Measurement of Testosterone in Human Sexuality Research: Methodological Considerations

Sari M. van Anders · Katherine L. Goldey · Sarah N. Bell

Abstract Testosterone (T) and other androgens are incorporated into an increasingly wide array of human sexuality research, but there are a number of issues that can affect or confound research outcomes. This review addresses various methodological issues relevant to research design in human studies with T; unaddressed, these issues may introduce unwanted noise, error, or conceptual barriers to interpreting results. Topics covered are (1) social and demographic factors (gender and sex; sexual orientations and sexual diversity; social/familial connections and processes; social location variables), (2) biological rhythms (diurnal variation; seasonality; menstrual cycles; aging and menopause), (3) sample collection, handling, and storage (saliva vs. blood; sialogogues, saliva, and tubes; sampling frequency, timing, and context; shipping samples), (4) health, medical issues, and the body (hormonal contraceptives; medications and nicotine; health conditions and stress; body composition, weight, and exercise), and (5) incorporating multiple hormones. Detailing a comprehensive set of important issues and relevant empirical evidence, this review provides a starting point for best practices in human sexuality research with T and other androgens that may be especially useful for those new to hormone research.

S. M. van Anders (✉)
Departments of Psychology and Women's Studies, Program in Neuroscience, Reproductive Sciences Program, Science, Technology, and Society Program, University of Michigan, 530 Church Street, Ann Arbor, MI 48109, USA
e-mail: smva@umich.edu

K. L. Goldey
Department of Psychology, University of Michigan, Ann Arbor, MI, USA

S. N. Bell
Departments of Psychology and Women's Studies, University of Michigan, Ann Arbor, MI, USA

Introduction

The body of research linking steroid hormones, such as testosterone (T) and other androgens, to human sexuality and social contexts is rapidly growing. For example, T has been linked with physiological and self-reported levels of arousal in women (Heiman et al., 2011) and men (Stoleru, Ennaji, Cournot, & Spira, 1993) as well as social and sexual behaviors in both men and women (Edelstein, Chopik, & Kean, 2011; van Anders & Goldey, 2010). Viewing sexual stimuli increases T in men (Redoute et al., 2000; Stoleru et al., 1999) but not women (van Anders, Brotto, Farrell, & Yule, 2009), whereas sexual thoughts have been shown to elicit T increases in women (Goldey & van Anders, 2011). In addition, research and theory indicate that T is positively linked with sexual intimacy and negatively linked with nurturant intimacy in men and women (van Anders, Goldey, & Kuo, 2011).

Given the increasing interest in measuring androgens in human sexuality research, evidence-based guidelines on appropriate methodological considerations are critical for empirically-supported research protocols. There are contemporary reviews on best practices for incorporating cortisol and other stress-related hormones into a variety of study designs with humans (e.g., Adam & Kumari, 2009). There have also been two recent reviews that addressed methodological issues for research with estrogens from a physical/biological anthropology approach (Jasienska & Jasienski, 2008; Vitzthum, 2009). In contrast, there are no methodological reviews on testosterone (T) or other androgens, except one foundational article published over two decades ago (Ellison, 1988). Furthermore, there are no reviews on best practices for incorporating T in sexuality studies, which can involve a

specific set of sexuality-related confounds and issues. Accordingly, our goal in this article is to provide a current and comprehensive review of methodological issues for human sexuality research with androgens, focusing on T. Though the considerations provided in this review are focused on human sexuality research, they will also be pertinent to more general human biobehavioral research that involves hormones. Given how critical best practices are to successful research, we also provide evidence-based practical suggestions that may be useful for avoiding confounds in sexuality studies with T (see Table 1). This article may thus be most useful to researchers who are considering or have recently begun incorporating T to better address their questions of interest.

Any methodological review might focus on myriad issues and the scope of this article is intended to be comprehensive rather than exhaustive, and is limited to *pre*-sampling issues, i.e., methodological concerns that affect study design. Though the majority of methodological research on androgens is conducted with men, studies with T increasingly include women, and this review focuses on both women and men where possible. Given the prevalence and utility of salivary measures in human sexu-

ality research, we focus more on methodological considerations for salivary rather than blood sampling of hormones (see also below: Saliva vs. Blood). Some of the issues we focus on in this review—such as gender/sex and other social location variables—will be relevant to the content of sexuality research as well as its methods, whereas other topics, like biological rhythms, are less likely to be of interest per se, but are critical to designing methodologically strong studies with hormones and conducting analyses that provide best chances for detecting weak to moderate effects (the range in which hormone-behavior associations tend to fall). There are a number of potential confounds and topics for research on T that are reviewed in this article; not all are relevant to each researcher or study.

We also want to note that the ease of salivary sampling has meant that incorporating T into sexuality research has become increasingly feasible. However, the theoretical rationales for measuring T, as well as the theoretical foundations for interpreting results with T, have not necessarily changed. Even in cause-effect studies, there may well be mediating variables that merit attention. And, theoretically-oriented research is as valuable in sexual social neuroendocrinology as it is in any field.

Table 1 Methodological issues in sexuality research with testosterone and other androgens in humans, and suggested practices for addressing them

Issue	Practices for addressing issue
Gender and sex[a]	Include women and men in studies; avoid attributing effects to sex rather than gender unless causality is demonstrated
Sexual orientations and diversity[a]	Measure sexual orientation via open-ended and Kinsey questions when appropriate; consider cultural relevancy of questions and terms; include sexual minorities in *all* studies, not just etiology research
Social/familial connections and processes[a]	Measure parental/relationship/sexual experience and status when relevant; control statistically or check for group differences. Exclude pregnant/lactating women unless questions are relevant, or check for group differences
Social location variables[a]	Record ethnicity, immigration status, and SES for characterizing samples via open-ended questions when possible
Diurnal variation[a]	Control for time statistically and via restrictions; choose afternoon for one-sample or experimental studies. Measure sleep length/quality if relevant
Seasonality	Control for date/season statistically and/or via restrictions
Menstrual cycles	If pertinent, identify phase via repeated hormone sampling or backwards counting. Otherwise, no need to control for phase
Aging and menopause[a]	Control for age statistically or via restrictions; consider menopausal status
Saliva versus blood	Choose saliva for most sexuality studies (especially behavioral or naturalistic studies); choose blood for clinical/medical audiences
Sialogogues, saliva, and tubes	Avoid sialogogues or introducing matter into mouth 1 h pre-testing. Use polystyrene tubes
Sampling frequency, timing, and context[a]	Increase *N* over sampling frequency. With pre/post designs: collect post samples ~15 min post-manipulation, use percent change scores, and ensure control activities are actually neutral. Identify gender(s) of experimenter(s)
Sample shipping	Use gel freezer packs, multiple layers of packaging, and 1-2 day shipping when participants ship saliva samples
Hormonal contraceptives[a]	Exclude women using HCs or check for group differences
Medications and nicotine	Query medication and nicotine use; exclude or statistically control for use
Health conditions and stress	Query conditions, exclude when they affect T
Body composition, weight, and exercise	Collect height and weight, control for BMI statistically when relevant. Collect information on recent and/or average exercise patterns (e.g., number and intensity); control statistically when relevant
Incorporating multiple hormones	Generally, consider similar issues to above. Consult relevant reviews for estrogens (Jasienska & Jasienski, 2008; Vitzthum, 2009) and cortisol (Adam & Kumari, 2009)

[a] Indicates issues that are especially important to consider and/or most widely controlled

Social and Demographic Factors

Gender and Sex

T is markedly higher in men than women with little overlap in distribution and researchers have hypothesized that women are more sensitive to fluctuations in T than men (i.e., in women, smaller differences in T may account for larger changes in behavior or desire compared with men whose T levels differ by the same amount) (Bancroft, 2002, 2009; Sherwin, 1988). According to this hypothesis, a gender/sex difference in sensitivity to T may occur because substantially higher levels of T in males during early development could de-sensitize them to the behavioral effects of T, which may function to offset potentially adverse behavioral effects of high levels of T (Bancroft, 2002, 2009). T is largely studied in men and this focus on males parallels non-human research where females are vastly understudied (Beery & Zucker, 2010), with some exceptions (see Dixson, 2012); accordingly, researchers may not appreciate that T is meaningfully studied in women (Halpern, Udry, & Suchindran, 1997; Hamilton & Meston, 2010; Singh, Vidaurri, Zambarano, & Dabbs, 1999; van Anders, Brotto, Farrell, & Yule, 2009; van Anders et al., 2011; van Anders, Hamilton, Schmidt, & Watson, 2007; Welling et al., 2007). In addition to basic research on T and women's sexuality, a large clinical literature has examined the effects of T deficiency and administration on women's sexual function, with nuanced results suggesting that some women are more sensitive to the behavioral effects of T than others (for reviews, see Bancroft, 2002, 2005, 2009). However, research that includes T typically focuses on male participants and/or addresses behaviors that are tied to cultural stereotypes about masculinity, including aggression and high/hypersexuality (Jordan-Young, 2010). In women, this has translated into either a deficit approach, where (low) T is generally studied in association with low sexual desire, or an over-focus on female biology, where T is only studied in relation to menopausal status, pregnancy, or menstrual cycles as opposed, for example, to the "general" behavior studied in men. But research on a host of topics beyond this limited range has provided important insights and T is tied to sexuality in nuanced ways that challenge cultural stereotypes; for example, associations between T and sexual desire may be positive, negative, or non-existent depending on context and gender/sex (van Anders, 2012b). Accordingly, researchers are increasingly including women in their research, even for those general questions that go beyond female-specific phenomena (though these issues are important and merit attention). And, researchers have moved beyond the more narrow culturally scripted proscription for T to topics on sexuality other than hyper/high sexuality (e.g., Exton et al., 2001; Goldey & van Anders, 2011; Sagarin, Cutler, Cutler, Lawler-Sagarin, & Matuszewich, 2009; van Anders & Dunn, 2009).

Sexuality differences between men and women related to T are often interpreted as evidence of *sex* differences, i.e., inborn, evolutionary, or nonplastic, because they are mediated by T and thus "biology." However, social contexts alter T in humans and non-human species (Gleason, Fuxjager, Oyegbile, & Marler, 2009; van Anders & Watson, 2006c), such that differential gender socializations could lead to ostensible "sex" differences that *appear* to be mediated by T. Or different patterns of T or T responsivity could reflect lifetimes of distinct social contexts and learning. Without experimental data, it would be difficult to tell whether hormone-behavior correlations that differ in men and women reflect gender or sex differences (Hines, 2005), so assumptions of biological causes for difference can be unfounded. Accordingly, many researchers use gender to avoid terminology that implies biological causation. Still, "gender" is sometimes used not to denote sociocultural construction, but rather as a placeholder for sex, so we use "gender/sex" to reflect that biological versus social causation cannot be clearly disentangled.

Sexual Orientations and Sexual Diversity

Sexual diversity may be incorporated as a major topic of investigation or an important individual difference variable by sexuality researchers, and research that examines links between sexuality and T is often conducted with an eye to questions that relate to sexual minorities. Human research on T and sexual minorities has a difficult past that can affect current research practices. For example, individuals from sexual minorities were pathologized by hormone researchers seeking to discover a T-related etiology or treatment for same-gender/sex fantasies, interests, and behaviors (Oudshoorn, 1994). This legacy has not surprisingly made LGBT (lesbian, gay, bisexual, trans-identified) individuals concerned about participating in hormone studies. In addition, LGB individuals may rightly conclude that *current* research on T and sexual diversity in humans is still largely concerned with etiologies of non-heterosexuality (van Anders, 2012a), even though the current frame is basic (e.g., how? why?) as opposed to medical (e.g., what went wrong?), and some researchers are LGBT-identified themselves and consider their work to, in part, promote acceptance/tolerance of LGBT individuals and communities.

Further, many researchers still exclude sexual minorities when that research is not explicitly about sexual minorities (e.g., about "general" sexual behaviors or relationships). This may be taken to suggest further to this community that their only epistemological value is related to their sexual minority status (van Anders, 2012a), and that researchers see them as too *other* to be included among the "regular" sample of everyday participants (i.e., "the" population). In contrast to this "etiology approach," researchers have begun to incorporate sexual minorities into their research in ways that do not position sexual diversity as a problem to be explained. For example, researchers have studied hormonal influences on sexual motivation in sexual minority women (Diamond & Wallen, 2010; Matteo & Rissman, 1984),

or how diverse approaches to relationships, including polyamory, are linked to T (van Anders, Hamilton, & Watson, 2007).

The way researchers position sexual diversity—as a problem to be explained, as a research-worthy sexual variation, or as just one of many possible individual difference variables—may impact the willingness of sexual minorities to participate in research, both at present and in the future, which makes these issues relevant to even the least politically-interested scientist. In addition, the inclusion or exclusion of sexual minorities in research has far reaching impacts in terms of the way it affects other researchers' ability to draw conclusions or build on past research. For example, the "etiology focus" on comparing individuals with same- versus other-sex interests means that research about bisexuality (the second most common sexual orientation/identity) (Herbenick et al., 2010) is largely absent from the literature and thus absent from scientific sexual knowledge.

Sexual diversity takes different forms between and within cultures and times, such that unitary understandings of sexual minorities cannot be taken for granted. For example, in some cultures, anal sex between men is seen as a part of gay sexuality whereas, in others, only receptive anal sex would qualify. "MSM" (men who have sex with men) or "WSW" (women who have sex with women) are terms used to describe sexual behaviors in contrast to sexual identities, though these terms are now informing identities in some places. Researchers who study links between T and sexuality typically focus on orientation, i.e., some mix of behavior, attraction, and fantasy typically via some iteration of the Kinsey questions of sexual practice (Kinsey, Pomeroy, & Martin, 1948), though some are interested in identity and therefore group individuals by self-report (e.g., heterosexual, queer, bisexual, lesbian, etc.). A large number of researchers simply ask participants to check a box indicating heterosexual, bisexual, or homosexual. All three possibilities (Kinsey questions, open-ended self-report, check boxes) have drawbacks and advantages. Given that sexual behavior, orientation, and identity reflect distinct aspects of an individual's sexuality, the method researchers use to categorize participants by sexual orientation or identity can affect the conclusions drawn from research.

The Kinsey questions are widely used and participants respond to questions of behavior and/or fantasy (and sometimes attraction) on scales of "0" to "6" that range from exclusively other gender/sex to exclusively same-gender/sex. Scholars have noted that there is no standardization of coding such that an array of practices and cut-offs are used to categorize individuals (Jordan-Young, 2010). Still, many researchers categorize participants who select "0" or "1" as heterosexual and participants who select "5" or "6" as homosexual (the orientation term used in the Kinsey questions) or gay/lesbian (identity terms preferred by the sexual minority community), with people selecting intermediate numbers as bisexual, though research does question collapsing 0's and 1's together (Chivers, Bouchard, Timmers, & Haberl, 2012). Some others will categorize participants as heterosexuals and non-heterosexuals, which collapses bisexual and same-

gender/sex oriented individuals together. This may be problematic as a large body of research demonstrates that bisexual, heterosexual, and lesbian/gay individuals differ on a number of dimensions in nonparallel ways, i.e., bisexual individuals are not "more" similar to either heterosexual or lesbian/gay individuals across measures (van Anders, 2012a). One disadvantage of the Kinsey questions is the "homosexual" terminology, which is seen as pejorative or outdated by many sexual minority individuals and intrinsically tied to the pathologizing interest of past medical practitioners and some "conversion" therapists (who claim to decrease same-gender sexual interests or at least practices); we sometimes consider using same-gender/sex oriented (which itself has gendered assumption problematics built in). Another disadvantage is that the Kinsey questions are rooted in gender, but some individuals are not interested in gender as a factor in sexual attraction (as with person-not-gender/sex), and others are genderqueer or trans-identified or interested in genderqueer/trans individuals (and thus may not have an expressed gender/sex to provide a basis for "same" or "other" gender/sex-attraction). In this case, bisexual and pansexual may appear to be the same (i.e., the midpoint on the scale) even though they imply very different things (i.e., attracted to women and men vs. attracted to people regardless of their gender). Some individuals may also be attracted to masculinity but not femininity in men *and* women (or the reverse), or may be attracted to masculine women but feminine men (or the reverse); obviously, the Kinsey questions become nonsensical for these individuals. The benefit of the Kinsey questions, however, is that they tap behavior and fantasy rather than identity, which can be important when identity is not the variable of interest. But this can be another drawback when current identity might matter more than life history of behavior (e.g., a woman who lived her life as heterosexual but fell in love and partnered with a woman in later life might have very heterosexual scores on this scale despite currently living and identifying as lesbian). Or, she might just respond to the scale in ways that affirm her current identity, which questions the utility of the scale at all if in practice it is used to reflect current identity anyway.

Open-ended self-report is also widely-used and is especially useful when identity is important and freely expressed. In contrast, in cultures where identifying as GLB or queer puts one at risk for violence or other repercussions, asking participants to report their sexual orientation or identity may be problematic and only certain individuals might participate (e.g., those who are already culturally stigmatized by public identities, those who are out, or those who make the choice to do so despite the potential cost). For example, changes in who is comfortable about identifying as lesbian or gay in public or in questionnaires may lead to perceptions of changes over time in the content of sexual diversity when, in reality, what has changed may be the visibility of specific parts of the community. In addition, people may identify one way but still have fantasies or behaviors that extend beyond this identity. The benefit of open-ended self-report is that

researchers can tap into existing and evolving linguistic norms (e.g., homosexual → gay → lesbian → dyke → queer) that can still be categorized in a post hoc fashion for analyses. That is, even self-report responses can be categorized for quantitative analysis purposes (van Anders & Goldey, 2010). However, providing examples (e.g., heterosexual, gay, lesbian, bisexual, queer) can be helpful, as some individuals are less confident with what "sexual orientation" actually means (and this is especially true for sexual majority individuals who have not been forced to contemplate their sexual identities). In our lab, participants have sometimes written "female" assuming the question asks what their sexual orientation is *to*, and others have written "regular" or "normal," which can obviously be offensive for sexual minority (or ally) experimenters to have to see repeatedly.

Asking participants to select from a preset list of options allows for quick, easily categorized responses, which can be especially useful for large online studies. However, because the options must be picked a priori, they may not reflect the community under study, and may be viewed as outdated or offensive. In our lab, we have found that only one participant has *ever* self-identified as "homosexual" and that sexual minority women rarely self-identify as "gay," yet both these terms appear regularly as pre-selected checklist options. As noted above, homosexual is both outdated and linked to conceptualizing sexual minorities as pathological; gay is a term that many use to refer to same-sex oriented individuals even though communities typically use it to refer to men only and terms like lesbian, dyke, queer (and others) are instead used by women (men also use other terms, including queer). The use of homo/hetero/bisexual or gay/heterosexual distinctions in pre-set checkboxes limits community members' abilities to self-define, and also imposes identity labels that sexual minorities have fought and still fight to both control and make public.

"Transsexual," "transgender" or similar terms are sometimes included in a list of sexual orientations, but many trans-identified individuals see their trans status as being a property of gender/sex rather than sexual orientation. However, natal sex may be seen as relevant to sexual orientation and/or sexual identity for some scientists and/or trans-identified individuals. Increasingly, though, many would argue that current gender/sex identification and community associations are what matters for sexual identity rather than natal sex or the junction of natal and transitioned gender/sex. This remains contentious and, likely, a study- or community-specific research and social issue.

Social and Familial Connections and Processes

Social and relational variables influence T levels and can also be linked to T in trait ways. For example, sexually active status—including current sexual activity levels or having ever engaged in sexual activity—are both linked with T. We have found that associations between T and partnering status are mediated by current sexual activity levels in women, such that the lower T in

women in long-term relationships relative to single women was explained by long-term partnered women's more frequent sexual activity (van Anders & Goldey, 2010). And, there is evidence that sexual experience can moderate associations between T and other social variables: Roney, Mahler, and Maestripieri (2003) (cf. Roney, Lukaszewski, & Simmons, 2007) found that only sexually experienced men showed T responses to conversations with women. Paradoxically, desisting from sexual activity for a period of time, i.e., abstinence, is itself linked to higher T as well (Exton et al., 2001). These factors may be meaningful confounds or explanations of other T-sexuality associations (e.g., it may be that anticipation is linked to higher T). Accordingly, some researchers assess sexual experience and frequency. However, a fascinating body of research highlights that "sex" and "sexuality" can be interpreted in different ways (e.g., sex might mean intercourse to some, but *any* sexual contact to others, including non-genital contact) (Sanders & Reinisch, 1999). Accordingly, many researchers define sexual experience and frequency of sexual activity by specifying what does and what does not count in the definition (e.g., consensual sexual contact with your or a partner's genitals; any sexual contact, including deep kissing but not friendly "pecks" or backrubs, etc.). The choice of an appropriate definition may depend on the outcome variable of the study; for example, a study on STI prevention may define sexual activity differently than a study on sexual desire. Given that more frequent masturbation and solitary orgasms are linked with higher T in women (van Anders, Hamilton, Schmidt, et al., 2007) and that T is differentially linked with solitary (positively) versus dyadic (negatively) desire in women (van Anders, 2012b), considering solitary sexual behaviors may be important as well.

In addition to sexual experience, a large body of research demonstrates that T differs by relationship status in women and men (Gray, Chapman et al., 2004; Gray & Campbell, 2009; van Anders, 2009; van Anders & Goldey, 2010; van Anders & Gray, 2007; van Anders & Watson, 2007b). This body of work suggests that monoamorously partnered individuals in romantic/sexual relationships characterized by commitment and nurturance have lower T. In contrast, individuals in multiple relationships, ostensibly monogamous relationships characterized by low commitment or cheating, or relationships characterized by a lack of nurturance and the presence of hostility, appear to have higher T as do single individuals. For women, partner presence seems to be a key variable, as same-city but not long-distance partnered women have lower T than single women. Casual relationships appear to differ by gender/sex, with lower T for women and higher T for men. Querying and recording relationship status can thus be a critical way to interpret potential third variable associations or other statistical issues. However, individuals have more complicated relationship profiles than might be expected and term definitions can be similarly helpful here, especially for research with populations that differ from researchers by age, social location, etc. Relationship characterizations and

terminologies can differ by culture and time, but also by generation. For example, "hook-ups" and "friends with benefits" are categories that do not neatly fit into single/committed dichotomies and might be more meaningful to younger versus older generations. Moreover, "single" might be interpreted as meaning free to pursue relationships *or* having no relationships; specifying the term's components can avoid misunderstanding (e.g., "single" = currently having no sexual or romantic contacts with anyone, meaning no hook-ups, one-night stands, etc.). As such, some researchers ask a variety of questions, including open-ended ones, to try and accurately assess relationship status given the complexities in nuance. Open-ended questions also allow researchers the opportunity to discover terms and relationship approaches they may not have previously encountered. Our lab's own research on polyamory stemmed from participants informing us that our limited checkboxes (which we previously used) did not fit their relationship approach.

Similar to relationship status, there is a growing body of evidence demonstrating that parents have lower T than non-parents and that T shows a birth-specific drop in parents (Gray & Campbell, 2009; Gray, Parkin, & Samms-Vaughan, 2007; Gray, Yang, & Pope, 2006; Kuzawa, Gettler, Huang, & McDade, 2010; Storey, Walsh, Quinton, & Wynne-Edwards, 2000). Gonadal steroids, including T, increase during pregnancy and are lower during lactation in women (Alder & Bancroft, 1988; Greenspan & Gardner, 2001) and T also changes among co-fathers (i.e., fathers who are involved in parental care together with their partners) over pregnancy with a decrease in T (with perhaps one brief increase) that stays low but slowly increases with infants' ages (Gettler, McDade, Feranil, & Kuzawa, 2011; Storey et al., 2000); this is similar to other biparental mammals (Wynne-Edwards, 2001). Accordingly, many researchers query participants about parental status, including age of offspring, and especially about pregnancy/lactation status, because these factors could introduce noise or confound other group differences in T.

Social Location Variables

Although sexuality researchers with training in Women's Studies or feminist psychology tend to consider intersections among sexuality and other social location variables, such as ethnicity, immigration status, or socioeconomic status (SES) (Blanc, 2005; Froyum, 2010; Phillips et al., 2011), sexuality research conducted from a physiological perspective (e.g., research with T or genital arousal) has devoted less attention to these identity variables. In contrast to research on sexuality and hormones, research on stress and cortisol increasingly takes these identity variables as critical to understanding the interplay between social location/experience and hormonal processes (e.g., DeSantis et al., 2007). Research on T is often unconcerned with these issues, yet characterizing samples is critical to understanding the phenomenon under investigation and to what extent the results generalize beyond the study's sample. Moreover, cultural variables have been shown to modify associations between gonadal hormones and other social and health variables (Gehlert et al., 2008). Accordingly, a few researchers use open-ended questions to address ethnicity and immigration status and include these in their demographic description. And, some use measures of participant (or parental, in the case of college students) income to crudely characterize SES.

Biological Rhythms

Diurnal Variation

Androgens show diurnal rhythms linked both to sleep patterns and time of day, with a near 50 % decrease from morning to evening (Dabbs, 1990b). Levels are highest upon waking and then steeply decline in the first 1–2 h post-waking, followed by a more moderate decline during the waking period until levels are lowest just before sleep, at which point androgens start to increase until their highest point immediately before waking (see Table 2) (Aedo, Nunez, Landgren, Cekan, & Diczfalusy, 1977; Axelsson, Ingre, Akerstedt, & Holmback, 2005; Boyar et al., 1974; Dabbs, 1990b; Piro, Fraioli, Sciarra, & Conti, 1973; van Anders & Hampson, 2005). These diurnal patterns are not necessarily standard, however, as older age (e.g., > 65 yrs) is associated with a flatter decrease over the day (Brambilla, Matsumoto, Araujo, & McKinlay, 2009; Bremner, Vitiello, & Prinz, 1983; Luboshitzky, 2003; Nicolau et al., 1985; Panico et al., 1990; Plymate, Tenover, & Bremner, 1989). Because of the strong and well-known diurnal variation in T, researchers have tended to restrict time of sampling and/or control for sampling time via statistical analyses.

In addition to these diurnal rhythms, sleep itself is linked to T: sleep duration is positively correlated with levels of T and sleep disruption is associated with altered levels of T (Goh & Tong, 2010; Luboshitzky, 2003). This may have implications for studies on androgens with parents of infants and young children, college students, depressed individuals, and shift workers, among other groups, because of the associated alterations in sleep patterns.

A number of researchers have reported associations between T and behavior that are stronger when T has been collected in the afternoon versus the morning (Berg & Wynne-Edwards, 2001; Gray, Kahlenberg, Barrett, Lipson, & Ellison, 2002; Muller & Wrangham, 2004; van Anders, Hamilton, Schmidt, et al., 2007; Worthman & Konner, 1987). This has not been consistently shown for any one behavior or across a range of behaviors, but the steep declines in T over the morning may add noise or variation that obscures underlying effects that are visible with afternoon sampling, when levels are less variable. The growing preference for afternoon sampling of T stands in stark contrast to cortisol and biomedical research that focus on waking samples

Table 2 Levels of testosterone over a 1-day period

Phase of day	Relative levels of androgens	Direction of change in androgens
Pre-waking sleep	High	Increasing
Waking point	Highest	–
Post-waking (1–2 h)	High-moderate	Decreasing rapidly
Awake	Moderate-low	Decreasing
Pre-sleep awake	Lowest	–
Sleep	Moderate-high	Increasing

Note See text for supporting citations

or daily slopes (O'Donnell, Badrick, Kumari, & Steptoe, 2008), yet this afternoon sampling approach with T has provided meaningful and consistent results. Many researchers successfully use a single sample to measure T if sampling time is restricted to the afternoon (see also below: Sampling Frequency, Timing, and Context), but researchers interested in T profiles across the day may sample T once in the morning and once in the afternoon or evening (Gettler et al., 2011; Gray et al., 2006).

Seasonality

In addition to diurnal rhythms, there is seasonal variation in androgens ranging up to twofold increases though the majority of evidence stems from cross-sectional rather than longitudinal studies. Moreover, evidence is somewhat variable and generally focused on seasonality in North America and Europe, limiting generalizability. Autumn tends to show the most consistent peaks in androgens in men and women (in men: Dabbs, 1990a; Moffat & Hampson, 2000; Reinberg et al., 1978; Reinberg, Lagoguey, Chauffournier, & Cesselin, 1975; Reinberg, Smolensky, Hallek, Smith, & Steinberger, 1988; Smals, Kloppenborg, & Benraad, 1976; Stanton, Mullette-Gillman, & Huettel, 2011; Svartberg, Jorde, Sundsfjord, Bonaa, & Barrett-Connor, 2003; van Anders, Hampson, & Watson, 2006; in women: Kauppila, Kivela, Pakarinen, & Vakkuri, 1987; Kauppila, Pakarinen, Kirkinen, & Markila, 1987; Stanton et al., 2011; van Anders et al., 2006; Wisniewski & Nelson, 2000). Though other peaks and no peaks have also been identified (see Table 3) (Brambilla, O'Donnell, Matsumoto, & McKinlay, 2007; Garde, Hansen, Skovgaard, & Christensen, 2000; Martikainen, Tapanainen, Vakkuri, Leppaluoto, & Huhtaniemi, 1985; Perry, Miller, Patrick, & Morley, 2000; Valero-Politi & Fuentes-Arderiu, 1998), there are data that are consistent with the autumn peak in male non-human primates, including Japanese macaques and rhesus monkeys (Gordon, Bernstein, & Rose, 1978; Muroyama, Shimizu, & Sugiura, 2007). Considering seasonal variation in androgens may be especially critical for longitudinal studies or when data collection spans a considerable time period, and some researchers have accordingly controlled for either season or day of testing in their analyses, or identified testing season/month in their methods.

Researchers have speculated that seasonal variation in androgens result from seasonal variations in nutrition, caloric intake, and work (Jasienska & Ellison, 2004; Vitzthum et al., 2009). It may also be possible that seasonal variation in exposure to light and weather patterns contribute directly to fluctuations in androgens as they do in some other species (Nelson, Denlinger, & Somers, 2009), and given that most research has focused on populations living in North America or Europe, it is unknown whether seasonal effects on T are limited to populations living at higher latitudes. There is no real body of literature investigating the *causes* of seasonality in humans, but seasonality can be an important methodological issue to consider in longitudinal studies or when data collection spans multiple seasons.

Menstrual Cycles

Menstrual cycles are characterized by large fluctuations in estrogens and progesterone, and T shows more moderate variation. T is low during the menstrual phase, but begins a gradual increase that continues over the follicular phase until a peak around ovulation, with a gradual decrease during the luteal phase until onset of menses (see Fig. 1) (Campbell & Ellison, 1992). Similar peaks in T around ovulation have been found in a number of non-human primate species, though experiments with rhesus monkeys suggest that the mid-cycle peak in T has little to no effect on sexual behavior (Dixson, 2012; Michael, Richter, Cain, Zumpe, & Bonsall, 1978). Researchers have addressed whether the magnitude of human menstrual variation in T is large enough such that studies should control for it, and concluded that (1) this is unnecessary unless menstrual variation in T is itself of interest and (2) menstrual variation in T is relatively small compared to other sources of variation like diurnicity or individual differences (Dabbs, 1990b; Dabbs & de La Rue, 1991). Accordingly, many researchers who incorporate T into their (non-phasic) behavioral research do not control for menstrual phase, while others who are specifically interested in menstrual or ovulatory phase do identify and analyze menstrual phase. Still, many researchers assume that menstrual status must be controlled in *any* hormonal studies with women, an assumption that likely stems from historical definitions of sex hormones that tied cyclicity and instability to women and females (Oudshoorn, 1994), especially in light of empirical evidence (as above) that consistently shows menstrual cycle adds less noise than time of day or other variables that routinely go uncontrolled.

Menstrual cycles relate to androgens in other ways. The variability in length and regularity of women's menstrual cycles can reflect differing hormonal contributions. Very long or short cycles can be related to altered T in anovulatory or even healthy women (Campbell & Ellison, 1992; van Anders & Watson, 2006a). Moreover, cycle length and regularity decrease with age (Vitzthum, 2009). And, there are population differences in menstrual cycle lengths, regularity, and hormones that are still being explored (Vitzthum, 2009).

Table 3 Seasonality and androgens in women and men

Peak	Women	Men
Autumn	Kauppila, Kivela, et al. (1987), Kauppila, Pakarinen, et al. (1987), Stanton et al. (2011), van Anders et al. (2006), Wisniewski & Nelson (2000)	Dabbs (1990a), Moffat & Hampson (2000), Reinberg et al. (1975, 1978, 1988), Smals et al. (1976), Stanton et al. (2011), Svartberg et al. (2003), van Anders et al. (2006)
May	–	Valero-Politi & Fuentes-Arderiu (1998)
Late Winter	–	Perry et al. (2000)
Null	–	Brambilla et al. (2007), Martikainen et al. (1985)
July–Sept.	Garde et al. (2000)	–

Fig. 1 Relative testosterone levels over a 30 day menstrual cycle, and phase breakdowns using Backwards Counting. An optional "ovulatory phase" could constitute two days around ovulation if necessary. *Note*: 30 days is closer to average cycle lengths than the traditional 28 day cycle
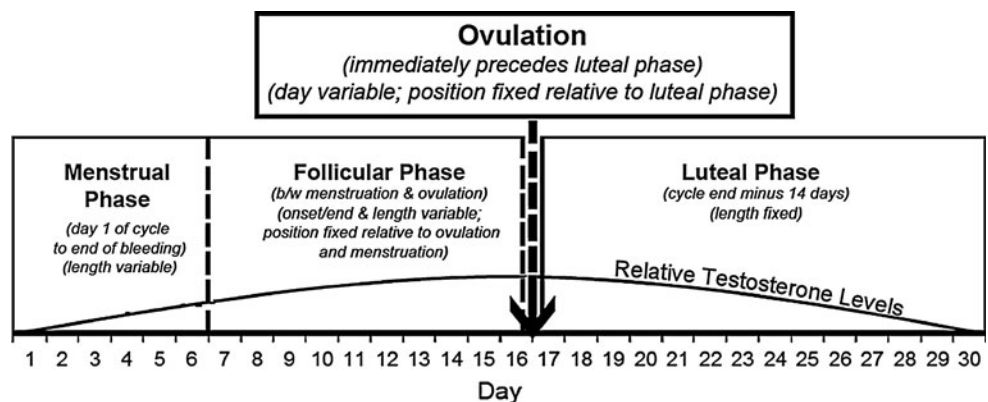
Identifying menstrual phase is most accurate with longitudinal monitoring of relative hormone levels (e.g., see Jasienska & Jasienski, 2008). However, researchers have sometimes used other methods to identify menstrual phase given that long-term monitoring is costly, invasive, and time-intensive (see also Vitzthum, 2009). Below, we detail three methods researchers have used as "shorthands" to menstrual phase identification: Forwards Estimation, the Two-Week Method, and Backwards Counting.

Forwards Estimation

Here, women's phase is identified by counting forward from the first day of the most recent period using a 28-day cycle such that Day 14 is ovulation. Despite the widespread use of Forwards Estimation (near-exclusively by non-hormonal researchers), it is unreliable for a number of reasons. It presumes a 28-day average that is known to be incorrect; average cycle length is instead 29.5 or 30 days. Secondly, it does not take into account the wide variation in cycle length between and within women (Vitzthum, 2009). Perhaps of most concern, there is neither consensus nor standardization of phase breakdown, which has translated into very loose determination of phases with no discernable decision rule.

The Two-Week Method

Researchers interested in quasi-experimental approaches to studying T-behavior links might consider incorporating menstrual phase into their research. Some of these researchers might use menstrual phase as a proxy for different hormone levels rather than because of interest in menstrual phase per se. Given the difficulties of accurately identifying phase, some researchers have instead used the Two-Week Method, in which women are simply tested at two points separated by 2 weeks (van Anders, Chernick, Chernick, Hampson, & Fisher, 2005; Welling et al., 2007). Here, women can be tested close to expected onset of menses or during menses and then 2 weeks following, because T levels should be higher midcycle than earlier or later. Accordingly, this approach is useful when hormone variation, rather than menstrual variation per se, is of interest. Obviously, it is of limited use when menstrual phase *is* specifically of interest. Moreover, it still requires two testing points separated by weeks where hormones must be sampled, which can be difficult for some research designs.

Backwards Counting

A third method for estimating menstrual phase is Backwards Counting (Harvey, 1987), which, like Forwards Estimation,

uses a counting method to identify women's menstrual phase. Here, researchers calculate the actual length of each woman's menstrual cycle via reports of the first days of two consecutive menstrual periods. Menstrual phases are thus more reliably estimated than the Forwards Estimation method, because the luteal phase is close to 14 days in healthy women and is less variable relative to other phases (Ellison, 2001). The menstrual phase is assigned to those days that contain menstrual bleeding. The follicular phase is situated between the menstrual and luteal phases. Ovulation should occur in between the follicular and luteal phases, so researchers could assign the 2 days around ovulation as an ovulatory phase if this is important—though this is less accurate because of the short window. An advantage of Backwards Counting is that it is relatively noninvasive, inexpensive, and though women need to be contacted post-study, this can be done remotely. Figure 1 provides a breakdown of phases by cycle day.

Aging and Menopause

Older ages are associated with lower T, but the cross-sectional nature of the majority of this research makes it difficult to definitively conclude whether these changes are indeed age-related or instead are due to cohort effects (i.e., variation due to birth period) (cf. Feldman et al., 2002; Morley, Perry, Patrick, Dollbaum, & Kells, 2006). Evidence shows lower T with older ages in women (Zumoff, Strain, Miller, & Rosner, 1995) and men (Burger, Dudley, Cui, Dennerstein, & Hopper, 2000; Ellison et al., 2002; Feldman et al., 2002; Ferrini & Barrett-Connor, 1998; Gray, Berlin, McKinlay, & Longcope, 1991; Morley et al., 1997; Nahoul & Roger, 1990; Uchida et al., 2006). Effects of aging on T in men may be a Western phenomenon related to atypically high T early in adulthood; some studies of non-Western populations show *no* significant age differences in T (Campbell, Gray, & Ellison, 2006; Ellison et al., 2002; Ellison & Panter-Brick, 1996) though some do (Ellison et al., 2002; Lukas, Campbell, & Ellison, 2004), depending on the specific population studied. Note that there is no research with women on aging and T in non-Western populations, so effects of aging that appear to be culturally-specific for men may also be so for women. Because of the variation in T by age, many researchers control for age in statistical analyses, and it may be that age is of varying importance depending on populations.

Though menopause brings a marked change in levels of many hormones in women, there are no specific menopause-related decreases in T (Burger et al., 2000). Instead, there is a change in relative levels of androgens and estrogens due to the large decline in estradiol that results from menopause (Vermeulen, 1980). Given this change in T and other hormones, many researchers limit their participants to premenopausal women and similarly-aged men, though this introduces serious issues about the generalizability of findings beyond this age range; given that menopause does not introduce large changes in T itself, excluding postmenopausal women may not really have any justification in theory or evidence.

Sample Collection, Handling, and Storage

Saliva Versus Blood

Androgens are most commonly measured via saliva or blood (serum) in humans, and these result in comparable though different measures. Some fraction of circulating T is bound to albumin or sex hormone binding globulin (SHBG), and the portion of interest to behavioral researchers is typically the unbound fraction available to travel throughout the blood and bind to receptors. Serum results in one of two measures: total T (a direct measure) or "free T" (generally an estimate based on the ratio of total T relative to SHBG). Data indicate that some estimates of free T from total T and SHBG are only moderately correlated with actual levels of free T, especially in men, but methods that physically separate the free and bound portions of serum T (e.g., equilibrium dialysis) are rarely used due to their high costs, labor intensiveness, and inadequate sensitivity for measuring free T in women (Ellison, 1988; Kapoor, Luttrell, & Williams, 1993; Morris, Malkin, Channer, & Jones, 2004; Rosner, Auchus, Azziz, Sluss, & Raff, 2007). Salivary T results in only one measure that is referred to as T, salivary T, or bioavailable T. Free T and salivary T are not the same—free T is an estimate whereas salivary T is a direct measure—though they both represent the unbound and potentially weakly-bound portion of T and are thus especially useful to researchers (Quissell, 1993). In addition, salivary measures may not reflect all of T circulating in the blood, though measures are usually highly correlated (Ellison, 1988). Additional androgens that can be measured include dehydroepiandrosterone (DHEA) and its sulphate (DHEAS), which are released from the adrenal gland and thus useful in comparison to other adrenal hormones like cortisol, as well as two that are not commonly measured: dihydrotestosterone (DHT), which is tied to physical virilization but less often measured in behavioral studies, and androstenedione, a weaker androgen and precursor to testosterone. Other androgens are uncommon in sexuality research or other behavioral research with humans.

Sampling androgens via blood or saliva brings method-specific advantages and disadvantages (see also Ellison, 1988; Vitzthum, 2009). Advantages of salivary sampling over blood include low or no biohazard implications, low invasiveness, high compliance from participants, ease of collection, storage, and shipment of samples, and ability to postpone freezing samples if needed. A special and important bonus of salivary sampling, particularly relevant to sexuality researchers, is the ability of participants to self-collect and self-store samples, allowing for the study of hormone-sexuality links in private and/or naturalistic settings. In addition, little is known about the effects of venipuncture on

subsequent hormone levels or measures relevant to sexuality, though blood sampling (actual or anticipated) seems prima facie to interfere with sexual arousal much more than saliva sampling. Moreover, blood loss itself (i.e., that accompanies blood sampling) is a signal to the body of physical damage or fluid loss (Garrioch, 2004) in a way that spitting is not; the widespread assumption that this emergency signal is nonreactive in terms of research questions is based on faith rather than evidence. The pulsatile fashion of gonadal steroid release may also make saliva a more accurate option, since saliva represents a sort of averaging of hormone release over a short period, thus reflecting both the highs and lows associated with pulsatile release. Furthermore, and countering some questions about whether hydration would matter, salivary flow rate does not affect the measurement of T in saliva (Arregger, Contreras, Tumilasci, Aquilano, & Cardoso, 2007). These advantages make salivary measurement of T easier, more practical, and more amenable to a large array of study designs. However, though salivary assays have been conducted for decades, some biomedical researchers still question the validity of using saliva, with blood being the gold standard.

Saliva T measures have also been validated for research, with studies demonstrating their internal validity (i.e., accuracy, precision, linearity of dilution, sensitivity, and specificity), reliability across time and different laboratories, and external validity (e.g., expected associations with time of day, age, pubertal status, gender/sex, pharmacological manipulations, and clinical conditions) (Dabbs, 1990b; Dabbs et al., 1995; Granger, Schwartz, Booth, & Arentz, 1999; Johnson, Jopling, & Burrin, 1987; Luisi et al., 1980; Walker, Wilson, Read, & Riad-Fahmy, 1980). They are still more controversial than blood as some studies point to nonsignificant or low correlations between salivary and free T in women (Granger, Shirtcliff, Booth, Kivlighan, & Schwartz, 2004; Shirtcliff, Granger, & Likos, 2002; Swinkels, Meulenberg, Ross, & Benraad, 1988), while other studies show good correlations (Khan-Dawood, Choe, & Dawood, 1984; Magrini, Chiodoni, Rey, & Felber, 1986; Swinkels et al., 1988). Some studies also show good correlations between salivary and *total* T (Granger et al., 2004; Shirtcliff et al., 2002). How problematic are these data for studies with T in women? Salivary T measurements in women may add noise and lead to underestimation of effects; thus, sufficient and large sample sizes of women should overcome these problems and do (van Anders, 2010b; van Anders & Dunn, 2009; van Anders & Watson, 2006b). Additionally, serum assays of T in women may be problematic (Taieb et al., 2003), so low correlations between salivary and serum measures of T may be due to issues with accuracy in serum or saliva measures. In men, correlations between salivary and free T are high (Goncharov et al., 2006; Granger et al., 2004; Khan-Dawood et al., 1984; Shirtcliff et al., 2002; Walker et al., 1980; Wang, Plymate, Nieschlag, & Paulsen, 1981). Researchers should recognize that all hormone research involves an *estimation* of the hormone level of interest, and exact measures of some true hormone level are never available. This is most simply reflected in

the accepted reporting of hormone levels that have been averaged from duplicate or triplicate assays of the same sample, along with intra-assay coefficients of variation; i.e., even the same assay provides slightly different estimates of the same hormone level from the same sample. Accordingly, hormone measurement (like all measurement) always involves compromise, and saliva holds many advantages in terms of validity and research design.

Sialogogues (Saliva Stimulants), Saliva, and Tubes

Many researchers use sialogogues to speed up saliva production. Though there are few empirical studies demonstrating any time benefit, there is one study showing that chewing gum speeds up saliva production by 3–6 min depending on type of gum (van Anders, 2010a). Countering this benefit, sialogogues affect the assay process for T: cotton artificially inflates readings of gonadal steroids and candy also alters results (Lipson & Ellison, 1989; Shirtcliff, Granger, Schwartz, & Curran, 2001). Despite widespread use, chewing gum also affects assays of gonadal steroids, including T (Granger et al., 2004; Lipson & Ellison, 1989; Paton, Lowe, & Irvine, 2010; Shirtcliff et al., 2000; cf. Dabbs, 1991), and this includes even popular choices like Trident sugar-free Original flavor. Research has shown that six variations of gum artificially inflate assay readings of T in women and men by up to 150 %, with larger effects for women (van Anders, 2010a). There is also conflicting evidence of whether time since chewing affects assays, as time spent chewing would potentially reflect the amount of chemicals leached from gum into the sample (Dabbs, 1991; Granger et al., 2004). Accordingly, unless saving 3–6 min is a critical design consideration, the detractions of sialogogues appear to vastly outweigh their benefits.

Collection of saliva samples for T typically involves spitting into tubes. However, saliva can bring impurities in the form of blood, from sores or recent tooth brushing, or particulate from food, tobacco, drink, gum, etc. These impurities are known to affect the quality of assays and therefore the fidelity of results (Granger et al., 2007). Some researchers request that participants avoid introducing matter into their mouths prior to the study by avoiding smoking, eating, drinking non-water fluids, or brushing their teeth. Some researchers ask participants to rinse their mouths with water to remove loose or detachable detritus. Other researchers use assay kits that are designed to test for blood contamination or ask participants to report on oral state or taste of blood (Hamilton, van Anders, Cox, & Watson, 2009), which might be especially helpful for studies with athletes or individuals with oral diseases or hygiene issues where mouth injuries or sores might be common.

Most researchers will want to use some sort of plastic tube rather than glass (the gold standard of inertness) to collect saliva samples because of cost and breakage considerations. Other collection vehicles like salivettes and material-based swabs adversely affect steroid assays (Kozaki, Hashiguchi, Kaji, Yasukouchi, & Tochihara, 2009; Kruger, Breunig, Bis-

kupek-Sigwart, & Dorr, 1996; Shirtcliff et al., 2001); how-ever, not all plastic tubes are acceptably inert. Though some recommend polypropylene tubes (Vitzthum, 2009), an empir-ical study demonstrated that they were problematic for steroidal assays (Banerjee, Levitz, & Rosenberg, 1985). Polystyrene tubes are more acceptable, and no studies to our knowledge have shown interference. Instead, T assays from samples collected in glass and polystyrene tubes are highly and significantly correlated at $r = .92$ (Lipson & Ellison, 1989).

Sampling Frequency, Timing, and Context

T shows high test-rest reliability over days ($r = .64$) and months ($r = .52$), suggesting that unitary measures of T can be mean-ingful for inferring trait levels of T, even though T does fluctuate over time and in response to social contexts (Dabbs & de La Rue, 1991). Indeed, research has indicated that between-person vari-ability in T is much larger than within-person variability (Bain, Langevin, D'Costa, Sanders, & Hucker, 1988; Dabbs, 1991; Dabbs & de La Rue, 1991; Rowe et al., 1974), suggesting that a stronger approach to detecting signals is to include additional participants over additional samples per participant. Indeed, one investigation concluded that one sample was as good as the mean of three samples for estimating trait T, which may be an exag-geration but is nonetheless instructive. A number of researchers do use one sample to ascertain trait levels of T in behavioral research (Carré & Putnam, 2010; Gray, Campbell, Marlowe, Lipson, & Ellison, 2004; Mehta & Josephs, 2006; Roney et al., 2007; Schultheiss et al., 2005; van Anders et al., 2009; van Anders, Hamilton, & Watson, 2007; van Anders & Goldey, 2010).

Though some sexuality studies assess correlations between T and behaviors or attitudes, others investigate changes in T. In these experimental designs, samples are typically taken before and after a manipulation, leading to the following meaningful measures of T: baseline T, which can be used to assess antici-patory changes; change in T, which can be absolute (post minus pre) or percent change (post minus pre, all divided by pre); and stimulated levels, i.e., post-manipulation levels. Some research-ers have found that percent changes in T are more sensitive than absolute changes, because the relative measures control for the large variability in baseline/trait T (van Anders, 2010a; van Anders & Watson, 2007a).

Despite an increasing body of research on experimentally manipulated T, the time course of T changes (e.g., time to T response, length of T response) is unknown. In spite of this, researchers have successfully used 15 min as the post-manip-ulation sampling time point, suggesting that 15 min is at least one time point at which changes are measurable (Mehta & Josephs, 2006; Schultheiss et al., 2005; van Anders, Hamilton, Schmidt, et al., 2007; van Anders & Watson, 2007a). Addi-tionally, researchers have sampled 10 min and 20 min post-manipulation for T levels (e.g., Carré & Putnam, 2010). These timeframes are typically measured from the end of the manipu-lation to the sample itself, but whether the length of the manipu-lation itself matters remains an open question. Researchers must also consider whether sampling times for T are the same or different as the most appropriate times for participants to complete questionnaires about psychological responses (e.g., mood and arousal) in experimental studies. It is unclear at what time points T and psychological responses to sexual stimuli are correlated, if they are correlated at all (e.g., Goldey & van Anders, 2011, 2012), so multiple measures of psychological responses (e.g., immediately post-manipulation and 15-min post-manipulation) may be helpful.

Given the 15 min gap between the manipulation's end and the second saliva sample in pre/post experimental studies on T, researchers use a variety of activities to fill this otherwise empty space. These filler activities should ideally be neutral, to avoid any additional confounding and contributing to changes in T. Research clearly demonstrates that a wide range of activities *do* affect T (for reviews, see van Anders & Watson, 2006c; van Anders et al., 2011), including just thinking sexual thoughts (Goldey & van Anders, 2011) or anything competitive (Archer, 2006; Carré & Putnam, 2010; Mehta & Josephs, 2006; van Anders & Watson, 2006c). Thus, filler activities like leafing through fashion magazines, with their sexualized images, responding to sexuality-related questionnaires, or completing cognitive tasks that may result in feelings of success or victory, may be prob-lematic as neutral tasks. Researchers have thus turned to activ-ities like somewhat boring travel videos as neutral/control filler activities (Goldey & van Anders, 2011).

In addition to activities, lab studies necessitate testers and experimenters. Evidence demonstrates that interactions with women increase T in heterosexual men (who were the only group studied) (Roney et al., 2003), which may confound gen-der/sex difference analyses if all testers are women but partici-pants are both women and men. Similarly, sexual orientation/ identity analyses on T may be confounded if participants of only one sexual orientation and gender are tested by the gender/sex they find most (or least) attractive. No researchers, so far as we are aware, counterbalance gender of tester unless this is a spe-cific research question. While doing so would be ideal, it seems difficult to reasonably accomplish, and some researchers instead choose to match the gender of testers and participants because this may increase participant comfort in some sexuality studies (but, of course, may not always: it seems reasonable to conjec-ture that gay-identified male participants might be more uncom-fortable with a heterosexual-identified male tester than a female tester due to high heterosexual-male homophobia). Some researchers therefore identify the gender of tester so that at least this potential confound can become apparent.

Sample Shipping

Some sexuality researchers may wish to recruit participants over a wide geographic area to diversify their sample or to

target a specific population that is underrepresented in the researcher's locale (e.g., polyamorous individuals, LGBT parents). Thus, the ability to have participants ship saliva samples to the researchers via mail is desirable. One previous study found that mailing saliva samples resulted in a small but significant decline in T for men's samples and a substantial elevation in T for women's samples (Dabbs, 1991). In addition, mailing also introduced random error into women's T measurements as evidenced by a relatively low correlation between mailed and frozen samples from the same individual. The reason that mailing affected women's and men's T differently is unknown, although Dabbs speculated that the elevation and greater introduction of error in women's T measurements could be due to women's already low levels of T. In this study, the samples were mailed unrefrigerated and spent an average of 8 days in transit.

We tested the viability of shipping *frozen* saliva samples using gel freezer packs by having three volunteers (two women, one man) provide eight pairs of samples (some volunteers provided multiple pairs of samples). Each pair of samples was provided by the same individual at the same time, who immediately froze both samples. Then, one sample of each pair (shipped sample) was delivered on dry ice to a location outside of the university and shipped back to the laboratory frozen with gel freezer packs (see Appendix for detailed packaging and shipping procedures). The other sample of each pair (control sample) remained in our lab freezer for the duration of the study. All shipped samples arrived frozen at the lab within 24–38 h, and the shipped and control samples were assayed for T via radioimmunoassay in the Core Assay Facility at the University of Michigan. Our results indicated that T remained stable during frozen storage and frozen shipping, in contrast to the previous efforts where samples were shipped at room temperature. Shipped and control samples within a pair did not significantly differ from one another, $t(7) < 1$ (Cohen's $d = 0.08$), and were highly correlated, Pearson's $r(6) = 0.90$, $p = .003$; Spearman's $\rho(6) = 0.79$, $p = .02$ (see Fig. 2). This suggests that shipping with freezer packs over a 24–38 h period preserves the integrity of saliva samples and allows researchers and participants to bypass the time requirements of traveling to pick up or drop off samples and the safety and regulatory concerns of using dry ice for studies involving salivary T. However, we certainly cannot guarantee that samples will be shipped in the promised amount of time, and time variability would certainly affect whether samples remained frozen—and, therefore, the integrity of the samples.

Health, Medical Issues, and the Body

Hormonal Contraceptives

Hormonal contraceptive (HC) use alters gonadal steroid levels markedly, since HCs themselves are synthetic estrogens and
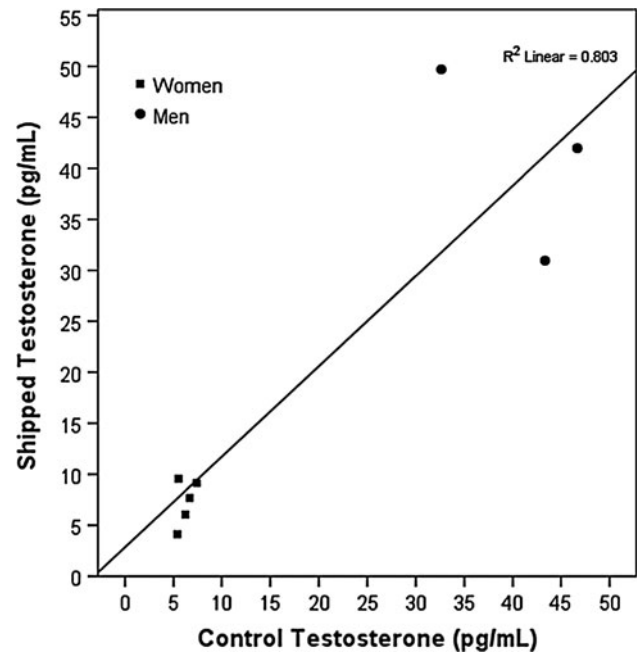

Fig. 2 Correlation between control (stored) and shipped levels of salivary testosterone from the same individual at the same time. *Note*: The five lowest data points are samples provided by two female volunteers, and the three highest data points are samples provided by a male volunteer

progesterone. Though HCs increase circulating estrogens and progesterone levels, they can decrease endogenous production of these hormones due to negative feedback loops. Accordingly, ovarian activity may be downregulated, and HC use appears to decrease T (Bancroft, Sherwin, Alexander, Davidson, & Walker, 1991; van Anders & Watson, 2006a). This effect may not be universal, as one prospective study found that HC use decreased T in some, but not all, women (Graham, Bancroft, Doll, Greco, & Tanner, 2007; Greco, Graham, Bancroft, Tanner, & Doll, 2007). Though oral contraceptives (the Pill) are the most well-known HC, there are a variety of others, including hormonal intrauterine devices (IUDs), injections, patches, etc. There is variation between HCs but also within HC type (e.g., pills can contain differing degrees of hormones on different regimens). In addition to affecting trait levels of T, there is mixed evidence of whether HCs moderate T *responsiveness*. HC users and naturally cycling women have been shown to exhibit similar T responses to both sexual and athletic activity (Edwards & O'Neal, 2009; van Anders, Hamilton, Schmidt, et al., 2007). Based on some of these data, Josephs (2009) suggested that HC use not be an exclusionary criterion and instead be analyzed as a source of variance. However, more recent research (Goldey & van Anders, 2011) suggests that HC use can lead to *opposing* T responses. Accordingly, most researchers continue to exclude HC users from research on T responsivity, or at the very least check for HC moderation of effects. However, researchers do exclude HC users from studies of trait T. In addition to hormonal

effects, there is some evidence that HCs have variable effects on sexual interest, with similar proportions of women reporting increases, decreases, or no change in sexual interest (Graham et al., 2007), which might be relevant to some sexuality studies.

Medications and Nicotine

Both nicotine and a variety of prescription or non-prescription medications can affect T. Nicotine has strong effects on T, with nicotine users having higher T than non-users (Ponholzer et al., 2005), such that many researchers either exclude nicotine users or statistically control for its use. Anabolic–androgenic steroids, treatments for polycystic ovary syndrome, and steroids (or their antagonists) used by individuals who have transitioned sex all affect circulating T levels and endogenous T production. Some medications, including those that do not affect T, may have sexual side effects. We have found that recruiting for "healthy" individuals does not lead to self-screening, as up to 27 % of our participants recruited with inclusion criteria that specified "healthy" reported using some medication. Accordingly, merely recruiting for healthy participants is not an effective strategy for studies with individuals who have unaffected levels of T. However, it would be difficult to pre-identify all possible contraindicated medications or to ask participants to complete a lengthy checklist of all medications. Some researchers use the strategy of asking participants to report medication usage and name of substance and then examining potential effects on T through drug databases. However, as the use of medications and "lifestyle" drugs continue to grow in Western nations, the inclusion of individuals using some substance may become unavoidable.

Health Conditions and Stress

As noted above, the adrenal gland releases androgens (DHEA, DHEAS) that are precursors to T. Accordingly, stress and stress-related psychological health conditions may affect T levels through adrenal activation, and these include very common conditions like anxiety, depression, eating disorders, and chronic stress (Burke, Davis, Otte, & Mohr, 2005; Hellhammer, Wust, & Kudielka, 2009). Moreover, because adrenal androgens make up a larger proportion of androgen levels in women relative to men, it is possible that women with high stress may demonstrate *elevated* androgen levels even while ovarian hormone output is suppressed (Cruess et al., 2001; Weimann, 2002). Men may show a different response to chronic stress, i.e., decreased T because chronic adrenal activation suppressed testicular output of hormones and the adrenals contribute only a low proportion of T (Aakvaag et al., 1978; Opstad & Aakvaag, 1982; Rose et al., 1969). As with medication use, the widespread prevalence of these conditions in Western nations makes the exclusion of people diagnosed with relevant conditions increasingly difficult.

In addition to stress and psychological conditions, there are physical conditions that can affect T. One of these is polycystic ovary syndrome (PCOS), which leads to higher levels of androgens in women (DeVane, Czekala, Judd, & Yen, 1975). Another of these is any medical attention to the ovaries or testes (e.g., cysts, etc.). In addition, illnesses may lead to altered gonadal output including T, as demonstrated in non-human species (e.g., Besedovsky & del Rey, 1996). Evidence from influenza vaccinations suggests that immune challenges decrease T (Simmons & Roney, 2009). Some researchers query and record the presence of psychological and physical health conditions in their samples, and either include these data in the sample description or use these data to exclude participants from analyses if the conditions are known to interfere with T.

Body Composition, Weight, and Exercise

Gonadal steroids can be synthesized from hormone precursors in fatty tissues and fat content can, therefore, affect steroid hormone levels (Deslypere, Verdonck, & Vermeulen, 1985; Nimrod & Ryan, 1974). T is negatively correlated with both weight and fat deposition in men (Fejes et al., 2006; van den Beld, de Jong, Grobbee, Pols, & Lamberts, 2000), but positively correlated with each in women (Leenen, van der Kooy, Seidell, Deurenberg, & Koppeschaar, 1994; Lukanova et al., 2004). Negative links between T and both weight and fat may occur because peripheral conversion of steroids in fatty tissue increases aromatization of T to E. The positive links between T and both weight and fat in women are likely attributable to lower levels of SHBG in women with higher fat content, and perhaps also to increased peripheral synthesis of T in fatty tissues (Leenen et al., 1994; Lukanova et al., 2004). Given links between weight and T, many researchers measure height and weight to compute body mass index (BMI), and assess its utility as a statistical control. Exercise itself affects T in ways that differ depending on activity and intensity level. Many studies find increased T post-exercise in women and men (Copeland, Consitt, & Tremblay, 2002; Kraemer et al., 1999), but very high intensity exercise (e.g., intense long-term running) can decrease T (Kuoppasalmi, Naveri, Harkonen, & Adlercreutz, 1980). Some researchers therefore control for exercise frequency and/or intensity given its mixed associations with T.

Incorporating Multiple Hormones

We have focused our review on T due to widespread interest in T among sexuality researchers; however, sexuality researchers sometimes measure other hormones such as cortisol (C), estradiol (E), and progesterone (P), often in combination with androgens (e.g., Heiman et al., 2011; van Anders et al., 2009). Unpublished data from our lab show that changes in T, E, and C in response to erotic stimuli in women are moderately to strongly correlated. C is sometimes included in sexuality research when

interactions between sexuality and stress or anxiety are of interest (e.g., Hamilton & Meston, 2011; van Anders, 2012b). The near-exclusive focus on E and P as markers of female reproductive cycles has limited the range of research topics investigated in relation to these hormones. Nonetheless, researchers have demonstrated that E responds to visual erotic stimuli in women, and that E is linked with perceptions of solitary orgasms in women but not men (van Anders & Dunn, 2009; van Anders et al., 2009). In addition, researchers have found that sexual identity moderates links between E and same-sex sexual desires, such that around the time of ovulation when E levels are at their highest, women who consistently identify as lesbian have more motivation to act on their attractions to the same sex as compared to women who identify as bisexual or who have changed their identity labels at some point (Diamond & Wallen, 2010). Sexuality research on P is mainly limited to studies examining menstrual variation in women's sexuality (e.g., Rupp et al., 2009), and research with P in men is scarce, despite an important role for P in male sexual behavior in non-human species (Wagner, 2006). However, one study demonstrated that non-sexual affiliative stimuli increased P in women *and* men (Schultheiss, Wirth, & Stanton, 2004).

Generally, the methodological issues reviewed above in relation to T are similarly important to consider when measuring other steroids such as C, E, and P. Adam and Kumari (2009) have provided a thorough review of methodological issues to consider when sampling C. However, research is mixed for E and scarce for P on some issues, specifically diurnality and seasonality (Bao et al., 2004; Bjornerem, Straume, Oian, & Berntsen, 2006; Brambilla et al., 2007; Goji, 1993). When sampling E or P in women, researchers will likely need to control for menstrual phase in some way given the large changes in E and P over the menstrual cycle (Nelson, 2005), or perhaps incorporate very large numbers of women. Sampling multiple hormones may require researchers to make some adjustments to sample collection schedules; for example, researchers studying C may be interested in the awakening response or daily slopes (O'Donnell et al., 2008), and one study showed that multiple samples per menstrual cycle over several cycles are optimal for measuring P in women, though the number of samples required may depend on the population (Jasienska & Jasienski, 2008). However, single samples for trait levels and one set of pre-post samples for state responses to experimental manipulations have also yielded meaningful results with C and E in sexuality studies (Goldey & van Anders, 2012; van Anders & Dunn, 2009; van Anders et al., 2009).

Conclusion

Clearly, there are a large number of methodological issues to consider in sexuality research design with T in humans. Some of these issues become more or less relevant depending on the question under investigation, and accordingly this review does *not* suggest or recommend that researchers attend to every possible confound and issue. Instead, this review has provided context for a variety of confounds along with some notions of how researchers currently address them methodologically, focusing especially on pragmatic and resource-efficient methods. There are a large host of issues this review did not cover, including post-collection methodological concerns. Furthermore, we did not address the relative effect sizes of the various methodological issues on T, and an examination of these issues via meta-analysis remains an area for future research (although a potentially challenging one given that few studies report effect sizes). Still, this review attempted to fill the gap in methodological reviews on T for human sexuality research and for human behavioral research in general, given that no contemporary examples exist in contrast to other hormones. Addressing relevant issues may only entail the addition of a few short questions or very minor adjustments to research design. The benefit of doing so is the increase in methodological rigor that should result in lower variation and increased ability to detect effects. Additional benefits of addressing potential confounds include the possibility of identifying under-researched or unknown effects or groups, and also engendering positive regard among participants. Attending to issues in design will help novices engage in evidence-based best practices for research with T, and strengthen the field of sexuality research with T in humans.

Appendix

Procedures for Packaging and Shipping Saliva Samples

1. $-12°$ C reusable gel freezer packs (Pelton Shepherd Industries, Stockton, CA) were used to keep the saliva samples frozen during shipping. These packs were kept frozen until samples were ready to ship.
2. Saliva samples were packaged inside Styrofoam tube inserts within a Ziploc bag, and this Ziploc bag containing saliva samples was frozen until ready to ship.
3. Just prior to shipping, the Ziploc bag containing the saliva samples was packaged with freezer packs in a Styrofoam box. The freezer packs were placed under, on top of, and surrounding all sides of the bag with saliva samples.
4. The Styrofoam box was then packaged within a cardboard box (ThermoSafe, Arlington Heights, IL) and a larger shipping box, and the package was shipped to our laboratory using 24-h express shipping.

References

Aakvaag, A., Bentdal, Ø., Quigstad, K., Walstad, P., Rønningen, H., & Fonnum, F. (1978). Testosterone and testosterone binding globulin (TeBG) in young men during prolonged stress. *International Journal of Andrology, 1*, 22–31.

Adam, E. K., & Kumari, M. (2009). Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology, 34*, 1423–1436.

Aedo, A. R., Nunez, M., Landgren, B. M., Cekan, S. Z., & Diczfalusy, E. (1977). Studies on the pattern of circulating steroids in the normal menstrual cycle: Circadian variation in the peri-ovulatory period. *Acta Endocrinologica, 84*, 320–332.

Alder, E., & Bancroft, J. (1988). The relationship between breast feeding persistence, sexuality and mood in postpartum women. *Psychological Medicine, 18*, 389–396.

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews, 30*, 319–345.

Arregger, A. L., Contreras, L. N., Tumilasci, O. R., Aquilano, D. R., & Cardoso, E. M. (2007). Salivary testosterone: A reliable approach to the diagnosis of male hypogonadism. *Clinical Endocrinology, 67*, 656–662.

Axelsson, J., Ingre, M., Akerstedt, T., & Holmback, U. (2005). Effects of acutely displaced sleep on testosterone. *Journal of Clinical Endocrinology and Metabolism, 90*, 4530–4535.

Bain, J., Langevin, R., D'Costa, M., Sanders, R. M., & Hucker, S. (1988). Serum pituitary and steroid hormone levels in the adult male: One value is as good as the mean of three. *Fertility and Sterility, 49*, 123–126.

Bancroft, J. (2002). Sexual effects of androgens in women: Some theoretical considerations. *Fertility and Sterility, 77*(Suppl. 4), 55–59.

Bancroft, J. (2005). The endocrinology of sexual arousal. *Journal of Endocrinology, 186*, 411–427.

Bancroft, J. (2009). *Human sexuality and its problems* (3rd ed.). Edinburgh: Churchill Livingstone, Elsevier.

Bancroft, J., Sherwin, B. B., Alexander, G. M., Davidson, D. W., & Walker, A. (1991). Oral contraceptives, androgens, and the sexuality of young women: II. The role of androgens. *Archives of Sexual Behavior, 20*, 121–135.

Banerjee, S., Levitz, M., & Rosenberg, C. R. (1985). On the stability of salivary progesterone under various conditions of storage. *Steroids, 46*, 967–974.

Bao, A. M., Ji, Y. F., Van Someren, E. J., Hofman, M. A., Liu, R. Y., & Zhou, J. N. (2004). Diurnal rhythms of free estradiol and cortisol during the normal menstrual cycle in women with major depression. *Hormones and Behavior, 45*, 93–102.

Beery, A. K., & Zucker, I. (2010). Sex bias in neuroscience and biomedical research. *Neuroscience and Biobehavioral Reviews, 35*, 565–572.

Berg, S. J., & Wynne-Edwards, K. E. (2001). Changes in testosterone, cortisol, and estradiol levels in men becoming fathers. *Mayo Clinic Proceedings, 76*, 582–592.

Besedovsky, H. O., & del Rey, A. (1996). Immune-neuro-endocrine interactions: Facts and hypotheses. *Endocrine Reviews, 17*, 64–102.

Bjornerem, A., Straume, B., Oian, P., & Berntsen, G. K. (2006). Seasonal variation of estradiol, follicle stimulating hormone, and dehydroepiandrosterone sulfate in women and men. *Journal of Clinical Endocrinology and Metabolism, 91*, 3798–3802.

Blanc, M. (2005). Social construction of male homosexualities in Vietnam: Some keys to understanding discrimination and implications for HIV prevention strategy. *International Social Science Journal, 57*, 661–673.

Boyar, R. M., Rosenfeld, R. S., Kapen, S., Finkelstein, J. W., Roffwarg, H. P., Weitzman, E. D., & Hellman, L. (1974). Human puberty: Simultaneous augmented secretion of luteinizing hormone and testosterone during sleep. *Journal of Clinical Investigation, 54*, 609–618.

Brambilla, D. J., Matsumoto, A. M., Araujo, A. B., & McKinlay, J. B. (2009). The effect of diurnal variation on clinical measurement of serum testosterone and other sex hormone levels in men. *Journal of Clinical Endocrinology and Metabolism, 94*, 907–913.

Brambilla, D. J., O'Donnell, A. B., Matsumoto, A. M., & McKinlay, J. B. (2007). Lack of seasonal variation in serum sex hormone levels in middle-aged to older men in the Boston area. *Journal of Clinical Endocrinology and Metabolism, 92*, 4224–4229.

Bremner, W. J., Vitiello, M. V., & Prinz, P. N. (1983). Loss of circadian rhythmicity in blood testosterone levels with aging in normal men. *Journal of Clinical Endocrinology and Metabolism, 56*, 1278–1281.

Burger, H. G., Dudley, E. C., Cui, J., Dennerstein, L., & Hopper, J. L. (2000). A prospective longitudinal study of serum testosterone, dehydroepiandrosterone sulfate, and sex hormone-binding globulin levels through the menopause transition. *Journal of Clinical Endocrinology and Metabolism, 85*, 2832–2838.

Burke, H. M., Davis, M. C., Otte, C., & Mohr, D. C. (2005). Depression and cortisol responses to psychological stress: A meta-analysis. *Psychoneuroendocrinology, 30*, 846–856.

Campbell, B. C., & Ellison, P. T. (1992). Menstrual variation in salivary testosterone among regularly cycling women. *Hormone Research, 37*, 132–136.

Campbell, B. C., Gray, P. B., & Ellison, P. T. (2006). Age-related patterns of body composition and salivary testosterone among Ariaal men of northern Kenya. *Aging Clinical and Experimental Research, 18*, 470–476.

Carré, J. M., & Putnam, S. K. (2010). Watching a previous victory produces an increase in testosterone among elite hockey players. *Psychoneuroendocrinology, 35*, 475–479.

Chivers, M. L., Bouchard, K. N., Timmers, A. D., & Haberl, M. (2012, July). *Specificity of women's sexual arousal varies with degree of same-gender attraction*. Poster session presented at the meeting of the International Academy of Sex Research, Estoril, Portugal.

Copeland, J. L., Consitt, L. A., & Tremblay, M. S. (2002). Hormonal responses to endurance and resistance exercise in females aged 19–69 years. *Journals of Gerontology Series A, 57*, B158–B165.

Cruess, D. G., Antoni, M. H., Kumar, M., McGregor, B., Alferi, S., Boyers, A. E., … Kilbourn, K. (2001). Effects of stress management on testosterone levels in women with early-stage breast cancer. *International Journal of Behavioral Medicine, 8*, 194–207.

Dabbs, J. M. (1990a). Age and seasonal variation in serum testosterone concentration among men. *Chronobiology International, 7*, 245–249.

Dabbs, J. M. (1990b). Salivary testosterone measurements: Reliability across hours, days, and weeks. *Physiology & Behavior, 48*, 83–86.

Dabbs, J. M. (1991). Salivary testosterone measurements: Collecting, storing, and mailing saliva samples. *Physiology & Behavior, 49*, 815–817.

Dabbs, J. M., Campbell, B. C., Gladue, B. A., Midgley, A. R., Navarro, M. A., Read, G. F., … Worthman, C. M. (1995). Reliability of salivary testosterone measurements: A multicenter evaluation. *Clinical Chemistry, 41*, 1581–1584.

Dabbs, J. M., & de La Rue, D. (1991). Salivary testosterone measurements among women: Relative magnitude of circadian and menstrual cycles. *Hormone Research, 35*, 182–184.

DeSantis, A. S., Adam, E. K., Doane, L. D., Mineka, S., Zinbarg, R. E., & Craske, M. G. (2007). Racial/ethnic differences in cortisol diurnal rhythms in a community sample of adolescents. *Journal of Adolescent Health, 41*, 3–13.

Deslypere, J. P., Verdonck, L., & Vermeulen, A. (1985). Fat tissue: A steroid reservoir and site of steroid metabolism. *Journal of Clinical Endocrinology and Metabolism, 61*, 564–570.

DeVane, G. W., Czekala, N. M., Judd, H. L., & Yen, S. S. (1975). Circulating gonadotropins, estrogens, and androgens in polycystic

ovarian disease. *American Journal of Obstetrics and Gynecology, 121*, 496–500.

Diamond, L. M., & Wallen, K. (2010). Sexual minority women's sexual motivation around the time of ovulation. *Archives of Sexual Behavior, 40*, 237–246.

Dixson, A. F. (2012). *Primate sexuality: Comparative studies of the prosimians, monkeys, apes, and humans* (2nd ed.). Oxford, UK: Oxford University Press.

Edelstein, R. S., Chopik, W. J., & Kean, E. L. (2011). Sociosexuality moderates the association between testosterone and relationship status in men and women. *Hormones and Behavior, 60*, 248–255.

Edwards, D. A., & O'Neal, J. L. (2009). Oral contraceptives decrease saliva testosterone but do not affect the rise in testosterone associated with athletic competition. *Hormones and Behavior, 56*, 195–198.

Ellison, P. T. (1988). Human salivary steroids: Methodological considerations and applications in physical anthropology. *American Journal of Physical Anthropology, 31*, 115–142.

Ellison, P. T. (2001). *On fertile ground*. Cambridge, MA: Harvard University Press.

Ellison, P. T., Bribiescas, R. G., Bentley, G. R., Campbell, B. C., Lipson, S. F., Panter-Brick, C., & Hill, K. (2002). Population variation in age-related decline in male salivary testosterone. *Human Reproduction, 17*, 3251–3253.

Ellison, P. T., & Panter-Brick, C. (1996). Salivary testosterone levels among Tamang and Kami males of Central Nepal. *Human Biology, 68*, 955–965.

Exton, M. S., Kruger, T. H., Bursch, N., Haake, P., Knapp, W., Schedlowski, M., & Hartmann, U. (2001). Endocrine response to masturbation-induced orgasm in healthy men following a 3-week sexual abstinence. *World Journal of Urology, 19*, 377–382.

Fejes, I., Koloszar, S., Zavaczki, Z., Daru, J., Szollosi, J., & Pal, A. (2006). Effect of body weight on testosterone/estradiol ratio in oligozoospermic patients. *Archives of Andrology, 52*, 97–102.

Feldman, H. A., Longcope, C., Derby, C. A., Johannes, C. B., Araujo, A. B., Coviello, A. D., … McKinlay, J. B. (2002). Age trends in the level of serum testosterone and other hormones in middle-aged men: Longitudinal results from the Massachusetts Male Aging Study. *Journal of Clinical Endocrinology and Metabolism, 87*, 589–598.

Ferrini, R. L., & Barrett-Connor, E. (1998). Sex hormones and age: A cross-sectional study of testosterone and estradiol and their bioavailable fractions in community-dwelling men. *American Journal of Epidemiology, 147*, 750–754.

Froyum, C. M. (2010). Making 'good girls': Sexual agency in the sexuality education of low-income black girls. *Culture, Health & Sexuality, 12*, 59–72.

Garde, A. H., Hansen, A. M., Skovgaard, L. T., & Christensen, J. M. (2000). Seasonal and biological variation of blood concentrations of total cholesterol, dehydroepiandrosterone sulfate, hemoglobin A1c, IgA, prolactin, and free testosterone in healthy women. *Clinical Chemistry, 46*, 551–559.

Garrioch, M. A. (2004). The body's response to blood loss. *Vox Sanguinis, 87*(Suppl. 1), 74–76.

Gehlert, S., Sohmer, D., Sacks, T., Mininger, C., McClintock, M., & Olopade, O. (2008). Targeting health disparities: A model linking upstream determinants to downstream interventions. *Health Affairs, 27*, 339–349.

Gettler, L. T., McDade, T. W., Feranil, A. B., & Kuzawa, C. W. (2011). Longitudinal evidence that fatherhood decreases testosterone in human males. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 16194–16199.

Gleason, E. D., Fuxjager, M. J., Oyegbile, T. O., & Marler, C. A. (2009). Testosterone release and social context: When it occurs and why. *Frontiers in Neuroendocrinology, 30*, 460–469.

Goh, V. H. H., & Tong, T. Y. Y. (2010). Sleep, sex steroid hormones, sexual activities, and aging in Asian men. *Journal of Andrology, 31*, 131–137.

Goji, K. (1993). Twenty-four-hour concentration profiles of gonadotropin and estradiol (E2) in prepubertal and early pubertal girls: The diurnal rise of E2 is opposite the nocturnal rise of gonadotropin. *Journal of Clinical Endocrinology and Metabolism, 77*, 1629–1635.

Goldey, K. L., & van Anders, S. M. (2011). Sexy thoughts: Effects of sexual cognitions on testosterone, cortisol, and arousal in women. *Hormones and Behavior, 59*, 754–764.

Goldey, K. L., & van Anders, S. M. (2012). Sexual thoughts: Links to testosterone and cortisol in men. *Archives of Sexual Behavior, 41*, 1461–1470.

Goncharov, N., Katsya, G., Dobracheva, A., Nizhnik, A., Kolesnikova, G., Herbst, V., & Westermann, J. (2006). Diagnostic significance of free salivary testosterone measurement using a direct luminescence immunoassay in healthy men and in patients with disorders of androgenic status. *Aging Male, 9*, 111–122.

Gordon, T., Bernstein, I., & Rose, R. (1978). Social and seasonal influences on testosterone secretion in the male rhesus monkey. *Physiology Behavior, 21*, 623–627.

Graham, C. A., Bancroft, J., Doll, H. A., Greco, T., & Tanner, A. (2007). Does oral contraceptive-induced reduction in free testosterone adversely affect the sexuality or mood of women? *Psychoneuroendocrinology, 32*, 246–255.

Granger, D. A., Cicchetti, D., Rogosch, F. A., Hibel, L. C., Teisl, M., & Flores, E. (2007). Blood contamination in children's saliva: Prevalence, stability, and impact on the measurement of salivary cortisol, testosterone, and dehydroepiandrosterone. *Psychoneuroendocrinology, 32*, 724–733.

Granger, D. A., Schwartz, E. B., Booth, A., & Arentz, M. (1999). Salivary testosterone determination in studies of child health and development. *Hormones and Behavior, 35*, 18–27.

Granger, D. A., Shirtcliff, E. A., Booth, A., Kivlighan, K. T., & Schwartz, E. B. (2004). The "trouble" with salivary testosterone. *Psychoneuroendocrinology, 29*, 1229–1240.

Gray, A., Berlin, J. A., McKinlay, J. B., & Longcope, C. (1991). An examination of research design effects on the association of testosterone and male aging: Results of a meta-analysis. *Journal of Clinical Epidemiology, 44*, 671–684.

Gray, P. B., & Campbell, B. C. (2009). Human male testosterone, pair-bonding, and fatherhood. In P. B. Gray & P. T. Ellison (Eds.), *Endocrinology of social relationships* (pp. 270–293). Cambridge, MA: Harvard University Press.

Gray, P. B., Campbell, B. C., Marlowe, F. W., Lipson, S. F., & Ellison, P. T. (2004). Social variables predict between-subject but not day-to-day variation in the testosterone of US men. *Psychoneuroendocrinology, 29*, 1153–1162.

Gray, P. B., Chapman, J. F., Burnham, T. C., McIntyre, M. H., Lipson, S. F., & Ellison, P. T. (2004). Human male pair bonding and testosterone. *Human Nature, 15*, 119–131.

Gray, P. B., Kahlenberg, S. M., Barrett, E. S., Lipson, S. F., & Ellison, P. T. (2002). Marriage and fatherhood are associated with lower testosterone in males. *Evolution and Human Behavior, 23*, 193–201.

Gray, P. B., Parkin, J. C., & Samms-Vaughan, M. E. (2007). Hormonal correlates of human paternal interactions: A hospital-based investigation in urban Jamaica. *Hormones and Behavior, 52*, 499–507.

Gray, P. B., Yang, C.-F. J., & Pope, H. G. (2006). Fathers have lower salivary testosterone levels than unmarried men and married non-fathers in Beijing, China. *Proceedings of the Royal Society B, 273*, 333–339.

Greco, T., Graham, C. A., Bancroft, J., Tanner, A., & Doll, H. A. (2007). The effects of oral contraceptives on androgen levels and their relevance to premenstrual mood and sexual interest: A comparison of two triphasic formulations containing norgestimate and either 35 or 25 microg of ethinyl estradiol. *Contraception, 76*, 8–17.

Greenspan, F. S., & Gardner, D. G. (Eds.). (2001). *Basic and clinical endocrinology*. New York: McGraw-Hill.

Halpern, C. T., Udry, J. R., & Suchindran, C. (1997). Testosterone predicts initiation of coitus in adolescent females. *Psychosomatic Medicine, 59*, 161–171.

Hamilton, L. D., & Meston, C. M. (2010). The effects of partner togetherness on salivary testosterone in women in long distance relationships. *Hormones and Behavior, 57*, 198–202.

Hamilton, L. D., & Meston, C. M. (2011). The role of salivary cortisol and DHEA-S in response to sexual, humorous, and anxiety-inducing stimuli. *Hormones and Behavior, 59*, 765–777.

Hamilton, L. D., van Anders, S. M., Cox, D. N., & Watson, N. V. (2009). The effect of competition on salivary testosterone in elite female athletes. *International Journal of Sports Physiology and Performance, 4*, 538–542.

Harvey, S. M. (1987). Female sexual behavior: Fluctuations during the menstrual cycle. *Journal of Psychosomatic Research, 31*, 101–110.

Heiman, J. R., Rupp, H., Janssen, E., Newhouse, S. K., Brauer, M., & Laan, E. (2011). Sexual desire, sexual arousal and hormonal differences in premenopausal US and Dutch women with and without low sexual desire. *Hormones and Behavior, 59*, 772–779.

Hellhammer, D., Wust, S., & Kudielka, B. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology, 34*, 163–171.

Herbenick, D., Reece, M., Schick, V., Sanders, S. A., Dodge, B., & Fortenberry, J. D. (2010). Sexual behavior in the United States: Results from a national probability sample of men and women ages 14–94. *Journal of Sexual Medicine, 7*(Suppl. 5), 255–265.

Hines, M. (2005). *Brain gender*. New York: Oxford University Press.

Jasienska, G., & Ellison, P. T. (2004). Energetic factors and seasonal changes in ovarian function in women from rural Poland. *American Journal of Human Biology, 16*, 563–580.

Jasienska, G., & Jasienski, M. (2008). Interpopulation, interindividual, intercycle, and intracycle natural variation in progesterone levels: A quantitative assessment and implications for population studies. *American Journal of Human Biology, 20*, 35–42.

Johnson, S. G., Jopling, G. F., & Burrin, J. M. (1987). Direct assay for testosterone in saliva: Relationship with a direct serum free testosterone assay. *International Journal of Chemistry, 163*, 309–318.

Jordan-Young, R. M. (2010). *Brain storm: The flaws in the science of sex differences*. Cambridge, MA: Harvard University Press.

Josephs, R. A. (2009). Moving beyond dichotomies in research on oral contraceptives: A comment on Edwards and O'Neal. *Hormones and Behavior, 56*, 193–194.

Kapoor, P., Luttrell, B. M., & Williams, D. (1993). The free androgen index is not valid for adult males. *Journal of Steroid Biochemistry and Molecular Biology, 45*, 325–326.

Kauppila, A., Kivela, A., Pakarinen, A., & Vakkuri, O. (1987). Inverse seasonal relationship between melatonin and ovarian activity in humans in a region with a strong seasonal contrast in luminosity. *Journal of Clinical Endocrinology and Metabolism, 65*, 823–828.

Kauppila, A., Pakarinen, A., Kirkinen, P., & Markila, U. (1987). The effect of season on the circulating concentrations of anterior pituitary, ovarian and adrenal cortex hormones and hormone binding proteins in the subarctic area: Evidence of increased activity of the pituitary-ovarian axis in spring. *Gynecological Endocrinology, 1*, 137–150.

Khan-Dawood, F. S., Choe, J. K., & Dawood, M. Y. (1984). Salivary and plasma bound and "free" testosterone in men and women. *American Journal of Obstetrics and Gynecology, 148*, 441–445.

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

Kozaki, T., Hashiguchi, N., Kaji, Y., Yasukouchi, A., & Tochihara, Y. (2009). Effects of saliva collection using cotton swab on cortisol enzyme immunoassay. *European Journal of Applied Physiology, 107*, 743–746.

Kraemer, W. J., Hakkinen, K., Newton, R. U., Nindl, B. C., Volek, J. S., McCormick, M., … Evans, W. J. (1999). Effects of heavy-resistance training on hormonal response patterns in younger vs. older men. *Journal of Applied Physiology, 87*, 982–992.

Kruger, C., Breunig, U., Biskupek-Sigwart, J., & Dorr, H. G. (1996). Problems with salivary 17-hydroxyprogesterone determinations using the Salivette device. *European Journal of Clinical Chemistry and Clinical Biochemistry, 34*, 926–929.

Kuoppasalmi, K., Naveri, H., Harkonen, M., & Adlercreutz, H. (1980). Plasma cortisol, androstenedione, testosterone and luteinizing hormone in running exercise of different intensities. *Scandinavian Journal of Clinical and Laboratory Investigation, 40*, 403–409.

Kuzawa, C. W., Gettler, L. T., Huang, Y., & McDade, T. W. (2010). Mothers have lower testosterone than non-mothers: Evidence from the Philippines. *Hormones and Behavior, 57*, 441–447.

Leenen, R., van der Kooy, K., Seidell, J. C., Deurenberg, P., & Koppeschaar, H. P. (1994). Visceral fat accumulation in relation to sex hormones in obese men and women undergoing weight loss therapy. *Journal of Clinical Endocrinology and Metabolism, 78*, 1515–1520.

Lipson, S. F., & Ellison, P. T. (1989). Development of protocols for the application of salivary steroid analysis to field conditions. *American Journal of Human Biology, 1*, 249–255.

Luboshitzky, R. (2003). Middle-aged men secrete less testosterone at night than young healthy men. *Journal of Clinical Endocrinology and Metabolism, 88*, 3160–3166.

Luisi, M., Bernini, G. P., Genovese, A. D., Birindelli, R., Barletta, D., Gasperi, M., & Franchi, F. (1980). Radioimmunoassay for "free" testosterone in human saliva. *Journal of Steroid Biochemistry, 12*, 513–516.

Lukanova, A., Lundin, E., Seleniuch-Jacquotte, A., Muti, P., Mure, A., Rinaldi, S., … Kakks, R. (2004). Body mass index, circulating levels of sex-steroid hormones, IGF-I and IGF-binding protein-3: A cross-sectional study in healthy women. *European Journal of Endocrinology, 150*, 161–171.

Lukas, W. D., Campbell, B. C., & Ellison, P. T. (2004). Testosterone, aging, and body composition in men from Harare, Zimbabwe. *American Journal of Human Biology, 16*, 704–712.

Magrini, G., Chiodoni, G., Rey, F., & Felber, J. P. (1986). Further evidence for the usefulness of the salivary testosterone radioimmunoassay in the assessment of androgenicity in man in basal and stimulated conditions. *Hormone Research, 23*, 65–73.

Martikainen, H., Tapanainen, J., Vakkuri, O., Leppaluoto, J., & Huhtaniemi, I. (1985). Circannual concentrations of melatonin, gonadotrophins, prolactin and gonadal steroids in males in a geographical area with a large annual variation in daylight. *Acta Endocrinologica, 109*, 446–450.

Matteo, S., & Rissman, E. F. (1984). Increased sexual activity during the midcycle portion of the human menstrual cycle. *Hormones and Behavior, 18*, 249–255.

Mehta, P. H., & Josephs, R. A. (2006). Testosterone change after losing predicts the decision to compete again. *Hormones and Behavior, 50*, 684–692.

Michael, R. P., Richter, M. C., Cain, J. A., Zumpe, D., & Bonsall, R. W. (1978). Artificial menstrual cycles, behaviour and the role of androgens in female rhesus monkeys. *Nature, 275*, 439–440.

Moffat, S. D., & Hampson, E. (2000). Salivary testosterone concentrations in left-handers: An association with cerebral language lateralization? *Neuropsychology, 14*, 71–81.

Morley, J. E., Kaiser, F. E., Perry, H. M., Patrick, P., Morley, P. M. K., Stauber, P. M., … Garry, P. J. (1997). Longitudinal changes in testosterone, luteinizing hormone, and follicle-stimulating hormone in healthy older men. *Metabolism, 46*, 410–413.

Morley, J. E., Perry, H. M., I. I. I., Patrick, P., Dollbaum, C. M., & Kells, J. M. (2006). Validation of salivary testosterone as a screening test for male hypogonadism. *The Aging Male, 9*, 165–169.

Morris, P. D., Malkin, C. J., Channer, K. S., & Jones, T. H. (2004). A mathematical comparison of techniques to predict biologically available testosterone in a cohort of 1072 men. *European Journal of Endocrinology, 151*, 241–249.

Muller, M. N., & Wrangham, R. W. (2004). Dominance, aggression and testosterone in wild chimpanzees: A test of the 'challenge hypothesis'. *Animal Behaviour, 67*, 113–123.

Muroyama, Y., Shimizu, K., & Sugiura, H. (2007). Seasonal variation in fecal testosterone levels in free-ranging male Japanese macaques. *American Journal of Primatology, 69*, 603–610.

Nahoul, K., & Roger, M. (1990). Age-related decline of plasma bioavailable testosterone in adult men. *Journal of Steroid Biochemistry, 35*, 293–299.

Nelson, R. J. (2005). *An introduction to behavioral endocrinology*. Sunderland, MA: Sinauer Associates.

Nelson, R. J., Denlinger, D. L., & Somers, D. E. (Eds.). (2009). *Photoperiodism: The biological calendar*. New York: Oxford University Press.

Nicolau, G. Y., Haus, E., Lakatua, D. J., Bogdan, C., Sackett-Lundeen, L., Popescu, M., … Robu, E. (1985). Circadian and circannual variations of FSH, LH, testosterone, dehydroepiandrosterone-sulfate (DHEA-S) and 17-hydroxy progesterone (17 OH-prog) in elderly men and women. *Endocrinologie, 23*, 223–246.

Nimrod, A., & Ryan, K. J. (1974). Aromatization of androgens by human fat tissue. *Journal of Steroid Biochemistry, 5*, 317.

O'Donnell, K., Badrick, E., Kumari, M., & Steptoe, A. (2008). Psychological coping styles and cortisol over the day in healthy older adults. *Psychoneuroendocrinology, 33*, 601–611.

Opstad, P. K., & Aakvaag, A. (1982). Decreased serum levels of oestradiol, testosterone and prolactin during prolonged physical strain and sleep deprivation, and the influence of a high calorie diet. *European Journal of Applied Physiology, 49*, 343–348.

Oudshoorn, N. (1994). *Beyond the natural body: An archaeology of sex hormones*. Abingdon, Oxon: Routledge.

Panico, S., Pisani, P., Muti, P., Recchione, C., Cavalleri, A., Totis, A., & Berrino, F. (1990). Diurnal variation of testosterone and estradiol: A source of bias in comparative studies on breast cancer. *Journal of Endocrinological Investigation, 12*, 423–426.

Paton, C. D., Lowe, T., & Irvine, A. (2010). Caffeinated chewing gum increases repeated sprint performance and augments increases in testosterone in competitive cyclists. *European Journal of Applied Physiology, 110*, 1243–1250.

Perry, H. M., Miller, D. K., Patrick, P., & Morley, J. E. (2000). Testosterone and leptin in older African-American men: Relation ship to age, strength, function, and season. *Metabolism, 49*, 1085–1091.

Phillips, G. II, Outlaw, A. Y., Hightow-Weidman, L., Jones, K. C., Wohl, A. R., Futterman, D., … Hidalgo, J. (2011). Sexual behaviors of racial/ethnic minority young men who have sex with men. *AIDS Patient Care and STDs, 25*, S47–S53.

Piro, C., Fraioli, F., Sciarra, F., & Conti, C. (1973). Circadian rhythm of plasma testosterone, cortisol and gonadotropins in normal male subjects. *Journal of Steroid Biochemistry, 4*, 312–329.

Plymate, S. R., Tenover, J. S., & Bremner, W. J. (1989). Circadian variation in testosterone, sex hormone-binding globulin, and calculated non-sex hormone-binding globulin bound testosterone in healthy young and elderly men. *Journal of Andrology, 10*, 366–371.

Ponholzer, A., Plas, E., Schatzl, G., Struhal, G., Brossner, C., Mock, K., … Madersbacher, S. (2005). Relationship between testosterone serum levels and lifestyle in aging men. *The Aging Male, 8*, 190–193.

Quissell, D. O. (1993). Steroid hormone analysis in human saliva. *Annals of the New York Academy of Sciences, 694*, 143–145.

Redoute, J., Stoleru, S., Gregoire, M. C., Costes, N., Cinotti, L., Lavenne, F., … Pujol, J. F. (2000). Brain processing of visual sexual stimuli in human males. *Human Brain Mapping, 11*, 162–177.

Reinberg, A., Lagoguey, M., Cesselin, F., Touitou, Y., Legrand, J. C., Delassalle, A., … Lagoguey, A. (1978). Circadian and circannual rhythms in plasma hormones and other variables of five healthy young human males. *Acta Endocrinologica, 88*, 417–427.

Reinberg, A., Lagoguey, M., Chauffournier, J. M., & Cesselin, F. (1975). Circannual and circadian rhythms in plasma testosterone in five healthy young Parisian males. *Acta Endocrinologica, 80*, 732–743.

Reinberg, A., Smolensky, M. H., Hallek, M., Smith, K. D., & Steinberger, E. (1988). Annual variation in semen characteristics and plasma hormone levels in men undergoing vasectomy. *Fertility and Sterility, 49*, 309–315.

Roney, J. R., Lukaszewski, A. W., & Simmons, Z. L. (2007). Rapid endocrine responses of young men to social interactions with young women. *Hormones and Behavior, 52*, 326–333.

Roney, J. R., Mahler, S. V., & Maestripieri, D. (2003). Behavioral and hormonal responses of men to brief interactions with women. *Evolution and Human Behavior, 24*, 365–375.

Rose, R. M., Bourne, P. G., Poe, R. O., Mougey, E. H., Collins, D. R., & Mason, J. W. (1969). Androgen responses to stress. *Psychosomatic Medicine, 31*, 418–436.

Rosner, W., Auchus, R. J., Azziz, R., Sluss, P. M., & Raff, H. (2007). Utility, limitations, and pitfalls in measuring testosterone: An endocrine society position statement. *Journal of Clinical Endocrinology and Metabolism, 92*, 405–413.

Rowe, P. H., Lincoln, G. A., Racey, P. A., Lehane, J., Stephenson, M. J., Shenton, J. C., & Glover, T. D. (1974). Temporal variations of testosterone levels in the peripheral blood plasma of men. *Journal of Endocrinology, 61*, 63–73.

Rupp, H. A., James, T. W., Ketterson, E. D., Sengelaub, D. R., Janssen, E., & Heiman, J. R. (2009). Neural activation in the orbitofrontal cortex in response to male faces increases during the follicular phase. *Hormones and Behavior, 56*, 66–72.

Sagarin, B. J., Cutler, B., Cutler, N., Lawler-Sagarin, K. A., & Matuszewich, L. (2009). Hormonal changes and couple bonding in consensual sadomasochistic activity. *Archives of Sexual Behavior, 38*, 186–200.

Sanders, S. A., & Reinisch, J. M. (1999). Would you say you "had sex" if…? *Journal of the American Medical Association, 281*, 275–277.

Schultheiss, O. C., Wirth, M. M., & Stanton, S. J. (2004). Effects of affiliation and power motivation arousal on salivary progesterone and testosterone. *Hormones and Behavior, 46*, 592–599.

Schultheiss, O. C., Wirth, M. M., Torges, C. M., Pang, J. S., Villacorta, M. A., & Welsh, K. M. (2005). Effects of implicit power motivation on men's and women's implicit learning and testosterone changes after social victory or defeat. *Journal of Personality and Social Psychology, 88*, 174–188.

Sherwin, B. B. (1988). A comparative analysis of the role of androgen in human male and female sexual behavior: Behavioral specificity, critical thresholds, and sensitivity. *Psychobiology, 16*, 416–425.

Shirtcliff, E. A., Granger, D. A., & Likos, A. (2002). Gender differences in the validity of testosterone measured in saliva by immunoassay. *Hormones and Behavior, 42*, 62–69.

Shirtcliff, E. A., Granger, D. A., Schwartz, E., & Curran, M. J. (2001). Use of salivary biomarkers in biobehavioral research: Cotton-based sample collection methods can interfere with salivary immunoassay results. *Psychoneuroendocrinology, 26*, 165–173.

Shirtcliff, E. A., Granger, D. A., Schwartz, E. B., Curran, M. J., Booth, A., & Overman, W. H. (2000). Assessing estradiol in biobehavioral studies using saliva and blood spots: Simple radioimmunoassay protocols, reliability, and comparative validity. *Hormones and Behavior, 38*, 137–147.

Simmons, Z. L., & Roney, J. R. (2009). Androgens and energy allocation: Quasi-experimental evidence for effects of influenza vaccination on men's testosterone. *American Journal of Human Biology, 21*, 133–135.

Singh, D., Vidaurri, M., Zambarano, R. J., & Dabbs, J. M., Jr. (1999). Lesbian erotic role identification: Behavioral, morphological, and hormonal correlates. *Journal of Personality and Social Psychology, 76*, 1035–1049.

Smals, A. G. H., Kloppenborg, P. W. C., & Benraad, T. J. (1976). Circannual cycle in plasma testosterone levels in man. *Journal of Clinical Endocrinology Metabolism, 42*, 979–982.

Stanton, S. J., Mullette-Gillman, O. A., & Huettel, S. A. (2011). Seasonal variation of salivary testosterone in men, normally cycling women, and women using hormonal contraceptives. *Physiology & Behavior, 104*, 804–808.

Stoleru, S., Ennaji, A., Cournot, A., & Spira, A. (1993). LH pulsatile secretion and testosterone blood levels are influenced by sexual arousal in human males. *Psychoneuroendocrinology, 18*, 205–218.

Stoleru, S., Gregoire, M. C., Gerard, D., Decety, J., Lafarge, E., Cinotti, L., … Comar, D. (1999). Neuroanatomical correlates of visually evoked sexual arousal in human males. *Archives of Sexual Behavior, 28*, 1–21.

Storey, A. E., Walsh, C. J., Quinton, R. L., & Wynne-Edwards, K. E. (2000). Hormonal correlates of paternal responsiveness in new and expectant fathers. *Evolution and Human Behavior, 21*, 79–95.

Svartberg, J., Jorde, R., Sundsfjord, J., Bonaa, K. H., & Barrett-Connor, E. (2003). Seasonal variation of testosterone and waist to hip ratio in men: The Tromso study. *Journal of Clinical Endocrinology Metabolism, 88*, 3099–3104.

Swinkels, L. M., Meulenberg, P. M., Ross, H. A., & Benraad, T. J. (1988). Salivary and plasma free testosterone and androstenedione levels in women using oral contraceptives containing desogestrel or levonorgestrel. *Annals of Clinical Biochemistry, 25*, 354–359.

Taieb, J., Mathian, B., Millot, F., Patricot, M. C., Mathieu, E., Queyrel, N., … Boudou, P. (2003). Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women, and children. *Clinical Chemistry, 49*, 1381–1395.

Uchida, A., Bribiescas, R. G., Ellison, P. T., Kanamori, M., Ando, J., Hirose, N., & Ono, Y. (2006). Age related variation of salivary testosterone values in healthy Japanese males. *The Aging Male, 9*, 207–213.

Valero-Politi, J., & Fuentes-Arderiu, X. (1998). Annual rhythmic variations of follitropin, lutropin, testosterone and sex-hormone-binding globulin in men. *Clinica Chimica Acta, 271*, 57–71.

van Anders, S. M. (2009). Androgens and diversity in adult human partnering. In P. B. Gray & P. T. Ellison (Eds.), *Endocrinology of social relationships* (pp. 340–363). Cambridge, MA: Harvard University Press.

van Anders, S. M. (2010a). Chewing gum has large effects on salivary testosterone, estradiol, and secretory immunoglobulin A assays in women and men. *Psychoneuroendocrinology, 35*, 305–309.

van Anders, S. M. (2010b). Social modulation of hormones. In E. R. Smith (Ed.), *The mind in context* (pp. 65–80). New York: Guilford Press.

van Anders, S. M. (2012a). From one bioscientist to another: Guidelines for researching and writing about bisexuality for the lab and biosciences. *Journal of Bisexuality, 12*, 393–403.

van Anders, S. M. (2012b). Testosterone and sexual desire in healthy women and men. *Archives of Sexual Behavior, 41*, 1471–1484.

van Anders, S. M., Brotto, L., Farrell, J., & Yule, M. (2009). Associations among physiological and subjective sexual response, sexual desire, and salivary steroid hormones in healthy premenopausal women. *Journal of Sexual Medicine, 6*, 739–751.

van Anders, S. M., Chernick, A. B., Chernick, B. A., Hampson, E., & Fisher, W. A. (2005). Preliminary clinical experience with androgen administration for pre- and postmenopausal women with hypoactive sexual desire. *Journal of Sex and Marital Therapy, 31*, 173–185.

van Anders, S. M., & Dunn, E. J. (2009). Are gonadal steroids linked with orgasm perceptions and sexual assertiveness in women and men? *Hormones and Behavior, 56*, 206–213.

van Anders, S. M., & Goldey, K. L. (2010). Testosterone and partnering are linked via relationship status for women and 'relationship orientation' for men. *Hormones and Behavior, 58*, 820–826.

van Anders, S. M., Goldey, K. L., & Kuo, P. X. (2011). The steroid/peptide theory of social bonds: Integrating testosterone and peptide responses for classifying social behavioral contexts. *Psychoneuroendocrinology, 36*, 1265–1275.

van Anders, S. M., & Gray, P. B. (2007). Hormones and human partnering. *Annual Review of Sex Research, 18*, 60–93.

van Anders, S. M., Hamilton, L. D., Schmidt, N., & Watson, N. V. (2007). Associations between testosterone secretion and sexual activity in women. *Hormones and Behavior, 51*, 477–482.

van Anders, S. M., Hamilton, L. D., & Watson, N. V. (2007). Multiple partners are associated with higher testosterone in North American men and women. *Hormones and Behavior, 51*, 454–459.

van Anders, S. M., & Hampson, E. (2005). Waist-to-hip ratio is positively associated with bioavailable testosterone but negatively associated with sexual desire in healthy premenopausal women. *Psychosomatic Medicine, 67*, 246–250.

van Anders, S. M., Hampson, E., & Watson, N. (2006). Seasonality, waist-to-hip ratio, and salivary testosterone. *Psychoneuroendocrinology, 31*, 895–899.

van Anders, S. M., & Watson, N. V. (2006a). Menstrual cycle irregularities are associated with testosterone levels in healthy premenopausal women. *American Journal of Human Biology, 18*, 841–844.

van Anders, S. M., & Watson, N. V. (2006b). Relationship status and testosterone in North American heterosexual and non-heterosexual men and women: Cross-sectional and longitudinal data. *Psychoneuroendocrinology, 31*, 715–723.

van Anders, S. M., & Watson, N. V. (2006c). Social neuroendocrinology: Effects of social contexts and behaviors on sex steroids in humans. *Human Nature, 17*, 212–237.

van Anders, S. M., & Watson, N. V. (2007a). Effects of ability- and chance-determined competition outcome on testosterone. *Physiology & Behavior, 90*, 634–642.

van Anders, S. M., & Watson, N. V. (2007b). Testosterone levels in women and men who are single, in long-distance relationships, or same-city relationships. *Hormones and Behavior, 51*, 286–291.

van den Beld, A. W., de Jong, F. H., Grobbee, D. E., Pols, H. A. P., & Lamberts, S. W. J. (2000). Measures of bioavailable serum testosterone and estradiol and their relationships with muscle strength, bone density, and body composition in elderly men. *Journal of Clinical Endocrinology and Metabolism, 85*, 3276–3282.

Vermeulen, A. (1980). Sex hormone status of the postmenopausal woman. *Maturitas, 2*, 81–89.

Vitzthum, V. J. (2009). The ecology and evolutionary endocrinology of reproduction in the human female. *American Journal of Physical Anthropology, 140*(Suppl. 49), 95–136.

Vitzthum, V. J., Worthman, C. M., Beall, C. M., Thornburg, J., Vargas, E., Villena, M., … Spielvogel, H. (2009). Seasonal and circadian variation in salivary testosterone in rural Bolivian men. *American Journal of Human Biology, 21*, 762–768.

Wagner, C. K. (2006). The many faces of progesterone: A role in adult and developing male brain. *Frontiers in Neuroendocrinology, 27*, 340–359.

Walker, R. F., Wilson, D. W., Read, G. F., & Riad-Fahmy, D. (1980). Assessment of testicular function by the radioimmunoassay of testosterone in saliva. *International Journal of Andrology, 3*, 105–120.

Wang, C., Plymate, S., Nieschlag, E., & Paulsen, C. A. (1981). Salivary testosterone in men: Further evidence of a direct correlation with free serum testosterone. *Journal of Clinical Endocrinology and Metabolism, 53*, 1021–1024.

Weimann, E. (2002). Gender-related differences in elite gymnasts: The female athlete triad. *Journal of Applied Physiology, 92*, 2146–2152.

Welling, L. L. M., Jones, B. C., DeBruine, L. M., Conway, C. A., Law Smith, M. J., Little, A. C., … Al-Dujaili, E. A. S. (2007). Raised salivary testosterone in women is associated with increased attraction to masculine faces. *Hormones and Behavior, 52*, 156–161.

Wisniewski, A. B., & Nelson, R. J. (2000). Seasonal variation in human functional cerebral lateralization and free testosterone concentrations. *Brain and Cognition, 43*, 429–438.

Worthman, C. M., & Konner, M. J. (1987). Testosterone levels change with subsistence hunting effort in !Kung San men. *Psychoneuroendocrinology, 12*, 449–458.

Wynne-Edwards, K. E. (2001). Hormonal changes in mammalian fathers. *Hormones and Behavior, 40*, 139–145.

Zumoff, B., Strain, G. W., Miller, L. K., & Rosner, W. (1995). Twenty-four-hour mean plasma testosterone concentration declines with age in normal premenopausal women. *Journal of Clinical Endocrinology and Metabolism, 80*, 1429–1430.