ORIGINAL PAPER

Sex Differences in Patterns of Genital Sexual Arousal: Measurement Artifacts or True Phenomena?

Kelly D. Suschinsky · Martin L. Lalumière · Meredith L. Chivers

Abstract Sex differences in patterns of sexual arousal have been reported recently. Men's genital arousal is typically more category-specific than women's, such that men experience their greatest genital arousal to stimuli depicting their preferred sex partners whereas women experience significant genital arousal to stimuli depicting both their preferred and non-preferred sex partners. In addition, men's genital and subjective sexual arousal patterns are more concordant than women's: The correlation between genital and subjective sexual arousal is much larger in men than in women. These sex differences could be due to low response-specificity in the measurement of genital arousal in women. The most commonly used measure of female sexual arousal, vaginal photoplethysmography, has not been fully validated and may not measure sexual arousal specifically. A total of 20 men and 20 women were presented with various sexual and non-sexual emotionally laden short film clips while their genital and subjective sexual arousal were measured. Results suggest that vaginal photoplethysmography is a measure of sexual arousal exclusively. Women's genital responses were highest during sexual stimuli and absent during all non-sexual stimuli. Sex differences in degree of category-specificity and concordance were replicated: Men's genital responses were more category-specific than women's and men's genital and subjective sexual arousal were more strongly correlated than women's. The results from the current study support the continued use of vaginal photoplethysmography in investigating sex differences in patterns of sexual arousal.

Introduction

Two commonly used objective measures of sexual arousal are vaginal photoplethysmography (VPP) and phallometry. Recent research using these measures suggests that there are several intriguing sex differences in patterns of sexual arousal. One such sex difference pertains to the nature of stimuli that elicit genital and subjective sexual arousal (i.e., feelings of sexual arousal). Men's arousal patterns, both genital and subjective, are category-specific, such that men show the highest level of arousal to stimuli involving their declared preferred sex partners and preferred sexual activities. For example, gay men typically show the highest genital and subjective sexual arousal to erotica depicting two men engaging in sexual activities and little arousal to stimuli involving only women, whereas heterosexual men typically show the highest genital and subjective sexual arousal to stimuli involving women and little arousal to erotica involving only men (e.g., Freund, Langevin, Cibiri, & Zajac, 1973; Rieger, Chivers, & Bailey, 2005).

In comparison, women's patterns of genital arousal are markedly different. Chivers, Rieger, Latty, and Bailey (2004) presented heterosexual and homosexual men and women, as well as post-operative male-to-female transsexuals, with a variety of sexual films. Unlike men and male-to-female transsexuals, a substantial proportion (35%) of women experienced their greatest genital arousal to sexually explicit

K. D. Suschinsky (✉) · M. L. Lalumière
Department of Psychology, University of Lethbridge,
Lethbridge, AB, Canada T1K 3M4
e-mail: kelly.suschinsky@uleth.ca

M. L. Chivers
Law and Mental Health Program, Centre for Addiction and
Mental Health, Toronto, ON, Canada

films depicting actors that did not correspond with their self-reported sex partner preferences. For subjective sexual arousal, however, women reported their greatest arousal to films that corresponded with their self-reported sexual orientation.

Chivers and Bailey (2005) presented heterosexual men and women with the same sexual stimuli from Chivers et al. (2004), as well as a sexual stimulus depicting male and female bonobos engaging in repeated penile–vaginal intercourse. Men did not respond genitally or subjectively to the non-human sexual stimulus. Women's genital responses indicated that they were significantly more aroused to the non-human sexual stimulus than to the neutral stimuli, but they did not report feeling as such. Women showed similarly high genital responses to all three forms of human sexual interactions presented (i.e., male–male, female–female, and male–female), but their highest subjective responses were to the male–female stimuli.

A second sex difference in arousal patterns concerns the concordance between genital and subjective measures of sexual arousal. Men show higher correlations between genital and subjective measures of sexual arousal than women. A recent meta-analysis of all studies reporting a correlation between genital and subjective sexual arousal in men and women indicated a positive and significant association in both sexes, but a higher association in men ($r = .68$ and $r = .31$, respectively). Lower concordance in women is also evident when participants are asked to attend to their genital sensations (Chivers, Seto, Lalumière, Laan, & Grimbos, 2007). Low concordance in women is not strictly a function of poor genital category-specificity because low concordance is evident in women who are presented with only stimuli involving their preferred sex partners (Wincze, Venditti, Barlow, & Mavissakalian, 1980).

One plausible explanation for the sex differences noted above is that VPP may not measure genital arousal as accurately as phallometry. Low category-specificity of genital responses in heterosexual women (i.e., the population of women most often studied using VPP) suggests poor discriminative validity, and low concordance suggests poor concordant validity. In an extensive review of measures of sexual arousal, Rosen and Beck (1988) concluded that measures of penile response are both reliable and valid, but that the reliability and validity of VPP had not been satisfactorily established. Although some research has examined the validity of VPP since Rosen and Beck's review (e.g., Laan, Everaerd, & Evers, 1995), there remain outstanding concerns that may compromise interpretations of research using this device (e.g., Levin, 1998).

It is clear that penile responses are specific to sexual stimuli because these responses only occur in the presence of stimuli involving sexual images or activity (for a review, see Zuckerman, 1971). It is unclear, however, if the same is true for VPP. Perhaps genital responses measured by VPP

show little category-specificity or concordance with subjective sexual arousal because subjectively non-preferred sexual stimuli (e.g., films depicting bonobos having sex) generate a non-sexual physiological response (i.e., increase in blood flow) that is detected by VPP.

Little research has been conducted to assess the response-specificity of VPP. Determining the response-specificity of VPP to sexual stimuli entails examining whether only sexual stimuli, versus other emotionally arousing stimuli, evoke increases in vaginal blood flow. The first study relevant to this question was conducted by Geer, Morokoff, and Greenwood (1974). Fourteen women were presented with two 8-min films, one erotic film depicting a heterosexual couple involved in a sexual interaction that included undressing, oral sex, and penile–vaginal intercourse, and the other a non-erotic film depicting battles and court life scenes from the Crusades. Compared to baseline levels of genital response, the erotic film elicited significant increases in vaginal pulse amplitude (VPA) and vaginal blood volume (VBV); the former reflects short-term changes in vaginal vasocongestion due to the amount of blood present in the vaginal tissue during each heart beat, whereas the latter reflects slow changes in the pooling of vaginal blood (Hatch, 1979). The non-erotic film also elicited significant increases in both VPA and VBV responses, but the erotic film produced significantly higher responses in VPA and VBV relative to the non-erotic film.

Hoon, Wincze, and Hoon (1976) assessed the response-specificity of several physiological measures of arousal using a sexually explicit film depicting heterosexual foreplay and control films. Control films included a neutral film of an oceanographic lecture and an emotionally dysphoric film depicting Nazi war crimes with images of dismembered bodies. The results indicated that VBV responses were significantly influenced by stimulus category: They were the largest during the sexual film, and absent during the two non-sexual films. Self-reported emotions corresponded with the content of the films: The sexual film produced the most subjective sexual arousal and the dysphoric film produced the most anxiety.

The most recent research directly examining the response-specificity of VPP was conducted by Laan et al. (1995). They presented 49 women with four film clips: a sexual clip depicting petting, cunnilingus, and intercourse; a sexually threatening clip depicting a man chasing a woman up a set of stairs and kissing her and touching her against her will; a non-sexual, anxiety-inducing clip depicting a woman being chased by a rabid dog; and a neutral clip depicting old buildings from a Dutch village. Participants rated themselves as most sexually aroused during the sexual film clip, somewhat aroused during the sexually threatening film, and not at all aroused during the non-sexual films. Both VPA and VBV responses were highest during the sexual stimuli, but VBV responses did not differentiate between the neutral stimulus

and the sexually threatening stimulus. VPA responses did not increase at all during the neutral or the non-sexual anxiety-inducing film clip, but VBV responses increased slightly to the neutral film clip. VPA and VBV responses were both significantly correlated with subjective sexual arousal during the sexual and sexually threatening stimuli. VBV responses, however, were also related to other emotions, including anxiety and threat.

These studies provide partial support for the notion that VPP (or at least the VPA component) is a measure specific to sexual arousal. One important limitation of these studies is that they included a very narrow range of control (non-sexual) stimuli, and none of the stimuli were selected to elicit positive affect. The problem associated with using only neutral and negatively valenced control stimuli is that the emotional responses they elicit do not approximate the positive emotionally valenced responses women usually experience during sexual arousal (e.g., enjoyment; Heiman, 1980). Because women typically experience positive affective responses during sexual arousal, a stringent test of the response-specificity of VPP would be to compare women's responses to sexual stimuli with responses to non-sexual stimuli of a positive valence, in addition to non-sexual stimuli of a negative valence.

Chivers et al. (2004) noted that stimuli depicting homosexual sexual interactions may be unusual or aversive to some heterosexual people and may produce a negative emotional response, such as anxiety or disgust, which could influence VPP readings and, in turn, could result in low category-specificity of sexual arousal. In early validation studies, however, anxiety-inducing (Laan et al., 1995) or dysphoric films (Hoon et al., 1976) did not elicit genital responses in women. It is possible, however, that the heterosexual women watching homosexual sexual interactions did not react negatively, but rather with amusement or other positive emotions, resulting in increased VPA responses to the films of homosexual sex. There is some evidence to support this possibility. Using clitoral ultrasonography, Kukkonen et al. (2006) found that the clitoral blood flow responses of women who were presented with a humorous film could not be distinguished from the responses of women who were presented with a sexual film. Perhaps then, women are experiencing a physiological response related to amusement or humor when viewing some films which results in increased blood flow to the genital area, and thus lesser category-specificity and concordance compared to men.

The purpose of the current study was twofold. The first objective was to determine whether VPP, the most commonly used measure of female sexual arousal, is a specific measure of sexual arousal or a measure of general arousal related to positive affect by exposing women to sexual and non-sexual stimuli of a positive and negative valence. It was expected that if VPP measures only sexual arousal, then only the sexual films will elicit a genital response. The second

objective was to attempt to replicate past research describing differences in arousal patterns between men and women. We expected that men would show more genital category-specificity and greater concordance between genital and subjective sexual arousal than women. We did not expect any sex differences in subjective sexual arousal patterns.

Method

Participants

A total of 23 men and 23 women were recruited from the University of Lethbridge and the community of Lethbridge using newspaper advertisements, posters, and visits to university classes. To be eligible for the study, participants were required to be predominantly or exclusively heterosexual, between 18 and 28 years of age, with no history of sexual arousal problems, sexually transmitted diseases, mental illnesses, or substance abuse, and not using medications to treat high blood pressure or a mental illness during the study period. Participants were also required to be in an intimate relationship for at least 6 months, sexually experienced (i.e., must have engaged in sexual intercourse before and have been exposed to erotica prior to the experiment), and nulliparous. Also, women must have reported a regular menstrual cycle and could not be pregnant or menstruating at the time of testing. Women using hormonal contraceptives were included. Data from six participants were excluded for one of the following reasons: self-reported bisexual sexual orientation on questionnaire items (two women), genital responses to the neutral stimuli that were higher than to any other stimulus category (i.e., sexual or non-sexual; one woman), not in an intimate relationship (two men), and technical failure (one man).[1]

The mean ages of the 20 men and 20 women with valid data were 22.0 ($SD = 2.5$) and 21.9 ($SD = 3.0$) years, respectively. Most men and women reported being in a dating relationship (85% and 80%, respectively), with the remainder reporting being engaged to be married (10% and 5%, respectively) or in a common-law relationship (5% and 15%, respectively). On average, women reported being in slightly longer relationships ($M = 21.9$ months, $SD = 17.7$ months) than men ($M = 19.4$ months, $SD = 15.0$ months). The majority of men and women were Caucasian (85% and 100%, respectively) and most of the men and

---

[1] Three participants did not fully meet one of the inclusion criteria but their data were used for analyses. One man reported having a history of mental illness but was not on medication at the time of testing. One man reported being in a monogamous relationship for three months. One woman reported using medication to treat a minor mental illness, but she had been on the same low dose of medication for a very long period of time.

women tested had completed or were completing post-secondary education at the time of testing (80% of both men and women). There were no statistically significant sex differences on any of these biographic factors. Sixty percent of the women reported using hormonal contraceptives.

The Kinsey Scale was used to assess both men and women's sexual partner preferences (Kinsey, Pomeroy, & Martin, 1948; Kinsey, Pomeroy, Martin, & Gebhard, 1953). Using a 7-point Likert scale, participants were asked to report the sex of the people they were most often romantically and sexually attracted to, as well as the sex of those people they had sexual fantasies about. Response options ranged from exclusive interest in same sex partners to exclusive interest in opposite sex partners. Based on their self-reported responses, both men and women were classified as predominantly or exclusively heterosexual. Women's scores, however, appeared to be more variable than men's scores. Most men and women reported that their romantic attractions were exclusively toward members of the opposite sex ($n = 19$ for both men and women), but reports for sexual attraction and sexual fantasy were somewhat different: Only one man reported any sexual attraction or fantasy involving same-sex individuals, but a substantial proportion of women reported some sexual attraction or fantasy involving same-sex individuals ($n = 8$ and $n = 13$, respectively). In addition to the Kinsey Scale, participants were also asked to report their sexual identity as gay, lesbian, bisexual, heterosexual, asexual, or other. Thirty-nine participants reported their sexual identity as heterosexual; one man reported his sexual identity as "other."

Measures

Audiovisual Stimuli

Audiovisual presentations were used because past research has indicated that they elicit the highest genital and subjective responses, in comparison to other forms of media, such as slides or audiotapes (e.g., Abel, Barlow, Blanchard, & Mavissakalian, 1975; Heiman, 1980). Film clips were selected based on the results obtained from a pilot study that was conducted to ensure that the clips elicited the intended emotional responses (e.g., anxiety, exhilaration, sexual arousal). The film clips were all approximately 90 s long and focused on a female character, when appropriate. All clips were taken from commercially available films (titles are available from the corresponding author).

There were 10 stimulus categories and two exemplars per category: neutral (e.g., beach scene of waves lapping at the shore); happy (e.g., woman singing and dancing with children); exhilarating (e.g., first person perspective of roller coaster ride); sad (e.g., woman being told her husband has died); anxiety-inducing (e.g., woman being chased by a rabid dog); low intensity male–female sexual interaction (e.g., partially clothed man and woman kissing and embracing); male–female sexual threat (e.g., a woman being raped); male–male sexual interactions (e.g., two men engaging in fellatio); female–female sexual interactions (e.g., two women engaging in cunnilingus); and male–female sexual interactions (e.g., a man and woman engaging in oral sex). Two additional clips (one neutral and one low intensity male–female sexual interaction) were used as adaptation stimuli.

Genital Measures

All psychophysiological data were sampled continuously throughout each film clip using a Limestone Technologies Inc. (Kingston, ON) DataPac_USB and Preftest software, Version 10. Women's genital arousal was assessed with changes in VPA using a vaginal photoplethysmograph equipped with an orange-red spectrum light source (Technische Handelsonderneming Coos, The Netherlands). The photoplethysmograph signal was sampled at a rate of 10 samples/s, band-pass filtered (.5–10 Hz), and digitized (40 Hz). Baseline was captured at the beginning of each stimulus and the peak response corresponded to the largest peak-to-trough distance during the stimulus presentation. A placement device made of flexible silicone was attached to the cable of the vaginal gauge and placed at a distance of 5 cm from the phototransistor (i.e., light detector). The placement device was used to control the depth and the orientation of the gauge (Laan et al., 1995). Movement artifacts were detected through visual inspection of the waveforms and removed prior to data analysis.[2]

Men's genital arousal was measured using mercury-in-rubber strain gauges of three sizes (75, 80, and 85 mm; D. M. Davis, New Jersey). The signal was sampled at a rate of 10 samples/s, low-pass filtered (to .5 Hz), and digitized (40 Hz). The signal was transformed into mm of circumference. Baseline was captured at the beginning of each stimulus and the peak response corresponded to the highest circumference value during the stimulus presentation. The gauges were calibrated over six 5 mm steps for each participant. Movement artifacts were detected through visual inspection of the response curves and removed prior to data analysis.

---

[2] Removal of movement artifacts from all physiological data was conducted by the first author, blind to stimulus categories. On average, the number of trials requiring editing for women ($M = 13.6$, $SD = 5.03$) was significantly higher than the number of trials requiring editing for men ($M = .5$, $SD = 1.60$), $t(38) = 11.09$, $p < .001$, $d = 3.51$.

Subjective Sexual Arousal

During the presentation of the experimental stimuli, participants continuously rated their subjective sexual arousal (i.e., how sexually aroused they felt) via a button press on a keypad. Participants saw their subjective ratings change throughout the film clip on a vertical bar on the left side of the same monitor that showed the films. Increases in subjective sexual arousal were indicated by an increase in the height of the bar and decreases in arousal were indicated by a decrease in the bar's height. Continuous subjective sexual arousal was reported as percentage of arousal, from 0 (no sexual arousal) to 100 (the feeling experienced immediately before orgasm).[3]

Post-stimulus Questions

After the presentation of each film, participants viewed questions on a computer monitor. Participants were asked to rate how pleasant or unpleasant each film clip was, how much attention they paid to the film, the intensity of the film, and the extent to which they experienced various emotions (e.g., sexual arousal, happiness, sadness, exhilaration, anxiety, boredom, calmness) using a button press on a keypad. Participants were asked to respond to the questions using a scale of 1–9, with 1 being the lowest level of response (i.e., the emotion in question was definitely not present, or the film was not at all intense) and 9 being the highest level of response (i.e., the emotion was definitely present or the film was definitely intense). Questions were presented in a fixed order, due to the nature of the software being used. The very first question asked participants to rate their subjective sexual arousal.

Procedure

All experimental procedures were approved by the University of Lethbridge's Human Subject Research Committee.

Screening

Prospective participants responded to advertisements via telephone or e-mail and received preliminary information about the study. Prior to participating in the experiment, the first author asked prospective participants a series of questions during a telephone interview to determine their eligibility.

Those who met the inclusion criteria and were interested in participating after receiving a description of the experimental procedure scheduled an appointment. Women were scheduled such that they were participating when they were not menstruating. Prior to testing, participants were asked to refrain from sexual activity of all types for 24 h, physical exercise of all types for 60 min, because exercise results in sympathetic nervous system arousal that can influence genital responses (Meston & Gorzalka, 1996), and using alcohol, tobacco, caffeine, cold medications, and recreational drugs on the day of testing, because these substances may influence both physiological and psychological sexual arousal.

Experimental Session

Participants were tested individually. Upon arrival at the laboratory, a female experimenter greeted the participant and explained the experimental procedure and how to insert or attach the genital gauge. After obtaining consent, the experimenter left the participant alone in the dimly lit room, where the participant undressed from the waist-down and inserted or attached the genital gauge while seated in a comfortable recliner. The experimenter communicated with the participant using text messages.

Participants watched the adaptation stimuli and all experimental stimuli on a 17 in computer monitor positioned five feet away at eye-level. Adaptation stimuli (i.e., one neutral and one low intensity sexual stimulus) were presented to familiarize the participant with the experimental setting and data collected during this time were not used in the analyses reported below. The experimental stimuli were presented in a quasi-random order that differed for each participant. One stimulus from each category was presented in the first half of trials and the other was presented in the second half, and stimuli from the same category were separated by at least one other stimulus. Participants were asked to respond to the films as naturally as possible, and to avoid contracting their muscles, manipulating their responses, touching their genitals, moving, or talking during the films. During the presentation of each film, participants rated their subjective sexual arousal continuously. Likewise, participants rated their subjective sexual arousal (in addition to other emotions) after each film clip had ended. Both continuous and static measures of subjective sexual arousal were assessed because some research indicates that sex differences in concordance depend on which measure of subjective sexual arousal is used (e.g., Chivers et al., 2007; Wincze et al., 1980).

Stimuli were separated by an interval of at least 30 s. Participants whose genital arousal levels did not return to baseline within 2 min were instructed to complete a distracter task in an attempt to lower arousal levels (e.g., read out loud from a nature magazine or count out loud backwards

---

[3] Only one woman reported that she had never experienced an orgasm during any form of sexual activity, with or without a partner. She did report slightly increased levels of subjective sexual arousal for both continuous and post-stimulus subjective sexual arousal for some sexual stimuli.

from 100 in different multiples). If arousal did not return to baseline levels after a maximum of 5 min, the next stimulus was presented regardless of genital arousal level. This did not occur often (19 out of 800 trials or 2.4%).

After all experimental stimuli were presented, participants were asked to remove the genital gauge and place it in a sealable plastic bag, re-dress, and complete the questionnaire. Upon completion of the questionnaire, the experimenter debriefed the participant, thanked him or her for participating, and compensated the participant with $50. The entire experimental session took approximately two and a half hours. No participants reported fatigue or boredom during the debriefing session, and there was no evidence to indicate any participants had fallen asleep during the testing procedure.

Data Analysis

Peak minus baseline scores for each stimulus were calculated for genital responses and continuous subjective sexual arousal responses. These scores were then standardized (i.e., transformed into z-scores within-subjects) for each type of response separately in order to eliminate effects of individual variation in responsiveness (Harris, Rice, Quinsey, Chaplin, & Earls, 1992). The z-scores were then averaged to produce mean scores for each stimulus category (e.g., happy, sad, etc.). Post-stimulus responses to questions were averaged for each category and not standardized. To test whether vaginal photoplethysmography measures sexual arousal or general arousal, separate 2 (sex) $\times$ 10 (stimulus category) analyses of variance (ANOVA) were conducted for genital and continuous subjective arousal responses. Planned contrasts were used to compare differences in responses to the sexual, non-sexual, and neutral stimuli.

To test whether men exhibited greater specificity in their responses compared to women, two indices were computed. The first index was a measure of category-specificity derived from Chivers et al. (2004). The category-specificity index, or *female–male contrast*, was computed separately for genital and subjective sexual arousal by subtracting the average response to the male–male stimuli from the average response to the female–female stimuli. Greater arousal to female–female stimuli resulted in a positive score, and greater arousal to male–male stimuli resulted in a negative score.

The second index was a measure of the amount of sexual arousal produced by one sexual stimulus category relative to all other sexual stimulus categories. The *stimulus-specificity index* was computed separately for genital and subjective sexual arousal by taking the response to the sexual stimulus category that generated the highest responses and then subtracting the average of the responses to the remaining sexual stimulus categories. For example, if a male participant

responded the most to the female–female stimuli, the average of his responses to the male–male, male–female, low intensity, and sexually threatening categories would be subtracted from his response to the female–female stimuli. Higher scores indicate greater specificity, such that one category of stimuli captured more of the overall arousal output produced by the participant. The sex difference in category-specificity was examined by comparing the average of the female–male contrast score and stimulus-specificity index for each sex using independent samples t-tests.

Non-parametric correlations (Spearman's rho) were used to test whether men exhibited higher concordance between genital and subjective responses than women. Each participant provided one correlation between genital and continuous subjective sexual arousal, and one correlation between genital and post-stimulus subjective sexual arousal. These correlations were based on 20 pairs of data, one pair for each stimulus presentation. The sex difference was examined by comparing the average correlations for each sex using an independent samples t-test.

Results

Manipulation Check

To ensure that the film clips elicited the intended emotions, participants were asked questions regarding their emotional reactions following the presentation of each film clip (see Table 1). Due to technical failure, post-stimulus responses were not available for one female participant. Repeated measures ANOVAs on the post-stimulus ratings revealed that the non-sexual stimulus categories produced different ratings on all six questions shown in Table 1, $F_{(4, 152)} = 16.02–109.31$, ps $< .001$. Bonferroni tests showed that both the happy and the exhilarating clips were rated as significantly more pleasant than both the sad and anxiety-inducing clips (ps $< .001$). Similarly, the sad and anxiety-inducing clips were rated as significantly more unpleasant than the happy and exhilarating clips (ps $< .001$). The happy film clips elicited significantly higher self-reported levels of happiness relative to all other non-sexual film clips, except for the exhilarating film clips (ps $< .01$). The sad film clips elicited significantly higher reports of sadness compared to all other non-sexual film clips (ps $< .001$). The exhilarating film clips resulted in significantly higher levels of self-reported exhilaration compared to all non-sexual film clips, except for the anxiety-inducing film clips (ps $< .001$). Finally, the anxiety-inducing film clips produced significantly higher levels of anxiety than all other non-sexual film clips (ps $< .01$). Figure 1 shows that only sexual film clips elicited subjective sexual arousal.

Table 1 Means and SDs of post-stimulus ratings of various emotions

Questions	Positive stimulus categories						Negative stimulus categories			
	Happy		Exhilarating		Neutral		Sad		Anxiety	
	M	SD	M	SD	M	SD	M	SD	M	SD
How pleasant did you find the film?	6.97	1.54	6.31	1.78	6.32	1.95	2.47	1.27	2.37	1.35
How unpleasant did you find the film?	1.35	.78	1.45	.83	1.45	1.18	5.47	1.85	5.37	1.90
How happy do you feel?	6.92	1.47	6.40	1.65	5.81	1.66	2.58	1.72	3.10	1.90
How sad do you feel?	1.27	.46	1.38	.87	1.33	.52	5.33	1.95	3.04	1.58
How exhilarated do you feel?	3.23	1.65	4.85	1.82	2.28	1.40	2.18	1.35	4.13	2.01
How anxious do you feel?	2.23	1.69	2.88	1.78	2.29	1.66	3.06	1.49	4.29	2.22

Note: Mean ratings were calculated by computing an average of the responses to the two films belonging to each stimulus category. Ratings were made on a scale of 1 (emotion is not at all present) to 9 (emotion is definitely present)
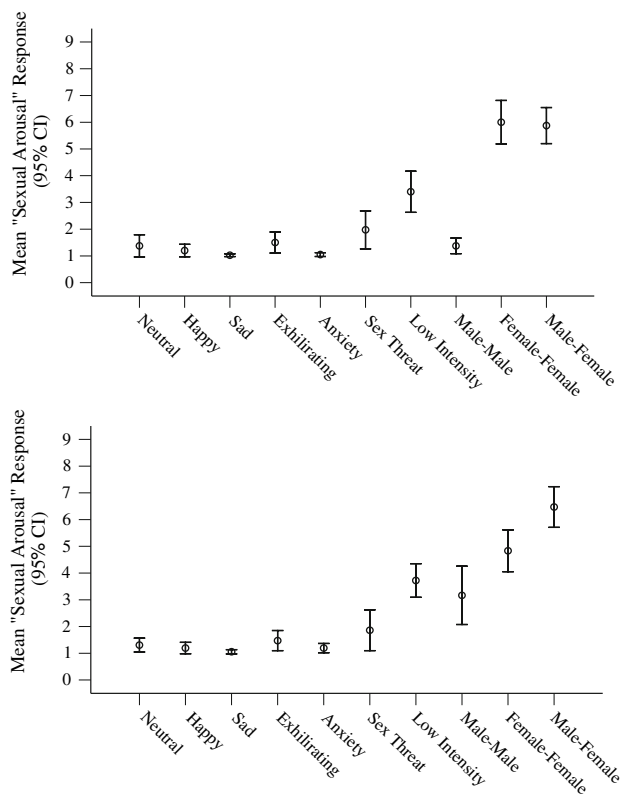
Fig. 1 Mean post-stimulus subjective sexual arousal responses (raw data) in men (top portion) and women (bottom portion)

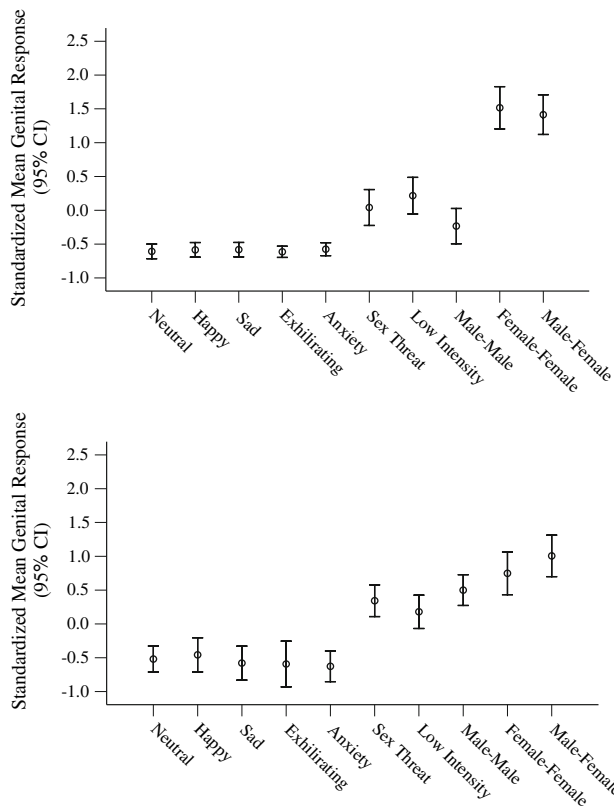## Genital and Subjective Sexual Arousal Responses to Sexual Stimuli

Genital Responses

Figure 2 presents the standardized mean genital responses for men and women. The figure shows that participants responded equally to the emotional non-sexual stimulus categories and the neutral stimuli, and showed higher responses to the sexual stimulus categories. A 2 (sex) × 10 (stimulus category) mixed ANOVA with repeated measures on the second factor confirmed this visual inspection. There was a significant main effect of stimulus category, $F(9, 342) = 69.10$, $p < .001$, $\eta_p^2 = .63$, and a significant sex by stimulus category interaction, $F(9, 342) = 5.44$, $p < .001$, $\eta_p^2 = .13$. No significant differences in overall genital responses were found between men and women because scores were standardized within subjects (comparisons of absolute levels of genital response between sexes is not appropriate here because the genital response units differ markedly).[4]

Simple effects were examined separately for men and women. Stimulus category was significant for both men and women, $F(9, 171) = 60.29$, $p < .001$, $\eta_p^2 = .76$ and $F(9, 171) = 22.37$, $p < .001$, $\eta_p^2 = .54$, respectively. Planned contrasts revealed that, for men, the sexual stimuli (i.e., sexual threat, low intensity, male–male, female–female, and male–female) produced greater genital responses compared to both the neutral stimuli, $F(1, 171) = 17.58$, $p < .001$, Cohen's $d = 2.63$, and the non-sexual stimuli (i.e., happy, sad, exhilarating, and anxiety-inducing), $F(1, 171) = 103.96$, $p < .001$, $d = 2.62$. The non-sexual and neutral stimuli did not produce significantly different genital responses, $F(1, 171) < 1$, $d = .08$. Similarly, for women, planned contrasts revealed that the sexual stimuli produced greater genital responses compared to both the neutral stimuli, $F(1, 171) = 10.27$, $p < .005$, $d = 1.90$, and the non-sexual stimuli, $F(1, 171) = 64.19$, $p < .001$, $d = 1.73$. Again, the non-sexual and neutral stimuli did not produce significantly different genital responses, $F(1, 171) < 1$, $d = .08$. An examination of unstandardized scores (peak minus baseline) showed that men and women exhibited zero or near-zero responses to non-sexual (including neutral) stimuli.

---

[4] The Greenhouse-Geisser correction was used when necessary (i.e., when the assumption of sphericity was not met) for all ANOVAs and always yielded the same results. Uncorrected values are reported.
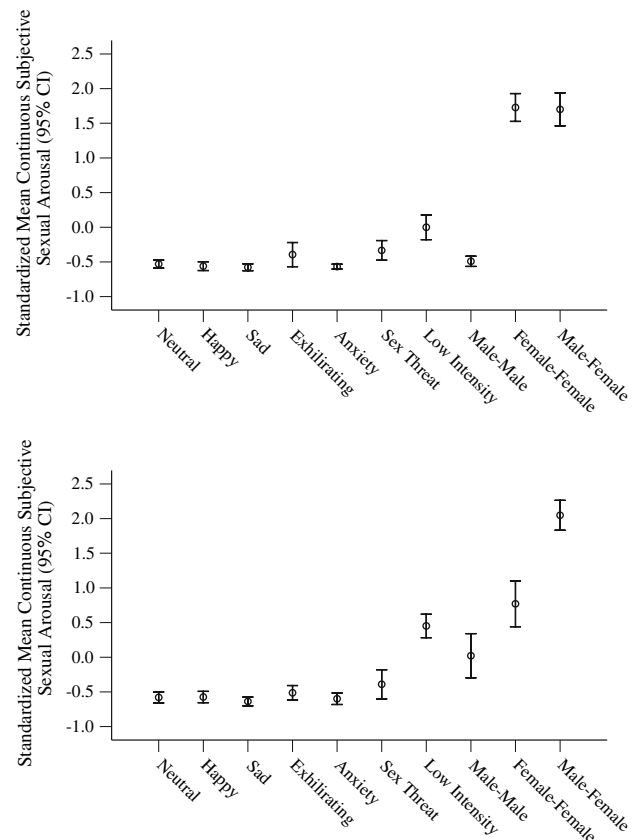
Fig. 2 Standardized mean genital responses in men (top portion) and women (bottom portion) as a function of stimulus category

Continuous Subjective Sexual Arousal

Two participants (one man and one woman) did not report any changes in subjective sexual arousal during any film clip. Data from these participants were not included in analyses of continuous subjective sexual arousal because *z*-scores could not be calculated (due to the fact that their self-reported sexual arousal showed no variation). As mentioned above, data from one female participant were unavailable due to technical failure during the experimental procedure. Figure 3 presents the standardized mean subjective sexual arousal scores for men and women. The figure suggests that both men and women reported their highest level of subjective sexual arousal to the sexual stimuli. A 2 (sex) × 10 (stimulus category) ANOVA revealed a main effect of stimulus category, $F(9, 315) = 220.19$, $p < .001$, $\eta_p^2 = .86$, and a significant sex by stimulus category interaction, $F(9, 315) = 12.26$, $p < .001$, $\eta_p^2 = .26$. No significant sex difference was found because scores were standardized within subjects.

Simple effects were conducted for each sex and indicated a significant effect of stimulus category for both men, $F(9, 162) = 172.06$, $p < .001$, $\eta_p^2 = .91$, and women, $F(9, 153) = 83.34$, $p < .001$, $\eta_p^2 = .83$. Planned contrasts revealed that, for men, the sexual stimuli produced higher subjective

Fig. 3 Standardized mean continuous subjective sexual arousal responses in men (top portion) and women (bottom portion) as a function of stimulus category

sexual arousal than the neutral stimuli, $F(1, 162) = 37.27$, $p < .001$, $d = 3.87$, and the non-sexual stimuli, $F(1, 162) = 24.72$, $p < .001$, $d = 3.55$. Non-sexual and neutral stimulus categories did not produce significantly different subjective sexual arousal, $F(1, 162) < 1$, $d = .02$. Similarly, for women, planned contrasts revealed that the sexual stimuli produced higher subjective sexual arousal than the neutral stimuli, $F(1, 153) = 23.58$, $p < .001$, $d = 2.23$, and the non-sexual stimuli, $F(1, 153) = 15.74$, $p < .001$, $d = 2.22$. The non-sexual and neutral stimulus categories did not produce significantly different subjective sexual arousal, $F(1, 153) < 1$, $d = .01$.

When considering the unstandardized subjective sexual arousal responses, independent samples *t*-tests revealed significant differences in reported sexual arousal between men and women for the male–male and female–female stimulus categories only. Women reported significantly more subjective arousal ($M = 13.09$, $SD = 19.89$) in response to the male–male stimuli than men did ($M = 1.97$, $SD = 4.41$), $t(37) = 2.44$, $p < .05$, $d = .78$. Men, however, reported significantly more subjective sexual arousal ($M = 49.08$, $SD = 19.23$) in comparison to women ($M = 20.05$, $SD = 15.01$) when viewing the female–female stimuli, $t(37) = 5.24$, $p < .001$, $d = 1.68$.

Category-specificity of Men and Women's Genital and Subjective Sexual Arousal

The patterns of genital arousal shown in Fig. 2 suggest that men's responses were focused on certain sexual stimulus categories, whereas women's responses were more distributed over all sexual stimulus categories. Men showed a clear preference for female–female and male–female stimuli, whereas women showed only a slight preference for these categories. In contrast, both men and women showed clear preferences with regard to subjective sexual arousal (Fig. 3). These patterns suggest that men were more category-specific than women with regard to their genital arousal, but not for subjective sexual arousal.

Genital Arousal Category-specificity

Two independent samples *t*-tests were conducted to assess the sex differences in category-specificity for genital responses.[5] For female–male contrast scores (i.e., response to female–female stimuli relative to male–male stimuli), men had significantly higher scores than women ($M = 1.75$, $SD = 1.08$ and $M = .31$, $SD = .86$, respectively), $t(38) = 4.84$, $p < .001$, $d = 1.48$, indicating that men's responses to the female–female stimuli were much higher than their responses to the male–male stimuli. It is notable that both sexes responded more genitally, on average, to female–female stimuli than male–male stimuli.

The second *t*-test, for the stimulus-specificity index (i.e., response to the sexual stimulus category that generated the highest responses relative to the average of responses to all other sexual stimulus categories), also revealed a sex difference, $t(37) = 3.57$, $p = .001$, $d = 1.17$. Overall, men responded the most to one sexual stimulus category ($M = 1.58$, $SD = .58$) compared to women ($M = .99$, $SD = .42$). Overall, 95% of men ($n = 19$) showed their highest arousal to a category that involved their stated preferred partner engaging in explicit sexual activities: Fifty percent of the men responded the most to female–female stimuli ($n = 10$) and 45% responded the most to male–female stimuli ($n = 9$). Fewer women (55%; $n = 11$) experienced their highest level of response to stimuli depicting their preferred sex partner engaging in explicit sexual activities: Forty-five percent of the women responded the most to the male–female stimuli ($n = 9$) and 10% responded the most to the male–male stimuli ($n = 2$). Fewer men (5%; $n = 1$) responded the most to stimuli depicting their non-preferred sex partner engaging in explicit sexual

---

5 The assumption of equality of variances was examined for all *t*-tests reported in this study. The assumption was not met for one test but the results did not change when using the corrected value. Uncorrected values are reported.

activities than women (25%; $n = 5$). One woman exhibited her greatest genital response to the low intensity sexual stimulus category. The remaining three women exhibited equally high genital responses to two or more sexual stimulus categories (male–female/female–female; male–female/sexual threat; sexual threat/low intensity/male–male).

Subjective Sexual Arousal Category-specificity

Two independent samples *t*-tests were conducted to assess whether sex differences in category-specificity were present for continuous subjective sexual arousal. The *t*-test assessing category-specificity using the female–male contrast score revealed a significant sex difference, $t(35) = 5.44$, $p < .001$, $d = 1.79$. Similar to the results for genital arousal, men's subjective sexual arousal was significantly more category-specific than women's ($M = 2.22$, $SD = .48$ and $M = .75$, $SD = 1.07$, respectively). Both sexes reported greater subjective sexual arousal in response to the female–female stimuli relative to the male–male stimuli. The *t*-test comparing men's ($M = 1.87$, $SD = .35$) and women's ($M = 1.90$, $SD = .46$) stimulus-specificity index revealed no significant sex difference for subjective sexual arousal, $t(35) = -.23$, *ns*, $d = -.08$. All men who reported changes in their continuous subjective responses throughout the experiment exhibited their highest subjective responses to stimuli depicting their preferred sex partners: Fifty-five percent of men ($n = 11$) reported their greatest subjective sexual arousal in response to the male–female stimulus category and the remaining 45% ($n = 8$) reported their greatest subjective sexual arousal in response to the female–female stimulus category. Women's subjective reports were more concentrated on the male–female stimulus category: A substantial proportion of the women (80%; $n = 16$) with valid data reported their greatest subjective sexual arousal in response to the male–female stimulus category, and the remaining 10% reported their greatest response to either the male–male stimulus category ($n = 1$) or the female–female stimulus category ($n = 1$). Results were identical when using post-stimulus subjective sexual arousal data.

Concordance Between Genital and Subjective Sexual Arousal

Three different within-subjects, non-parametric (Spearman's rho) correlations were calculated to determine the presence of sex differences in concordance between measures of sexual arousal. The first two correlations assessed the relationship between genital arousal and the two measures of subjective sexual arousal (i.e., continuous subjective sexual arousal and post-stimulus subjective sexual arousal). The third correlation

Table 2 Mean correlations (Spearman's rho) between genital and subjective measures of sexual arousal for all stimuli (top portion) and sexual stimuli only (bottom portion)

Type of correlation	All stimulus categories	
	Men	Women
Genital-continuous subjective	.59	.48
Genital-post subjective*	.67	.50
Continuous subjective-post subjective	.82	.87

Type of Correlation	Only sexual stimuli	
	Men	Women
Genital-continuous subjective**	.66	.29
Genital-post subjective***	.72	.33
Continuous subjective-post subjective	.84	.86

Note: Sex difference, * $p < .10$, ** $p < .01$, *** $p < .001$

assessed the relationship between the two subjective sexual arousal measures, in order to determine if participants were responding in a similar manner to both types of subjective measures.

The mean correlations across all stimuli for each sex are presented in Table 2 (top portion). Men tended to exhibit higher sexual concordance, although only the correlations between genital responses and post-stimulus subjective sexual arousal approached significance, $t(36) = 2.00$, $p = .055$, $d = .65$. The two measures of subjective sexual arousal (i.e., continuous and post-stimulus) were highly correlated in both men and women.

Because there was little genital and subjective sexual response to the non-sexual stimuli, it is possible that the low variance in genital and subjective responses to these stimuli could affect the size of the correlations. Thus, correlations between genital and subjective responses to the sexual stimuli alone were calculated. The means of these correlations are presented in Table 2 (lower portion). Results show that men were more sexually concordant than women when correlating both their genital and continuous subjective sexual arousal, $t(35) = 3.03$, $p < .01$, $d = 1.00$, and their genital and post-stimulus subjective sexual arousal, $t(36) = 3.96$, $p < .001$, $d = 1.29$. Again, the correlation between the two measures of subjective sexual arousal (i.e., continuous and post-stimulus) was high in both men and women.

Subsidiary Analyses

Because not all of the women who participated in the study reported sexual attractions or fantasies toward men exclusively, exploratory analyses were conducted to compare differences in category-specificity for two groups of women,

those who were predominantly heterosexual (i.e., women who reported *sometimes* experiencing sexual attractions or engaging in fantasies involving other women; $n = 15$) and those who were exclusively heterosexual (i.e., women who reported *never* experiencing sexual attractions or engaging in fantasies involving other women; $n = 5$). These analyses were not conducted for men because only one man reported any sexual attraction or fantasies directed toward men.

Independent samples t-tests revealed no significant difference for genital responses between the two groups of women when using the female–male contrast score, $t(18) = -.97$, $d = -.50$. Although no significant difference was found when using the stimulus-specificity index, $t(18) = 1.67$, the effect size of .86 can be considered to be large, with predominantly heterosexual women exhibiting more stimulus-specific genital responses ($M = 1.05$, $SD = .43$) than exclusively heterosexual women ($M = .70$, $SD = .32$). There was no significant difference for subjective sexual arousal when using the female–male contrast score, $t(16) -.68$, $d = -.36$. For the stimulus-specificity index, a significant difference was found between the two groups of women, $t(16) = 3.11$, $p < .01$, $d = 1.63$, and in the direction opposite to what was reported for genital arousal: Predominantly heterosexual women exhibited *less* stimulus-specific subjective sexual arousal responses ($M = 1.73$, $SD = .33$) than exclusively heterosexual women ($M = 2.35$, $SD = .49$). No significant differences were found between the two groups of women when examining concordance, although there was a tendency for predominantly heterosexual women to be more concordant than exclusively heterosexual women (e.g., $M = .38$, $SD = .39$ and $M = .06$, $SD = .20$, $d = .89$, respectively, for the correlation between genital response and continuous subjective sexual arousal).

Discussion

Results from the current study support the use of vaginal photoplethysmography as a measure of physiological sexual arousal in women. Women, similar to men, exhibited significantly higher genital responses to sexual stimuli compared to a variety of non-sexual stimuli. Women's pattern of genital responses differed from men's, in that women's genital responses were less category- and stimulus-specific. Subjectively, men and women were both fairly category- and stimulus-specific, in that they reported higher levels of arousal to one or two stimulus categories compared to all other sexual stimulus categories, although women's subjective reports were less discriminating than men's on the female–male contrast score. Women's sexual responses were also less concordant than men's, because they exhibited lower correlations between genital and subjective sexual arousal. These sex differences replicate previous findings by Chivers et al. (2004) using different audiovisual stimuli.

Vaginal Photoplethysmography as a Measure of Sexual Arousal

The current study provides further evidence of the construct validity of vaginal photoplethysmography. Phallometry is accepted as a well-validated measure of sexual arousal in men (for reviews, see Geer & Janssen, 2000; Rosen & Beck, 1988) but, by comparison, relatively little research had been conducted on the validity of VPP. Both men and women exhibited their greatest genital responses to the sexual stimuli and only sexual stimuli produced genital responses. This pattern of genital response mirrored subjective responses; only sexual stimuli generated subjective sexual arousal. The current study improves on previous research by assessing genital responses to a wide range of non-sexual, positively and negatively valenced stimuli, and a variety of sexual stimuli.

Although it appears that vaginal photoplethysmography accurately assesses sexual arousal in women, the degree of accuracy with which VPP can assess sexual *preference* in women is less clear. *Sexual arousal* refers to genital or subjective responses whereas *sexual preference* refers to some sort of ranking of arousal across different stimulus categories. As noted by Chivers (2005), men's sexual arousal reflects their sexual preferences: Men genitally respond the most to stimuli that involve their preferred partner, based on gender (e.g., Chivers et al., 2004; Chivers & Bailey, 2005; Freund et al., 1973) and age (e.g., Blanchard, Klassen, Dickey, Kuban, & Blak, 2001; Freund et al., 1973). Men's genital responses are used to infer sexual preference in forensic populations because phallometric responses are often more valid than self-report in this context (they match sexual history, whereas self-reported arousal sometimes does not). In women, the relationship between sexual arousal and preference is less clear: Women experience similar increases in genital arousal to subjectively preferred and non-preferred stimuli in this and other studies (e.g., Chivers et al., 2004; Chivers & Bailey, 2005). It is important to note here, however, that category-specific patterns of genital arousal have been measured with VPP in transsexual and natal women. Chivers et al. (2004) reported that post-operative male-to-female transsexuals tested in their study demonstrated highly category-specific patterns of genital arousal, yet the natal heterosexual and lesbian women did not. Category-specific patterns of arousal have recently been attained with VPP in lesbian women, but not heterosexual women, using less sexually-intense stimuli (Chivers, Seto, & Blanchard, 2008); however, the degree of category-specificity exhibited by lesbian women was significantly lower than men's. Because women's sexual preferences may not be as closely linked to their genital arousal patterns as men's, the usefulness of VPP as an assessment or diagnostic tool of sexual preferences in forensic or clinical settings is likely limited, at least until women's patterns of genital arousal are better understood.

Category-specificity of Genital and Subjective Sexual Arousal

The results of the current study replicated past research indicating that men's genital responses are more category-specific than women's (e.g., Chivers et al., 2004, 2008; Chivers & Bailey, 2005; Laan, Sonderman, & Janssen, 1995). The vast majority of men exhibited their highest level of genital response to stimuli involving their preferred sex partners (i.e., either the male–female or female–female stimuli); none of the men responded equally to two stimulus categories. Women, however, were more variable in their responses: Some women responded the most genitally to either the male–female or female–female stimuli and one third of the women either responded the most to one of the other sexual stimulus categories or equally to two or more sexual stimulus categories. Fewer women experienced their greatest arousal in response to stimuli depicting their preferred sex partners.

Results from this study are relevant to the question of which features of sexual stimuli are needed to produce a genital response in women. Several studies have shown that women produced the highest and equal response to film clips of couples engaged in sex, regardless of the gender of the actors (Chivers et al., 2004, 2008; Chivers & Bailey, 2005). All of the clips used in these studies have involved penetrative sex (e.g., for one of the female–female clips, a strap-on dildo was used). The female–female clips used in the current study did not include any sort of penetration and still produced very high levels of arousal, even though all female participants were heterosexual. This suggests that male features (e.g., a penis, a male body type) are not necessary to produce genital arousal in heterosexual women. Our results, along with the recent findings of Chivers et al. (2008), suggest that the quality of, and the activity depicted in, a stimulus are probably more important to women's genital responding.

Previous research has indicated that, similar to genital responses, men's subjective responses are category-specific. Women's subjective responses are also considered to be category-specific, because they typically report their highest level of arousal to their preferred stimulus category (Chivers et al., 2004, 2008; Chivers & Bailey, 2005). The results from the current study indicate that both men and women are category-specific in terms of their subjective sexual response. When considering the gender of the actors in the stimuli (i.e., using the female–male contrast score), men were much more specific than women, in that they reported higher levels of arousal to the female–female stimuli than the male–male stimuli. Although women's reports of subjective sexual arousal were similar (i.e., women also reported greater arousal to the female–female stimuli than male–male stimuli), the difference in reported subjective sexual arousal between the two types of stimuli was much

smaller. The stimulus-specificity index, however, showed no sex difference in subjective category-specificity.

Clearly, women's genital responses are more variable than men's. Perhaps this variability is the result of a greater erotic plasticity of female sexuality. Baumeister (2000) suggested that the degree to which an individual's sexuality (in terms of general opinions, feelings of arousal, and behaviors) can be formed and modified by external factors, such as cultural and situational pressures, is greater in women than in men. One of the basic tenets of Baumeister's theory is particularly relevant to the concept of category-specificity. Baumeister suggested that women show more intra-individual variation in sexual behavior over time than men. In swinging populations, for example, same-sex sexual activities are common for women but rare in men (Fang, 1976). Harrison, Bennett, Globetti, and Alsikafi (1974) examined the influence of dating experience on sexual standards and found that as women accumulated more dating experience, they became more permissive in their sexual practices whereas men exhibited no change.

Intra-individual variation has also been demonstrated in women's sexual identity. Diamond (2003) reported data collected from interviews with 80 non-heterosexual women between 18 and 25 years of age. Forty-eight percent of the women interviewed changed their sexual identity label within the five-year study period. Slightly less than half of those women who relinquished their sexual minority identities adopted a heterosexual sexual identity and the remainder adopted an unlabeled identity.

Further demonstrating the greater flexibility of women's sexuality, and consistent with the results of the present study, Rullo, Kinnish, and Strassberg (2006) found that heterosexual women were significantly more likely to report engaging in homosexual fantasies and experiencing homosexual romantic attractions than heterosexual men. The exclusion criteria employed in the current study were intended to select heterosexual participants. Nevertheless, a fair proportion of the women who participated in the study reported that they had experienced sexual attractions or engaged in sexual fantasies involving women (75%). In comparison, only one man reported experiencing any sexual attractions or engaging in any fantasies involving men. Perhaps because men's sexual attitudes and behaviors are more rigid (or organized earlier in development), men only genitally respond to the stimuli that they prefer, whereas women genitally respond to a variety of stimuli categories because they are more flexible in their preferences.

Due to the fact that there appeared to be two different groups of women who participated in the experiment (i.e., those who were exclusively heterosexual and did not report any sexual attraction or fantasies directed toward women, and those who were predominantly heterosexual and reported occasional sexual attraction or fantasies involving women), exploratory analyses were conducted to determine if "degree"

of heterosexuality was related to category-specificity. Exploratory analyses revealed an intriguing finding: In terms of genital responses, women who reported being exclusively heterosexual tended to be less category- and stimulus-specific than women who were classified as predominantly heterosexual. Both groups of women responded more to female–female stimuli than male–male stimuli, but exclusively heterosexual women's genital responses were more distributed across sexual stimulus categories compared to predominantly heterosexual women. These findings were similar to Chivers et al.'s (2008) recent finding that non-specificity of genital arousal is likely more characteristic of heterosexual than lesbian women. It is intriguing that exclusively heterosexual women were less category- and stimulus-specific with regard to their genital responses, because they were *more* category- and stimulus-specific than predominantly heterosexual women in terms of subjective sexual arousal. Further research examining the sexual arousal patterns of larger samples of exclusively and predominantly heterosexual women should provide insight into what factors, if any, distinguish exclusively heterosexual women from those with same-sex attractions.

Concordance of Genital and Subjective Sexual Arousal

Both men and women exhibited a positive relationship between genital and subjective sexual arousal. Similar to previous findings (e.g., Chivers et al., 2007), women were less concordant than men overall. This sex difference was obtained for both types of subjective measures. The two measures of subjective sexual arousal correlated highly with each other for both sexes, indicating that both men and women responded subjectively in a consistent manner. Exclusively heterosexual women tended to be less concordant overall than predominantly heterosexual women. The difference was not significant, likely because the sample size was quite small for the group of exclusively heterosexual women; future studies should attempt to replicate this group difference.

Low sexual concordance in women has recently come into question. Using hierarchical linear modeling (HLM), Rellini, McCall, Randall, and Meston (2005) found that genital responses predict subjective sexual arousal responses and vice versa, suggesting greater concordance among women than has previously been reported in the literature. It was concluded that low concordance among women may be the result of statistical artifact and that HLM may be a superior method for assessing concordance. Although Rellini et al. also reported a wide variability in concordance (range: $r = .08$ to $r = .79$) and over half (i.e., 16 of 22) of the women exhibited significant correlations between genital and subjective sexual arousal, responses to non-sexual

stimuli were included in their analyses when calculating the correlations. As suggested by our data, the inclusion of responses to non-sexual stimuli in the calculation of sexual concordance can significantly increase, perhaps spuriously, concordance estimates.

Why are women less concordant overall than men? Research outside of the realm of sexuality suggests that men and women rely on different strategies when assessing internal states. Pennebaker and Roberts (1992) reviewed sex differences in the perception of various bodily states, such as blood glucose levels, heart rate, and blood pressure. Two different paradigms are used in research on visceral perception: experimental settings in which a participant is tested in a laboratory relatively free from external situational cues and naturalistic settings in which a participant is required to estimate his or her physiological states throughout a normal day. The review indicated that men and women were equally good at estimating their physiological states in naturalistic settings. In experimental settings devoid of external situational cues, however, a clear sex difference emerged, in that men were more accurate at assessing bodily states than women when there was little information available from the environment. Perhaps this sex difference in perception of internal states in the laboratory is responsible for the fact that men are more concordant than women in laboratory assessments of sexual arousal. The experimental setting used in this study was designed to be as "natural" as possible, in that the participant was reclined in a comfortable chair and surrounded by a tasteful décor. Despite the attempt to make the experimental setting as comfortable and realistic as possible, it did not contain relevant situational cues, such as a sexual partner. Studies examining concordance in more naturalistic settings are needed to examine this question further.

Another possible explanation for the sex difference in concordance is that men and women are designed differently in terms of their physiological and psychological sexual response systems. As alluded to by Symons (1979), women's reproductive interests may not be best served by a psychology that is subordinate to physiology, due to the high costs of poor mating choices for women. Men and women have always had different minimal parental investment, defined as the minimum amount of resources or time necessary to produce an offspring. Over evolutionary history, women's minimal investment has consisted of the production of an ovum, 9 months of gestation, and childbirth, whereas men's minimal investment has consisted of a single ejaculate. These fundamental differences in parental investment resulted in women having much lower potential reproductive rates than men (Clutton-Brock & Vincent, 1991).

Based on these differences, there was greater selection pressure on men than on women to increase the *quantity* of their sex partners, and greater selection pressure on women than men to increase the *quality* of their sex partners. To increase *quantity* of sexual partners, men would have benefited by being easily aroused at the sight of women with high reproductive value. To increase *quality* of sexual partners, women would have benefited from being judicious in their selection of potential mates. In fact, increasing *quantity* would have had little benefits for women, but potentially great costs. As such, selection would have likely favored women whose motivation to engage in sexual activity was less dependent upon physiological responses, at least in terms of sexual arousal, in an attempt to promote greater control over mate choice. Similarly, women may choose to engage in sexual activity despite little or no physiological response or genuine sexual interest in her mate, in exchange for other benefits, such as social or physical resources. Lesser concordance between genital and subjective sexual arousal may have led to better decision-making in the context of mate choice and negotiation of her sociopolitical environment.

The nature of female sexual response may also play a role in producing low concordance among women. Female genital response to sexual stimuli has been hypothesized to be automatic, or reflexively activated by sexual stimuli (Chivers, 2005; Chivers et al., 2008; Laan et al., 1995), whereas feeling sexually aroused, that is, appraising one's state as sexual arousal, involves controlled cognitive processes and elaboration upon multiple sources of internal and external information (e.g., cues of sexual response, social context of sexual arousal, appraisal of potential mate value). Reflexive vaginal responding may have been beneficial because vaginal vasocongestion results in lubrication of the genital tract, reducing the likelihood of injury and subsequent infection resulting from vaginal penetration. Ancestral women who did not reflexively lubricate and who experienced unwanted sex would have been more likely to experience injuries or infections that could have rendered them reproductively sterile or resulted in their deaths after sustaining injury during genital penetration. Thus, there might have been selective pressure on genital automaticity to any cues of sexual activities. Discordance may, therefore, be the result of low subjective arousal despite reflexive genital response to a sexual stimulus.

Future research should test the possibility that there may be certain contexts in which concordance should be expected to be as high in women as it is in men. For example, although it may be beneficial for a woman to be less aware of her sexual arousal when *judging* potential mates, no such benefits are expected after she has reached a decision and is involved in an intimate relationship. In fact, it would likely be in a woman's best interests to be aware of sexual arousal in response to her current mate, in order to promote sexual activities in the relationship. Studies manipulating the content of sexual stimuli (e.g., a couple involved in a committed relationship versus a man and woman meeting and having sex without being in a committed relationship) could provide a test of this hypothesis.

Limitations

The limitations of the current study include the stimuli that were used. Although the current study improved upon past research by incorporating positively and negatively valenced non-sexual stimuli, as well as several types of sexual stimuli, it is possible that non-sexual stimuli that elicit other emotions could have produced a genital response. It would be beneficial in the future to use non-sexual stimuli that elicit other emotions, such as anger, disgust, or humor, to determine whether women experience genital responses to these stimuli. For instance, Kukkonen et al. (2006) found no significant differences in peak systolic and peak diastolic velocity of clitoral blood flow in women who were presented with a humorous versus erotic film. Humorous clips were not used in the current study because laughter is likely to introduce movement artifacts which would render the VPP data uninterpretable.

Another limitation concerns the presentation of several stimuli meant to elicit strong emotional responses over a fairly short period of time. Some research suggests that emotional arousal of any kind (positive or negative) can facilitate sexual arousal (e.g., Dutton & Aron, 1974; Meston & Frolich, 2003). One could argue that the emotional or physiological responses elicited by one stimulus could have influenced subsequent responses through excitation transfer. This possibility is unlikely to explain the current results, however, for two reasons. First, all participants viewed the stimuli in a random order. Because there were no genital responses to the non-sexual and neutral stimulus categories, it is unlikely that viewing non-sexual or neutral stimulus categories influenced responses to sexual stimuli. Conversely, it is clear that genital responses produced by the sexual stimuli did not result in responses to the non-sexual or neutral stimuli. Second, an examination of the post-stimulus responses indicated that there were no carry-over effects in terms of self-reported emotions: Participants did not report high levels of positive emotions (e.g., happiness, exhilaration) in response to the negative non-sexual films or vice versa.

Another limitation involves the degree to which we can be certain that the emotional manipulation produced changes in general arousal. Although it is clear that participants experienced changes in their emotions in response to the various film clips, it is unclear whether physiological changes occurred beyond the genital region in response to the film clips. It would be beneficial to verify the fact that VPP remains impervious to physiological changes related to non-sexual, yet emotionally charged, stimuli by examining peripheral physiological changes (e.g., skin conductance) along with genital responses.

Lastly, the generalizability of the results is questionable. A well-known selection bias exists in sexual psychophysiology research, such that individuals who volunteer to participate in sexual psychophysiology experiments are typically more sexually experienced, less concerned about their performance, and have been exposed to more erotica than non-volunteers' (Wolchik, Braver, & Jansen, 1985). Chivers et al. (2004) mentioned that although volunteers differ from non-volunteers in terms of sexual experience, this does not imply that volunteers' patterns of arousal are significantly different from non-volunteers'. Chivers et al. found that levels of cooperation were not associated with arousal patterns in women, providing indirect evidence that arousal patterns of volunteers and non-volunteers are likely not different. Despite this finding, it is still important to note that the results from the current study may only be generalizable to individuals similar to those who participate in these types of studies.

Conclusion

The results of this study indicate that vaginal photoplethysmography is an accurate and valid measure of sexual arousal in women. VPP has the ability to differentiate between various types of stimuli; the highest arousal is observed in response to sexual stimuli, and non-sexual but emotionally charged stimuli do not produce more genital arousal than neutral stimuli. Low category-specificity of genital arousal and low concordance between genital and subjective sexual arousal in women appear to be more than simple by-products of the difficulty in measuring female sexual arousal; they are valid phenomena in need of explanation.

References

Abel, G. G., Barlow, D. H., Blanchard, E. B., & Mavissakalian, M. (1975). Measurement of sexual arousal in male homosexuals: Effects of instructions and stimulus modality. *Archives of Sexual Behavior, 4*, 623–629.

Baumeister, R. F. (2000). Gender differences in erotic plasticity: The female sex drive as socially flexible and responsive. *Psychological Bulletin, 126*, 347–374.

Blanchard, R., Klassen, P., Dickey, R., Kuban, M. E., & Blak, T. (2001). Sensitivity and specificity of the phallometric test for pedophilia in nonadmitting sex offenders. *Psychological Assessment, 13*, 118–126.

Chivers, M. L. (2005). A brief review and discussion of sex differences in the specificity of sexual arousal. *Sexual and Relationship Therapy, 20*, 377–390.

Chivers, M. L., & Bailey, J. M. (2005). A sex difference in features that elicit genital response. *Biological Psychology, 70*, 115–120.

Chivers, M. L., Rieger, G., Latty, E., & Bailey, J. M. (2004). A sex difference in the specificity of sexual arousal. *Psychological Science, 15*, 736–744.

Chivers, M. L., Seto, M. C., & Blanchard, R. (2008). Gender and sexual orientation differences in sexual response to sexual activities versus gender of actors in sexual films. *Journal of Personality and Social Psychology, 93*, 1108–1121.

Chivers, M. L., Seto, M. C., Lalumière, M. L., Laan, E., & Grimbos, T. (2007). *Agreement of self-reported and genital measures of sexual arousal among men and women: A meta-analysis.* Manuscript submitted for publication.

Clutton-Brock, T. H., & Vincent, A. C. J. (1991). Sexual selection and the potential reproductive rates of males and females. *Nature, 351*, 58–60.

Diamond, L. M. (2003). Was it a phase? Young women's relinquishment of lesbian/bisexual identities over a 5-year period. *Journal of Personality and Social Psychology, 84*, 352–364.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology, 30*, 510–517.

Fang, B. (1976). Swinging: In retrospect. *Journal of Sex Research, 12*, 220–237.

Freund, K., Langevin, R., Cibiri, S., & Zajac, Y. (1973). Heterosexual aversion in homosexual males. *British Journal of Psychiatry, 122*, 163–169.

Geer, J., & Janssen, E. (2000). The sexual response system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 315–341). New York: Cambridge University Press.

Geer, J. H., Morokoff, P., & Greenwood, P. (1974). Sexual arousal in women: The development of a measurement device for vaginal blood volume. *Archives of Sexual Behavior, 3*, 559–564.

Harris, G. T., Rice, M. E., Quinsey, V. L., Chaplin, T. C., & Earls, C. M. (1992). Maximizing the discriminant validity of phallometric assessment data. *Psychological Assessment, 4*, 502–511.

Harrison, D. A., Bennett, W. H., Globetti, G., & Alsikafi, M. (1974). Premarital sexual standards of rural youth. *Journal of Sex Research, 10*, 266–277.

Hatch, J. P. (1979). Vaginal photoplethysmography: Methodological considerations. *Archives of Sexual Behavior, 8*, 357–374.

Heiman, J. R. (1980). Female sexual response patterns: Interactions of physiological, affective, and contextual cues. *Archives of General Psychiatry, 37*, 1311–1316.

Hoon, P. W., Wincze, J. P., & Hoon, E. F. (1976). Physiological assessment of sexual arousal in women. *Psychophysiology, 13*, 196–204.

Kinsey, A. C., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male*. Philadelphia: Saunders.

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., & Gebhard, P. H. (1953). *Sexual behavior in the human female*. Philadelphia: Saunders.

Kukkonen, T. M., Paterson, L., Binik, Y. M., Amsel, R., Bouvier, F., & Khalifé, S. (2006). Convergent and discriminant validity of clitoral color Doppler ultrasonography as a measure of female sexual arousal. *Journal of Sex and Marital Therapy, 32*, 281–287.

Laan, E., Everaerd, W., & Evers, A. (1995). Assessment of female sexual arousal: Response specificity and construct validity. *Psychophysiology, 32*, 476–485.

Laan, E., Sonderman, M., & Janssen, E. (1995, September). Straight and lesbian women's sexual responses to straight and lesbian erotica: No sexual orientation effects. *Poster presented at the annual meeting of the International Academy of Sex Research*, Provincetown, MA.

Levin, R. J. (1998). Assessing human female sexual arousal by vaginal photoplethysmography: A critical examination. *Sexologies: European Journal of Medical Sexology, 6*, 25–31.

Meston, C. M., & Frolich, P. F. (2003). Love at first fright: Partner salience moderates roller-coaster-induced excitation transfer. *Archives of Sexual Behavior, 32*, 537–544.

Meston, C. M., & Gorzalka, B. B. (1996). The effects of immediate, delayed, and residual sympathetic activation on sexual arousal in women. *Behaviour Research and Therapy, 34*, 143–148.

Pennebaker, J. W., & Roberts, T. (1992). Toward a his and hers theory of emotion: Gender differences in visceral perception. *Journal of Social and Clinical Psychology, 11*, 199–212.

Rellini, A. H., McCall, K. M., Randall, P. K., & Meston, C. M. (2005). The relationship between women's subjective and physiological sexual arousal. *Psychophysiology, 42*, 116–124.

Rieger, G., Chivers, M. L., & Bailey, J. M. (2005). Sexual arousal patterns of bisexual men. *Psychological Science, 16*, 579–584.

Rosen, R. C., & Beck, J. G. (1988). *Patterns of sexual arousal: Psychophysiological processes and clinical applications*. New York: Guilford Press.

Rullo, J., Kinnish, K., & Strassberg, D. (2006). Sex differences in the specificity of sexual behavior, fantasy, and romantic attraction. *Poster presented at the annual meeting of the International Academy of Sex Research*, Amsterdam, The Netherlands.

Symons, D. (1979). *The evolution of human sexuality*. New York: Oxford University Press.

Wincze, J. P., Venditti, E., Barlow, D., & Mavissakalian, M. (1980). The effects of a subjective monitoring task in the physiological measure of genital response to erotic stimulation. *Archives of Sexual Behavior, 9*, 533–545.

Wolchik, S. A., Braver, S. L., & Jensen, K. (1985). Volunteer bias in erotica research: Effects of intrusiveness of measure and sexual background. *Archives of Sexual Behavior, 14*, 93–107.

Zuckerman, M. (1971). Physiological measures of sexual arousal in the human. *Psychological Bulletin, 75*, 297–329.