/00926230490247183.

Dunn, K. M., Croft, P. R., & Hackett, G. I. (1999). Association of sexual problems with social, psychological, and physical problems in men and women: A cross sectional population survey. *Journal of Epidemiology and Community Health, 53*, 144–148. doi:10.1136/jech.53.3.144.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–298. doi:10.1037/1082-989X.4.3.272.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453. doi:10.1037/1082-989X.3.4.424.

Kalmbach, D. A., Ciesla, J. A., Janata, J. W., & Kingsberg, S. A. (2012). Specificity of anhedonic depression and anxious arousal with sexual problems among sexually healthy young adults. *Journal of Sexual Medicine, 9*, 505–513. doi:10.1111/j.1743-6109.2011.02533.x.

Kennedy, S. H., Dickens, S. E., Eisfeld, B. S., & Bagby, R. M. (1999). Sexual dysfunction before antidepressant therapy in major depression. *Journal of Affective Disorders, 56*, 201–208. doi:10.1016/S0165-0327(99)00050-6.

Kingsberg, S. A., & Althof, S. E. (2011). Satisfying sexual events as outcome measures in clinical trial of female sexual dysfunction. *Journal of Sexual Medicine, 8*, 3262–3270. doi:10.1111/j.1743-6109.2011.02447.x.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher & C. Schatsschneider (Eds.), *Applied quantitative analysis in the social sciences* (pp. 171–207). New York: Routledge.

Laumann, E. O., Paik, A., & Rosen, R. C. (1999). Sexual dysfunction in the United States: Prevalence and predictors. *Journal of the American Medical Association, 281*, 537–544. doi:10.1001/jama.281.6.537.

Levine, S. B. (2003). The nature of sexual desire: A clinician's perspective. *Archives of Sexual Behavior, 32*, 279–285. doi:10.1023/A:1023421819465.

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 and Marital Therapy, 32*, 255–266. doi:10.1080/00926230600575348.

Luzzi, G. A., & Law, L. A. (2006). The male sexual pain syndromes. *International Journal of STD and AIDS, 17*, 720–726. doi:10.1258/095646206778691220.

Masheb, R. M., Lozano-Blanco, C., Kohorn, E. I., Minkin, M. J., & Kerns, R. D. (2004). Assessing sexual function and dyspareunia with the Female Sexual Function Index (FSFI) in women with vulvodynia. *Journal of Sex and Marital Therapy, 30*, 315–324. doi:10.1080/00926230490463264.

McHorney, C. A., Rush, J., Golombok, S., Davis, S., Bouchard, C., Brown, C., … DeRogatis, L. (2004). Profile of Female Sexual Function: A patient-based, international, psychometric instrument for the assessment of hypoactive sexual desire in oophorectomized women. *Menopause, 11*, 474–483. doi:10.1080/00926230490247183.

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 and Marital Therapy, 29*, 39–46. doi:10.1080/713847100.

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

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.

Rosen, R. C. (2000). Prevalence and risk factors of sexual dysfunction in men and women. *Current Psychiatry Reports, 2*, 189–195. doi:10.1007/s11920-996-0006-2.

Rosen, R. C., & Bachmann, G. A. (2008). Sexual well-being, happiness, and satisfaction, in women: The case for a new conceptual paradigm. *Journal of Sex and Marital Therapy, 34*, 291–297. doi:10.1080/00926230802096234.

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

Rosen, R. C., Riley, A., Wagner, G., Osterloh, I. H., Kirkpatrick, J., & Mishra, A. (1997). The International Index of Erectile Function (IIEF): A multidimensional scale for assessment of erectile dysfunction. *Urology, 49*, 822–830. doi:10.1016/S0090-4295(97)00238-0.

Velicer, W. F., & Fava, J. L. (1998). Affects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*, 231–251. doi:10.1037/1082-989X.3.2.231.

Verit, F. F., & Verit, A. (2007). Validation of the Female Sexual Function Index in women with chronic pelvic pain. *Journal of Sexual Medicine, 4*, 1635–1641. doi:10.1111/j.1743-6109.2007.00604.x.

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