

# Hormonal Changes and Couple Bonding in Consensual Sadomasochistic Activity

Brad J. Sagarin · Bert Cutler · Nadine Cutler ·  
Kimberly A. Lawler-Sagarin · Leslie Matuszewich

Received: 19 December 2006 / Revised: 13 June 2007 / Accepted: 22 December 2007 / Published online: 19 June 2008  
© Springer Science+Business Media, LLC 2008

**Abstract** In two studies, 58 sadomasochistic (SM) practitioners provided physiological measures of salivary cortisol and testosterone (hormones associated with stress and dominance, respectively) and psychological measures of relationship closeness before and after participating in SM activities. Observed activities included bondage, sensory deprivation, a variety of painful and pleasurable stimulation, verbal and non-verbal communication, and expressions of caring and affection. During the scenes, cortisol rose significantly for participants who were bound, receiving stimulation, and following orders, but not for participants who were providing stimulation, orders, or structure. Female participants who were bound, receiving stimulation, and following orders also showed increases in testosterone during the scenes. Thereafter, participants who reported that their SM activities went well showed reductions in physiological stress (cortisol) and increases in relationship closeness. Among participants who reported that their SM activities went poorly, some showed decreases in relationship closeness whereas others showed increases. The increases in relationship closeness combined with the displays of caring and affection observed as part of the SM activities offer support for the modern view that SM,

when performed consensually, has the potential to increase intimacy between participants.

**Keywords** Sadomasochism · Cortisol · Testosterone · Sexuality

## Introduction

Historical records suggest that humans have incorporated bondage, pain, and power exchange into their sexual practices for thousands of years (Weinberg, Williams, & Moser, 1984), and surveys suggest that a substantial minority of people engage in or fantasize about such practices today (Moser & Levitt, 1987). Nevertheless, sadomasochism has received limited attention from researchers. Early theorists viewed sadomasochism (SM) as a symptom of pathology (Freud, 1938; Krafft-Ebbing, 1886/1965) and SM-based relationships as pathological (Krafft-Ebbing, 1886/1965). These perspectives were echoed in the revised third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 1987), in which sadistic or masochistic behavior, in and of itself, provided sufficient evidence for a diagnosis of a psychosexual disorder. As Weinberg et al. (1984) noted, the perspective that SM represented a pathological condition may have stemmed from the fact that clinical practitioners and researchers provided the early understanding of the subject: “The traditional model generalizes to the whole of sadomasochism the activities and experiences of those persons most likely to come to the attention of clinicians” (p. 388).

Subsequent researchers have adopted a less negative perspective, viewing SM not as a symptom of mental illness but as a relatively healthy social phenomenon (Gebhard, 1969; Kamel, 1980; Weinberg, 1994) performed by well educated people (Breslow, Evans, & Langley, 1985; Moser & Levitt,

---

B. J. Sagarin (✉) · L. Matuszewich  
Department of Psychology, Northern Illinois University, DeKalb,  
IL 60115, USA  
e-mail: bsagarin@niu.edu

B. Cutler  
Department of Motion Pictures/Television Production, Scottsdale  
Community College, Scottsdale, AZ, USA

N. Cutler  
Tempe, AZ, USA

K. A. Lawler-Sagarin  
Department of Chemistry, Elmhurst College, Elmhurst, IL, USA

1987) with positive attitudes about themselves and their behaviors (Breslow, Evans, & Langley, 1986; Spengler, 1977; for a comprehensive review, see Weinberg, 1994). These researchers have focused on non-clinical samples, including readers of SM magazines and customers at a store selling SM equipment (Breslow et al., 1985), attendees at SM workshops and an SM convention (Myers, 1992), patrons at a bar and sex club in New York City (Brodsky, 1993), and members of SM clubs in the U.S. (Moser & Levitt, 1987; Weinberg et al., 1984), West Germany (Spengler, 1977), and Finland (Alison, Santtila, Sandnabba, & Nordling, 2001; Sandnabba, Santtila, & Nordling, 1999; Santtila, Sandnabba, Alison, & Nordling, 2002). The *DSM-IV-TR* (American Psychiatric Association 2000) reflects this new perspective, requiring that “the [SM] fantasies, sexual urges, or behaviors cause clinically significant distress or impairment in social, occupational, or other important areas of functioning” (p. 573) before a diagnosis of a psychosexual disorder is appropriate.

The writings of the SM community also provide detailed information about SM activities (Baldwin, 1993; Bannon, 1993; Brame, Jacobs, & Brame, 1993; Califa, 1993; Miller & Devon, 1995; Rinella, 1994; Wiseman, 1996). These books describe a number of physiological and psychological effects of SM scenes (defined periods of time during which participants engage in a series of SM activities), including the couple bonding that takes place during and after a scene (Rinella, 1994), the stress of an interrupted scene (Baldwin, 1993), and the possible health benefits of SM activities, including stress reduction (Mains, 1984). To our knowledge, none of these observations have undergone formal testing.

The present studies were designed to test these observations by examining the effects of SM activities on a physiological measure of stress (salivary cortisol, a stress-related hormone; Kirschbaum & Hellhammer, 1994) and psychological measures of relationship closeness. To this end, participants provided saliva samples and completed questionnaires before and after participating in SM scenes. In addition, because the SM community literature describes a number of activities that do not appear in the current taxonomies of SM behaviors in the academic literature (e.g., Alison et al., 2001), we attempted to document the behaviors that occurred during the observed scenes. Finally, because SM activities are often associated with dominance and submission, the present studies offered an exploratory examination of the effects of these activities on levels of testosterone, a hormone associated with dominance in women and men (Dabbs & Dabbs, 2000).

### Physiological Stress

Salivary cortisol has been used as a measure of physiological stress across a wide range of stressors. Cortisol levels rise in response to negative experiences, such as dental examinations (Benjamins, Asscheman, & Schuur, 1992), insults (Cohen,

Nisbett, Bowdle, & Schwarz, 1996), and experimentally induced pain (Dixon, Thorn, & Ward, 2004), neutral experiences, such as cognitive tasks (Bohnen, Houx, Nicolson, & Jolles, 1990; Bohnen, Nicolson, Sulon, & Jolles, 1991), and positive experiences, such as parachute jumping (Cook, Read, Walker, Harris, & Riad-Fahmy, 1992; for a review, see Kirschbaum & Hellhammer, 1994). Cortisol levels also rise in response to physical exercise, peaking somewhat after the conclusion of the stressor (20–30 min after in Cook et al., 1992). As a result, Kirschbaum and Hellhammer (1994) recommend collecting saliva samples for at least 30 min after participants finish exercising.

According to Kirschbaum and Hellhammer (1994), rises in cortisol levels will be seen particularly “in situations with high ego-involvement, low predictability, low controllability, and novelty” (p. 318). Although participants in our studies adopting the bottom role (participants who were bound, receiving stimulation, following orders, etc.) did so consensually, it seems likely that their experience would include the factors identified by Kirschbaum and Hellhammer, particularly for scenes involving bondage and sensory deprivation. Based on these psychological aspects of the bottom role, as well as the physical pain experienced by bottoms in some scenes, we predicted a rise in cortisol levels for bottoms, possibly followed by a drop in cortisol after the conclusion of the scene. We did not anticipate a similar rise in cortisol levels for tops (participants who provided stimulation, orders, or structure), as their role implies high levels of predictability and controllability and the administration, rather than the receipt, of physical pain.

### Relationship Closeness

The effects of SM activities on the perceived closeness between partners may be quite complex due to the number of emotions generated by each scene. On the one hand, to the extent that the activities represent an intense, shared, positive sexual experience, they might bring the participants closer together. Along these lines, Cutler (2003) interviewed SM couples and found that many used SM to increase bonding in the relationship. On the other hand, given that SM activities often involve one participant causing pain or distress to the other participant, it seems possible that the activities might push the participants further apart. Two factors might mitigate this latter possibility. First, unlike many situations of interpersonal violence and aggression, the participants in SM activities participate consensually, including the participant receiving the pain or distress. Second, mechanisms such as safewords (Wiseman, 1996) enable participants to end the SM activities if the experience becomes overly aversive. We predicted that participants would report an increase in relationship closeness after their scenes, consistent with the

writings of experienced SM practitioners (Baldwin, 1993; Mains, 1984).

### SM Behaviors

Alison et al. (2001) empirically derived four sexual scripts related to SM activities: hypermasculinity (e.g., rimming, using dildos, cockbinding, etc.), administration of pain (e.g., applying clothespins, spanking, caning, etc.), humiliation (e.g., flagellation, verbal humiliation, gags, etc.), and physical restriction (e.g., bondage, handcuffs, chains, etc.). The activities comprising these scripts also appear in SM-community taxonomies (e.g., Wiseman, 1996), although the SM-community literature typically describes flagellation as a method of administering pain rather than as a method of causing humiliation (Bean, 2000). Indeed, the SM community literature makes the point that the same activity can have different meanings for different people (Bean, 2000). In addition, the SM-community taxonomies contain additional activities not included in Alison et al. (2001).

### Testosterone

Because testosterone is associated with dominance in women and men (Dabbs & Dabbs, 2000), it seemed a particularly interesting hormone to examine in the context of SM scenes in which participants adopted dominant and submissive roles. Past research on testosterone has examined its function as a predictor of individual differences (e.g., violence among male and female prisoners, Dabbs, Frady, Carr, & Besch, 1987, Dabbs, Ruback, Frady, Hopper, & Sgoutas, 1988; occupations, Dabbs, de LaRue, & Williams, 1990; “butch” and “femme” lesbian roles, Singh, Vidaurri, Zambarano, & Dabbs, 1999), an antecedent to behavior (e.g., tennis players’ testosterone levels rose just before a match, Booth, Shelley, Mazur, Tharp, & Kittok, 1989), and a consequence of behavior (e.g., winning and losing, Booth et al., 1989, Gladue, Boechler, & McCaul, 1989, Mazur, Booth, & Dabbs, 1992; having sex, Dabbs & Mohammed, 1992; being insulted, Cohen et al., 1996; for reviews, see Dabbs, 1993; Dabbs & Dabbs, 2000). As with cortisol, testosterone changes appear in saliva approximately 20 min after the corresponding stimulus (Cohen et al., 1996; Dabbs, 1993; Veldhuis et al., 1987). We measured testosterone primarily for exploratory purposes, but the past research discussed above suggests at least one specific prediction: With respect to individual differences, because testosterone has been shown to predict position in dominance hierarchies (Dabbs et al., 1987), it is possible that baseline differences in testosterone might predict SM participants’ preferred roles, with participants higher in testosterone preferring the top role and participants lower in testosterone preferring the bottom role.

## Study 1

### Method

#### *Participants*

A total of 13 attendees (6 women, 7 men; *M* age, 45 years; range, 30–57) at a party hosted by the Arizona Power Exchange volunteered for the study. The participants took part in seven two-person scenes. All scenes involved one man and one woman. Men performed the top role in six scenes, while a woman performed the top role in the seventh scene. All participants had pre-existing relationships with their scene partners, including three pairs of spouses, two pairs of friends, one pair of play partners (i.e., people who participate in SM scenes together), and one pair of partners (type of partnership unspecified).

#### *Procedure*

Data collection took place at a party hosted by the Arizona Power Exchange (APEX), a Phoenix-based BDSM organization established in 1988. (“BDSM” is used within the SM community as a synthesis of the categories of Bondage and Discipline, Dominance and Submission, and Sadomasochism.) APEX, on its business cards, states that it is a support organization and “a non-profit organization which explores consensual Dominance/Submission, S & M, and Bondage & Discipline relationships with acceptance, caring, respect and dignity.” The group started as an informal organization and in 2001 became a members-only 501(c)(7) corporation. APEX has over 400 active members and 2,500 dues paying members over its history. The organization hosts weekly educational meetings and parties as well as periodic special events. APEX defines itself as pansexual, with support for all sexual orientations. More information regarding APEX can be found at <http://www.arizonapowerexchange.org>.

The APEX space consists of a main room with a large area designated for scenes and a smaller sitting area for those watching the scenes, as well as a separate entrance area, a library, and an orientation room. One researcher recruited participants at the entrance area while two researchers administered surveys and took saliva samples in the orientation room. All participants were recruited before the first scene began, and the researcher performing recruitment also acted as the observer during scenes. During data collection, all scenes were being observed by multiple people, one of whom was the study’s observer. APEX partygoers know that their scenes will be observed and, as such, participant reactivity was not anticipated to be a major factor.

Upon arrival at the party, attendees were asked if they planned to participate in a scene that evening. Those who

said yes were invited to participate in the study. We obtained a waiver of a signed informed consent form from the Northern Illinois University Institutional Review Board. This waiver was requested because a signed consent form would have eliminated the anonymity of the participants. Instead, one researcher read the approved informed consent statement and participants gave their consent verbally.

Participants were asked to complete two surveys, one immediately upon joining the study and one at the conclusion of their scene. Participants were also asked to provide four saliva samples: one immediately upon joining the study (baseline), one within 10 min prior to beginning their scene (pre-scene), one approximately 20 min after the end of the main scene activities, typically denoted by the bottom being untied, lowered to the floor, etc. (during-scene, given the previously discussed 20 min delay between a stimulus and the corresponding changes in salivary hormone levels), and one approximately 40 min after the scene (post-scene). Because the baseline samples were collected at the party, they include the effects of attending a party, anticipating SM activities, etc. As such, they probably differ from a true resting baseline. Nevertheless, we believe they represent the appropriate baseline to use when examining the hormonal changes that occur as a result of SM activities. In addition, because scenes varied in length, we timed the saliva samples to correspond to the beginning and ending of the main scene activities rather than to specific time intervals. As a result, the intervals between the samples and the nature and intensity of the activity preceding each sample differed from participant to participant, but the samples represent the same portion of a scene for each participant.

All 13 participants gave baseline and during-scene samples, and 10 participants gave post-scene samples. In addition, all participants filled out the pre-scene questionnaire, and 12 participants filled out the post-scene questionnaire. Unfortunately, most participants began their scenes without providing pre-scene samples, and SM community norms prohibited us from interfering with scenes once they had begun. Thus, pre-scene samples were collected from relatively few participants. As a result, the analyses below include only the baseline, during-scene, and post-scene samples.

One researcher observed the seven scenes, taking notes on the activities and reactions of the participants. The observer also notified another researcher of the times to administer the during-scene and post-scene vials. This other researcher then offered these vials to participants 20 and 40 min after the end of the main scene activities.

The observer was familiar with academic taxonomies of SM behaviors (e.g., Alison et al., 2001) and SM-community taxonomies (e.g., Wiseman, 1996) and had observed SM scenes on other occasions prior to data collection. Observations were made according to the following protocol:

The observer took time stamped notes on each scene using longhand in a notebook. The start time for each scene was recorded. Thereafter, whenever the activity changed, the observer would record the current time and the new activity (e.g., “11:00 pm: Hand restraints placed” “11:02 pm: Bottom disrobes, faces north” “11:05 pm: Bottom kneels”). Between 5 and 45 timestamped entries were made for each scene, depending upon the length of the scene and the number of activities. The timestamps were typically recorded using 1-min gradations, but when activities were changing rapidly, timestamps were recorded using 15-s gradations. The scenes occurred throughout the evening, with three scenes occurring simultaneously for 18 min.

Because the behaviors included in academic SM taxonomies comprise only a subset of the behaviors included in SM-community taxonomies, we anticipated that an a priori list of behaviors based on these taxonomies might miss some of the behaviors performed by participants. As such, we decided not to restrict the observer to a predefined list. The observer used the classes of behaviors found in both types of taxonomies as a guide to the specific behaviors to record. After data collection was complete, we created categories of behaviors based on the notes, taking into account both types of taxonomies as well as additional previously undocumented behaviors.

### Measures

During the party, saliva samples were placed in an ice chest immediately after collection. Subsequently, samples were stored in a household freezer before being sent to Salimetrics, LLC (State College, PA) for analysis. Samples were shipped on dry ice via overnight delivery.

Saliva samples were tested for cortisol and testosterone using enzyme immunoassays performed by Salimetrics (see Salimetrics, 2000a, b, for technical information on the immunoassays; see also Chard, 1990). Duplicate tests were performed for each sample, and the final measures of cortisol and testosterone consist of the average of the two separate assays, cortisol:  $r(43) = .994, p < .001$ ; testosterone:  $r(43) = .985, p < .001$ . The assays provide indices of free cortisol and testosterone.

The pre-scene questionnaire consisted of 11 questions including demographics, typical and current BDSM roles, anticipation and comfort with the upcoming scene, and two measures of relationship closeness (Cialdini, Brown, Lewis, Luce, & Neuberg's, 1997, measure of the perception of the dyad as “we,” and Aron, Aron, & Smollan's, 1992, Inclusion of Other in Self scale, IOS). The post-scene questionnaire also consisted of 11 questions including perceptions of scene quality, reactions to the scene, and the same two measures of relationship closeness (see Appendix).

## Results

### Observed Activities

Scenes ranged in duration from 26 to 75 min, with a mean of 54.6 min. Some scenes involved bondage (hand restraints on x-cross and trapeze; hand-and-foot restraints on rack), whereas other scenes used no bondage (draping body on x-cross or on padded table). The behaviors observed during scenes included painful stimulation (e.g., spanking, flogging, whipping, biting), pleasurable stimulation (e.g., kissing, caressing, massaging), communication (e.g., talking, eye contact), and aftercare (see Table 1).

Aftercare refers to participant interaction after completing the main scene activities and before leaving the play area. It serves to re-orient the participants to regular (non-SM) interaction and movement. Aftercare included gentle contact (e.g., hugging, cuddling), pleasurable stimulation (e.g., caressing, massaging), and communication (e.g., talking, eye contact).

Acts denoting caring attention were present during more than just the aftercare segment of the scenes, although they were seen for longer duration during aftercare. All SM activities were expressed in a range of styles among the couples, ranging from light-hearted to more serious or passionate. Although a range of sexual activities are allowed by APEX rules, no such activities were seen in the observed scenes on this research occasion.

### Physiological and Psychological Effects

**Baseline Differences:** In the pre-scene questionnaire, bottoms expressed significantly greater anticipation of the

**Table 1** SM activities observed during and after scenes in Study 1

Activity	Number of scenes
<b>During scene</b>	
Bondage/physical restraint	3 of 6 <sup>a</sup>
Sensory deprivation (e.g., blindfolds)	1 of 6 <sup>a</sup>
Spanking, paddling	3 of 7
Flogging (i.e., striking with a multi-tailed instrument)	7 of 7
Whipping (i.e., striking with a single-tailed instrument)	2 of 7
Pinching, clamping (with fingers or clothespins)	2 of 7
Kissing, nibbling, biting	4 of 7
Caressing, stroking, massaging	7 of 7
Talking, laughing together, mutual eye contact	3 of 7
<b>Aftercare</b>	
Hugging, holding, cuddling	5 of 7
Caressing, stroking, massaging	5 of 7
Talking, mutual eye contact	2 of 7

<sup>a</sup> Incomplete records were kept for one scene; the bondage and sensory deprivation used in that scene (if any) were not recorded

evening's scene ( $M = 5.5$ ,  $SD = 0.5$ ) compared to tops ( $M = 3.9$ ,  $SD = 1.6$ ),  $t(7.63) = 2.59$ ,  $p = .03$  (equal variances not assumed, due to significant Levene's test for equality of variances). Both bottoms and tops felt comfortable about the upcoming scene (bottoms:  $M = 5.7$ ,  $SD = 1.2$ ; tops:  $M = 6.0$ ,  $SD = 1.4$ ),  $t(11) < 1$ .

At baseline, tops and bottoms did not differ significantly in their levels of cortisol,  $t(5.99) = 1.50$ ,  $p = .18$  (see Table 2). When examined by sex, women and men did not differ significantly in their levels of baseline cortisol,  $t(6.90) < 1$ .

Tops and bottoms also did not differ in levels of baseline testosterone (tops:  $M = 90.9$  pg/ml,  $SD = 43.3$  pg/ml, bottoms:  $M = 80.5$  pg/ml,  $SD = 45.3$  pg/ml),  $t(11) < 1$ . This latter result was of concern due to the substantially greater proportion of men adopting the top role (6/7) than the bottom role (1/6). Indeed, when examined by sex, women and men did not differ significantly in their levels of baseline testosterone (women:  $M = 76.5$  pg/ml,  $SD = 45.5$  pg/ml, men:  $M = 94.3$  pg/ml,  $SD = 41.8$  pg/ml),  $t(11) < 1$ .

After consulting with Salimetrics, the company that performed the assays, and checking our data coding and data entry, we were unable to determine why a sex difference in testosterone levels did not appear despite the expectation that large sex differences in testosterone should appear in most cases. The lack of a sex difference cast sufficient doubt on the testosterone data that we excluded these data from Study 1. We have retained the cortisol data for Study 1 because the cortisol results were consistent with the findings of Study 2 in which we performed our own assays and in which baseline testosterone levels displayed the expected sex difference.

**Table 2** Study 1: Changes in cortisol between baseline and 20 min after the scene

Group	Baseline		20 Min after scene		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
<b>Bottoms</b>					
Female ( $n = 5$ )	0.091	0.063	0.193	0.122	1.05 <sup>a</sup>
Male ( $n = 1$ )	0.127	–	0.195	–	– <sup>a</sup>
Total ( $n = 6$ )	0.097	0.058	0.193	0.109	1.10*
<b>Tops</b>					
Female ( $n = 1$ )	0.035	–	0.052	–	– <sup>a</sup>
Male ( $n = 6$ )	0.064	0.018	0.070	0.027	0.26 <sup>a</sup>
Total ( $n = 7$ )	0.060	0.020	0.068	0.026	0.34

*Note:* Cortisol levels are in units of  $\mu\text{g}/\text{dl}$ . Following the recommendations of Dunlap, Cortina, Vaslow, and Burke (1996), effect sizes (Cohen's *d*) were calculated using the original standard deviations for each group rather than the paired-samples *t*-test. Significance tests represent the simple effects within the context of the full ANOVA

<sup>a</sup> Because only one male participant took the bottom role and only one female participant took the top role, simple effects within each sex were not performed

\*  $p < .05$

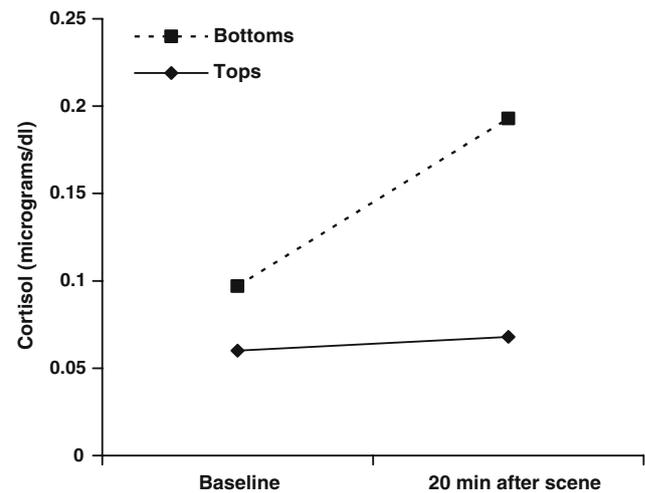
**During-scene Effects:** The physiological effects of taking part in an SM scene (during-scene effects) were examined using a 2 (role)  $\times$  2 (time: baseline vs. 20 min after the scene) mixed model ANOVA.<sup>1</sup> As discussed above, there was a 20-min delay between the stimulus and the corresponding hormonal changes in the saliva. Given this, the saliva sample taken 20 min after the scene represents the hormonal levels during the scene.

Between baseline and 20 min after the scene (again, corresponding to hormonal levels during the scene), cortisol rose significantly for bottoms,  $F(1, 11) = 5.33, p = .04$ , but not for tops,  $F(1, 11) < 1$  (see Fig. 1 and Table 2). These different patterns resulted in a significant main effect of role (bottoms showed significantly higher levels of cortisol than tops),  $F(1, 11) = 17.87, p = .001$ , a marginal main effect of time (bottoms and tops showed a marginal increase in cortisol between baseline and 20 min after the scene),  $F(1, 11) = 3.36, p = .09$ , and a non-significant Role  $\times$  Time interaction,  $F(1, 11) = 2.42, p = .15$ . Because only one male participant took the bottom role and only one female participant took the top role, separate analyses within each sex were not performed. However, an examination of the means revealed similar patterns for female and male bottoms and female and male tops (see Table 2).

**Post-scene Effects:** The after effects of participating in a scene were moderated by the participants' experience of the quality of the scene (i.e., how well the scene went). To code scene quality, we created a dichotomous variable that represented how participants felt about the scene (taken from the second question on the post-scene questionnaire: "How do you feel about the scene? Did it go well? Did it go poorly?").

After the scene was over, participants whose scenes went well reported that they felt significantly more comfortable with the scene ( $M = 5.9, SD = 1.0$ ) compared to participants whose scene went poorly ( $M = 4.0, SD = 0.0$ ),  $t(10) = -2.60, p = .03$ . In addition, participants whose scenes went well rated the scenes as marginally more intense ( $M = 4.3, SD = 0.9$ ) compared to participants whose scene went poorly ( $M = 3.0, SD = 0.0$ ),  $t(10) = -1.86, p = .09$ .

The physiological post-scene effects for cortisol were examined using a 2 (role)  $\times$  2 (scene quality)  $\times$  2 (time: 20 vs. 40 min after the scene) mixed model ANOVA. There was a significant Scene Quality  $\times$  Time interaction,  $F(1, 6) = 6.66, p = .04$ . Bottoms and tops whose scenes went



**Fig. 1** Study 1: Changes in cortisol between baseline and 20 min after the scene

well showed a reduction in cortisol from 20 to 40 min after the scene, whereas the bottom and top whose scene went poorly showed an increase (see Table 3). This effect was qualified by a marginal three-way interaction,  $F(1, 6) = 4.05, p = .09$ . An examination of the means suggests that the bottom whose scene went poorly showed an increase in cortisol compared to the decrease shown by the bottoms whose scenes went well, whereas the top whose scene went poorly showed a slight decrease in cortisol compared to the larger decrease shown by tops whose scenes went well. Because all bottoms who provided post-scene saliva samples were women and all tops who provided post-scene

**Table 3** Study 1: Changes in cortisol between 20 and 40 min after the scene

Group	20 Min after scene		40 Min after scene		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Scene went poorly					
Bottom (female; <i>n</i> = 1)	0.168	–	0.255	–	– <sup>a</sup>
Top (male; <i>n</i> = 1)	0.070	–	0.065	–	– <sup>a</sup>
Total ( <i>n</i> = 2)	0.119	0.069	0.160	0.134	0.38
Scenes went well					
Bottoms (all female; <i>n</i> = 4)	0.199	0.140	0.152	0.161	–0.31 <sup>a</sup>
Tops (all male; <i>n</i> = 4)	0.068	0.034	0.046	0.016	–0.83 <sup>a</sup>
Total ( <i>n</i> = 8)	0.133	0.118	0.099	0.120	–0.29*

*Note:* Cortisol levels are in units of  $\mu\text{g/dl}$ . Following the recommendations of Dunlap et al. (1996), effect sizes (Cohen's *d*) were calculated using the original standard deviations for each group rather than the paired-samples *t*-test. Significance tests represent the simple effects within the context of the full ANOVA

<sup>a</sup> Because all bottoms who provided post-scene saliva samples were women and all tops who provided post-scene saliva samples were men, simple effects within each sex were not performed

\*  $p < .05$

<sup>1</sup> Time was analyzed as a series of two-level factors (baseline vs. during-scene, during-scene vs. post-scene) instead of a three-level factor (baseline, during-scene, and post-scene) due to missing data. Treating time as a three-level factor would require either a method of imputation or the use of listwise deletion. Imputation seemed unwarranted, as standard imputation methods such as mean imputation lead to biased results and more sophisticated imputation methods such as Multiple Imputation require much larger sample sizes (West & Sagarin, 2000). Listwise deletion, on the other hand, would have reduced the sample size for some statistical tests.

saliva samples were men, an analysis of the effects of role and scene quality separated by sex was not possible.

Relationship closeness (represented by the question “to what extent would you use the word ‘we’ to describe your relationship with the other person or people in the scene?”) showed patterns similar to cortisol. A 2 (role)  $\times$  2 (scene quality)  $\times$  2 (time: pre-scene vs. post-scene questionnaire) mixed model ANOVA yielded a significant Scene Quality  $\times$  Time interaction,  $F(1, 8) = 8.33, p = .02$ . Bottoms and tops whose scenes went well reported increases in relationship closeness (pre-scene:  $M = 5.8, SD = 1.5$ , post-scene:  $M = 6.3, SD = 0.9$ ), whereas the bottom and top whose scene went poorly reported decreases (pre-scene:  $M = 7.0, SD = 0.0$ , post-scene:  $M = 5.5, SD = 0.7$ ).

One of the two participants whose scene went poorly did not complete the after-scene IOS measure. As a result, IOS analyses include only those participants whose scenes went well. For these participants, a 2 (role)  $\times$  2 (time) mixed model ANOVA revealed a non-significant main effect of role, such that bottoms ( $M = 4.4, SD = 0.5$ ) did not differ significantly from tops ( $M = 3.4, SD = 1.1$ ) in their levels of self-other overlap,  $F(1, 8) = 3.39, p = .10$ , and a non-significant main effect of time, such that bottoms and tops did not differ significantly in self-other overlap from before the scene ( $M = 3.7, SD = 1.2$ ) to after the scene ( $M = 4.1, SD = 0.9$ ),  $F(1, 8) = 2.91, p = .13$ . There was no Role  $\times$  Time interaction,  $F(1, 8) < 1$ .

Other post-scene questions did not differ by role or scene quality. Overall, participants responded near the midpoint on the scales when reporting how sexually arousing they found their scenes ( $M = 3.9, SD = 1.7$ ), whether they might be having sex later that evening ( $M = 3.8, SD = 2.1$ ), and how much “afterglow buzz” they felt ( $M = 3.9, SD = 1.4$ ).

## Discussion

Study 1 was designed to examine the physiological and psychological effects of SM activities on the participants. We predicted that (1) bottoms would show increases in physiological stress during their scenes due to the lack of control inherent in the bottom role and the pain experienced by the bottom in some scenes, (2) tops would not show increases in physiological stress, and (3) both bottoms and tops would show increases in relationship closeness and, possibly, decreases in physiological stress after their scenes.

Consistent with predictions, bottoms showed significant increases in cortisol during their scenes, but tops did not. After their scenes, bottoms and tops whose scenes went well showed decreases in cortisol and increases in relationship closeness. In contrast, the bottom and top whose scene went poorly showed decreases in relationship closeness and the bottom showed an increase in cortisol.

Study 2 was designed to replicate and extend the findings of Study 1 with a larger, more diverse sample.

## Study 2

### Method

#### Participants

A total of 61 attendees (29 women, 32 men; 27 heterosexual, 19 bisexual, 6 gay or lesbian, 9 did not report sexual orientation;  $M$  age, 40.5 years; range, 21–62) at Thunder in the Mountains volunteered for the study. Forty-five participants (19 women, 26 men) took part in scenes. In some cases, the scenes involved other study participants. In other cases, the scenes involved non-participants, although these non-participants were aware that their scene partners were participating in the study. Sixteen participants gave baseline samples but did not end up participating in a scene. Of the 19 women who participated in scenes, 7 performed the top role and 12 performed the bottom role.<sup>2</sup> Of the 26 men who participated in scenes, 14 performed the top role and 12 performed the bottom role. Most participants had pre-existing relationships with their scene partners, including 7 spouses, 11 romantic partners, 9 friends, 1 play partner, 12 with SM relationships (master, slave, 24/7 partner, etc.), and 1 with an unspecified 2 + year relationship. Three participants had little or no pre-existing relationship with their scene partners (“friend of a friend,” “I hardly know him at all,” “recent acquaintance”). One participant responded “?”.

#### Procedure

Data collection took place at two parties hosted by Thunder in the Mountains, an annual leather event held in Denver, Colorado. This particular conference for the Leather/BDSM subculture has a focus on SM skills and play and has a reputation of being a gathering of highly experienced individuals. The Saturday and Sunday daytime presentations consisted of four tracks, each of eight 90-min seminars on advanced SM techniques. On Friday and Saturday nights, the convention hosted dungeon play parties held in the host

<sup>2</sup> One of the 12 women who performed the bottom role had indicated on her pre-scene questionnaire that she would be performing the top role. On her post-scene questionnaire, she explained that she had, “planned on topping & wrote that on 1st form. [But] ended up meeting two friends and bottoming to them.” This participant was included as a top in the analyses of baseline data and included as a bottom in the analyses of during-scene and post-scene data. Alternative analyses that (a) include this participant exclusively as a top, (b) include this participant exclusively as a bottom, and (c) exclude this participant, do not substantively change the results.

hotel's convention center. At the opening ceremonies on Friday evening, the Director of Thunder in the Mountains announced the study to the participants and the researchers maintained an information table throughout the event.

The parties took place in a 33,000 square foot convention hall equipped with 40 Saint Andrews crosses for flogging, 40 bondage ladders, numerous spanking benches, suspension structures, wrestling space, and other dedicated SM equipment to accommodate many simultaneous SM scenes. As with the APEX space, all scenes that occurred at the parties took place within sight of other party attendees. The researchers set up a table in the middle of the room where participants completed surveys and provided saliva samples. Party attendees who approached the researchers' table were asked if they planned to participate in a scene that evening. Thereafter, the procedure was similar to Study 1 with the following exceptions: Because of the large number of simultaneous scenes as well as the large party space, participants were given fluorescent cards to indicate that their scenes were part of the study and the two observers recorded the end times of the scenes but did not take notes on the SM activities.

All participants gave baseline samples. Thereafter, 45 participants actually participated in a scene (21 tops, 24 bottoms). Forty-two gave during-scene samples (18 tops, 24 bottoms). Twenty-three participants gave post-scene samples (9 tops, 14 bottoms). In addition, all participants filled out the first questionnaire. Thirty-two participants filled out the second questionnaire (14 tops, 18 bottoms).

### Measures

Saliva samples for Study 1 were assayed by Salimetrics, LLC (State College, PA). We performed the assays for Study 2 using immunoassay kits purchased from Salimetrics, LLC. All saliva samples were stored on wet ice during the collection procedures and then at  $-65^{\circ}\text{C}$  for later analysis. Immediately prior to assaying, samples were thawed and centrifuged ( $1,500 \times g$  at 3,000 rpm for 15 min) to separate saliva from any other matter. Each sample was analyzed for both testosterone and cortisol, without freezing the sample again.

Each sample, standard or assay diluent, was pipetted into plates pre-coated with antibodies for either cortisol or testosterone. The standard for cortisol ranged from 0.007  $\mu\text{g}/\text{dl}$  (0.20 nmol/l) to 1.8  $\mu\text{g}/\text{dl}$  (49.7 nmol/l), with an average 3.60% intra-assay and 2.57% inter-assay coefficient of variation. The standard for testosterone ranged from 3.7 pg/ml (12.8 pmol/l) to 360.0 pg/ml (1248.3 pmol/l), with an average 3.68% intra-assay and 1.79% inter-assay coefficient of variation. All samples were assayed in duplicate: cortisol:  $r(142) = .97$ ,  $p < .001$ ; testosterone:  $r(135) = .96$ ,  $p < .001$ . Sample duplicates that exceeded the standard curve were diluted with sample buffer and re-assayed the same day. Likewise, sample duplicates that varied more than 20%

were re-assayed on another plate the same day and the new values reported (Kivlighan et al., 2004).

The immunoassays were conducted in accordance with the directions from the Salimetrics, LLC hormone kits. Briefly, the conjugate was added to each well with the sample, standard or assay diluent. After 60 min, the plate was rinsed and 200  $\mu\text{l}$  of tetramethylbenzidine (TMB) solution was added. The reaction was stopped after 30 min with sulfuric acid and the plate read with a BioRad E1A Reader Model 2550 within 10 min at wavelength 450 nm.

Two samples were not analyzed for cortisol and six samples were not analyzed for testosterone due to insufficient quantity of saliva. One participant reported using topical testosterone medication. Testosterone data from this participant (three samples) were excluded from statistical analyses. Cortisol levels for two (out of 144) samples and testosterone levels for five (out of 137) samples fell outside of the standard range and were excluded from statistical analyses.

The pre-scene questionnaire was similar to the pre-scene questionnaire used in Study 1 with three additions. Participants were asked to report their sexual orientation (gay/lesbian, bisexual, heterosexual, other), their consumption of foods and other behaviors that might impact saliva samples (e.g., alcohol, tooth brushing, major meals, dairy, and acidic or high sugar food or drink), and whether they had participated in this study before. No saliva samples or participants were excluded on the basis of these questions. The post-scene questionnaire was similar to the post-scene questionnaire used in Study 1 with the removal of one question: whether they anticipated having sex tonight. In both questionnaires, the IOS was expanded from five points to seven points.

## Results

### *Physiological and Psychological Effects*

**Baseline Differences:** In the pre-scene questionnaire, bottoms ( $M = 5.6$ ,  $SD = 1.1$ ) and tops ( $M = 5.6$ ,  $SD = 1.2$ ) did not differ in their anticipation of the evening's scene,  $t(46) < 1$ . Both bottoms and tops felt comfortable about the upcoming scene (bottoms:  $M = 5.6$ ,  $SD = 1.6$ ; tops:  $M = 5.7$ ,  $SD = 1.6$ ),  $t(46) < 1$ .

As in Study 1, at baseline, tops and bottoms did not differ significantly in levels of cortisol (tops:  $M = 0.070$   $\mu\text{g}/\text{dl}$ ,  $SD = 0.057$   $\mu\text{g}/\text{dl}$ ; bottoms:  $M = 0.113$   $\mu\text{g}/\text{dl}$ ,  $SD = 0.139$   $\mu\text{g}/\text{dl}$ ),  $t(34.08) = 1.44$ . Including participant sex as a factor produced no main effect of sex or a Role  $\times$  Sex interaction.

Baseline differences in testosterone were examined using a 2 (sex)  $\times$  2 (role) between-subjects ANOVA. At baseline, women ( $M = 174.8$  pg/ml,  $SD = 159.2$  pg/ml) had significantly lower levels of testosterone than men ( $M = 322.9$  pg/ml,  $SD = 308.4$  pg/ml),  $F(1, 41) = 4.72$ ,  $p = .04$ ,

and bottoms ( $M = 300.2$  pg/ml,  $SD = 310.9$  pg/ml) had marginally higher levels of testosterone than tops ( $M = 181.0$  pg/ml,  $SD = 130.6$  pg/ml),  $F(1, 41) = 3.69, p = .06$ .

**During-scene Effects:** During-scene effects were examined using  $2$  (role)  $\times$   $2$  (time) mixed model ANOVAs. Between baseline and 20 min after the scene, cortisol rose significantly for bottoms,  $F(1, 37) = 11.43, p = .002$ , but not for tops,  $F(1, 37) < 1$  (see Table 4). These different patterns resulted in a marginal main effect of role,  $F(1, 37) = 3.63, p = .06$ , a significant main effect of time,  $F(1, 37) = 6.74, p = .01$ , and a marginal Role  $\times$  Time interaction,  $F(1, 37) = 3.99, p = .05$ .

Because the sexes might differ in their hormonal responses, separate analyses were conducted for women and men for the during-scene and post-scene effects. For the during-scene effect of cortisol, similar patterns emerged for women and men. For women, between baseline and 20 min after the scene, cortisol rose significantly for bottoms,  $F(1, 17) = 5.37, p = .03$ , but not for tops,  $F(1, 17) < 1$  (see Table 4). For men, between baseline and 20 min after the scene, cortisol also rose significantly for bottoms,  $F(1, 18) = 5.55, p = .03$ , but not for

tops,  $F(1, 18) < 1$ . For men, there was also a significant main effect of time,  $F(1, 18) = 4.85, p = .04$ .

Between baseline and 20 min after the scene, testosterone levels did not change significantly for bottoms,  $F(1, 33) < 1$ , or tops,  $F(1, 33) < 1$  (see Table 4). Somewhat different patterns emerged for women and men. For women, between baseline and 20 min after the scene, testosterone levels rose significantly for bottoms,  $F(1, 17) = 8.50, p = .01$ , but not for tops,  $F(1, 17) < 1$ . The ANOVA revealed a significant main effect of time,  $F(1, 17) = 5.38, p = .03$ . For men, between baseline and 20 min after the scene, testosterone levels did not change significantly for bottoms,  $F(1, 14) = 1.20$ , or tops,  $F(1, 14) < 1$ .

**Post-scene Effects:** In Study 2, participants rated one scene as having gone poorly (“The worst scene we’ve ever had!” “went poorly”). These participants completed the post-scene questionnaire, but they did not provide saliva samples 40 min after their scene. Thus, analyses of post-scene questionnaire measures include these participants, but analyses of post-scene physiological measures do not.

After the scene was over, participants whose scenes went well reported that they felt significantly more comfortable with the scene and significantly more “afterglow buzz” (comfort:  $M = 6.0, SD = 1.1$ , afterglow:  $M = 5.2, SD = 1.5$ ) compared to participants whose scene went poorly (comfort:  $M = 2.0, SD = 0.0$ , afterglow:  $M = 1.5, SD = 0.7$ ),  $t(29) = -5.05, p < .001, t(30) = -3.49, p = .002$ , respectively. In addition, participants whose scenes went well rated the scenes as significantly more intense and sexually arousing (intense:  $M = 5.2, SD = 1.5$ , arousing:  $M = 4.6, SD = 1.8$ ) compared to participants whose scene went poorly (intense:  $M = 2.0, SD = 1.4$ , arousing:  $M = 1.5, SD = 0.7$ ),  $t(30) = -3.00, p = .005, t(30) = -2.38, p = .02$ , respectively.

The physiological post-scene effects were examined using  $2$  (role)  $\times$   $2$  (time) mixed model ANOVAs. For cortisol, there was a marginal main effect of time with bottoms and tops showing reductions in cortisol from 20 min to 40 min after the scene,  $F(1, 18) = 3.25, p = .09$  (see Table 5 and Fig. 2). There was also a marginal main effect of role with bottoms showing higher levels of cortisol than tops,  $F(1, 18) = 3.80, p = .07$ .

For testosterone, bottoms and tops did not show significant changes from 20 to 40 min after the scene,  $F(1, 17) = 1.31$  (see Table 5). When examined separately by sex, women showed a significant Role  $\times$  Time interaction,  $F(1, 7) = 6.06, p = .04$ . An examination of the means reveals that female bottoms showed a reduction in testosterone from 20 to 40 min after the scene, whereas female tops showed an increase. There were no significant effects for men.

In contrast to Study 1, the effects of SM scenes on relationship closeness were not moderated by scene quality: The Scene Quality  $\times$  Time interaction was non-significant for the perceptions of the dyad as “we,”  $F(1, 28) < 1$ , and for the

**Table 4** Study 2: Changes in cortisol and testosterone between baseline and 20 min after the scene

Group	Baseline		20 Min after scene		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
<i>Cortisol</i>					
Bottoms					
Female ( <i>n</i> = 12)	0.067	0.062	0.241	0.318	0.76*
Male ( <i>n</i> = 9)	0.100	0.120	0.289	0.348	0.73*
Total ( <i>n</i> = 21)	0.081	0.090	0.262	0.324	0.76**
Tops					
Female ( <i>n</i> = 7)	0.082	0.070	0.065	0.033	-0.31
Male ( <i>n</i> = 11)	0.075	0.061	0.124	0.087	0.65
Total ( <i>n</i> = 18)	0.078	0.063	0.101	0.076	0.33
<i>Testosterone</i>					
Bottoms					
Female ( <i>n</i> = 12)	122.0	143.7	239.0	169.5	0.74*
Male ( <i>n</i> = 8)	455.0	456.6	347.0	262.0	-0.29
Total ( <i>n</i> = 20)	255.2	341.7	282.2	211.8	0.09
Tops					
Female ( <i>n</i> = 7)	140.3	78.5	176.7	121.0	0.36
Male ( <i>n</i> = 8)	250.1	171.3	275.6	99.9	0.18
Total ( <i>n</i> = 15)	198.8	143.2	229.4	117.8	0.23

*Note:* Cortisol levels are in units of  $\mu\text{g/dl}$ . Testosterone levels are in units of pg/ml. Following the recommendations of Dunlap et al. (1996), effect sizes (Cohen’s *d*) were calculated using the original standard deviations for each group rather than the paired-samples *t*-test. Significance tests represent the simple effects within the context of the full ANOVAs

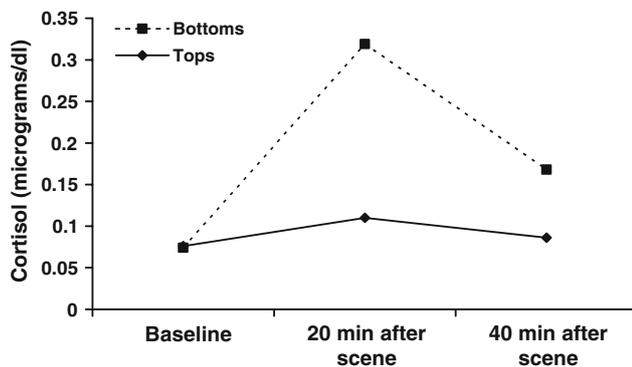
\*  $p < .05$ , \*\*  $p < .01$

**Table 5** Study 2: Changes in cortisol and testosterone between 20 and 40 min after the scene

Group	20 Min after scene		40 Min after scene		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
<i>Cortisol</i>					
Bottoms					
Female ( <i>n</i> = 6)	0.388	0.408	0.191	0.156	−0.64 <sup>†</sup>
Male ( <i>n</i> = 6)	0.252	0.197	0.182	0.175	−0.38
Total ( <i>n</i> = 12)	0.320	0.314	0.186	0.158	−0.54*
Tops					
Female ( <i>n</i> = 3)	0.072	0.045	0.057	0.016	−0.44
Male ( <i>n</i> = 5)	0.132	0.085	0.104	0.066	−0.37
Total ( <i>n</i> = 8)	0.110	0.076	0.086	0.056	−0.36
<i>Testosterone</i>					
Bottoms					
Female ( <i>n</i> = 6)	192.3	56.4	137.2	61.8	−0.93 <sup>†</sup>
Male ( <i>n</i> = 6)	297.6	287.3	181.8	83.6	−0.55
Total ( <i>n</i> = 12)	244.9	204.9	159.5	73.9	−0.55*
Tops					
Female ( <i>n</i> = 3)	176.8	88.4	227.5	79.7	0.60
Male ( <i>n</i> = 4)	226.2	82.0	207.6	82.7	−0.23
Total ( <i>n</i> = 7)	205.0	81.6	216.1	75.2	0.14

Note: Cortisol levels are in units of  $\mu\text{g}/\text{dl}$ . Testosterone levels are in units of  $\text{pg}/\text{ml}$ . Following the recommendations of Dunlap et al. (1996), effect sizes (Cohen's *d*) were calculated using the original standard deviations for each group rather than the paired-samples *t*-test. Significance tests represent the simple effects within the context of the full ANOVAs

\*  $p < .05$ , <sup>†</sup>  $p < .10$



**Fig. 2** Study 2: Changes in cortisol between baseline, 20 min after the scene, and 40 min after the scene. Because this Figure includes only those participants with cortisol data at all three time points, some of the means differ from the means presented in Tables 4 and 5

IOS,  $F(1, 27) < 1$ . Instead, both bottoms and tops showed a significant increase in the IOS from before the scene ( $M = 4.3$ ,  $SD = 1.9$ ) to after the scene ( $M = 4.8$ ,  $SD = 1.9$ ),

$t(28) = -2.71$ ,  $p = .01$ ,<sup>3</sup> but not for the perceptions of the dyad as “we” from before the scene.

## Discussion

Study 2 was designed to replicate and extend the findings of Study 1 by examining the physiological and psychological effects of SM activities on a larger, more diverse sample of participants. Consistent with predictions and the findings of Study 1, bottoms showed significant increases in cortisol during their scenes, but tops did not. After their scenes, bottoms and tops showed decreases in cortisol and increases in relationship closeness. Female bottoms also showed significant increases in testosterone during their scenes and marginal decreases in testosterone after their scenes.

## General Discussion

The goals of the present studies were to examine the physiological and psychological effects of consensual SM activities and to document SM behaviors that fall outside of the current taxonomies in the academic literature.

### Observed SM Activities

In Study 1, a wide variety of activities were observed across the seven scenes. Of particular note were caring behaviors (e.g., massaging, attention and focus to the physical and emotional state of the bottom) and acts of intimacy (e.g., eye contact, whispering) observed both during the scenes and as the primary activity of aftercare. These activities were observed in all scenes, including the scene that went poorly. During this latter scene, for example, the top noticed that the bottom, who was bent over a table, was moving one foot around, unable to rest it evenly on the floor. The top paused the scene activities and placed a pillow under the bottom's foot to relieve the discomfort.

Alison et al. (2001) empirically derived four sexual scripts related to SM activities: hypermasculinity (e.g., rimming, using dildos, cockbinding, etc.), administration of pain (e.g., applying clothespins, spanking, caning, etc.), humiliation (e.g., flagellation, verbal humiliation, gags, etc.), and physical restriction (e.g., bondage, handcuffs, chains, etc.). The activities observed in Study 1 fell predominantly into the categories

<sup>3</sup> A series of additional analyses were run to examine the effects of relationship duration on relationship closeness. These analyses revealed a main effect of relationship duration (prior relationships were associated with significantly greater degrees of relationship closeness), but relationship duration did not moderate the effects of SM activities on relationship closeness. Instead, across all levels of relationship duration, participants reported increases in relationship closeness from before to after their scenes.

of administration of pain and physical restriction. Participants also engaged in flagellation using floggers and whips, although it appeared from observation and casual conversation that the tops used flagellation as a method of administering pain rather than as a way to inflict humiliation—a distinction consistent with SM community literature, which typically describes flagellation as a method of administering pain (Bean, 2000).

It is also notable that a number of observed activities—pleasurable stimulation, some emotional and psychological techniques, caring behaviors, and acts of intimacy—were not represented within Alison et al.'s (2001) categories. The present data suggest that an additional category should be considered related to the caring behaviors. The presence of caring behaviors and acts of intimacy, in particular, highlight the concern that SM participants have for the well-being and positive experience of their scene partners and the positive relationship context in which the scenes occur.

### Relationship Closeness

Across both studies, bottoms and tops whose scenes went well reported increases in relationship closeness. In Study 1, the bottom and top whose scene went poorly reported decreases in relationship closeness, but this effect did not replicate for the bottom and top whose scene went poorly in Study 2. The increases in relationship closeness observed in scenes that went well align with the possible benefits of SM activities discussed in the SM community literature (Baldwin, 1993; Mains, 1984). The divergent results for the two scenes that went poorly suggest that the effects of negative SM experiences on relationship closeness may be moderated by a variety of factors such as the history of the relationship, the reason the scene went poorly, the experience level of the participants, etc.

In these studies, relationship closeness was assessed immediately after the scene. Thus, the observed changes in relationship closeness represent an immediate effect that is likely to be temporary. However, in line with Cutler's (2003) finding that SM couples use SM scenes to increase the intimacy of their relationships, the SM scenes in the present studies might have had a smaller lasting effect on the intimacy between the participants.

### Physiological Stress

A second goal of these studies was to assess physiological stress using salivary cortisol. Based on Kirschbaum and Hellhammer's (1994) statement that cortisol will rise "in situations with high ego-involvement, low predictability, low controllability, and novelty" (p. 318), we predicted that cortisol levels would rise for bottoms due to the lack of control inherent in the bottom role as well as the physical pain experienced by bottoms in some scenes. Consistent with predictions, between baseline and 20 min after the scene,

cortisol levels rose significantly for bottoms but not for tops. Although Kirschbaum and Hellhammer included four factors in their statement—ego-involvement, predictability, controllability, and novelty—it seems likely that the bottom and top roles differ more in terms of predictability and controllability than ego-involvement and novelty. Indeed, these results may reveal something about the complex issue of control within an SM scene.

With respect to control, some SM practitioners claim that the bottom is truly in control of the SM scene, and in the larger context in which the scene takes place, the bottom does wield considerable control. Before many scenes begin, the bottom and top negotiate the activities that will take place, with the bottom setting limits on the type and intensity of the activities. During the scene, the bottom can use safewords that indicate to the top that the activities must slow down or change or that the scene must end immediately. Thus, the bottom exerts control both before and during the scene. However, during the scene, the top controls the nature and intensity of the activities, and particularly for scenes involving bondage, the bottom must rely on the top to stay within the negotiated parameters of the scene and to respect any safewords the bottom uses. In addition, some bottoms avoid negotiation and safewords as a means of surrendering more control (Easton & Liszt, 1998). In the present studies, the rise in cortisol observed in bottoms but not tops might reflect the control the top exerts during the scene itself despite the control some bottoms exert in negotiation and in the use of safewords.

After the scene, bottoms and tops whose scenes went well showed reductions in cortisol from 20 to 40 min after the end of the main scene activities, whereas the bottom whose scene went poorly showed an increase in cortisol. The reductions in cortisol observed in scenes that went well represent a downward trajectory, although cortisol levels at 40 min still exceeded cortisol levels at baseline. SM community authors have discussed the stress-reducing benefits of SM activities (Baldwin, 1993; Mains, 1984), but sampling at later time points would be needed to determine whether cortisol levels eventually drop below baseline. Baldwin (1993) and Mains (1984) also recognize the risks inherent in SM activities—risks visible in Study 1 in the stress reactions of the bottom whose scene went poorly. It should be noted, however, that the increase in cortisol observed in the bottom whose scene went poorly represents the reactions of only one participant. As such, this effect should be interpreted with caution.

### Testosterone

Study 2 provided an exploratory analysis of the effects of SM activities on participants' levels of testosterone. Two effects emerged: a baseline difference between bottoms and

tops and a change in testosterone during and after the scene in female bottoms.

In Study 2, bottoms displayed marginally higher levels of baseline testosterone than tops. This effect was unexpected and it appears to contradict prior findings that higher levels of testosterone are associated with greater dominance (e.g., Dabbs et al., 1987, 1988; see also Dabbs & Dabbs, 2000). We offer three possible interpretations: (1) Most baseline samples were collected at the parties. Bottoms' higher levels of baseline testosterone might represent a physiological response to an anticipated physical, emotional, or endurance challenge, similar to the finding that tennis players' testosterone levels rose just before a match (Booth et al., 1989). (2) Bottoms' higher testosterone levels might represent dominance in other contexts. According to Baumeister (1997):

Within modern Western society, masochism is far from evenly distributed. The evidence is not entirely conclusive, in part because of methodological difficulties of obtaining reliable information about who engages in such unusual forms of sex play. Still, it seems that masochists are mainly drawn from among the upper and upper middle classes, especially including people who are well-off and successful or even powerful in their daily lives. Thus, prostitutes who cater to rich and powerful clients find that they are often asked to give spankings or similar services, and indeed their clients want to be spanked much more often than they desire to spank the prostitute... (pp. 142–143)

Baumeister's theory posits that some masochists take part in SM activities to relieve temporarily the stress of maintaining a self identity—stress that might be particularly intense for those with power and responsibility. If the bottoms in our sample represent a group of people who are dominant in other contexts, their higher levels of baseline testosterone might align with their dominant professions, dominant positions in social hierarchies, etc. (see Dabbs & Dabbs, 2000). In this case, their bottom roles in an SM context would represent a notable exception to their dominance elsewhere. (3) The baseline differences might be a spurious finding. Future studies could examine these baseline differences more precisely by obtaining saliva samples outside of a party context and by assessing non-SM dominance-related characteristics.

In addition, in Study 2, testosterone levels rose significantly in female bottoms between baseline and 20 min after the scene and then dropped marginally between 20 and 40 min after the scene. Male bottoms showed non-significant drops in testosterone between baseline and 20 min after the scene and between 20 and 40 min after the scene. A rise in testosterone levels in female bottoms between baseline and 20 min after the scene might indicate an aggressive response to the SM activities, and the subsequent reduction

in testosterone levels might then indicate a reduction in this aggression. Alternately, female bottoms' increase in testosterone between baseline and 20 min after the scene might be associated with increases in positive mood during this same period. Consistent with this, McCaul, Gladue, and Joppa (1992) found that winning a chance-based task caused increases in positive mood and testosterone in men, and that the increases in positive mood may have mediated the increases in testosterone. Future studies could test these speculations by assessing bottoms' self-reported experiences of aggression and other moods and correlating these reports with changes in testosterone levels. These findings should be interpreted with caution, however, as they were unpredicted and based on a relatively small sample size.

### Limitations and Future Directions

One of the major limitations of the present studies is the high proportion of missing data for some measures. Given the nature of the data collection environment and the corresponding need to minimize the interference with participants' activities as much as possible, we had greater control over some measures than others. As such, we collected complete data on the pre-scene questionnaire and the baseline saliva samples, which were administered after informed consent was verbally given. For the pre-scene measure, participants were asked to stop by the saliva sampling station approximately 10 min before beginning their scene. Because we did not know when participants would choose to begin their scenes (indeed, participants themselves may not have known upon enrolling in the study what time they would begin their scenes), and because SM community norms prohibit interfering with a scene once it has begun, we had much less control over this measure, and as a result, fewer participants provided it. Because we had few pre-scene samples, we cannot determine whether activities that occurred between baseline and the beginning of a scene might have impacted during-scene measures. In contrast to the pre-scene measure, we could observe when scenes ended and approach participants at that point. Thus, nearly complete during-scene data were obtained from participants. Thereafter, some participants chose not to complete the post-scene questionnaire and some participants chose to leave the party before another 20 min had passed. It would, of course, have been preferable to have complete data from all participants, but because missing data status was unrelated to role, these missing data are unlikely to threaten the internal validity of the statistical tests (Cook & Campbell, 1979; West & Sagarin, 2000). The major effect of the missing data was loss of statistical power, most notably in the pre-scene samples.

The present studies were conducted in a naturalistic setting with a procedure designed to minimize the impact of the

research on the participants' activities. One cost of these design decisions was variability in the timeframe and timing of the saliva samples and in the nature and intensity of the activities that preceded each sample. Future studies could reduce this variability by requesting saliva samples from participants at fixed intervals with baseline samples potentially collected outside of a party context. This would require prior consent of participants, of course, and the periodic interruption of scene activities runs the risk of affecting the activities and participants' responses to the activities. Nevertheless, this procedure would enable a more precise analysis of hormonal changes before, during, and after a scene as well as an analysis of hormone levels corresponding to specific observed activities.

The present studies are also limited in terms of the moderators we examined. We focused primarily on SM role and, to a lesser extent, on scene quality as moderators. A number of other factors are likely to moderate the observed effects. For example, participants' past experience with SM activities might moderate the increase in bottoms' cortisol levels during the scene. Experienced SM practitioners adopting the bottom role may perceive their scenes as more predictable, more controllable, and less novel than novice SM practitioners. Likewise, novice SM practitioners adopting the top role might show cortisol effects similar to bottoms if their lack of experience undermines the feelings of control implied by their role. Future studies could also examine the moderating effects of scene length and pace. A quick, intense scene might yield different physiological and psychological responses than a long, drawn out, leisurely paced scene. In addition, future studies should measure scene quality on a continuous scale to increase the precision and statistical power of the test of this moderator.

Future studies might also benefit from an assessment of additional outcome variables. In particular, measures of mood and affect changes that occur during and after a scene would provide additional insight into the effects of SM activities and into potential mediators of the changes in cortisol, testosterone, and relationship closeness. In addition, an assessment of the degree of pain experienced by bottoms during their scenes would enable a test of whether the increases in cortisol are associated with physical pain or loss of control or predictability. A number of options are available for assessing these variables: (1) After the scene, participants could give retrospective accounts of their mood, affect, and pain levels during the scene, (2) scenes could be videotaped, and after the scene, participants could watch the tape and report the mood, affect, and pain levels they were feeling during the scene, or (3) researchers could interrupt the scene at regular intervals to assess mood, affect, and pain levels. Clearly, these options represent tradeoffs between the accuracy of the measures and the unobtrusiveness of the

procedures. Invasive procedures increase the risk of impacting participants' experiences and, hence, the measured outcomes.

Finally, as suggested by an anonymous reviewer, the effects observed in the present studies might not be unique to SM. They might merely be the result of one person being dominant over another in front of an audience. Future studies could test this in a number of ways. First, researchers could examine participants' physiological and psychological reactions in analogous situations not involving SM (e.g., one person gives spoken social attention to another in front of an audience). Results similar to the present ones would suggest that the effects are not unique to SM. More generally, our understanding of SM could benefit from an attempt to reproduce facets of the SM experience in controlled environments. An alternative approach would be to collect data on private SM scenes. Participants trained in the data collection protocol could be given questionnaires and vials for saliva samples and instructed to complete the questionnaires and provide saliva samples before and after their next private SM scene. This would eliminate any audience effects as well as effects that emerge from being in a potentially unfamiliar environment (as was the case for participants in Study 2).

In conclusion, the present studies sought to examine the physiological and psychological effects of consensual sado-masochistic activity. Although the physiological reactions of bottoms and tops tended to differ, the psychological reactions converged, with bottoms and tops reporting increases in relationship closeness after their scenes. These results, combined with the displays of caring and affection observed in scenes, offer support for the modern view of SM as an expression of human sexuality that, when performed consensually, has the potential to increase the intimacy between participants.

**Acknowledgements** Some of the findings reported here were presented at the May 2002 meeting of the Midwestern Psychological Association in Chicago, Illinois and the February 2003 meeting of the Society for Personality and Social Psychology in Los Angeles, California. We thank the Board of Directors and Membership of the Arizona Power Exchange, and the Director, Staff, and Attendees of Thunder in the Mountains for their invaluable help in running the studies. We also thank the Board of Directors and Membership of the Leather Rose and the Membership of Kinky by Association for providing comments and feedback on the studies and pilot samples for the assay procedure used in Study 2. Finally, we thank Doug Granger and Jeremy Trexel for their advice regarding the physiological measures, and Evelyn Comber, Rich Dockter, and Doug Henderson for their comments on an earlier draft of this article.

## Appendix

Pre-scene questionnaire:

- (1) Gender (choices: female, male, other).
- (2) Age.

- (3) What prescription or over-the-counter drugs have you taken in the last 24 h?
- (4) Typical BDSM roles (choices: top, bottom, switch, dominant, submissive, sadist, masochist).
- (5) What role will you be playing in the upcoming scene?
- (6) What is your relationship to the other person or people in the scene?
- (7) How much have you played with this person or people before tonight?
- (8) How much are you anticipating tonight's scene? (on a 1–7 scale, with 1 indicating “not at all” and 7 indicating “very much”.)
- (9) How comfortable do you feel about the upcoming scene? (on a 1–7 scale, with 1 indicating “very uncomfortable” and 7 indicating “very comfortable”.)
- (10) To what extent would you use the word “we” to describe your relationship with the other person or people in the scene? (on a 1–7 scale, with 1 indicating “I definitely would not use the word ‘we’ to describe my relationship” and 7 indicating “I definitely would use the word ‘we’ to describe my relationship”.)
- (11) Which of the following pairs of circles best represents your relationship with the other person or people in the scene? (from a choice of five increasingly overlapping pairs.)

Post-scene questionnaire:

- (1) Please describe your activities during the scene.
- (2) How do you feel about the scene? Did it go well? Did it go poorly?
- (3) How comfortable did you feel about the scene? (on a 1–7 scale, with 1 indicating “very uncomfortable” and 7 indicating “very comfortable”.)
- (4) How sexually aroused were you during tonight's scene? (on a 1–7 scale, with 1 indicating “no arousal” and 7 indicating “orgasm”.)
- (5) Do you anticipate having sex tonight? (on a 1–7 scale, with 1 indicating “definitely not” and 7 indicating “definitely”.)
- (6) What activity was most sexually arousing for you?
- (7) How intense was tonight's scene? (on a 1–7 scale, with 1 indicating “not at all intense” and 7 indicating “very intense”.)
- (8) What activity was most intense for you?
- (9) How much of an afterglow buzz do you currently feel? (on a 1–7 scale, with 1 indicating “none” and 7 indicating “lots”.)
- (10) Now that the scene is over, to what extent would you use the word “we” to describe your relationship with the other person or people in the scene? (on a 1–7 scale, with 1 indicating “I definitely would not use the word ‘we’ to describe my relationship” and 7 indicating “I

definitely would use the word ‘we’ to describe my relationship”.)

- (11) Which of the following pairs of circles best represents your relationship with the other person or people in the scene? (from a choice of five increasingly overlapping pairs.)

## References

- Alison, L., Santtila, P., Sandnabba, N. K., & Nordling, N. (2001). Sadomasochistically-oriented behavior: Diversity in practice and meaning. *Archives of Sexual Behavior*, *30*, 1–12.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596–612.
- Baldwin, G. (1993). *Ties that bind: The SM/leather/fetish erotic style*. San Francisco: Daedalus Publishing Company.
- Bannon, R. (1993). *Learning the ropes: A basic guide to safe and fun s/m lovemaking*. Los Angeles: Daedalus Publishing Company.
- Baumeister, R. F. (1997). The enigmatic appeal of sexual masochism: Why people desire pain, bondage, and humiliation in sex. *Journal of Social and Clinical Psychology*, *16*, 133–150.
- Bean, J. W. (2000). *Flogging*. San Francisco: Greenery Press.
- Benjamins, C., Asscheman, H., & Schuur, A. H. (1992). Increased salivary cortisol in severe dental anxiety. *Psychophysiology*, *29*, 302–305.
- Bohnen, N., Houx, P., Nicolson, N., & Jolles, J. (1990). Cortisol reactivity and cognitive performance in a continuous mental task paradigm. *Biological Psychology*, *31*, 107–116.
- Bohnen, N., Nicolson, N., Sulon, J., & Jolles, J. (1991). Coping style, trait anxiety and cortisol reactivity during mental stress. *Journal of Psychosomatic Research*, *35*, 141–147.
- Booth, A., Shelley, G., Mazur, A., Tharp, G., & Kittok, R. (1989). Testosterone, and winning and losing in human competition. *Hormones and Behavior*, *23*, 555–571.
- Brame, G., Jacobs, J., & Brame, W. (1993). *Different loving: An exploration of the world of sexual dominance and submission*. New York: Villard Books.
- Breslow, N., Evans, L., & Langley, J. (1985). On the prevalence and roles of females in the sadomasochistic subculture: Report on an empirical study. *Archives of Sexual Behavior*, *14*, 303–317.
- Breslow, N., Evans, L., & Langley, J. (1986). Comparisons among heterosexual, bisexual, and homosexual male sadomasochists. *Journal of Homosexuality*, *13*, 83–107.
- Brodsky, J. I. (1993). The Mineshaft: A retrospective ethnography. *Journal of Homosexuality*, *24*, 233–251.
- Califia, P. (1993). *Sensuous magic: A guide for adventurous couples*. San Francisco: Cleis Press.
- Chard, T. (1990). *An introduction to radioimmunoassay and related techniques*. Amsterdam: Elsevier.
- Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy–altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, *73*, 481–494.
- Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the southern culture of honor: An “experimental

- ethnography". *Journal of Personality and Social Psychology*, 70, 945–960.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Cook, N. J., Read, G. F., Walker, R. F., Harris, B., & Riad-Fahmy, D. (1992). Salivary cortisol and testosterone as markers of stress in normal subjects in abnormal situations. In C. Kirschbaum, G. F. Read, & D. H. Hellhammer (Eds.), *Assessment of hormones and drugs in saliva in biobehavioral research* (pp. 147–162). Seattle: Hogrefe & Huber.
- Cutler, A. B. (2003). *Partner selection, power dynamics, and sexual bargaining in self-defined BDSM couples*. Unpublished doctoral dissertation, The Institute for Advanced Study of Human Sexuality, San Francisco.
- Dabbs, J. M. (1993). Salivary testosterone measurements in behavioral studies. *Annals of the New York Academy of Sciences*, 694, 177–183.
- Dabbs, J. M., & Dabbs, M. G. (2000). *Heroes, rogues, and lovers: Testosterone and behavior*. New York: McGraw Hill.
- Dabbs, J. M., de LaRue, D., & Williams, P. M. (1990). Testosterone and occupational choice: Actors, ministers, and other men. *Journal of Personality and Social Psychology*, 59, 1261–1265.
- Dabbs, J. M., Frady, R. L., Carr, T. S., & Besch, N. F. (1987). Salivary testosterone and criminal violence in young adult prison inmates. *Psychosomatic Medicine*, 49, 174–182.
- Dabbs, J. M., & Mohammed, S. (1992). Male and female salivary testosterone concentrations before and after sexual activity. *Physiology and Behavior*, 52, 195–197.
- Dabbs, J. M., Ruback, R. B., Frady, R. L., Hopper, C. H., & Sgoutas, D. S. (1988). Saliva testosterone and criminal violence among women. *Personality and Individual Differences*, 9, 269–275.
- Dixon, K. E., Thorn, B. E., & Ward, L. C. (2004). An evaluation of sex differences in psychological and physiological responses to experimentally-induced pain: A path analytic description. *Pain*, 112, 188–196.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170–177.
- Easton, D., & Liszt, C. A. (1998). *The bottoming book: How to get terrible things done to you by wonderful people*. San Francisco: Greenery Press.
- Freud, S. (1938). *The basic writings of Sigmund Freud* (A. A. Brill, Trans. and Ed.). New York: Modern Library.
- Gebhard, P. (1969). Fetishism and sadomasochism. In J. H. Masserman (Ed.), *Dynamics of deviant sexuality* (pp. 71–80). New York: Grune & Stratton.
- Gladue, B. A., Boechler, M., & McCaul, D. D. (1989). Hormonal response to competition in human males. *Aggressive Behavior*, 15, 409–422.
- Kamel, G. W. (1980). Leathersex: Meaningful aspects of gay sadomasochism. *Deviant Behavior*, 1, 171–191.
- Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology*, 19, 313–333.
- Kivlighan, K. T., Granger, D. A., Schwartz, E. B., Nelson, V., Curran, M., & Shirtcliff, E. A. (2004). Quantifying blood leakage into the oral mucosa and its effects on the measurement of cortisol, dehydroepiandrosterone, and testosterone in saliva. *Hormones and Behavior*, 46, 39–46.
- Krafft-Ebing, R. V. (1965). *Psychopathia sexualis* (F. S. Klaff, Trans.). New York: Stein & Day (Original work published 1886).
- Mains, G. (1984). *Urban aboriginals: A celebration of leather sexuality*. San Francisco: Gay Sunshine Press.
- Mazur, A., Booth, A., & Dabbs, J. M. (1992). Testosterone and chess competition. *Social Psychology Quarterly*, 55, 70–77.
- McCaul, K. D., Gladue, B. A., & Joppa, M. (1992). Winning, losing, mood, and testosterone. *Hormones and Behavior*, 26, 486–504.
- Miller, P., & Devon, M. (1995). *Screw the roses, send me the thorns: The romance and sexual sorcery of sadomasochism*. Fairfield, CT: Mystic Rose Books.
- Moser, C., & Levitt, E. E. (1987). An exploratory-descriptive study of a sadomasochistically oriented sample. *Journal of Sex Research*, 23, 322–337.
- Myers, J. (1992). Nonmainstream body modification: Genital piercing, branding, burning, and cutting. *Journal of Contemporary Ethnography*, 21, 267–306.
- Rinella, J. (1994). *The master's manual: A handbook of erotic dominance*. San Francisco: Daedalus Publishing Company.
- Salimetrics, LLC. (2000a, October 14). HS-cortisol high sensitivity salivary cortisol enzyme immunoassay kit. Retrieved from <http://www.salimetrics.com/pdf/CortisolKitInsert90601.pdf>.
- Salimetrics, LLC. (2000b, October 12). Salivary testosterone enzyme immunoassay kit. Retrieved from <http://www.salimetrics.com/pdf/TestosteroneKit.pdf>.
- Sandnabba, N. K., Santtila, P., & Nordling, N. (1999). Sexual behavior and social adaptation among sadomasochistically-oriented males. *Journal of Sex Research*, 36, 273–282.
- Santtila, P., Sandnabba, N. K., Alison, L., & Nordling, N. (2002). Investigating the underlying structure in sadomasochistically oriented behavior. *Archives of Sexual Behavior*, 31, 185–196.
- Singh, D., Vidaurri, M., Zambarano, R. J., & Dabbs, J. M. (1999). Lesbian erotic role identification: Behavioral, morphological, and hormonal correlates. *Journal of Personality and Social Psychology*, 76, 1035–1049.
- Spengler, A. (1977). Manifest sadomasochism of males: Results of an empirical study. *Archives of Sexual Behavior*, 6, 441–456.
- Veldhuis, J. D., King, J. C., Urban, R. J., Rogol, A. D., Evans, W. S., Kolp, S. A., et al. (1987). Operating characteristics of the male hypothalamo-pituitary-gonadal axis: Pulsatile release of testosterone and follicle-stimulating hormone and their temporal coupling with luteinizing hormone. *Journal of Clinical Endocrinology and Metabolism*, 65, 929–941.
- Weinberg, M. S., Williams, C. J., & Moser, C. (1984). The social constituents of sadomasochism. *Social Problems*, 31, 379–389.
- Weinberg, T. S. (1994). Research in sadomasochism: A review of sociological and social psychological literature. *Annual Review of Sex Research*, 5, 257–279.
- West, S. G., & Sagarin, B. J. (2000). Participant selection and loss in randomized experiments. In L. Bickman (Ed.), *Research design: Donald Campbell's legacy* (Vol. 2, pp. 117–154). Thousand Oaks, CA: Sage.
- Wiseman, J. (1996). *SM 101: A realistic introduction*. San Francisco: Greenery Press.