

The Development of In Vivo Measures to Assess the Impact of Sex-Drive Reducing Medications in an Offender with an Intellectual Disability

Jerry A. Rea^{1,2} · Michael R. Dixon¹ · Robert D. Zettle³ · Kasey L. Wright¹

Received: 1 February 2016 / Revised: 6 August 2016 / Accepted: 8 August 2016
© Springer Science+Business Media New York 2016

Abstract The ability to adequately evaluate medications in the treatment of paraphilias has been limited by reliance upon self-report as a measure of effectiveness over periods of time that may be too short to detect reoffending. One solution to this shortcoming is the development of valid, long-term, stable assessment measures. The purpose of this case study was to analyze the effects of Prozac and Provera on an array of behaviors germane to the successful treatment of paraphilias, including: (a) sexual arousal in the laboratory and natural environment, (b) sexual thoughts (deviant and nondeviant) accompanied by arousal in the natural environment, and (c) overt actions in the community associated with increased risk of reoffending over a 31-month period for an exhibitionist with an intellectual disability. Despite the ineffectiveness of the medications, the measures demonstrated long-term, differentiated significant clinical responding; further underscored the importance of assessing deviant sexual arousal and adherence to relapse-prevention procedures in the natural environment; and provided a new methodology to assess sexual preoccupations and sexual arousal. Use of these in vivo measures raises questions regarding their potential to improve the predictability of risk assessments, and serve as an aide in the analysis of whether a treatment procedure is effective for an individual.

Keywords Sexual preoccupation · Deviant sexual arousal · Paraphilias · Intellectual disability

Introduction

Medications such as selective serotonin reuptake inhibitors (e.g., Prozac) and antiandrogens (e.g., Provera, Lupron) have been recommended as a potential adjunctive therapy in the treatment of paraphilias (e.g., Association of the Treatment of Sexual Abusers, 2014; Thibaut et al., 2010). Antiandrogen medications have produced the greatest treatment effects in the research literature (Schmucker & Losel, 2008). The use of these medication classes is quite high in North American sexual-offender treatment programs. Reported use in Canada of antiandrogens in community and residential programs was 44 %, while selective serotonin reuptake inhibitors were 61 %. Use of one or both classes of medications in Canada was 75 and 63 % in residential and community programs, respectively (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010). Medication use in the USA in community and residential programs was 16 % for antiandrogens, while selective serotonin reuptake inhibitors were near 53 %. Data on the use of multiple medication classes within residential and community programs in the U.S. were not reported. Unfortunately and of greater relevance for the purpose of this study, there have been only a handful of randomized control studies and each contained few participants. In addition, these studies were published two decades ago and did not examine the newer selective serotonin reuptake inhibitors or antiandrogen medications (see Khan et al., 2015).

The above medication classes rely upon different biological mechanisms to reduce or eliminate unconventional sexual fantasies, urges, arousal, and behaviors (see Thibaut et al., 2010 for a comprehensive review of the biological mechanisms and the efficacy of these medications). There are several

✉ Jerry A. Rea
Jerry.Rea@pshtc.ks.gov

¹ Parsons State Hospital and Training Center, 2601 Gabriel, Parsons, KS 67357, USA

² University of Kansas Life Span Institute at Parsons, Lawrence, KS, USA

³ Department of Psychology, Wichita State University, Wichita, KS, USA

limitations of such medication studies. First, the vast majority of the medication studies have primarily used self-reports of masturbation, preoccupations, urges, fantasies, offending, and sexual arousal (e.g., Kafka & Prentky, 1992; Winder et al., 2014), and/or arrest records (e.g., Maletzky, Tolan, & McFarland, 2006) to evaluate their effectiveness. However, the unanonymous self-report of offenders can be unreliable as to what controls their sexual arousal (Hinton, O'Neil, & Webster, 1980; Quinsey, Steinman, Bergersen, & Holmes, 1975; Rea et al., 2003) and whether they have recently engaged in deviant behavior (Rosen & Kopel, 1977). In addition, arrest records severely underestimate rates of reoffending (e.g., Abel et al., 1987).

A second limitation of such medication-related treatment was that follow-up was 12 months or less (e.g., Cooper, Cernovsky, & Magnus, 1992; Wincze, Bansal, & Malamud, 1986). Research indicates that reoffending can occur over a period of 25–31 years (Hanson, Steffy, & Gauthier, 1993; Prentky, Lee, Knight, & Cerce, 1997) and that there is no known ceiling for recidivism rates (Doren, 1998). Numerous studies have indicated that once medications were discontinued, urges and fantasies returned (e.g., Schober et al., 2005) and recidivism occurred (e.g., Emory, Cole, & Meyer, 1995; Thibaut, Cordier, & Kuhn, 1996). Thus, some offenders may require treatment for a lifetime (Thibaut et al., 2010).

Finally, another drawback in using medication in the treatment of paraphilias is that its effectiveness may not generalize from the clinic to the natural environment. Rea et al. (2003) reported that a sexual offender with an intellectual disability, who was treated with covert sensitization, reduced his arousal to children in the laboratory as measured by a penile plethysmograph. He also reported elimination of arousal to children. Contrary to his self-report, his arousal to children in the natural environment, measured by a portable-penile plethysmograph, remained high and unaffected.

A measure that would ostensibly strengthen the assessment of a medication's effectiveness in decreasing deviant arousal is the penile plethysmograph. A number of studies reported that deviant arousal assessed in this manner was predictive of reoffending (Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2004; Malcolm, Andrews, & Quinsey, 1993; Marques, Wiederanders, Day, Nelson, & van Ommeren, 2005). Despite these findings, only a third of the studies that used antiandrogen treatment assessed arousal using the penile plethysmograph (Thibaut et al., 2010), while only 11 % of the studies that evaluated selective serotonin reuptake inhibitors (SSRIs) used the penile plethysmograph to assess sexual arousal (Adi et al., 2002).

The ability to adequately evaluate the potential for including medications in the treatment of paraphilias has been limited by reliance upon self-report as a measure of effectiveness over periods of time that may be insufficient to detect reoffending. One solution to this shortcoming may lie in the development of valid, long-term, stable assessment measures. Indeed, there have been recommendations for adequately designed, long-term stud-

ies using medication with offenders who have an intellectual disability in the natural environment (Clarke, 1989; Cooper, 1995), while others have advocated for research that includes multiple outcome measures of sexual responsiveness and other relevant behaviors (Bancroft, 1976; O'Donohue & Letourneau, 1992). The purpose of this case study was to analyze the effects of Prozac and Provera on an array of behaviors germane to the successful treatment of paraphilias including: (a) sexual arousal in the laboratory and natural environment, (b) sexual thoughts (deviant and nondeviant) accompanied by arousal, and (c) overt actions in the community associated with increased risk of reoffending over a 31-month period.

Method

Participant

Originally, four sexual offenders with intellectual disabilities began the study. However, three of the participants dropped out of the study in different phases. Two participants were discharged from the residential-treatment facility. The third participant dropped out of the study, but remained in the residential-treatment program.

The participant was a 24-year-old sexual offender who had been convicted of making obscene telephone calls. He self-reported (confirmed by polygraph examinations) that he had exposed himself to women, engaged in voyeurism, fellated his 9-year-old nephew (when offender was 12), cross-dressed, and masturbated at least 5–7 times per day. The participant's intellectual functioning was within the mild range (Full Scale IQ 67) as assessed by the Wechsler Adult Intelligence Scale—Third Edition (Wechsler, 1997). He scored 100 % in the categories of intimacy, dating, intercourse, and community/risks and hazards of the Socio-Sexual Knowledge and Attitudes Test (Wish, McCombs, & Edmonson, 1980), suggesting that offending was not a function of his lack of knowledge of appropriate sociosexual behaviors.

Procedures were approved prior to the offender's participation in this project by his treatment team, the facility's human rights committee, the administrator of the facility, his guardian, the University of Kansas Advisory Committee on Human Experimentation, and the participant. He was advised that he could withdraw his consent at any point during the study without it affecting his treatment program.

Design

The study followed a single-blind, reversal design across the conditions of baseline, Prozac-placebo, Prozac, Provera-placebo, Provera, and self-control. The participant, his psychologist, medication nurse, and direct-support staff were blind to these experimental conditions. Due to the medications ineffec-

tiveness, a reversal condition could not be implemented. In addition, the participant withdrew from the study at the completion of the Provera condition and left the residential program shortly thereafter. Therefore, the self-control procedure could not be implemented. This phase would have consisted of the portable-penile plethysmograph generating an auditory signal when sexual arousal reached 25 % of full erection. When reaching this threshold, the participant would have implemented a covert sensitization stimulus consisting of an empirically demonstrated, imagined negative consequence (Fernandez, Shingler, & Marshall, 2006). Our plan was to use a multiple-baseline design (see Barlow & Hersen, 1984) to assess the effects of this control procedure across the settings of: (a) laboratory, (b) community analogues, and (c) during the all-day measures.

Settings, Apparatuses, and Stimuli

Laboratory

The apparatus used for laboratory assessment of arousal was a Park's indium-gallium circumference strain gauge that transduced penile responses that were recorded by a Zenith 386 SX20 computer. The participant wore a set of headphones while he listened to laboratory assessment stimuli. Laboratory assessment stimuli consisted of 11, approximately 3-min audiotaped descriptions of: (a) a man exposing himself to a boy and a girl in three settings each, (b) consensual sexual interaction with a woman, (c) rape of a woman, (d) a compliant sexual interaction with a girl, (e) girl rape, and (f) a compliant sexual interaction with a boy (Association of the Treatment of Sexual Abusers, 1993).

Community Analogues

These assessment sessions were conducted in three settings that had all been used in previous research (Rea, DeBriere, Butler, & Saunders, 1998; Rea et al., 2003): (a) playgrounds, (b) a college campus, and (c) a vacant parking lot. The playgrounds/college-campus settings allowed for an assessment of the participant's arousal to women and girls (highest responses during laboratory assessments) in the natural environment, while the vacant parking lot setting provided a control condition in which only the data collectors were present. To eliminate the participant's future access to children and adults observed during the assessments, sessions were conducted in a town where he did not live and would not reside in the near future. In addition, these public areas kept the identities of the individuals and their residential addresses unknown.

In order to verify that treatment was effective, and not a function of the absence of the stimuli that controlled arousal during an assessment, data were collected on the presence of women, girls, boys, and men using a handheld computer (Psion Organiser II) synchronized with the portable-penile plethysmograph. For instance, if arousal was decreased in the laboratory and in the

community to adult-female stimuli, one could verify that female stimuli were present in the community, thus the reductions were not a function of the absence of such stimuli. Penile responses were measured by a strain gauge, and data were recorded by a portable-penile plethysmograph (Rea et al., 1998). This device consisted of a microprocessor that had a 32-K memory and an onboard analog-to-digital converter, a battery pack, and an instrument grade amplifier.

All-Day Sexual Arousal

The offender wore the portable-penile plethysmograph upwards of 13-h-a-day while at work, in the community during leisure activities, and on his residential unit. He wore a wrist watch and carried a data sheet in a folder (that remained confidentially concealed) in which he recorded the occurrence of sexual thoughts, with and without sexual arousal, every 60 min (on the hour).

Relapse-Prevention Probes

The offender was taken by a male-staff member unfamiliar to the participant to his place of employment at a car detail service in the community that was owned and operated by the treatment facility's endowment association. One of three unfamiliar-female-staff members from the facility, who were not part of the participant's treatment team, was also present during each probe session. However, female-staff member number 2 was unable to continue participation and was replaced by a fourth unfamiliar female.

Assessment Procedures

Laboratory Assessments

Assessment sessions of sexual arousal were originally scheduled once a week in the laboratory and once a month in the community during baseline. However, due to numerous scheduling issues that arose during the study (e.g., weather, illnesses, staffing schedules), this schedule was adjusted so that a sufficient number of sessions were implemented in each assessment environment in order to evaluate the effects of the independent variables of Prozac and Provera on sexual arousal.

In the laboratory, each stimulus tape was presented during a session and randomized across sessions. There was a minimum of 30 s and a return to baseline of 30–35 % full erection between the presentations of each tape. The assessment of all-day sexual arousal and deviant and nondeviant sexual thoughts occurred 6 months prior to the assessment of sexual stimuli in the laboratory and the community analogue assessments. During these assessments the participant often was aroused 30–35 % of full erection or greater for upwards of 75 min. In addition, despite maintaining 30–35 % of full erection for prolonged periods of time, responding increased numerous times during assessment

sessions, and reached a peak of 82 % of full erection during an assessment session (full erection was 4.4 cm). Thus, due to the participant's difficulty in detumescing below 30 % full erection, 30–35 % of full erection was determined to be his baseline level of full erection. During preassessment sessions, he reported that he discriminated becoming aroused at 25 % full erection (which was verified by a linear increase in plethysmograph tracings).

Community-Sexual Arousal Assessments

These assessment sessions consisted of 5, 5-min trials (one for each assessment category). The assessment categories were: (a) fantasy of consensual sexual interaction with a woman, (b) fantasy of exposing to a woman, (c) fantasy of a compliant sexual interaction with a girl, (d) fantasy of exposing to a girl, and (e) a vacant parking lot. These categories were implemented based on the laboratory stimulus categories that produced responding (with the exception of the isolation/neutral condition that was used as a control condition) in order to evaluate whether the medications reduced arousal in the laboratory and community. The order in which the assessment categories were presented was randomized across sessions. Each trial was conducted in a different location.

As mentioned above, a trial began when the participant's sexual arousal was at or below 30–35 % of full erection. If this criterion was reached, research staff drove the participant to a college, playground, or vacant parking lot according to the prearranged schedule. If this criterion was not reached within 30 min, the assessment session was terminated for the day.

The participant remained in the vehicle wearing the portable-penile plethysmograph during all community sessions. Research staff left the vehicle, but maintained visual contact of the participant during the assessment. At the playground, the participant was told to look at a girl and to either imagine exposing himself to her or of having a compliant sexual interaction with her, depending upon the assessment session. At the college, he was instructed to look at women and to either imagine exposing himself to a woman or of having a consensual sexual interaction with her, depending upon the assessment session. In order to determine fidelity of each respective analogue, the participant audiotaped his sexual fantasies to verify he was following the research staff instructions. After the stimulus category, the participant was taken to the next setting. Due to the scarcity of parks and playgrounds in the community where the women analogue was conducted (thus few children out in public areas), the girl analogues were conducted on a separate day in a different town where there were more parks and playgrounds. If people were not present in a particular setting after 15 min, that assessment category was terminated for the day. During the vacant parking lot sessions, the participant was instructed to allow any arousal he experienced to occur.

During assessments, research staff recorded the presence of women, girls, boys, and men on a second-by-second basis using handheld computers. This provided verification that the setting contained women and girls that were being assessed (e.g., in the "woman setting" it only contained women and did not contain girls). The handheld computers and the portable-penile plethysmograph were synchronized so that a second-by-second analysis could be performed on both the presence of people and levels of the offender's sexual arousal. Reliability checks on the presence of women and girls were conducted during 27 % of all sessions. An agreement was scored if both observers recorded the presence of a particular category for that second during the session. Overall reliability was 94 % (range, 60–100 %), 93 % for women (range, 60–100), and 94 % for girls (range, 81–98 %).

All-Day Sexual Arousal and Deviant and Nondeviant Sexual Thoughts Assessments

Bancroft and Wu (1983) reported using weekly self-ratings of the frequency of sexual thoughts (on a scale of 0–4) and the extent to which these thoughts were accompanied by sexual excitement (on a 0–2 scale). In order to systematically replicate this data system while providing more frequent and objective measurement of sexual excitement with deviant and nondeviant thoughts, the participant wore the portable-penile plethysmograph and recorded sexual thoughts with and without arousal while at his job, on the living unit (which was comprised of all males), and in the community. In addition, he recorded his deviant thoughts (e.g., urges and/or fantasies to expose) with arousal (25 % full erection) and without arousal every 60 min and non-deviant sexual thoughts (e.g., urges and fantasies of consensual sexual interactions with a woman). In order to ensure that the participant was recording data on the hour, a research staff member intermittently checked his data sheet to verify that data were recorded at these intervals.

Relapse-Prevention Generalization Assessments in the Community

O'Donohue and Letourneau (1992) recommended the assessment of other relevant sexual behaviors, such as verbalizations, approach behavior, touching, and cognitions in the natural environment to evaluate treatment. The authors indicated that the assessment of such behaviors has been ignored for understandable reasons. To address this deficit in a safe manner, Rea, Dixon, and Zettle (2014) designed procedures to assess the generalization of relapse-prevention behaviors of sexual offenders. Specifically, the degree to which key actions such as avoiding approaching, talking to, and touching children generalized from treatment staff to nontreatment individuals in the community was evaluated. We accordingly followed similar procedures in this study.

The participant's relapse-prevention plan included assessment of whether he: (a) stayed with male staff while on community outings, (b) stayed as far away as possible from women when male staff were not present, and (c) refrained from staring at, talking to, or having physical contact with women within close range in the absence of male staff. Assessment of these behaviors occurred at a car detail shop where the participant worked during the week. Data collection sessions occurred over a 24-month period, lasted 5–10 min per session, and occurred during various times and days of the week. An unfamiliar-male-staff member drove him to the car detail shop and then went to another section of the building, leaving him alone with an unfamiliar-female-staff member who was also unfamiliar to the participant. During this time, she provided no feedback to the participant regarding his performance of the relapse-prevention behaviors. The three unfamiliar-female-staff members were rotated for each assessment session.

Prior to the study, if the participant failed to comply with any of the identified behaviors with familiar treatment staff, the activity was terminated and he was returned to the agency. Additional consequences included sanctions of future trips, restrictions on the living unit, and increased supervision on the living unit for a period of time until compliance with rules was demonstrated. Each unfamiliar companion was trained on the participant's relapse-prevention plan and how to score his compliance. Post-trip interviews were conducted that verified the unfamiliar companions did so with 100 % accuracy. Both the male and female unfamiliar staff were instructed not to provide compliance feedback to the participant. However, the female companions were instructed to redirect the participant if touched by him.

Independent Variables

Each independent variable was implemented until data across the dependent measures became stabilized according to somewhat different decision rules. For measures of deviant arousal and thoughts targeted for reductions in the laboratory, community analogues, and all-day condition, the next phase was not implemented until they were within the range of responding at the end of an independent-variable phase (and not on a downward trend). In addition, the relapse-prevention behaviors targeted for acceleration had to also be within the range of responding (and not on an increasing trend). Both data decision rules had to be met before the next independent-variable phase was introduced.

Prozac-Placebo Condition

This condition lasted 16 weeks and consisted of the administration of a sugar capsule that resembled the Prozac capsule.

Prozac Condition

Kafka and Prentky (1992) reported that 20–60 mg/day was effective in reducing deviant sexual behavior. Prozac was first administered at a dosage of 20 mg for 2 weeks. When there was not a desired effect on sexual arousal and related thoughts, as well as relapse-prevention behaviors, the medication was titrated until reaching a maximum dosage of 80 mg. The dosage was gradually titrated by 20 mg (i.e., 40 mg for 2 weeks, 60 mg for 1 month, and 80 mg for 1 month).

Provera-Placebo Condition

This condition lasted 15 weeks and consisted of 150 cc of saline administered once every 2 weeks.

Provera Condition

Dosage levels of Provera from 60 mg/day for 10–18 months (Gottesman & Schubert, 1993) to 100–400 mg/week intramuscularly for 16 weeks (Kiersch, 1990) have been effective in reducing deviant sexual behavior. Provera was administered at a dosage of 150 mg intramuscularly every 2 weeks for 30 weeks.

Dependent Measures

Percentage of Full Erection

This measure was derived by taking the maximum penile response during a given assessment period, divided by the participant's full erection. Full erection was defined as the largest response obtained during the entire study (including a masturbation session conducted prior to the first assessment session) minus the flaccid measure that was the lowest level of plethysmograph output obtained during the study. His range of full erection was 4.4 cm.

Percentage of Correct Identification of All-Day Sexual Arousal

A laboratory technician analyzed each 60-min interval of portable-penile plethysmograph output. The participant's data sheet was then compared to determine whether there was an accurate recording of the occurrence (25 % full erection or greater) or non-occurrence of arousal for that 60-min interval. The percentage of correct recordings of arousal was derived by the number of intervals in which arousal was recorded correctly by the participant and then divided by the total number of 60-min intervals during a session.

Percentage of Session Intervals Sexual Arousal Was Equal to or Greater Than 25 % Full Erection During All-Day Measures

The same portable-penile plethysmograph data as above were used to determine the percentage of 60-min intervals in which arousal occurred. The number of 60-min intervals in which arousal was at or above 25 % full erection was divided by the number of 60-min intervals during a data collection session.

Deviant Sexual Thoughts With and Without Arousal During All-Day Measures

The offender recorded his deviant thoughts (e.g., rape, sexual interactions with someone under 18 years of age, window peeking, exposure), with and without arousal, every 60 min. Scoring deviant thoughts with and without arousal by the participant were not mutually exclusive categories during each 60-min interval. The percentage of time deviant thoughts with arousal occurred was obtained by dividing the number of intervals in which deviant thoughts with arousal occurred (equal to or greater than 25 % full erection), by the number of 60-min intervals during a data collection session. Similar calculations were used to determine the percentage of time deviant thoughts without arousal (less than 25 % full erection) occurred.

Nondeviant Sexual Thoughts With and Without Arousal During All-Day Measures

The participant also recorded his nondeviant thoughts (e.g., consensual sexual interaction with someone at least 18 years of age) with and without arousal every 60 min. Scoring of each respective category was determined in the same manner as for deviant thoughts by dividing the number of intervals in which they were reported by the total number of 60-min intervals during a data collection session. Nondeviant thoughts with and without arousal were not mutually exclusive with respect to each other or to deviant thoughts. Thus, it was possible for the participant to report both deviant and nondeviant thoughts during both the presence and absence of arousal within a given 60-min interval.

No Sexual Thoughts/No Arousal During the All-Day Measures

Midway through the Prozac-placebo phase (session 53), we decided it would be useful to add a mutually exclusive code as a means of assessing the stimulus control properties of the participant's scoring thoughts with and without arousal. Doing so provided another verification of intervals during which no arousal occurred as well as indirect support of the accuracy of scoring sexual thoughts. The participant began to record no sexual thoughts/no arousal during each 60-min interval during session 53. This

scoring category was a mutually exclusive category from the categories of nondeviant and deviant thoughts with and without arousal.

Relapse-Prevention Behaviors

The occurrence or nonoccurrence of each of the five relapse-prevention behaviors served as another dependent measure. These behaviors were: (a) did the offender stay with staff; and if potential victims were within close range did he, (b) not stare, (c) not talk to them, (d) not touch them, and (e) stay as far away as possible from them?

Results

Percentage Correct Identification of Sexual Arousal

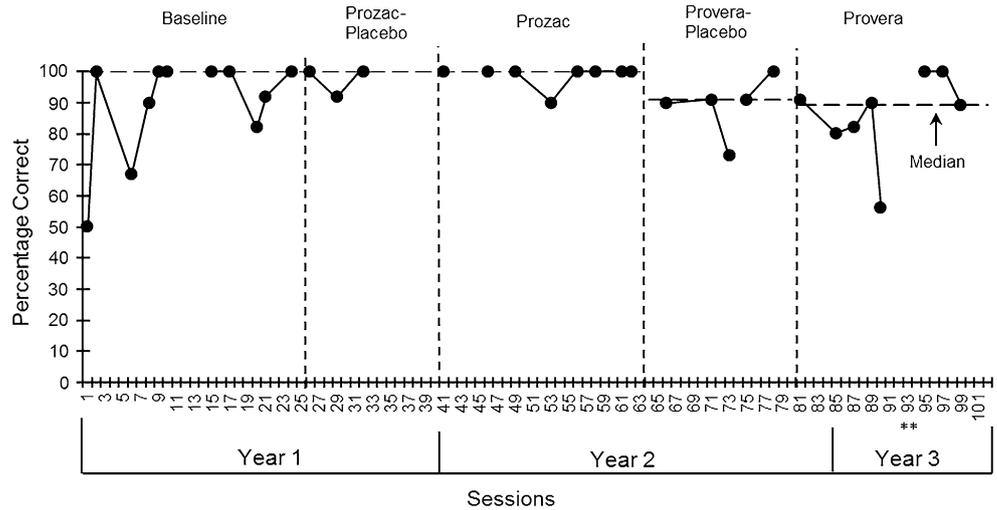
The median percentage correct identification of sexual arousal by the offender across the conditions of baseline, Prozac-placebo, Prozac, Provera-placebo, and Provera was 100 % (range, 50–100 %), 100 % (range, 92–100 %), 100 % (range, 90–100 %), 91 % (range, 73–100 %), and 90 % (range, 56–100 %), respectively (see Fig. 1). During the Provera condition, the accuracy of scoring decreased to 56 and 33 % correct on sessions 90 and 93, respectively. During session 93, the gauge fell off after 165 min. The participant continued to score the occurrence and nonoccurrence of sexual arousal for 360 min. He received feedback after this session regarding his inaccurate scoring of sexual arousal that resulted in an increased accuracy of 100, 100, and 89 % during the last 3 sessions of the condition, respectively. However, when all data were checked for accuracy in preparation of this article, the session 93 data sheet was missing. Although we were confident that his accuracy was only 33 % at the time of the study, we removed these data from Figs. 1, 3, and 4 due to our inability to verify this session's data.

Percentage of Session Time During which Sexual Arousal Occurred (25 % or Greater)

In general, arousal increased from baseline, but was variable across the conditions (see Fig. 2). The baseline median percentage of session time that the participant was aroused was 44 % (range, 22–95 %). During the Prozac-placebo condition, the median percentage of session time that arousal occurred increased to 70 % (range, 19–99 %). The median percentage time that he was sexually aroused was 61 % during the Prozac (range, 24–75 %) and Provera-placebo (range, 38–95 %) conditions.

In the Provera condition, despite the strain gauge being off during the majority of the session 93, we were still able to calculate data based upon the 165 min the participant had the gauge on. The median percentage of time that arousal occurred reached its peak at 85 % (range, 53–100 %) despite testosterone

Fig. 1 Percentage correct identification of sexual arousal by participant. **Given feedback about incorrect reporting of arousal



dropping to a low of 36 ng/dL from a high of 509 ng/dL (93 % reduction) during the Provera condition. In 25 of the 36 sessions (69 %), the participant was aroused 50 % (or greater) of the time.

Percentage of Session Intervals During Which Deviant Thoughts With/Without Arousal Occurred

As shown in Fig. 3, the median percentage of session intervals that contained deviant thoughts with arousal: (a) during baseline was high (100 %) with some variability (range, 0–100 %); (b) during Prozac-placebo was high (100 %) with no variability, and (c) during the Prozac condition was high (100 %) with little variability (range, 85–100 %). During the Provera-placebo condition, median responding remained high (82 %) with moderate variability (range, 58–100 %). The Provera condition produced the lowest median responding (71 %) and the greatest variability (range, 11–100 %).

The median percentage of intervals (along with the range of responding) for deviant thoughts without arousal was nearly identical to deviant thoughts with arousal during the first three conditions. Median responding was 100 % during baseline with some variability (range, 0–100 %), high (100 %) with no variability during Prozac-placebo, and also high (100 %) with little variability (range, 82–100 %) during the Prozac condition. Median responding during the Provera-placebo was 73 %, while variability was moderate (range, 55–91 %). As was the case for deviant thoughts with arousal, the Provera condition produced the lowest median responding (59 %) and the greatest across-session variability (range, 33–100 %).

The variability in recording deviant thoughts with and without arousal during the Provera condition may at least, in part, be attributed to the decrease in the percentage of correct identification of arousal during sessions 90 and 93. Data for session 93 were removed from the analysis due to the missing data sheet,

Fig. 2 Percentage of session intervals during which sexual arousal occurred (25 % or greater). **Given feedback about incorrect reporting of arousal

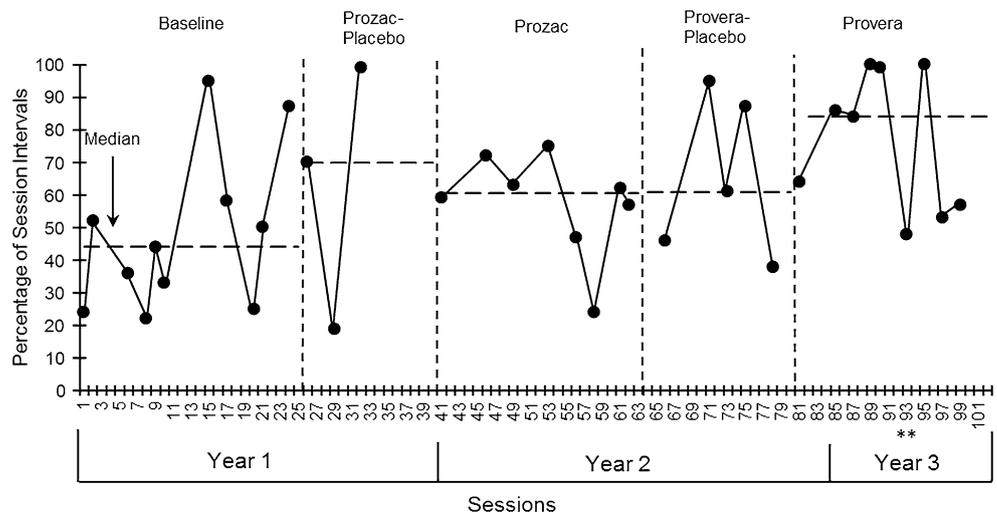
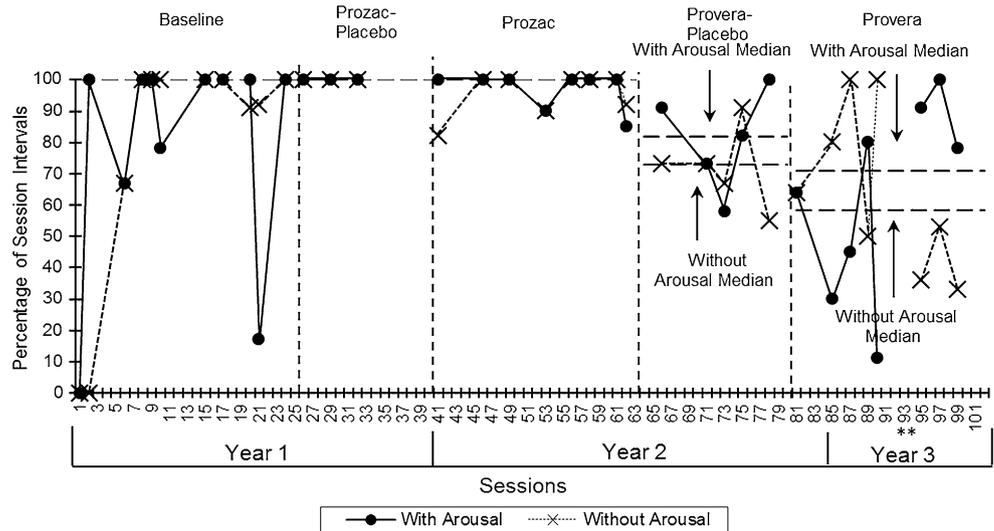


Fig. 4 Percentage of session intervals during which nondeviant thoughts with/without arousal occurred. **Given feedback about incorrect reporting of arousal



(b) Provera-placebo, 73 % (range, 67–91 %); and (c) Provera, 30 % (range, 0–82 %). A 43 % decrease in accuracy and an increase in variability occurred throughout the Provera condition.

Laboratory Assessment of Sexual Arousal

Eleven audiotaped assessment stimuli were used during laboratory assessments of arousal. Due to space limitations, the results have been categorized into two general patterns of responding across the five conditions.

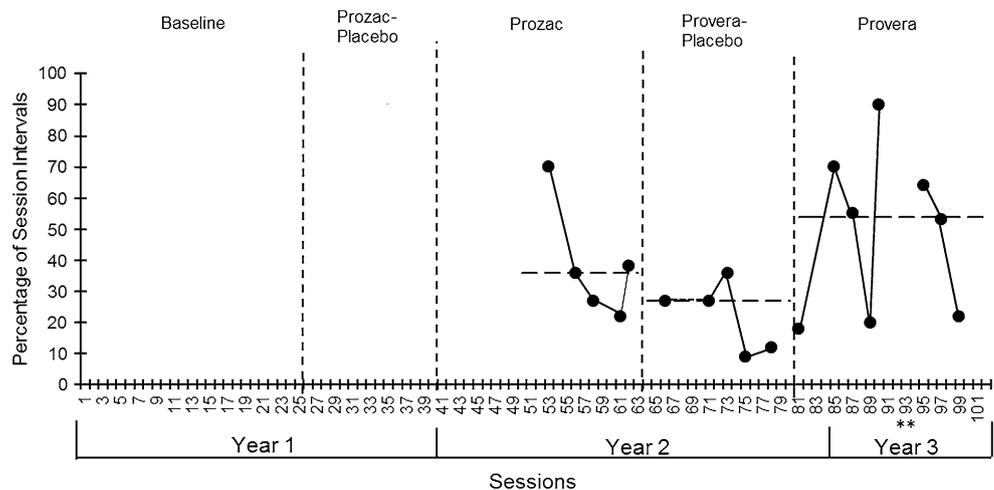
The first pattern was lower median responding (below 40 %) across the five conditions than for the other eight audiotapes (see Table 1, Pattern 1) with variability as much as 50 % full erection for three tapes (girl rape, boy compliant, woman rape). The second pattern occurred for the remaining eight tapes. Median responding was in general above 50 % across the five conditions (with the exception of girl compliant and woman consent during baseline) with variability for: (a) girl complaint; (b) woman consent; (c) exposing to women at a laundromat tennis court, or

while in a car talking to a woman; and (d) exposing to a girl at a department store, while in a car, and in a park. Medications had no effect on sexual arousal.

Community-Sexual Arousal Analogue Assessments

Four community assessment analogues were designed to assess the concurrence of the four laboratory stimulus categories that produced high arousal and the effects of medication(s) on sexual arousal in the natural environment. The analogues were: (a) woman consent, (b) exposure to woman, (c) girl compliant, and (d) girl exposure. Although the girl compliant and woman consent-laboratory tapes didn't produce the overall highest median responding during baseline, they did demonstrate an increasing trend during the last two sessions. Thus, they were added as community-sexual arousal assessment analogues. A fifth analogue, isolation scenario (i.e., parking lot) was added as a control condition.

Fig. 5 Percentage of session intervals in which no sexual thoughts/no sexual arousal occurred. **Given feedback about incorrect reporting of arousal



Due to high, near equivalent responding to the three laboratory audiotaped depictions of a man exposing to a woman and to the three laboratory audiotaped depictions of a man exposing to a girl, arousal during the laboratory audiotaped descriptions of a man exposing to a girl while in a car and to a woman while in a car at the park, is graphically displayed within its corresponding community analogues of exposing to a woman and a girl. The overall medians for arousal during these laboratory tapes are also displayed within the same graphs to evaluate the concurrence with the community analogue protocols and if the medication(s) had a differential effect within the assessment procedures (see Figs. 6, 7, 8, 9).

Sexual arousal to fantasies of exposing to a girl and girl compliant was lower when assessed in the community than in the laboratory. Median arousal to the laboratory tapes of a man exposing to a girl was 35, 43, 35, 15, and 14 % higher in the laboratory than in the community during baseline, Prozac-placebo, Prozac, Provera-placebo, and Provera conditions, respectively (see Fig. 6). With the exception during baseline, in which arousal was nearly identical, sexual arousal to the fantasy of girl compliant was 23–58 % higher when assessed in the laboratory than in the community (see Fig. 7). Prozac and Provera did not have an effect on sexual arousal in the community.

Sexual arousal was high and relatively stable during the community analogues of a man exposing himself to a woman while in a car and during the woman consent scenarios and, apart from baseline, was comparable to levels assessed in the laboratory (see Figs. 8, 9, respectively). As was the case in the laboratory, Prozac and Provera had no effect on sexual arousal in the community. Finally, sexual arousal during the isolation condition of

the parking lot was similar to the exposing to a girl and girl compliant community analogues across all five conditions (see Fig. 10).

Relapse-Prevention Probes with Unfamiliar Staff

Prebaseline compliance of the participant with his relapse-prevention plan in the presence of eight familiar treatment staff was near 100 %. During the baseline probes to the three unfamiliar staff members, the participant avoided staring at and initiating physical contact with potential victims. However, with these unfamiliar staff, he did not demonstrate the three relapse-prevention behaviors of: (a) staying with staff, as well as (b) avoiding talking to, and (c) staying as far away as possible from potential victims during all conditions (see Table 2).

With unfamiliar-staff member one, during the Prozac-placebo, the first Prozac-probe, Provera-placebo, and Provera conditions, the participant behaviors of staying with staff, and avoiding talking to and staying as far away as possible from potential victims failed to generalize (see Table 2, unfamiliar-staff member one). However, the participant did avoid talking to her during the second probe during the Prozac condition.

A similar lack of generalization also occurred in the presence of unfamiliar-staff member two, who only conducted probes during the baseline, Prozac-placebo, and Prozac conditions before leaving the agency (see Table 2, unfamiliar-staff member two). With unfamiliar-staff member 3, only the baseline, Prozac-placebo, Prozac, and Provera-placebo conditions were conducted due to the participant's withdrawal from the study prior to a probe being conducted in the Provera condition.

Table 1 Median responding and ranges for the 11 “Association for the Treatment for the Sexual Abusers” tapes

	Baseline	Prozac-placebo	Prozac	Provera-placebo	Provera
<i>Pattern 1 Lower median/variable responding</i>					
Girl-rape tape	21 % (15–51 %)	28 % (20–58 %)	37 % (30–48 %)	36 % (34–40 %)	29 % (6–45 %)
Boy compliant	25 % (8–40 %)	33 % (32–39 %)	38 % (29–40 %)	32 % (22–43 %)	36 % (20–75 %)
Woman rape	23 % (15–32 %)	33 % (24–34 %)	33 % (29–41 %)	35 % (28–38 %)	27 % (4–45 %)
<i>Pattern 2 Higher median/variable responding</i>					
Girl compliant	20 % (13–53 %)	66 % (60–67 %)	68 % (45–76 %)	69 % (63–72 %)	51 % (17–65 %)
Woman consent	17 % (12–50 %)	68 % (65–68 %)	71 % (64–80 %)	65 % (63–65 %)	64 % (26–86 %)
Exposing to women:					
At laundromat	55 % (50–60 %)	65 % (63–77 %)	68 % (65–80 %)	67 % (62–70 %)	69 % (62–78 %)
At tennis court	55 % (45–58 %)	68 % (65–76 %)	70 % (63–81 %)	64 % (63–72 %)	67 % (60–80 %)
In car talk to/follow	50 % (43–51 %)	65 % (64–76 %)	72 % (65–80 %)	63 % (33–73 %)	67 % (35–82 %)
Exposing to girl:					
At department store	50 % (48–55 %)	64 % (64–77 %)	74 % (39–95 %)	48 % (43–64 %)	44 % (12–80 %)
In car	60 % (55–75 %)	65 % (40–72 %)	67 % (61–75 %)	63 % (35–70 %)	44 % (11–73 %)
In park	55 % (45–55 %)	48 % (45–74 %)	66 % (59–73 %)	63 % (55–65 %)	43 % (16–75 %)

Median (range)

Fig. 6 Arousal during fantasy of exposing to a girl while in a car: Laboratory tape and community analogue

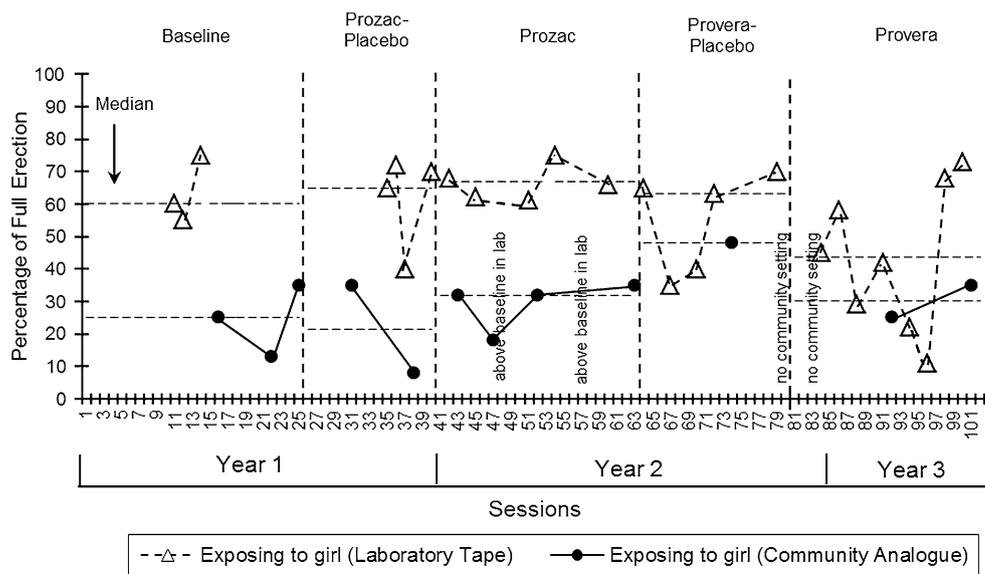
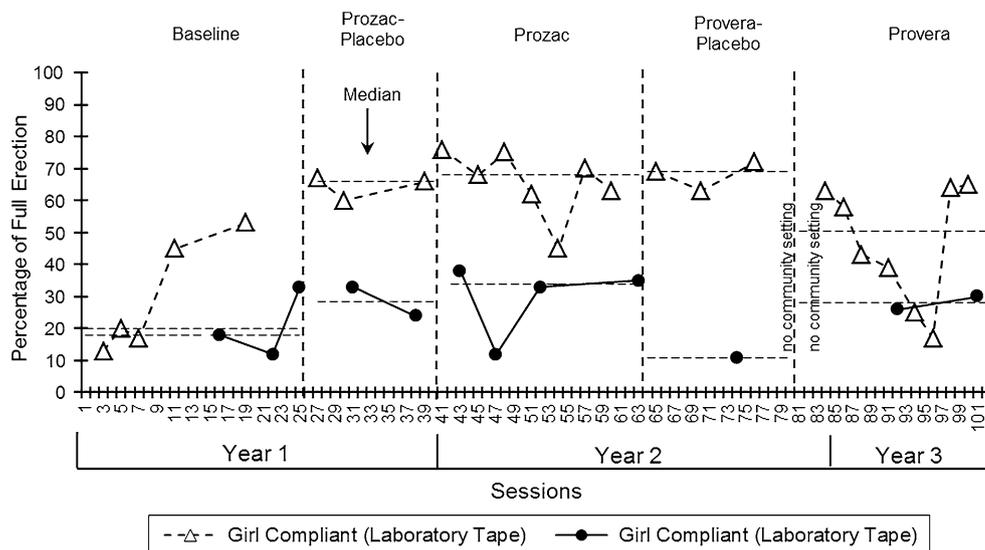


Fig. 7 Arousal during fantasy of girl compliant: Laboratory and community analogue



The behaviors of staying with staff, and avoiding talking to and staying as far away as possible from her failed to generalize across the three conditions (see Table 2, unfamiliar-staff member three).

A fourth female-staff member was recruited during the Provera-placebo condition to replace unfamiliar-staff member two. As with the other three unfamiliar staff members, the participant did not demonstrate staying with staff, avoiding talking to unfamiliar staff, and staying as far away as possible from her. However, during the first probe in the Provera condition, he stayed with the staff member and as far away as possible from the unfamiliar-female-staff member (see Table 2, unfamiliar-staff member four), but did not avoid talking to her. During the second probe in the Provera condition, the participant did not stay with staff and as far away as possible, and he also stared at her, which was not exhibited during the Provera-placebo condition or the

first Provera probe. One possible explanation of the lack of generalization is that responding was not maintained with familiar staff during the course of the study. However, responding to familiar treatment staff was near 100 % throughout the course of the study.

Recidivism

The participant ultimately left the residential-treatment program against the advice of his treatment team. At the conclusion of the study, he reported that he was still masturbating 5 or more times per day. Four years post-discharge he was arrested for felony lewd and lascivious behavior that was reduced to a misdemeanor in a plea agreement. However, he reported exposing himself within days after his discharge from the program.

Fig. 8 Arousal during fantasy of exposing to a woman in car: Laboratory and community analogue

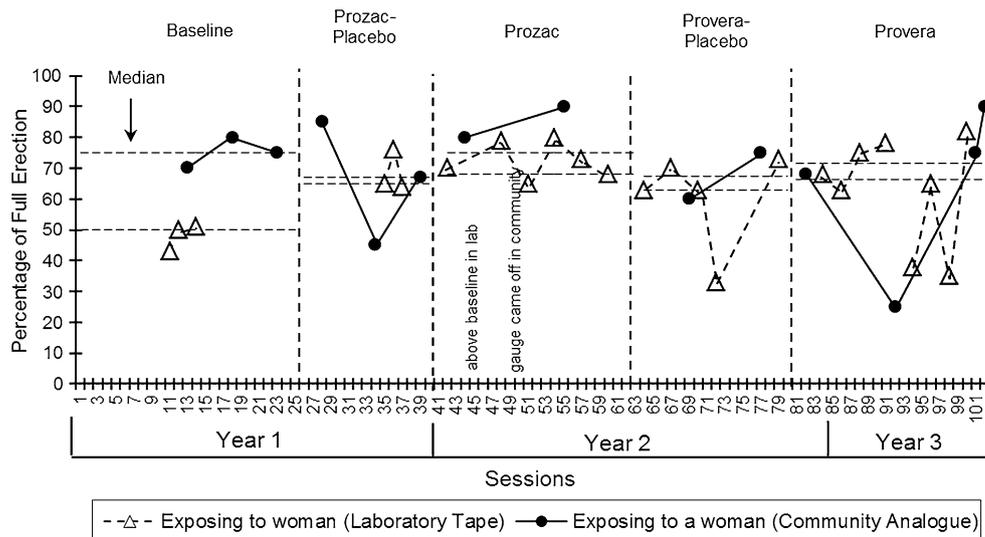


Fig. 9 Arousal during fantasy of woman consent: Laboratory and community analogue

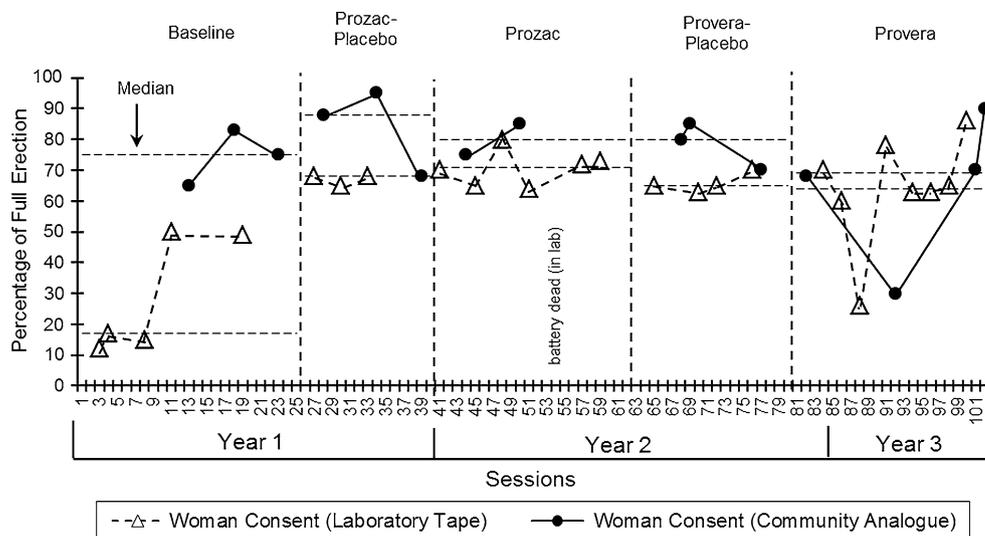


Fig. 10 Isolation (community analogue)

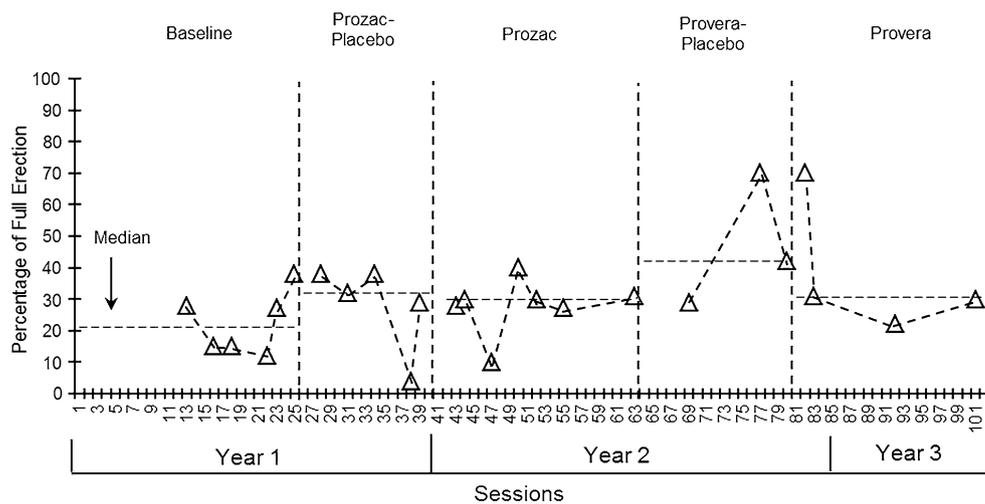


Table 2 Correct/incorrect responses across the five relapse-prevention behaviors by condition for unfamiliar staff 1, 2, 3, and 4

	Unfamiliar staff #1					Unfamiliar staff #2		
	Baseline	Prozac-placebo	Prozac	Provera-placebo	Provera	Baseline	Prozac-placebo	Prozac
Did participant stay with staff?	○	○	○	○	○	○	○	○
If potential victims in close range								
Did participant avoid staring at them	●	●	●	●	●	●	●	●
Did participant avoid talking to them	○	○	○	●	○	○	○	○
Did participant avoid physical contact	●	●	●	●	●	●	●	●
Did participant stay as far away as possible	○	○	○	○	○	○	○	○
	Unfamiliar staff #3				Unfamiliar staff #4			
	Baseline	Prozac-placebo	Prozac	Provera-placebo	Provera-placebo	Provera	Provera	
Did participant stay with staff?	○	○	○	○	○	●	○	
If potential victims in close range								
Did participant avoid staring at them	●	●	●	●	●	●	○	
Did participant avoid talking to them	○	○	○	○	○	○	○	
Did participant avoid physical contact	●	●	●	●	●	●	●	
Did participant stay as far away as possible	○	○	○	○	○	●	○	

● = Correct response

○ = Incorrect response

Discussion

This study was designed to assess whether medication(s) would reduce sexual arousal in the laboratory and natural environment, impact deviant thoughts and arousal, and facilitate relapse-prevention behaviors in the natural environment. Although Prozac and Provera were ineffective in reducing sexual fantasies and arousal in the laboratory and natural environment, and in facilitating relapse-prevention behaviors, the findings of this case report may be useful in contributing to the development of a strategic approach to long-term assessment of deviant sexual behaviors in the natural environment. In particular, we believe at least four valuable, albeit tentative, conclusions may be derived from the results of this study and related research.

First, sexual arousal assessment demonstrated long-term, differentiated significant clinical responding between stimulus categories, despite response variability within categories in the laboratory and the natural environment. In general, these data replicated earlier findings of the congruence between the stimulus elements that controlled sexual responding in the laboratory and natural environment and offending history (Rea et al., 1998, 2003). In addition, deviant sexual arousal in the laboratory and natural environment was contiguous to his self-report of reoffending shortly after his discharge from the residential-treatment program, which is similar to previous findings that deviant arousal was predictive of reoffending (e.g., Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2004).

Reyes et al. (2006) also demonstrated long-term, discriminated responding between categories with variability within categories across 45–90 successive sessions under laboratory conditions with offenders with intellectual disabilities. Despite the current findings and that of Reyes et al. repeated presentation of the same stimuli during laboratory assessments may produce within session (O'Donohue & Geer, 1985) and across sessions (Schaefer & Colgan, 1977) habituation that will increase the difficulty of discerning reductions of responding due to habituation and/or treatment. Given the possible treatment implications, it is imperative for clinicians to identify and control the variables responsible for such variability in sexual arousal. Deviant arousal in conjunction with other measures can: (a) identify which offenders and behaviors are in need of treatment, (b) determine whether treatment interventions have reduced deviant arousal, and (c) estimate the likelihood that an individual might reoffend (Association of the Treatment of Sexual Abusers, 2014; Merdian & Jones, 2011; Thibaut et al., 2010).

A second implication of this study's findings is that they further underscore the importance of assessing deviant sexual arousal in the natural environment rather than relying upon self-report measures in evaluating treatment effectiveness. For example, Rea et al. (2003) reported a reduction of sexual arousal to children in the laboratory and self-reported elimination of arousal to children for a sexual offender with an intellectual disability. However, sexual arousal in the natural environment remained unchanged and high (nearly 70 % full erection). Similarly, the offender in this study self-reported near elimination of deviant thoughts

with arousal during the Provera condition (sessions 90 and 93). However, his percentage correct identification of sexual arousal was low. When feedback was provided regarding his inaccurate self-report of sexual arousal, the percentage correct identification of sexual arousal increased to nearly 100 %, and his report of deviant thoughts with arousal increased to nearly 80 % or greater for the remaining three sessions, representing a recovery to previous high levels of reporting.

A third implication of our findings involves the importance of evaluating both relative and absolute levels of deviant arousal in natural settings. As previously mentioned, due to the participant's inability to detumescence below 30–35 % of full erection for prolonged periods of time (and that he discriminated arousal at 25 % full erection) during the all-day sexual arousal and sexual thoughts assessments, his baseline measure of arousal was considered 30–35 %. Sexual arousal in the community to girl scenarios was much lower than similar laboratory stimuli (i.e., girl compliant and exposure tapes). In fact, with the exception of one assessment session, responding above 30–35 % did not occur. Conversely, responding increased from 30 to 35 % during the woman assessment sessions (laboratory and community assessment sessions) and reached peaks of nearly 90 % of full erection. These data are somewhat congruent with his sexual-offending history in that he had not offended against children (verified by polygraph examination). However, with the exception of the girl compliant community analogue in baseline and Provera-placebo conditions, the overall median and individual responses would be considered clinically significant and should be addressed, 20 % full erection (Howes, 1995) or 9.4–10.6 mm (.94–1.06 cm) in the absence of knowing full erection (Howes, 2003). His range of full erection was 44 mm (4.4 cm), and 30–35 % of full erection was 1.32–1.54 cm.

The fourth implication is the demonstration of high-sex drive during a majority of the all-day sexual arousal and sexual thoughts assessments in this participant and the other three participants who began but did not complete this study, as well as with other offenders in our research program (Rea et al., 1998, 2003). Offenders in our previous research program were sexually aroused a median of 100 % of the time (unpublished raw data). In addition, the median percentage of time that the four offenders in this study (this number includes the three offenders who didn't complete the study) were engaged in sexual fantasies was also 100 %. These findings provide critical, empirical support that hypersexual behavior is a characteristic of sexual offenders as demonstrated by a propensity for increased sexual fantasizing or "preoccupation" (Kafka, 2008) and serves as a predictor of reoffending (Hanson & Morton-Bourgon, 2004; Klein, Schmidt, Turner, & Briken, 2015).

Hanson, Harris, Scott, and Helmus (2007) in their analysis of risk factors relied upon structured interviews of offenders and a review of their community supervision files. The authors reported that such an assessment methodology was only moderately predictive of reoffending and that the development of

more comprehensive evaluations and specialized testing (e.g., phallometric assessments) could improve the reliability and validity of risk factors. Using Howes' (1995, 2003) criteria of clinically significant sexual arousal and Kafka's (1997) criteria of 1–2 h/day/week engaged in sexual preoccupation, the all-day sexual arousal and sexual thoughts measures in this study may have greater and/or additive-predictive influence than relying upon self-report and case-file information.

Given the above findings, and that the variables of deviant sexual arousal, sexual preoccupations, and a lack of self-control are predictive of reoffending (Hanson & Morton-Bourgon, 2004), a predisposition toward "crossover offending" is created. Indeed, research has demonstrated that a majority of offenders have committed "crossover crimes" that include victims who were adults and children, male and female, known and unknown victims, as well as multiple relationship crimes that were unknown prior to a polygraph examination (Elements of Change, 2000; Heil, Ahlymeyer, & Simon, 2003). Therefore, assessment of an offender's control of sexual fantasies and arousal in the natural environment and generalization of relapse-prevention skills are critical in reducing such "crossover offending." One clear limitation of this study involved assessment of the participant's thoughts during the presence versus absence of sexual arousal. In particular, the accuracy of the no thoughts/no arousal category was called into question. This code was introduced as a measure of the nonoccurrence of sexual thoughts and arousal that provided an additional verification of the stimulus control of the sexual thoughts with and without arousal codes. In addition, the administrative error in the discontinuation of the collection of arousal/no sexual thoughts hampers the ability to determine the relationship between arousal/no sexual thoughts and the arousal that was associated with sexual thoughts. One possible explanation for the inaccuracy of scoring no thoughts/no arousal code is that the addition of this code may have been too difficult to score with so many other codes. In terms of clinical importance, it may be unnecessary to score this code. However, the larger question would seem to be whether the scoring of any private behaviors is accurate (e.g., reporting of deviant arousal as nondeviant) and how the validity of self-report could be increased through the use of polygraph assessments. Schober et al. (2005) assessed five child molesters regarding their strong urges to reoffend with the polygraph while deviant sexual arousal was assessed by the plethysmograph every 3 months. The authors reported that Lupron (a GnRH analogue used to reduce testosterone) eliminated strong urges to reoffend and masturbation to thoughts of individuals under 18 years of age for the majority of offenders (as indicated by nondeceptive responses during a polygraph examination) in a majority of the offenders. The plethysmograph was able to detect increases in sexual arousal that correlated with minimal rises in serum testosterone levels. However, when Lupron was discontinued, strong urges to reoffend and masturbation to thoughts of individuals under 18 reoccurred (as indicated by deceptive responses during a polygraph examination). These data suggest

that the validity of the self-report of appropriate and deviant thoughts with arousal could be increased by random polygraph examinations after an assessment session. In addition, these procedures may increase the accuracy of reporting such private behaviors across time.

Schober et al. (2005) indicated that the 3-month assessment schedule used in their study was too burdensome for clinicians. However, Kafka (2008) reported that 66 % of males seen in their clinic for paraphilias (including sexual offending) had hypersexual desire (7 or more orgasms per week) and sexual preoccupations for more than 1–2 h/day/week. In addition, the anonymous, confidential-online report of a community sample of 8718 German men demonstrated that the variables of sexual fantasizing involving children and antisociality were associated with sexual contact with children (Klein et al., 2015). Given the degree of “crossover” offending and the high rate of reoffending, a more frequent, comprehensive assessment schedule than once every 3 months can be recommended. Indeed, Levenson and Prescott (2014) suggest that the field of sexual-offender treatment should shift from searching for treatments that work to empirically demonstrating that a treatment procedure is effective for an individual. Thus, the importance of the evaluation of the acquisition and generalization of cognitive-behavioral skills in the natural environment becomes paramount (Fernandez et al., 2006).

To address the administrative time commitment concern for clinicians, while still providing more frequent, comprehensive assessment and monitoring of sexual-offender treatment based on Risk-Need-Responsivity principles (Hanson, Bourgon, Helmus, & Hodgson, 2009), a portable-penile plethysmograph containing a global positioning satellite (GPS) component that simultaneously tracks location and sexual arousal could be utilized. In addition, the device could provide an auditory-biofeedback signal to the offender for location-exclusion zone violations and significant sexual arousal responses (e.g., 20 % full erection or 10.6 mm in the absence of knowing full erection), while sending these data immediately or delayed to the clinician and community supervisor. These measures would enhance the containment model by assisting the offender in controlling their sexual arousal and adherence to relapse-prevention plans (English, 2004). Pilot work is being conducted at this time with a portable-penile plethysmograph/GPS device possessing these capabilities.

Although our results did not support the use of medication-adjunctive therapy, we would be remiss not to point out methodological and clinical weaknesses of our protocols. First, Prozac has produced mixed results in the treatment literature. Some studies have reported clinically significant results (e.g., Kafka & Prentky, 1992), while others have not (Perilstein, Lipper, & Friedman, 1991; Stein et al., 1992). Other SSRIs may be more effective than Prozac for some individuals. Bradford, Greenberg, Gojer, Martindale, and Goldberg (1995) reported that Zolofit decreased deviant arousal by 53 %, while appropriate sexual arousal was maintained or increased.

For the participant in this study, Provera was also ineffective in reducing deviant sexual arousal in the laboratory, deviant thoughts and sexual arousal in the natural environment, increased the percentage of time he was sexually aroused during the day, and did not increase generalization of relapse-prevention behaviors to unfamiliar staff, despite testosterone levels dropping to that of a prepubescent girl. These findings may be explained in part by research that has demonstrated sexual arousal to erotic visual stimuli is independent of androgens (Bancroft, Tennent, Lucas, & Cass, 1974; Bancroft & Wu, 1983; Carani, Bancroft, Granata, Del Rio, & Marrama, 1992; Kwan, Greenleaf, Mann, Crapo, & Davidson, 1983).

While sexual arousal to erotic visual stimuli is independent of androgens, the dependence of androgens to produce sexual arousal to sexual fantasies has been mixed. For instance, Kwan et al. (1983) reported that half of the hypogonadal participants in their study produced greater arousal and percentage of time above 1.5 cm during assessment of sexual fantasies than all of the nonhypogonadal participants. Others have reported hypogonadal participants had lower arousal to sexual fantasies than nonhypogonadal men (Bancroft & Wu, 1983) and in experimentally induced hypogonadism in nonoffenders (Bagatell, Heiman, Rivier, & Bremner, 1994). Bancroft et al. (1974) reported that an antiandrogen lowered sexual arousal in sexual offenders to three types of stimuli (fantasy, slide, and an erotic film) from baseline levels. However, there was no statistically significant interaction between the antiandrogen and type of stimulus.

These mixed findings suggest that some men despite testosterone reductions below pretreatment levels have enough circulating testosterone for maintenance of normal sexual functioning (Bancroft, 2005). Indeed, as many as 5 % of surgically castrated sexual offenders reoffended after 10 or more years post-surgery (Schumaker & Losel, 2008), and some castrated men reported some preservation of sexual function (Davidson, Kwan, & Greenleaf, 1982). However, these findings do not explain the reduction of arousal to child stimuli in the natural environment while responding to child stimuli in the laboratory remained high to auditory stimuli.

Finally, there are additional issues this case report raises. First, additional replication of these procedures for offenders, with and without intellectual disabilities, is needed to determine the generalizability and limitations of such measures. Second, with the advancement of new technologies that allow the measurement of many biobehavioral measures and location, the field must debate the balance of civil liberties and the use of such technologies in the natural environment. The studies in which we have implemented the described assessment measures have been vetted by numerous approval processes to ensure public safety and human rights safeguards that in turn provide an examination of the procedure’s social validity (Winett, Moore, & Anderson, 1991). Merdian and Jones (2011) acknowledged

the ecological validity of the sexual arousal assessment methodology within the natural environment used in our work, although they state that “unsurprisingly, this has not been widely applied” (p. 154). Thus, future replications of these procedures are necessary to further elucidate the social validity of the assessment measures.

This study and previous work would appear to set the formative framework of a long-term, comprehensive assessment methodology for sexual offenders with and without an intellectual disability. Future research will illuminate whether these measures can assist the treatment provider in evaluating the acquisition and generalization of treatment effectiveness to the natural environment. Further, these measures can contribute to risk assessments, assist offenders in better managing their self-control strategies, and assist therapists and community supervisors in better managing long-term risk in the natural environment.

Acknowledgments The authors would like to thank Dr. Joseph Spradlin, Professor Emeritus of the University of Kansas, and Dr. David Lindeman, Senior Scientist of the University of Kansas Lifespan Institute at Parsons, for their invaluable feedback on previous drafts of this article.

Compliance with Ethical Standards

Conflict of interest The authors declare they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

- Abel, G. G., Becker, J. V., Mittelman, M., Cunningham-Rathner, J., Rouleau, J. L., & Murphy, W. D. (1987). Self-reported sex crimes of nonincarcerated paraphiliacs. *Journal of Interpersonal Violence*, *2*, 3–25.
- Adi, Y., Ashcroft, D., Browne, K., Beech, A., Fry-Smith, A., & Hyde, C. (2002). Clinical effectiveness and cost-consequences of selective serotonin reuptake inhibitors in the treatment of sexual offenders. *Health Technology Assessment*, *6*, 13–35.
- Association for the Treatment of Sexual Abusers. (1993). *Auditory stimuli for penile plethysmography (11 cassette recordings)*. Beaverton: ATSA.
- Association for the Treatment of Sexual Abusers. (2014). *Practice standards and guidelines for members of the association for the treatment of sexual abusers*. Beaverton, OR.
- Bagatell, C. J., Heiman, J. R., Rivier, J. E., & Bremner, W. J. (1994). Effects of endogenous testosterone and estradiol on sexual behavior in normal young men. *Journal of Clinical Endocrinology and Metabolism*, *78*, 711–716.
- Bancroft, J. H. (1976). Evaluation of the effectiveness of drugs on sexual behavior. *British Journal of Clinical Pharmacology*, *3*(Suppl. 1), 83–90.
- Bancroft, J. H. (2005). The endocrinology of sexual arousal. Starling review. *Journal of Endocrinology*, *186*, 411–427.
- Bancroft, J., Tennent, G., Lucas, K., & Cass, J. (1974). The control of deviant sexual behavior by drugs: Behavioural changes following oestrogens and anti-androgens. *British Journal of Psychiatry*, *125*, 310–315.
- Bancroft, J., & Wu, F. C. W. (1983). Changes in erectile responsiveness during androgen replacement therapy. *Archives of Sexual Behavior*, *12*, 59–66.
- Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change*. Needham Heights, MA: Simon & Schuster.
- Bradford, J. M. W., Greenberg, D., Gojer, J., Martindale, J. J., & Goldberg, M. (1995, May). *Sertraline in the treatment of pedophilia: An open labeled study*. Paper presented at the meeting of the American Psychiatric Association, Miami, FL.
- Carani, C., Bancroft, J., Granata, A., Del Rio, G., & Marrama, P. (1992). Testosterone and erectile function, nocturnal penile tumescence and rigidity, and erectile response to visual erotic stimuli in hypogonadal and eugonadal men. *Psychoneuroendocrinology*, *17*, 647–654.
- Clarke, D. (1989). Antilibidinal drugs and mental retardation: A review. *Medical Science and the Law*, *29*, 136–146.
- Cooper, A. J. (1995). Review of the role of two antilibidinal drugs in the treatment of sex offenders with mental retardation. *Mental Retardation*, *33*, 42–48.
- Cooper, A. J., Cernovsky, Z., & Magnus, R. V. (1992). The long-term use of cyproterone acetate in pedophilia: A case study. *Journal of Sex and Marital Therapy*, *18*, 292–302.
- Davidson, J. M., Kwan, M., & Greenleaf, W. J. (1982). Hormonal replacement and sexuality in men. *Clinical Endocrinology Metabolism*, *11*, 599–623.
- Doren, D. M. (1998). Recidivism base rates, predictions of sex offender recidivism, and the “sexual predator” commitment laws. *Behavioral Sciences and the Law*, *16*, 97–114.
- Elements of Change. (2000). Combo of polygraph & treatment reveals many sex offenders offend across multiple relationship, age, gender, & crime type categories. *Office of Research and Statistics, Division of Criminal Justice, Colorado Department of Public Safety*, *5*, 1–8.
- Emory, L. E., Cole, C. M., & Meyer, W. J. (1995). Use of Depo-Provera to control sexual aggression in persons with traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *10*, 47–58.
- English, K. (2004). The containment approach to managing sex offenders. *Seton Hall Law Review*, *34*, 1255–1272.
- Fernandez, Y. M., Shingler, J., & Marshall, W. L. (2006). Putting “behavior” back into the cognitive-behavioral treatment of sexual offenders. In W. L. Marshall, Y. M. Fernandez, L. E. Marshall, & G. A. Serran (Eds.), *Sexual offender treatment: Controversial issues* (pp. 211–223). Hoboken, NJ: Wiley.
- Gottesman, H. G., & Schubert, D. S. (1993). Low-dose oral medroxyprogesterone acetate in the management of the paraphilias. *Journal of Clinical Psychiatry*, *54*, 182–188.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correction treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, *36*, 865–891.
- Hanson, R. K., & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66*, 348–362.
- Hanson, R. K., Harris, A., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The dynamic supervision project*. Ottawa: Public Safety Canada.
- Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis* (Corrections Research User Report No. 2004-02). Ottawa, Ontario, Canada: Public Safety and Emergency Preparedness Canada.
- Hanson, R. K., Steffy, R., & Gauthier, R. (1993). Long-term recidivism of child molesters. *Journal of Consulting and Clinical Psychology*, *61*, 646–652.
- Heil, P., Ahlymeyer, S., & Simon, D. (2003). Crossover sexual offenses. *Sexual Abuse: A Journal of Research and Treatment*, *15*, 221–236.

- Hinton, J. W., O'Neil, M. T., & Webster, S. (1980). Psychophysiological measurement of sex offenders in a security hospital. *Archives of Sexual Behavior*, 9, 205–216.
- Howes, R. J. (1995). A survey of plethysmographic assessment in North America. *Sexual Abuse: A Journal of Research and Treatment*, 7, 9–24.
- Howes, R. J. (2003). Circumferential change scores in phallometric assessment: Normative data. *Sexual Abuse: A Journal of Research and Treatment*, 15, 365–375.
- Kafka, M. P. (1997). Hypersexual desire in males: An operational definition and clinical implications for men with paraphilias and paraphilia-related disorders. *Archives of Sexual Behavior*, 26, 505–526.
- Kafka, M. P. (2008). Neurobiological processes and comorbidity in sexual deviance. In D. R. Laws & W. T. O'Donohue (Eds.), *Sexual deviance: Theory, assessment, and treatment* (pp. 571–593). New York: Guilford.
- Kafka, M., & Prentky, R. (1992). Fluoxetine treatment of nonparaphilic sexual addictions and paraphilias in men. *Journal of Clinical Psychiatry*, 53, 351–358.
- Khan, O., Ferriter, M., Huband, N., Powney, M. J., Dennis, J. A., & Duggan, C. (2015). Pharmacological interventions for those who have sexually offended or are at risk of reoffending. *Cochrane Database of Systematic Reviews*, 2, 1–109. doi:10.1002/14651858.CD007989.pub2.
- Kiersch, T. A. (1990). Treatment of sex offenders with Depo Provera. *Bulletin of the American Academy of Psychiatry and the Law*, 18, 179–187.
- Klein, V., Schmidt, A. F., Turner, D., & Briken, P. (2015). Are sex drive and hypersexuality associated with pedophilic interest and child sexual abuse in a male community sample? *PLoS One*, 10(7), e0129730. doi:10.1371/journal.pone.0129730
- Kwan, M., Greenleaf, W. J., Mann, J., Crapo, L., & Davidson, J. M. (1983). The nature of androgen action on male sexuality: A combined laboratory-self-report study on hypogonadal men. *Journal of Clinical Endocrinology and Metabolism*, 57, 557–562.
- Levenson, J., & Prescott, D. S. (2014). Déjà vu: From Furby to Langstrom and the evaluation of sex offender treatment effectiveness. *Journal of Sexual Aggression*, 20, 257–266.
- Malcolm, P., Andrews, D., & Quinsey, V. L. (1993). Discriminate and predictive validity of phallometrically measured sexual age and gender preference. *Journal of Interpersonal Violence*, 8, 486–501.
- Maletzky, B., Tolan, A., & McFarland, B. (2006). The oregon depo-Provera program: A five-year follow-up. *Sexual Abuse: A Journal of Research and Treatment*, 18, 303–316.
- Marques, J. K., Wiederanders, M., Day, D. M., Nelson, C., & van Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment*, 17, 79–107.
- McGrath, R., Cumming, G., Burchard, B., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management: The Safer Society 2009 North American Survey*. Brandon, VT: Safer Society Press.
- Merdian, H., & Jones, D. (2011). Phallometric assessment of sexual arousal. In D. Boer, R. Eher, L. Craig, M. Miner, & F. Pfafflin (Eds.), *International perspectives on the assessment and treatment of sexual offenders: Theory, practice, and research* (pp. 221–231). West Sussex, England: Wiley.
- O'Donohue, W. T., & Geer, J. H. (1985). The habituation of sexual arousal. *Archives of Sexual Behavior*, 14, 233–246.
- O'Donohue, W. T., & Letourneau, E. (1992). The psychometric properties of the penile tumescence assessment of child molesters. *Journal of Psychopathology and Behavioral Assessment*, 14, 123–174.
- Perilstein, R., Lipper, S., & Friedman, L. (1991). Three cases of paraphilias responsive to fluoxetine treatment. *Journal of Clinical Psychiatry*, 52, 169–170.
- Prentky, R. A., Lee, A. F., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior*, 21, 635–659.
- Quinsey, V. L., Steinman, C. M., Bergersen, S. G., & Holmes, T. F. (1975). Penile circumference, skin conductance, and ranking responses of child molesters and "normals" to sexual and nonsexual visual stimuli. *Behavior Therapy*, 6, 213–219.
- Rea, J. A., DeBriere, T., Butler, K., & Saunders, K. J. (1998). An analysis of four sexual offenders' arousal in the natural environment through the use of a portable penile plethysmograph. *Sexual Abuse: A Journal of Research and Treatment*, 10, 239–255.
- Rea, J. A., Dixon, M., & Zettle, R. D. (2014). Assessing the generalization of relapse-prevention behaviors of sexual offenders diagnosed with an intellectual disability. *Behavior Modification*, 38, 23–42.
- Rea, J. A., Williams, D., Saunders, K. J., Dixon, M., Wright, K., & Spradlin, J. E. (2003). Covert sensitization: A generalization analysis in the laboratory and natural environment through the use of a portable-penile plethysmograph. *The Behavior Analyst Today*, 4, 190–199.
- Reyes, J. R., Vollmer, T. R., Sloman, K. N., Hall, A., Reed, R., Jansen, G., & Stoutimore, M. (2006). Assessment of deviant arousal in adult male sex offenders with developmental disabilities. *Journal of Applied Behavior Analysis*, 39, 173–188.
- Rosen, R. C., & Kopel, S. A. (1977). Penile plethysmograph and biofeedback in the treatment of a transvestite-exhibitionist. *Journal of Consulting and Clinical Psychology*, 45, 908–916.
- Schaefer, H. H., & Colgan, A. H. (1977). The effect of pornography on penile tumescence as a function of reinforcement and novelty. *Behavior Therapy*, 8, 938–946.
- Schmucker, M., & Losel, F. (2008). Does sexual offender treatment work? A systematic review of outcome evaluations. *Psychothema*, 20, 10–19.
- Schober, J. M., Kuhn, P. J., Kovacs, P. G., Earle, J. H., Byrne, P. M., & Fries, R. A. (2005). Leuprolide acetate suppresses pedophilic urges and arousability. *Archives of Sexual Behavior*, 34, 691–705.
- Stein, D., Hollander, E., Anthony, D., Schneier, F., Fallon, B., Liebowitz, M., & Klein, D. (1992). Serotonergic medications for sexual obsessions, sexual addictions, and paraphilias. *Journal of Clinical Psychiatry*, 53, 267–271.
- Thibaut, F., Cordier, B., & Kuhn, J. M. (1996). Gonadotrophin hormone releasing hormone agonist in cases of severe paraphilia: A lifetime treatment. *Psychoneuroendocrinology*, 21, 411–419.
- Thibaut, F., De La Barra, F., Gordon, H., Cosyns, P., Bradford, J. M., & WFSBP Task Force on Sexual Disorders. (2010). The World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the Biological Treatment of Paraphilias. *World Journal of Biological Psychiatry*, 11, 605–655.
- Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wincze, J. P., Bansal, S., & Malamud, M. (1986). Effects of medroxyprogesterone acetate on subjective arousal, arousal to erotic stimulation, and nocturnal penile tumescence in male sex offenders. *Archives of Sexual Behavior*, 15, 293–305.
- Winder, B., Lievesley, R., Kaul, A., Elliott, H. J., Thorne, K., & Hocken, K. (2014). Preliminary evaluation of the use of pharmacological treatment with convicted sexual offenders experiencing high levels of sexual preoccupation, hypersexuality and/or sexual compulsivity. *Journal of Forensic Psychiatry & Psychology*, 25, 176–194.
- Winett, R. A., Moore, J. F., & Anderson, E. S. (1991). Extending the concept of social validity: Behavior analysis for disease prevention and health promotion. *Journal of Applied Behavior Analysis*, 24, 215–230.
- Wish, J. R., McCombs, K. F., & Edmonson, B. (1980). *The socio-sexual knowledge and attitudes test* (1st ed.). Wood Dale, IL: Stoelting.