Darwin's Legacy: An Evolutionary View of Women's Reproductive and Sexual Functioning

Amy L. Harris [a] & Virginia J. Vitzthum [a] [b]

[a] Anthropology Department , Indiana University

[b] The Kinsey Institute for Research in Sex, Gender, and Reproduction , Indiana University
Published online: 12 Mar 2013.

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

ARSR REVIEW ARTICLES

Darwin's Legacy: An Evolutionary View of Women's Reproductive and Sexual Functioning

Amy L. Harris
Anthropology Department, Indiana University

Virginia J. Vitzthum
Anthropology Department, and The Kinsey Institute for Research in Sex, Gender, and Reproduction, Indiana University

On the Origin of Species, *published just over 150 years ago, has deeply influenced thinking in both scientific and wider communities. Darwin's legacy includes recognition of the fact that all organisms evolve; that variation within and between species is natural and normal; and that an evolutionary approach to understanding the sources and consequences of this variation comprises theoretical frameworks, testable hypotheses, and rigorously collected evidence. With an eye toward facilitating communication and productive collaboration among researchers from different intellectual traditions who nonetheless share a common interest in women's reproductive and sexual functioning, we discuss evolutionary concepts and models, summarize the known variability in ovarian functioning and consider the implications of this variability for conducting sex research, and evaluate the relative merits of various biomarkers that serve as proxy measurements of a woman's reproductive and hormonal status. With these perspectives and methods from reproductive ecology at hand, we examine several contentious issues: the links between hormones and sexuality in premenopausal and perimenopausal women, the causes of premenstrual syndrome, and the existence (or not) of menstrual synchrony. In none of these cases is as much known as is often claimed. In each, there are abundant opportunities for innovative, albeit challenging, research.*

Nothing in biology makes sense except in the light of evolution.

—Theodosius Dobzhansky (1973)

Worldwide celebrations in 2009 marking the 200th birthday of Charles Darwin and the 150th anniversary of the publication of *On the Origin of Species* rightly acknowledged the transformative impact of Darwin's work on a host of scholarly fields including, of course, biology. Nonetheless, the use of an evolutionary lens to understand variation in women's reproductive functioning is a relatively recent development, arguably beginning only about 40 years ago. Happily, human reproductive ecology has bloomed rapidly despite its late start, and nowadays scholars from the various disciplines that study human reproduction have begun to incorporate an evolutionary perspective and the methods of reproductive ecology into their work.

Darwin's legacy is not only a widespread recognition of the fact that organisms, including humans, have evolved. Darwin spent decades amassing and analyzing the evidence that gave credence to evolutionary theory, and many hold him in high regard as much for his empirical contributions as for his ideas. To avoid the trap of being nothing more than "wind sauce and air pudding," an evolutionary perspective on any aspect of an organism's biology and behavior must be grounded in both an explicit theoretical framework *and* rigorously collected empirical data. Sex researchers have also long held comparable views regarding the centrality of empiricism in their field (witness The Kinsey Institute and other centers). In the practice of science, data trumps theory (i.e., hypotheses must be tested with evidence). Hence, with an eye toward facilitating a productive exchange between sex researchers and reproductive ecologists, in this review we examine both the theoretical arguments and the currently available empirical evidence that inform an evolutionary view of women's reproductive and sexual functioning.

We begin with a brief history of the different scientific approaches to human reproduction and discuss the Flexible Response Model (FRM), an evolutionary model that helps to explain variation in women's reproductive functioning. Because it is often taken for granted that some specific aspect or another of reproductive functioning is "adaptive," we also note the strengths and limitations of different concepts of adaptation. A third signature feature of Darwin's legacy is the recognition that variation among individuals is natural rather than aberrant. This evolutionary view contrasts starkly with platonic ideals and biomedical concepts of "normal" and "pathological." In keeping with this Darwinian insight, we next examine the substantial variability in ovarian functioning now known to occur temporally (i.e., from cycle to cycle within an individual), among women, and among populations, and consider the implications of this variability for conducting sex research. Within this section we also evaluate the relative merits of various biomarkers that serve as proxy measurements of a woman's reproductive and hormonal status. In the final section we apply perspectives and methods from reproductive ecology to help shed light on several contentious issues: the links between hormones and sexuality in premenopausal and perimenopausal women, the causes of premenstrual syndrome, and the existence (or not) of menstrual synchrony.

The breadth of sex research makes it impossible to be comprehensive in this single work, and not all questions within sex research are necessarily amenable to evolutionary approaches. As well, space and time precluded drawing more extensively than we did from the rich literature on nonhuman primates. Nor have we undertaken an in-depth critique of related work in evolutionary psychology, although we do draw attention to those methods and data presented herein that are directly relevant to

critically evaluating published literature and informing future studies in that field. Nonetheless, despite these omissions, with this contribution we hope to build bridges that facilitate communication and productive research among those researchers focused primarily on sexual functioning and behavior and those focused principally on understanding human reproductive functioning and behavior within an evolutionary framework.

Evolutionary Models of Reproductive Functioning

Most textbooks discussing the history of evolutionary theory note that Thomas Robert Malthus's "Essay on the Principle of Population," written in 1798, was a significant factor shaping Charles Darwin's thinking on the relationship between species and their environments. Much less recognized, however, is the impact of Malthus on conceptualizations of human reproductive functioning during the subsequent two centuries. In his view, the reproductive system was much like an unflagging machine without any internal controls or capacity to respond to changing conditions. Population growth was held in check only by external factors (famine, disease, war, and homicide), old age, and conscious restraint. This view continues to underpin contemporary assumptions of healthy ovarian functioning (e.g., most premenopausal women "should have" regular, 28-day, ovulatory menstrual cycles). Too often it is presumed that if a woman's reproductive machinery deviates from the relentless productivity envisioned by Malthus, she must be malfunctioning. This pathology framework underpins the medicalization of women's normal, if variable, reproductive functioning (Ginsburg & Rapp, 1991; Lock, 2001; Meyer, 2001; Santos, 1997).

Challenges from evolutionary scientists to this Malthusian model began about 40 years ago. Rose Frisch's hypothesis (Frisch, 1978; Frisch & McArthur, 1974; Frisch & Revelle, 1970, 1971a, 1971b) that a critical fat threshold must be exceeded to achieve normal ovarian functioning was pivotal in sparking wide debate on the ecology of human reproduction. Although the details of her model were criticized by demographers (Bongaarts, 1980; Menken, Trussell, & Watkins, 1981; Trussell, 1980) and human biologists (Cameron, 1976; Malina, 1983; Quandt, 1984; Scott & Johnston, 1985), Frisch's core argument regarding the importance of energy intake and expenditure in human reproductive functioning is the foundation of energetics models of ovarian functioning. In 1976, "The Evolution of Human Reproduction" by physiologist Roger Short squarely placed human fertility patterns within an evolutionary framework.

Since then, bioanthropologists and human biologists have contributed to a rapidly growing body of evidence on the extent and causes of natural (nonpathological) variation in human reproductive *functioning* (e.g., Bribiescas,

2001; Ellison, Peacock, & Lager, 1989; Holman, 1996; Konner & Worthman, 1980; Leslie & Fry, 1989; Nepomnaschy, Welch, McConnell, Strassmann, & England, 2004; Strassmann, 1997; Vitzthum, 1989, 2008a; Vitzthum, Worthman, et al., 2009; Wood, Johnson, & Campbell, 1985). Concurrently, and somewhat independently, other anthropologists have focused on the application of "evolutionary ecology models and concepts to the study of human behavioral diversity" (Winterhalder & Smith, 2000, p. 51), including modeling and measuring variation in reproductive *behavior* (i.e., mating strategies and parental investment) and associated outcomes (e.g., reproductive success). Although intellectually vigorous and highly productive, behavioral ecologists are less concerned with elucidating biological mechanisms per se, rather assuming their necessary existence.

Over time, human reproductive ecology (HRE) has coalesced into a field of inquiry that has adopted, if not quite integrated, intellectual contributions from biological anthropology, behavioral ecology, demography, physiology, medicine, and evolutionary biology. Reflecting this diversity, there are several definitions of HRE (e.g., Campbell & Wood, 1994; Ellison, 2001; Hill & Hurtado, 1996; Leslie & Little, 2003; Morbeck, Galloway, & Zihlman, 1997; Winterhalder & Smith, 2000), some of which emphasize a concern with how proximate mechanisms cause variation in fertility and others which stress the importance of evolutionary mechanisms for understanding why these proximate mechanisms exist.

For the purpose of the present discussion, we define human reproductive ecology as the study of the mechanisms that link variation in local habitats with variation in reproductive traits (Vitzthum, 2008b, 2009). *Habitats* comprise the physical, biological, and social conditions that an individual must accommodate and exploit for survival and reproduction. We emphasize local habitats because these are the contexts within which natural selection acts on individuals, and because human biological uniformitarianism across these varied habitats is a demonstrably incorrect assumption (Leslie & Little, 2003).

Mechanisms subsume both physiological and behavioral processes that modulate reproductive traits. These two sets of mechanisms have rarely been studied in tandem, in part because of the technical challenges of measuring biomarkers outside of a clinical setting. Nonetheless, reproductive behavior, as a manifestation of the neurohormonal mechanics of the human brain (which may be as much or more a product of environmental input as of genetic *pre-wiring*) is necessarily physiological, at least in part. Moreover, reproductive behaviors exert much of their influence on fertility by modulating more proximate physiological determinants. Unlike others, this definition stresses the study of mechanisms over that of relationships to encourage a more specific examination of *how* differing habitats generate variation in reproductive traits. Theoretical frameworks, elegant models, and sophisticated statistics are powerful tools for generating and testing hypotheses and for quantifying associations among variables. But if the biology doesn't work that way, then it's the model that must be adjusted.

Reproductive traits are the suite of physiological, morphological, developmental, and behavioral phenotypes (i.e., traits) concerned with the production and nurturance of offspring. Collectively, these traits determine the magnitude and temporal patterning of an individual's lifetime reproductive investment. Human reproductive phenotypes that have been investigated within one or another explanatory framework include age and size at puberty, at first live birth, and at peak reproductive maturity; mating strategies including sexual behaviors; number, size, quality, spacing, and sex ratio of offspring; age-specific fecundability (the monthly probability of conception); probability of pregnancy loss (specific to parental age and conceptus age); offspring provisioning including lactation and age at weaning; and age at menopause. This listing is not exhaustive but rather exemplifies the foci of the several theoretical models used by reproductive ecologists to investigate human reproduction.

Proximate and evolutionary explanations are complementary, rather than alternative, approaches to understanding variation in human reproductive patterns. Some mechanisms that translate habitat variation into reproduction variation may operate as they do because of evolutionary processes (e.g., natural selection, genetic drift, phylogenetic constraints), others because of cultural innovations (e.g., contraception), and others because of novel changes in a habitat (e.g., introduction of endocrine disruptors). In addition, not all evolutionary processes generate adaptations. For example, genetic drift is a stochastic process that can lead to a high frequency of nonadaptive (even maladaptive) phenotypes in a population and has its greatest impact in species that produce few offspring per parental couple, such as humans and other primates.

Concepts of Adaptation

. . . adaptation is a special and onerous concept that should only be used where it is really necessary.

—George C. Williams (1966)

Debates regarding the nature of evolutionary adaptation and how to recognize it have a long and sometimes acrimonious history (Rose & Lauder, 1996; Vitzthum, 2008b). Many contemporary evolutionary scientists have adopted the position that adaptations are only those features that have evolved through selection on heritable phenotypes for a specific function as the result of an associated increase in fitness (Gould & Lewontin, 1979; Gould & Vrba, 1982; Williams, 1966). These are recognizable as well-engineered solutions for specific challenges to survival and reproduction (i.e., the "argument from design") (Lauder, 1996; Williams,

209

1966). This restricted concept of adaptation suffers from the sometimes insurmountable difficulties of acquiring the data necessary to assess a trait by these criteria (Leroi, Rose, & Lauder, 1994) and from its idealized notion of evolutionary processes. Natural selection is more like a tinkerer than an engineer (Jacob, 1977). Any single morphological, physiological, or behavioral component of an organism may serve multiple purposes, and many features are not readily perceivable by an investigator as having been shaped by selection to meet a specific challenge.

Another approach, often used in studies of behavioral variation, considers a phenotypic variant to be adaptive if its current possessor has a higher fitness compared with some other variant in the same population (Caro & Borgerhoff Mulder, 1987; Clutton-Brock & Harvey, 1979; Fisher, 1985). Defining adaptation by its effects (on fitness) rather than its causes (selection among heritable variants) avoids faulty perceptions of design and the misleading notion of "goal" implicit in an argument from design, dispels speculations about unknown genetic influences and past selection pressures, and specifies a clear (albeit difficult to measure) criterion for testing hypotheses. However, this approach to identifying an adaptation usually begs the question of whether the phenotype under study is heritable, and hence subject to natural selection, and ignores rather than solves the difficulties encountered in evaluating the evolutionary origins and maintenance of a character (Vitzthum, 2008b).

Many human biologists have adopted an ecological concept of adaptation, predicated on the assumption that selection has favored the phenotypes associated with beneficial responses to environmental challenges and emphasizing an individual's ability to "surmount the challenges to life" (Lasker, 1969; Mazess, 1975; Thomas, Gage, & Little, 1989). In this view, any biological or behavioral response that affords a beneficial adjustment to physical or social conditions is considered adaptive, even if it is not known to be the direct consequence of natural selection on a heritable phenotype. Thus, in addition to genetic change, physiological acclimatization, learning, and other mechanisms of adjustment are all considered adaptive mechanisms. Adaptations can be identified by measuring fitness or its presumed proxies (health, energy efficiency, longevity) (Thomas et al., 1989; Wiley, 1992).

However, natural selection need not enhance individual well-being or congruity with the environment. For example, despite the fact that the allele for sickle cell hemoglobin is lethal when homozygous (barring biomedical intervention), it can reach high frequencies in populations with endemic malaria because of heterozygote advantage (Livingstone, 1958). In this case, natural selection inexorably operates to maximize mean (not individual) fitness. The pool of heterozygous parents cannot prevent the production of homozygous offspring of lower fitness (referred to as genetic load; Rose & Lauder, 1996).

Although the sickle-cell adaptation is wasteful (lost offspring), the inefficiency is an inevitable outcome of the interaction between this particular genetic variant and a specific environment. Such a genetic adaptation to malaria is clearly *not* the manifestation of any engineered design, nor is it energetically efficient.

The assumption that selection yields an optimal solution to some challenge can promote the mistaken expectation that a genuinely adaptive response is efficient and not detrimental to the individual. But selection favors those phenotypes with the highest total lifetime reproductive success (LRS) in a specific population, even if this greater fitness comes at substantial cost to individuals. For this and other reasons, there is merit in distinguishing adaptations (based on fitness) from individual adaptability (responsiveness to challenge) (Vitzthum, 2008b). Yet while there is no necessary relationship between reproductive fitness and measures of individual well-being or efficiency (Voland, 1998), neither is there any reason to assume that such a relationship is rare (Wiley, 1992). Rather, life history theory (discussed in the next section) provides a framework for elucidating the fitness benefits and costs of allocating resources to self-maintenance over the course of a lifetime. With regard to individual well-being, studies of adaptation and adaptability can yield valuable insights into the patterns of morbidity, growth, fertility, and mortality in human populations across different environments (for examples, see Chisholm, 1993; Chisholm & Coall, 1998; Geronimus, Bound, & Waidmann, 1999; and Wiley, 1992).

Because there are several fundamentally different understandings of the terms *adaptation* and *adaptive*, each with its strengths and limitations, any investigator or author should state clearly what is meant when they use these terms in their studies and writings. In our research, and in this article, we concur with a criterion of adaptation based on life history theory (discussed in more detail in the next section). Specifically, a given instance of investing in reproduction is rightly considered adaptive if (and only if) LRS is greater than would be the case if no effort at investment in that instance had occurred (Caro & Borgerhoff Mulder, 1987; Clutton-Brock & Harvey 1979; Fisher, 1985). Likewise, a given instance of *not* investing in reproduction is also rightly considered adaptive if (and only if) the total LRS is greater than would be the case if some effort at investment in that instance had occurred. Natural selection maximizes LRS (and over time, multigenerational reproductive success), not fertility (Stearns, 1992; Williams, 1966). High fertility is evolutionarily inconsequential if all the offspring die before reproducing themselves.

Life History Theory

Life history theory (LHT) is an analytical framework within evolutionary theory that is concerned with the fundamental challenge faced by all life forms: What is

the optimal allocation of finite resources within a finite lifetime? (Borgerhoff Mulder, 1992; Charnov, 1993; Fisher, 1930; Gadgil & Bossert, 1970; Hill & Kaplan, 1999; Lessells, 1991; Stearns, 1992; Vitzthum, 2008b). Each alternative potential solution to this question, referred to as a life history strategy (LHS), comprises a suite of reproductive and developmental traits (e.g., age and size at initiating reproduction; the number, quality, and timing of offspring; age at death) and the associated timing of resource allocation to growth, reproduction, and somatic maintenance. LHT may be thought of as evolutionary economics and LHS as investment schedules subject to natural selection.

Reproductive effort (RE) or investment is the allocation of resources to producing a live offspring that survives to maturity, and comprises mating effort and pre- and postnatal parental investment. Modulating reproductive effort refers to any physiological or behavioral change that varies the opportunities for reproduction and/or the probability of a live offspring reaching maturity. Reproductive success (inclusive fitness) refers to the number of an individual's offspring that survive to reproductive maturity plus those of its kin discounted by the degree of relationship. Studies of humans are fortunate if they can estimate LRS, but it is the multigenerational inclusive fitness of a LHS that is evolutionarily significant (Chisholm & Coall, 2008). Depending on the species, resources comprise the availability of a suitable habitat, mate, and social group; the energy and other nutrients available to support a growing conceptus, live offspring, and one's own soma; and the physical and psychological status of the parents (Vitzthum, 2008b). Time (Hill & Kaplan, 1999; Promislow & Harvey, 1990) and information (Worthman, 2003) are also critical resources in limited supply during an individual's lifetime.

Under the assumption that any unit of resource may be allocated to only one function, the fitness advantage of any given investment is traded off against the fitness advantages of any other potential investments (Zera & Harshman, 2001). The relative optimality of a LHS is measured in the currency of reproductive success. Natural selection favors the LHS that results in an organism leaving relatively more copies of its genes in subsequent generations than would have been left by following some other strategy. With enough time in a given environment, natural selection will result in most members of a population following the optimal strategy, to varying degrees. Changes in environmental conditions can prompt selection for a different optimal LHS. Life history traits are often quite flexible in their expression, depending on the specific environmental conditions encountered by the organism.

LHT clarifies the fitness costs and benefits of modulating reproductive investment in any given opportunity. Perhaps the best studied example of such a trade-off in humans is the investment in breastfeeding one's youngest offspring (McDade & Worthman, 1998; Sellen & Smay, 2001; Vitzthum, 1994a, 1994b). Depending upon the frequency, duration, and intensity of breastfeeding, and on the age of the nursling, lactation suppresses ovulation, thus preventing investment in another pregnancy. In the absence of breastfeeding, human females ovulate, on average, about six weeks after giving birth (Gray, Campbell, Zacur, Labbok, & MacRae, 1987). Such a swift shift of investment from a newborn to a potential future progeny could, however, risk the survival of the newborn. On the other hand, a lengthy investment in lactation risks future reproductive potential. The trade-off between these two costs must be balanced to optimize reproductive success. Although it has been speculated that breastfeeding could be a long-term "natural contraceptive," LHT argues against such an expectation. In fact, efforts to extend the lactational amenorrhea method (LAM) much beyond six months postpartum have had mixed results (Cooney, Nyirabukeye, Labbok, Hoser, & Ballard, 1996; van der Wijden, Brown, & Kleijnen, 2003).

Just as there are fitness advantages to delaying investment in a new reproductive opportunity while still breastfeeding a current offspring, so might inclusive fitness be increased by suppressing reproductive investment in other circumstances. It is in a female's best interests to allocate resources to as many viable offspring as possible while apportioning just enough to each to ensure its survival to reproductive maturity. Allocating adequate resources to her own survival is also necessary, in so far as it increases her reproductive success. Hence, strategic modulation of reproductive effort is potentially adaptive because investment in a new conception may risk one's own survival, future reproductive opportunities, and/or current offspring survival (Vitzthum, 2008b).

An Application of LHT to Human Reproduction: The Flexible Response Model

As noted earlier, Frisch's proposal (Frisch, 1978; Frisch & McArthur, 1974; Frisch & Revelle, 1970, 1971a, 1971b) that human reproduction is sensitive to energetic conditions (i.e., energy intake and expenditure; body fat stores) ignited an often contentious debate.

Scientists with training in evolutionary biology or medicine supported her core hypothesis, if not always the specifics of her model. Their position was based on innumerable observations of the impact of energetic conditions on fertility in both wild and domestic animals, and on reports that even moderate exercise, dieting, or the stress of exam week disrupted ovarian functioning in women in industrialized countries (e.g., Ellison & Lager, 1986). Collectively, measurements of ovarian hormones were and are a particularly compelling body of evidence in support of energetic models. Several studies have demonstrated that concentrations

of progesterone (P4) and estrogens are lower in ovarian cycles occurring during periods of increased exercise and/or decreased caloric intake compared to hormone levels in cycles occurring in the absence of such stressors (Ellison et al., 1989; Ellison & Lager, 1986; Jasienska, 2001; Panter-Brick, Lotstein, & Ellison, 1993; Prior, 1985a, 1985b, 1987). Furthermore, a set of studies reported that P4 concentrations are lower in nonindustrialized populations (in which workloads are high and food intake is marginally adequate) compared to P4 concentrations in healthy Bostonian women not experiencing energetic stress (see Figure 1). Ellison and others have argued that one consequence of the lower hormone levels in these energetically stressed populations must be lower fecundity (the capacity to conceive) and hence lower fertility (the production of a live offspring). In addition, he argued that the lower fecundity is an adaptive mechanism that spares women the energetic demands of any offspring production that is potentially wasteful (because of inadequate resources that increase the risk of both offspring and maternal death). In this model, reproductive suppression is a function of current energetic status and/or any increase in energetic stress. Such suppression is argued to be adaptive because it is expected to lead to a higher LRS than would ill-spent investment in an offspring that dies in utero or before adulthood (Ellison, 1994).

In response to the energetic models, demographers and other social scientists pointed to substantial evidence and analyses (summarized in Bongaarts, 1980) that human fertility in populations worldwide is barely

affected by energetic factors except in cases of starvation. The compelling fact (even more so in the 20th century than today) is that most women in nonindustrialized populations engage in often arduous physical labor, are chronically undernourished, and often also experience periods of even greater food scarcity. Economic and psychosocial stressors that threaten their health and even their lives are commonplace. Yet despite these demands, each of these same women can experience a dozen or more pregnancies in her lifetime. This seeming paradox could not be readily countered by those supporting energetic models.

A review (Vitzthum & Smith, 1989) of the arguments and data on both sides of this debate revealed substantial scientific merit in both positions, leading me (VJV) to explore the possibility that this paradox might be solved if approached from the perspective of LHT. Drawing as well on what was known of genetics, developmental biology, and human adaptability, I developed the FRM of reproductive functioning to explain variation among women in reproductive suppression under seemingly comparable conditions (Vitzthum, 1990, 1992, 1997, 2001; Vitzthum & Smith, 1989).

I argued that the answer to the apparent paradox lies in the advantage gained by U.S. women (who usually experience good conditions) of delaying reproduction in the face of temporarily poor conditions: they can expect things to get better soon. Most women in poor countries, on the other hand, have lived all their lives in chronically poor environments and cannot expect conditions to improve, hence delaying would carry no benefit. In other words, a current reproductive decision is dependent on both the absolute quality of current conditions and the relative quality of these conditions compared to prior conditions and predicted future conditions.

I proposed that a woman's physiology "judges" current environmental conditions based on the conditions she experienced as she matured. The best conditions encountered before adulthood are likely to be about the best conditions that one will experience in a lifetime. Even if these habitats are mediocre on some absolute scale, it makes no sense to delay reproduction under the prevailing conditions if there is no reasonable prospect of improvement within a finite time period. Of course, if things become even more demanding, then women temporally experiencing even worse than usual conditions may also delay their reproductive investments for a time. However, as long as the probability of a successful conception in these poorer conditions is greater than zero, women can be expected to acclimate over time and to resume reproductive functioning. Individuals able to acclimate will have a reproductive advantage over those who do not. The FRM applies to both pre- and postconception reproductive decisions. Because poor conditions were the norm during their maturation,
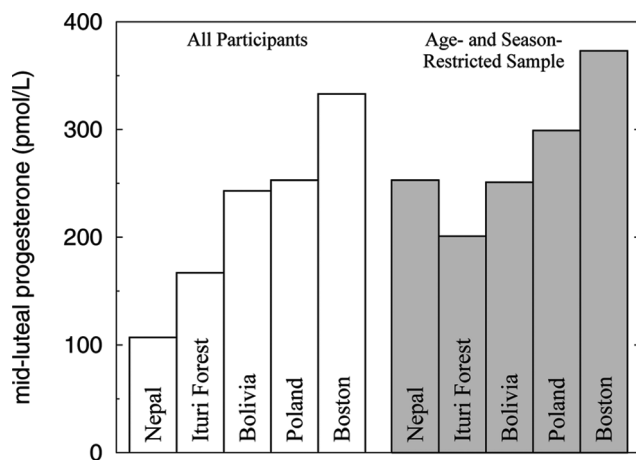
Figure 1. Comparison of midluteal progesterone [P4] concentrations in five samples assayed following the same protocol in the same laboratory. The left set of bars shows the comparison for all participants in each of the five samples, suggesting large interpopulational differences in P4. The right set of bars (shaded) shows the effect of selecting a comparable subsample (ages 25 to 35 years and measured during least energetically demanding season) from each of the five samples: when adjusted for the effects of age and season, the interpopulational differences are less but are still substantial (Data from Vitzthum, Ellison, Sukalich, Caceres, & Spielvogel, 2000).

women accustomed to these may maintain their pregnancies in the face of stressors that would prompt termination in better-off women.

In contrast to energetics models, the FRM proposes that fecundity is a function of current conditions (including one's own status) relative to long-term average conditions (viewed as an estimate of likely future conditions). Hence, the decision to reproduce depends on (a) the probability of successful reproduction (conception and live birth) in the present conditions, (b) the probability of conditions changing (for better or worse) within a finite period, (c) the risk to future reproductive opportunities, and (d) the expected duration until the end of the reproductive life span.

Life history theorists have hypothesized that age-specific fertility schedules will vary with mortality schedules (Charnov, 1993; Stearns, 1992; Williams, 1966). For example, if mortality risk is high, it can be selectively advantageous to begin reproduction early (Chisholm, 1993; Geronimus et al., 1999; Migliano, Vinicius, & Lahr, 2007) and have many smaller offspring (Charnov & Ernest, 2006). Hence, in contrast to energetics models, the FRM (Vitzthum, 1990, 1997, 2001) predicts that fecundity is not necessarily reduced in women living in suboptimal conditions but would depend on the four probabilities noted previously.

Hypotheses regarding the relative fitness of alternative behavioral or physiological phenotypes (e.g., in conditions of resource scarcity, breastfeeding versus weaning, or ovulating versus not ovulating) are difficult to test in humans because of the daunting task of accurately measuring reproductive success (although some have surmounted this challenge, e.g., Borgerhoff Mulder, 2000; Strassmann & Gillespie, 2002). Rather, investigators have, of necessity, tested other predictions from these evolutionary models and, if supported, inferred that adaptive modulation of reproductive functioning is occurring. This tack has, for example, yielded substantial evidence of statistically significant associations between the occurrence of energetic stressors and changes in ovarian hormones, as was previously noted. But because we can rarely measure the true LRS of a human phenotype, it remains possible that such changes are nonadaptive disruptions of normal functioning.

Both the FRM and energetics models generate hypotheses that can be tested with empirical data. To test the predictions of these models, Vitzthum and Dr. Hilde Spielvogel at the Bolivian Institute for High Altitude Biology in La Paz implemented Project REPA (Reproduction and Ecology in Provincía Aroma) in 30 Bolivian agropastoral communities. In Bolivian women, P4 concentrations in ovulatory cycles averaged about 70% of those in a sample of Chicago women. Over the course of nearly two years, our team collected every-other-day saliva samples from almost 200 women, for up to eight sequential cycles per woman, to measure ovarian steroid levels. During the late third and subsequent weeks of each cycle, we also collected urine samples that were tested for hCG (human chorionic gonadotrophin), a biomarker of implantation. Thus we were able to correlate steroid levels with specific conceptions and their outcomes (Vitzthum, Spielvogel, & Thornburg, 2004).

In brief, as predicted by the FRM, conceptions occurred at the lower P4 levels typical of these Bolivian women, the conceptions had the same rate of early pregnancy loss as do U.S. women, and the live births were of normal gestational duration and of normal weight. Thus, these data refute the hypothesis that variation in steroid levels across populations is necessarily associated with variation in fecundity and fertility (Ellison, 1994; Ellison, Panter-Brick, Lipson, & O'Rourke, 1993; Lipson, 2001) and lend support to the Flexible Response Model (Vitzthum, 2009; Vitzthum et al., 2004; Vitzthum, Spielvogel, Thornburg, & West, 2006). In addition, when energetic stressors increased in the physically demanding planting and harvesting seasons, these Bolivian women experienced increased rates of anovulatory cycles, increased rates of pregnancy loss, and lower P4 concentrations than they did in less arduous seasons (Harris & Vitzthum, 2012; Vitzthum, Thornburg, & Spielvogel, 2009). These patterns suggest that it is not the absolute levels of ovarian hormones that determine fecundity and fertility in an individual but rather changes from her typical hormone levels that are acting as the salient biological signals influencing her fecundity and fertility.

These findings from Project REPA highlight both the usefulness of LHT for understanding variation in reproductive functioning and also the value of documenting the full range of variation in biomarkers of that functioning. A lack of appreciation for the extraordinary diversity across human populations and among women within those populations can easily lead to erroneous conclusions and (implicitly or explicitly) assuming that the characteristics of biomarkers in the United States are the norms against which to compare all other populations. The review of the relevant data in the next section clearly demonstrates that reality is otherwise.

Variation in and Measurement of Women's Ovarian Functioning

Das einzig Regelmässige an der Regel ist ihre unregelmässigkeit. (The only regularity of the menstrual cycle is its irregularity.)

—Ludwig Fraenkel (1926)

There is a growing interest among sex researchers in the dynamic relationships between an individual's internal hormonal milieu, especially those associated with ovarian cycling, and her psychosexual states and/or her externally manifested conscious or unconscious sexual signaling and/or behavior. Obviously, discovering
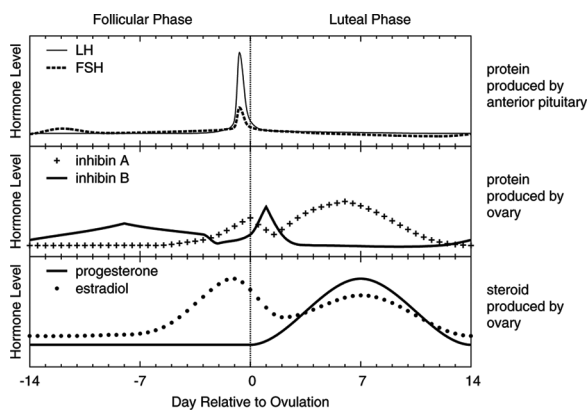
these relationships requires accurate measurements of the relevant variables, and sex researchers have directed considerable effort to measuring behaviors and psychological states and responses.

In contrast, there is less appreciation of the substantial variation within and between women in ovarian cycling and the associated changes in hormone concentrations. Implicitly or (more rarely) explicitly, investigators have assumed that nearly all healthy premenopausal women typically experience "normal" ovarian cycles substantially similar to those depicted in medical texts, a presumption that lends support to the use of nonhormonal biomarkers as proxies for hormonal biomarkers of ovarian functioning. In fact, the evidence now available clearly demonstrates that deviations from idealized cycles are normal and common—a fact which is consistent with the evolutionary history of women's reproductive biology and which we will see has significant implications for successfully addressing many questions in sex research.

The Ovarian Cycle and Fecundity

Figure 2 is a depiction of the principal hormonal changes that occur over the course of an *idealized* ovarian cycle. The ovary produces the steroids, estradiol (E2) and progesterone (P4), and the proteins, inhibin A and inhibin B. The gonadotrophins, follicule stimulating hormone (FSH) and luteinizing hormone (LH), are proteins produced by the anterior pituitary gland. Negative and positive feedback mechanisms between these ovarian and pituitary hormones and the hypothalamus

(which produces gonadotrophin-releasing-hormone, GnRH) regulate follicle maturation, the timing of ovulation, and the proliferation of the endometrium (Messinis, 2006). If fertilization occurs, the conceptus produces human chorionic gonadotrophin [hCG], which is molecularly similar to LH.

Ovulation (release of the egg from the follicle) occurs roughly at about midcycle, demarcating the *follicular* (preovulatory) and *luteal* (postovulatory) *phases* of the cycle (see Figure 2). At the transition from a previous luteal phase to the beginning of a cycle, FSH begins to rise slightly as a consequence of low E2, P4, and inhibin A levels. Rising E2 and inhibin B levels, produced by the growing dominant follicle, then suppress the FSH increase. Ovulation is triggered by a surge in LH, itself brought about by a sufficiently high level of E2. After ovulation, elevated LH causes the ruptured follicle to redifferentiate into the corpus luteum ("yellow body"), which produces P4, inhibin A, and some E2. By about the middle of the luteal phase, the falling LH levels are no longer sufficient to maintain the corpus luteum, which then regresses if no conception has occurred. Falling hormone levels eventually prompt the menstrual bleeding that denotes the beginning of the next cycle.

Fecundity (but in French, *fertilité*) is the biological capacity of an individual or couple to conceive, and *fertility* (in French, *fécondité*) is the production of a live offspring. In other words, fecundity refers to the potential to reproduce, and fertility refers to the realization of that potential. For example, a person who abstains from sexual intercourse may be fecund but is not fertile. French-English mistranslations aside, the two terms are often mistakenly treated as synonyms, leading to much confusion even within a single work. For the sake of the readers, it is advisable for authors to define the terms at the outset in any work.

In theory, fertility is relatively easy to measure if there is an observer and a record-keeping mechanism (i.e., hospital nurse, birth certificate) but can be difficult if relying on the recall of past events. For example, outside of a clinical setting, the birth of a child who died shortly thereafter may be forgotten or not mentioned or may be considered a stillbirth even if born alive.

Fecundity is difficult to operationalize. Premenarcheal and postmenopausal females are infecund, but the varying fecundity between these two transitions is nebulous. The postmenarcheal and premenopausal years, and the weeks after a live birth, are subfecund compared to periods of peak fecundity, but by how much? *Fecundability*, the monthly probability from 0 to 1 of conception, quantifies fecundity, but it cannot be measured directly because conceptions cannot be readily detected at their occurrence. It is estimated that about half of all conceptions are lost before implantation (which occurs about nine days after conception in successful pregnancies), and even with the most advanced laboratory methods currently available, conceptions cannot be detected until

Figure 2. Hormonal profiles during an idealized ovulatory cycle. Follicular and luteal phases are demarcated by ovulation (dominant follicle releases the ovum, day 0). At the transition between cycles, FSH rises slightly because of low estradiol [E2], progesterone [P4], and inhibin A levels. Rising E2 and inhibin B, produced by the dominant follicle, then suppress FSH. High E2 triggers the LH surge that induces ovulation and, subsequently, causes the dominant follicle to redifferentiate into the corpus luteum ("yellow body"), which produces P4, inhibin A, and some E2. Without conception, falling LH is insufficient to maintain the corpus luteum, which regresses. Dropping hormone levels eventually prompt endometrial shedding (menses), denoting the beginning of the next cycle.

214

about one week after implantation (Vitzthum, 2008a). *Apparent fecundability* is an approximation estimated from the conceptions that are detected (Wood, 1994).

Generally, sex researchers have been more focused on fecundity than on fertility. More specifically, there is an effort to understand the interplay (or at least describe the associations) between the different hormonal milieus that characterize varying degrees of fecundity and psychosexual states, sexual signaling, and/or sexual behaviors. The difficulty of directly measuring fecundity or hormonal concentrations often prompts investigators to rely on nonhormonal biomarkers of variation in fecundity and hormone concentrations.

Nonhormonal Biomarkers of Fecundity

Cycle (segment) length and menses duration. Cyclical vaginal bleeding and the duration between successive onsets are arguably the oldest and most common biomarkers of a woman's ovarian functioning. For analytical purposes, the World Health Organization (WHO) refers to a menstrual interval as a *segment* to avoid the presumption of normal cycling (Snowden & Christian, 1983). *Segment length* is defined as the first day of menses up to and including the day before the subsequent menses. In general, most studies have, intentionally or not, defined segment length thusly. Published data on segment length (see Figures 3 and 4) reveal the significant variation across populations, the even greater variation between women within a population, and the appreciable changes that occur over the reproductive life span.

Temporal variation from cycle to cycle (frequently mislabeled as "irregularity") is also greater and more common than widely appreciated. Most studies of regularity have been conducted with women who have described themselves as "regular." Despite this selection
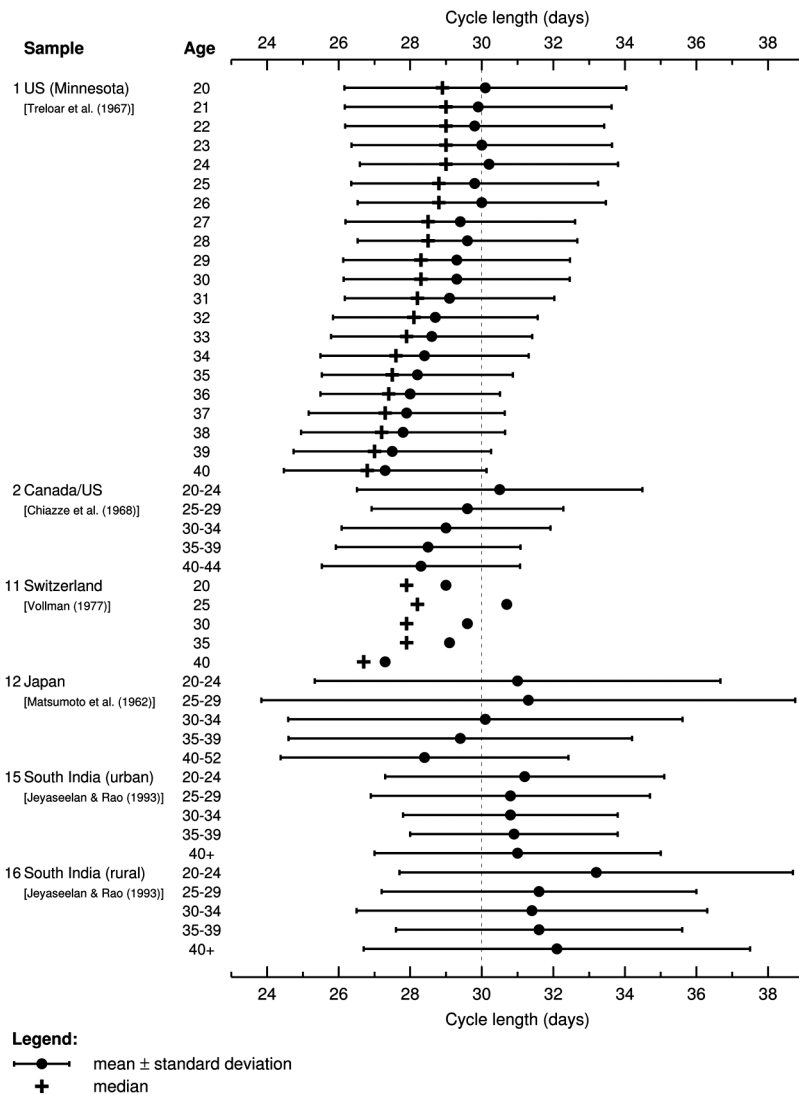
Figure 3. Cycle length distributions (based on recorded data) for selected subsamples of various age ranges. The numbers at the far left of each sample identify the corresponding sample and data, available in Vitzthum (2009).

215
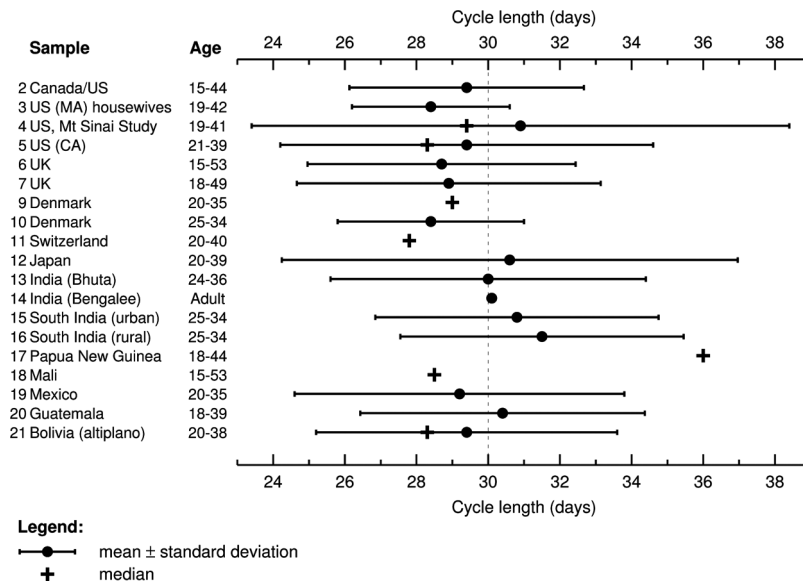
Figure 4. Cycle length distributions (based on recorded data) for selected samples from various human populations. The numbers at the far left of each sample identify the corresponding sample and data, available in Vitzthum (2009).

bias, the women's initial reports were not consistent with their own prospectively collected data, which suggests that segment length is even more variable in the larger population. For example, more than half of 303 Guatemalan women who self-identified as having regular cycles every 26 to 32 days were then observed to have one or more of three consecutive segments fall outside this range (Burkhart, de Souza, Salazar, & Hess, 1999). Bolivian, Indian, and U.S. women all have similarly high levels of variability (Creinin, Keverline, & Meyn, 2004; Vitzthum, Spielvogel, Caceres, & Gaines, 2000; Williams, 2006). The different indicators of variability used by each study make direct comparison difficult, but about half or more of the women in each sample have a range in segment length of at least six days, and about one-quarter have a range greater than two weeks, even if only a few cycles were recorded by each woman. In a recent five-year study of U.S. White women, the authors concluded, "Each woman has a wide range of cycle lengths that cannot be easily distinguished from other women's ranges of cycle lengths" (Ferrell et al., 2005, p. 574).

Despite this clear evidence of substantial variation within and between populations, and throughout the reproductive life span, most investigators appear to be unaware of these data. For example, many studies have specified a required cycle length (e.g., 25 to 30 days) and/or required that a study participant have "regular" cycles, thereby (assuming the study criteria were applied reliably) excluding the majority of normal healthy women from a study. Aside from the selection bias that may have been introduced by these restrictive criteria, because most studies have relied on a woman's recollection of her cycle patterns, it is unlikely that many

participants in any study have actually met these criteria. Comparative studies have demonstrated that recalled reports of segment length and variability are inconsistent with prospectively recorded data (Burkhart et al., 1999; Creinin et al., 2004; Steiner, Hertz-Picciotto, Taylor, Schoenbach, & Wheeless, 2001).

Based on extensive analyses by a working group, WHO defines a menses *episode* as the period in days from the first through the final appearance of blood during waking hours (Snowden & Christian, 1983). For example, if upon awakening on Monday morning, a woman detects bleeding for the first time and that bleeding continues through any time during waking hours on Thursday, but is no longer present on Friday upon awakening, then the episode would have been four days long. If she had been without bleeding on Wednesday, but then the bleeding had continued on Thursday, the whole episode would still be considered four days. If there had been two or more bleeding-free days, then the bleeding on Monday and on Thursday would be considered two separate episodes of one day each. Very few studies distinguish spotting from more substantial flow. As imperfect (or arbitrary) as the WHO criterion for menses episode may seem to some, it has the benefit of promoting comparability across studies and is relatively easy to implement in different populations.

Published data on episode durations are depicted in Figure 5. Although it has been hypothesized that episode duration may be positively correlated with hormone levels and endometrial development, and hence could be used as a proxy variable for fecundity, no studies have directly evaluated this suggestion. In fact, the reasons for and significance of striking interpopulational variation in menses duration (almost twofold worldwide) are

216
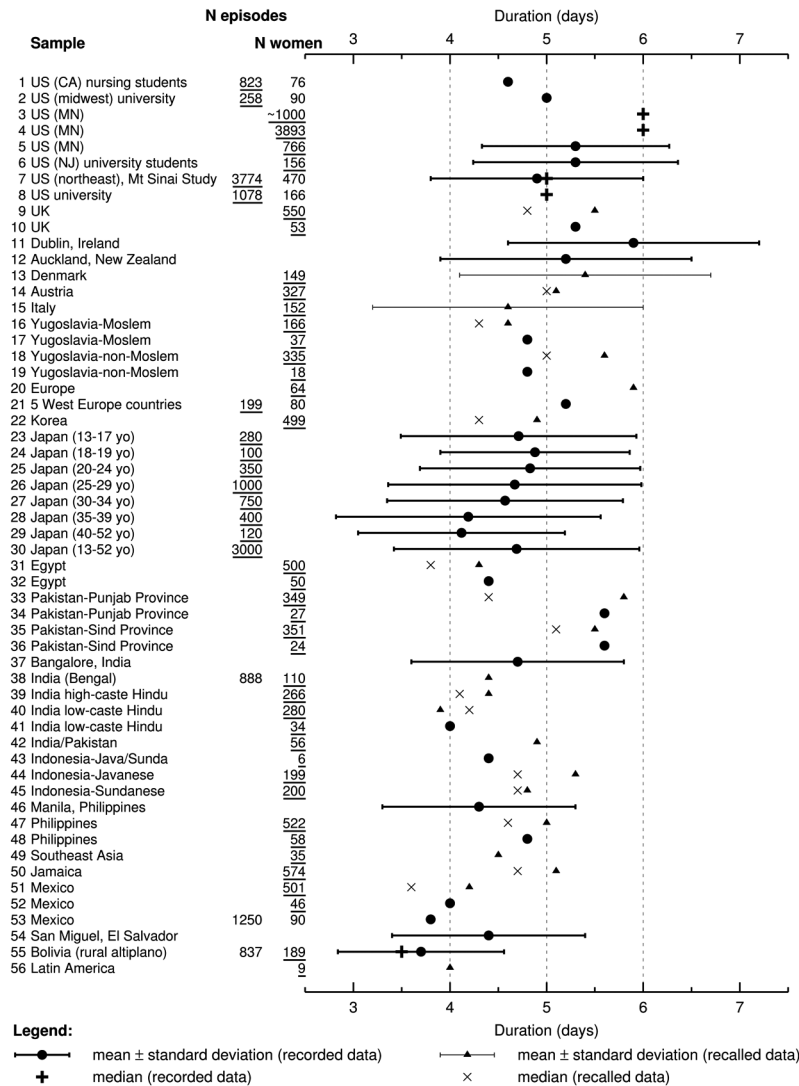
Figure 5. Menses duration distributions for selected samples from various human populations. For each sample, the *N* for the units of analysis (episodes or women) is underlined. Note the distinction between recorded data (which is quite reliable) and recalled data (which may be subject to various recall biases). The numbers at the far left identify the sample and data, available in Vitzthum (2009).

unknown. A multicountry WHO study concluded that geographical location is the best predictor of menses duration (Belsey & Peregoudov, 1988); hence the data in Figure 5 are arranged roughly from West to East, beginning in the Western hemisphere, and north to south. In general, those of European ancestry have the longest menses, although duration among Pakistani women is as long. The shortest menses are reported by women in South America and India. Samples from northern and southeast Asian populations fall between these extremes.

Early pregnancy loss is *not* associated with significantly ($p > 0.05$) longer vaginal bleeding. In both U.S. and Bolivian women, bleeding is no more than 0.4 days longer than a woman's average menses, primarily due to a slight increase in light bleeding (Promislow, Baird, Wilcox, & Weinberg, 2007; Vitzthum, Spielvogel, Caceres, & Miller, 2001).

Within-woman variability in menses duration appears low. Most studies report that a woman's episode rarely differs by more than a day or two from cycle to cycle. On the other hand, a single day is 25% of the duration in an event that lasts only four days. Thus one day is as variable as a 28-day cycle varying by seven days. In an analysis of a large subsample of a U.S. longitudinal data set, the median episode duration was six days, and less than 25% of women at a given age had an episode duration that ranged by more than one day during the year (Belsey & Pinol, 1997). For the entire analyzed sample, aged 15 to 43 years, 75% ranged less than or equal to three days. In Bolivia, for each woman observed for at least three bleeding episodes that were neither followed nor preceded by a conception ($n = 115$), more than half had a range of less than or equal to 1 day,

217

and only 6% had a range of more than 3 days (Vitzthum et al., 2001). This study defined regularity in a woman's menses as the ratio of the range in days (between the shortest and the longest episode duration) to her mean duration. Those women with mean episode durations above and below the median episode duration had comparable regularity in menses duration.

Phase durations, ovulation, and the fertile period. Whether a woman's given segment (cycle) length is known or presumed, a large number of studies have assumed that ovulation occurs on a specific day of that cycle (e.g., 14 days after menses begins) which is, in effect, assuming that cycle phase lengths are relatively invariable. As noted, the ovarian cycle consists of two phases (Figure 2): the follicular (preovulatory), defined by conventions (that predate knowledge of hormonal patterns) as beginning with the first day of menstrual bleeding and ending with ovulation, and the luteal (post-ovulatory), beginning with ovulation and ending on the day before the first day of menstrual bleeding.

Although the follicular phase varies somewhat more than the luteal (refer to Table 1), luteal phase length is also variable within and across individuals (Bailey & Marshall, 1970; Döring, 1969; Matsumoto, Nogami, & Ohkuri, 1962). Many large-scale studies have relied on changes in basal body temperature (which usually rises about $0.2°C$ to $0.5°C$ with ovulation) and/or cervical mucus to determine the timing of ovulation, but ultrasound and hormonal biomarkers give the best estimates. Of 68 Swedish women, slightly more than one-third of the E2 and LH peaks (which occur immediately prior to ovulation) occurred before day 12 or after day 18 of the cycle; 31% of the women had luteal phases shorter than 12 or longer than 15 days (Landgren, Undén, & Diczfalusy, 1980). Of 141 U.S. women (mean age 29 years), each contributing ≥ 3 cycles, 34% had a range of follicular-phase lengths >7 days and 9% had a range of luteal-phase lengths >7 days (Fehring, Schneider, & Raviele, 2006). Given such variation, there is very good reason to think any study that assumes ovulation is occurring on some specified day, absent any biological data to support this, would be on shaky ground.

In fact, whether ovulation has occurred in a given cycle cannot be taken for granted. The probability of ovulation is lower during the years following menarche up to about 18 years of age and also in the years prior to menopause (Apter, 1997; Döring, 1969; Lipson & Ellison, 1992; Mansfield & Emans, 1984). The data for women aged 20 to 40 years in industrialized countries suggest ovulation rates are typically about 85% (Table 2), but some groups of women may have substantially lower rates. For example, the probability of ovulation was only 63% in a sample of unmarried 20- to 29-year-old New Zealand women of European ancestry (Metcalf, 1983; Metcalf & Mackenzie, 1980). Even in a sample of Swedish women (aged 20 to 37 years, mean 27.6 years) who were recruited

for contraceptive testing and had met very strict selection criteria, 7 of the 43 (16%) subsequently did not ovulate (Landgren et al., 1980; Landgren & Diczfalusy, 1980). Ovulation rates were only 45% in a sample of poor peri-urban La Paz women (aged 23 to 35 years) sampled during the winter, a period of food scarcity (Vitzthum et al., 2002). Ovulation rates are unknown for poorer women or women in more stressful occupations in the United States, but it cannot be assumed that they are as high as those of women in less demanding circumstances, nor can it be assumed that this latter group would have ovulation rates near 100%.

Even when ovulation does occur, the length of the fertile window is not symmetric about the day of ovulation, nor is it particularly long. Rather, the fertile window lasts only about six days, ending at or shortly after ovulation (Dunson, Baird, Wilcox, & Weinberg, 1999; Dunson, Columbo, & Baird, 2002; Wilcox, Weinberg, & Baird, 1995). For example, in a study of intercourse and conceptions in 647 noncontracepting couples, the probability of pregnancy from intercourse did not rise above 5% until about five days before ovulation and dropped to near 0% within one day afterward (see Figure 6).

Several investigators (most notably evolutionary psychologists) have examined whether the frequency of some sexually linked signal or behavior (e.g., greater flirtatiousness) is higher during one period of the ovarian cycle than another, sometimes predicated on the argument that greater mating effort should be evident during periods of higher fecundity. But many of these studies seem to have defined periods of relative fecundity without much regard to what is known of the underlying biology (periods defined as having high fecundity have included the entire follicular phase, the first 14 days of the cycle, the days between the end of menses and the 14th cycle day, the two or four days centered on the day of ovulation, etc.).

The merits of theories and studies that predict substantial differences in sexually related variables between periods of the human ovarian cycle differing in fecundity remain much debated. It is hoped a resolution of at least some of these disputes may be achieved by defining variables that are a closer match to what is currently known of the relevant biology. We return to this issue again in the later discussion of hormones and sexuality.

Hormonal Biomarkers of Fecundity

Compared to nonhormonal constructs, hormonal biomarkers are unquestionably superior indicators of a woman's ovarian functioning. But absolute hormone levels are not perfect proxies for fecundity. There is substantial variation in hormonal biomarkers within and between women and populations (specifics on this subject follow), and the relationship of that variation to fecundity is not straightforward (as discussed in the earlier section on the FRM and

Table 1. *Descriptive Statistics for Phase Lengths*

Population [Reference]	Method	Age (years)	N Cycles	N Women	Unit of Analysis[a]	FOLLICULAR				LUTEAL			
						Median	Mean	SD		Median	Mean	SD	
U.S. (midwest) universities [Collett et al., 1954]	BBT	17–50	285	132	C						14.4		
Japan[b] [Matsumoto et al., 1962]	BBT	13–52	3000	3000	W		17.85	6.24			12.65	1.63	
United Kingdom[c] [Monari & Montanari, 1998]	BBT	16–55	18677		C	16	16.10	4.14		12	12.21	2.89	
		16–20	163		C	16	16.62	4.27		13	12.37	2.79	
		21–25	2382		C	17	17.15	4.20		12	12.22	2.87	
		26–30	4091		C	16	16.73	3.92		12	12.14	2.82	
		31–35	4942		C	16	16.16	3.95		12	12.26	2.66	
		36–40	3774		C	15	15.56	3.63		12	12.24	2.49	
		41–45	2264		C	15	15.00	4.17		12	12.06	3.09	
		46–50	993		C	14	15.15	5.69		12	12.46	4.71	
		51–55	68		C	15	17.28	9.23		13	12.69	2.27	
5 countries[d] [WHO, 1983b]	cervical mucus		6472	687		15	15	2.60	5th–95th percentile: 10.5–19.0	13.5	13.5	2.8	5th–95th percentile: 8.7–17.2
Ireland							15	2.80			13.2	2.1	
India							15.2	2.10			13.5	3.5	
New Zealand							15.8	3.00			12.8	2.2	
El Salvador[e]							13.6	2.10			14.5	1.8	
Philippines							15.6	2.60			13.3	3	
New Zealand [France et al., 1992]	cervical mucus						15.5	3.10	95% CI of mean: 9–22		12.6	1.8	95% CI of mean: 9–17
UK/Sweden [Lenton et al., 1984b]	LH surge; daily sampling	18–50	327		C						14.13[f]	1.41	68% CI of mean: 12.7–15.5
UK/Sweden [Lenton et al., 1984a]	LH surge; daily sampling	18–39	293	293	W		12.9[g]		68% CI of mean: 10.3–16.3				
		18–24	42		C		14.2		11.3–18				
		25–29	125		C		12.9		10.2–16.4				
		30–34	91		C		13.1		10.7–16				
		35–39	35		C		12.1		9.7–14.9				
		40–44	18		C		10.4		8.2–13.1				
		45–50	14		C		10.5		7.2–15				
US (4 cities)[h] [Fehring et al., 2006]	LH surge; daily sampling	21–44	1060	141		16.0	16.50	3.40	95% CI of mean: 10–22	13.0	12.40	2.00	95% CI of mean: 9–16
U.S. (CA) [Waller et al., 1998]	urinary hormones; daily samples	21–39	>370	>370	W	15.6	16.60	5.30		13.0	12.90	1.80	
5 West Europe countries [Ecochard & Gougeon, 2000]	ultrasound	19–42	199	80	C		14.6	0.35			13.6	0.25	

[a]C = cycles, W = women. [b]selected sample. [c]excluded "anomalous" segments; ≥6 consecutive segments/woman. [d]natural family planning trials; women selected for regularity. [e]significantly different lengths from other 4 countries. [f]geometric mean; excludes short phases; start defined as day after LH surge; $14.13 + 0.31$ time from ovulation $= 14.44$ days. [g]geometric mean; excluded cycles >40 days; phase end = LH surge-1; add 1 day for LH peak and 0.7 days to ovulation $= 14.6$ days. [h]selected for regular cycles 21–42 days long; follicular includes days of LH surge and ovulation.

Table 2. *Anovulation Fraction (%)*

Population [Source]	Method	Age (years)	N Cycles	N Women	% Anovulatory
New Zealand [Metcalf & MacKenzie 1980; Metcalf 1983]	weekly urine	20–24[a]	355	113	28
		25–29[a]	211	70	6
		30–39[a]	193	62	2
		40–44[a]	221	67	10
		20–24[b], by cycle length:			
		<21 days	13		100
		21–35 days	286		20
		36–42 days	25		48
		>42 days	16		63
		25–29[b,c], by cycle length:			
		<21 days	1		0
		21–35 days	154		5.2
		36–42 days	18		5.6
		>42 days	6		0
		5–8 years post-menarche:			
		living with parents	117	39	16
		not living with parents	142	45	54
		20–29:			
		married		60	1
		not living with relations		72	37
U.S.[d] [Collett et al., 1954]	BBT	17–18	81	59	31
		20–24	112	33	10
		25–29	45	21	4
		30–34	23	11	0
		35–39	33	13	12
		40–50	33	11	15
		17–50	327	146	14
Switzerland [Vollman, 1977]	BBT	29			1
		40–45			34
		total	14848	621	18
Germany[e] [Döring, 1969]	BBT	18–20	282		27
		21–25	287		13
		26–30	418		5
		31–35	822		7
		36–40	640		3
		41–45	275		12
U.S. (CA) [Waller et al., 1998]	daily urine	18–39		372	7[f]
Chicago[g] [Vitzthum et al., 2002]	salivary progesterone	23–38	22	22	9
Bolivia[g,h] [Vitzthum et al., 2002]	salivary progesterone	23–35, better-off	25	25	12
		23–35, poorer	29	29	55
5 West Europe nations [Ecochard et al., 2001]	ultrasound	18–45[i]	326	107	5

[a]Percent of all cycles; different number of cycles/woman. [b]women observed for 3 months. [c]similar rates in 30–39 year olds. [d]Students & staff at midwest/northeast colleges. [e]mean cycles/woman = 6.7. [f]percent of women w/ ≥1 anovulatory cycles >35 days long. [g]restricted to cycles 25–35 days long. [h]winter season. [i]sample selected for high fertility.

the findings from Project REPA). The simple (if somewhat daunting) truth is that "typical" cycles occur only rarely.

In idealized depictions (Figure 2), ovarian steroids (E2 and P4) are usually portrayed as being at their lowest concentrations on the first day of menses. In fact, the transition from luteal to follicular phase (of the next cycle) is a continuum of declining steroid concentrations that may still be elevated above baseline concentrations throughout much or all of menstrual bleeding and may not be at their lowest levels until about the middle of the follicular phase.

Although there is marked variation between cycles and among women, substantial rises in E2 do not typically occur until the second half of the follicular phase (the total duration of the follicular phase varies in normal healthy women from seven to 21 days; Wood, 1994). In an ovulatory cycle, E2 peaks about 36 hours before ovulation then declines sharply and rises again to a secondary peak (or plateau) during the middle days of the luteal phase. In the absence of conception, E2 declines over the next week or so. While the highest E2 peak usually occurs during the very late follicular phase, it is not necessarily the case in any given woman that the total E2 concentration, nor the average daily concentration, is higher during the follicular phase than during the luteal phase. In fact, E2 is likely to be higher
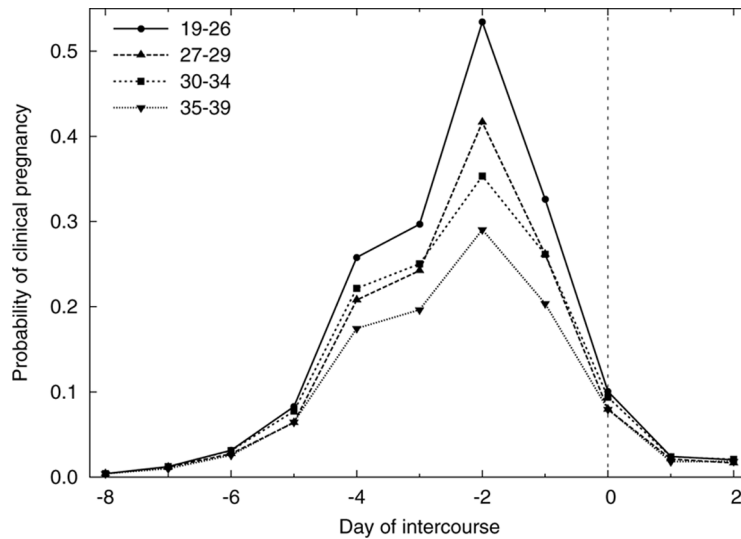
Figure 6. Fertile window for four age groups. Probability of conception is highest for an act of intercourse occurring two days prior to ovulation. Redrawn from Dunson et al. (2002). Changes with age in the level and duration of fertility in the menstrual cycle. *Human Reproduction*, *17*(5), 1399–1403; used with permission of Oxford University Press and the European Society of Human Reproduction and Embryology.

during most of the luteal phase than it was in the first half of the previous follicular phase.

Subsequent to ovulation, ovarian P4 is produced by the corpus luteum (the transformed ruptured follicle). During the luteal phase, P4 first rises and then, in the absence of conception, declines. While this pattern is often depicted as symmetric with a peak at about Day 7 of the luteal phase (closely resembling a Gaussian "normal" curve), this idealized P4 profile is only rarely achieved in reality. In some cycles, the P4 rise is rapid and then falls quickly; in others, the rise is slow and the fall is rapid. Thus the highest P4 concentrations may occur early or late or in the middle of the luteal phase, and may be short in duration (resembling a peak) or long (resembling a plateau).

Moreover, leaving aside the details of timing and patterns, the total and peak P4 concentrations in one cycle are only modestly correlated to those of the subsequent cycles. For example, the correlation of mean-peak-P4 (i.e., the average P4 over a five-day window centered on the observed peak P4 in the luteal phase) in 153 pairs of consecutive ovulatory cycles observed in Project REPA was only 0.376, and the intraclass correlation coefficient (ICC) was only 0.32 (ICC = between-subject variance/total variance) (Harris & Vitzthum, 2012). Thus, 68% of total variance in P4 in this sample is attributable to within-woman differences in P4 across cycles. This modest correlation and high within-woman variance was observed despite the fact that P4 was measured in saliva samples collected every other day throughout a cycle. In other words, the high variance is not easily attributable to measurement error, nor to low sampling density, but it is an inherent biological attribute of the ovarian cycle in healthy women.

In sum, ovarian hormones rise and fall over the course of the ovarian cycle in broadly predictable patterns, but there is so much variation from cycle to cycle within a woman and between women in the timing of these changes that one cannot predict with much precision the day on which E2 or P4 will be at their highest concentrations. To know this, one must measure hormone concentrations several times (preferably daily) over the course of an ovarian cycle.

Interpopulational Variation in Hormones

Studies over the past forty years have exposed unexpectedly large differences between populations in ovarian steroid levels in healthy premenopausal adult women but have failed to reveal any consistent patterns in this variation (Vitzthum, 2009).

Estrogen levels in samples of Asian women have been reported to be anywhere from 55% to about 90% of the levels observed in samples of "white" women (Dickinson, MacMahon, Cole, & Brown, 1974; Key, Chen, Wang, Pike, & Boreham, 1990; MacMahon et al., 1974; Shimizu, Ross, Bernstein, Pike, & Henderson, 1990; Trichopoulos, Yen, Brown, Cole, & MacMahon, 1984; Wang, Key, Pike, Boreham, & Chen, 1991). In a comparative study (Bernstein et al., 1990) that gave particularly careful attention to controlling for several confounders, luteal-phase E2 levels were about 20% higher in Los Angeles whites compared to women in Shanghai. Cycle length was comparable in the samples, and P4 levels in ovulatory cycles were not significantly different. In contrast, P4 concentrations in ovulatory cycles from Bolivian women were about 70% of those in a sample of U.S. women, and about 15% lower in poorer than in better-off Bolivian women (Vitzthum et al., 2002). Although these various findings may give the impression that white women in wealthier

221

industrialized countries have the highest ovarian steroid concentrations, in fact, Mongolian pastoralists have the highest P4 levels yet to be reported (Vitzthum, Thornburg, Schaebs, Deimel, & Deschner, 2011), and women whose childhood was in East Germany before reunification had *higher* P4 levels than their counterparts in wealthier West Germany (Vitzthum et al., 2011).

Within the United States, hormone concentrations vary with ethnicity, but the patterns from different studies are inconsistent. Among northern California women, aged 18 to 39 years, Asians had lower estrogen levels compared to whites, but the concentration in Hispanics was about 12% higher (Windham et al., 2002). Progesterone metabolites did not vary with ethnicity in this population. Another study in Los Angeles reported that African-American women had 25% higher luteal E2 and 36% higher P4 levels, and Latina women had 15% higher luteal E2 and 18% higher P4 levels, than non-Latina whites (Haiman et al., 2002). Matched samples from the Nurses' Health Study II observed elevated E2 levels in both African Americans and Asian Americans compared to Caucasians, but no variation in P4 by ethnicity (Pinheiro, Holmes, Pollak, Barbieri, & Hankinson, 2005).

Less is known of interpopulational variation in gonadotrophins (LH and FSH). In both the !Kung San (van der Walt, Wilmsen, & Jenkins, 1978) and Bolivians (Vitzthum, Worthman, Spielvogel, & Thornburg, 2007), ovarian steroids were low relative to comparative samples (urban South African Blacks and Chicago women, respectively), but gonadotrophin levels were similar. These patterns suggest that the hypothalamic-pituitary-ovarian axis was not impaired in these populations and that these lower ovarian steroid levels were adequate for successful reproductive functioning.

Covariation of Hormonal and Nonhormonal Biomarkers

It is generally assumed that variation in segment and phase lengths is associated with hormonal variation, but the limited data available for testing this hypothesis are far from consistent (refer to Table 3). The first detailed study was based on daily blood samples from 68 complete ovulatory cycles (selected for a length of 25 to 36 days) from Swedish women (19 to 39 years old) who self-reported having regular cycles (Landgren et al., 1980). The Women's Reproductive Health Study (Windham et al., 2002) was a population-based examination of 411 U.S. women (aged 18 to 39 years) self-collecting daily urine for up to six months (totaled 1,451 complete follicular phases and 1,459 complete luteal phases, all from ovulatory cycles). Such differences in sample selection among the studies no doubt explain some of disparate findings in Table 3.

In the Swedish study, both total cycle and follicular phase lengths were most strongly correlated with mean

LH during days -3 to -7 before LH peak ($r = 0.45$ and 0.43, respectively) and mean E2 during days 1 through 6 from menses onset ($r = -0.44$ and -0.53, respectively). Similarly, Windham and colleagues (2002) also observed higher baseline and daily mean follicular-phase estrogen metabolite levels (but lower total follicular estrogen) in short follicular phases than in cycles with follicular phases 12 to 23 days long. In long (>23 days) follicular phases, the total follicular-phase estrogen was greater, and the luteal-phase daily mean was also higher. In contrast, neither Broom and colleagues (1981) nor Harlow, Baird, Weinberg, and Wilcox (2000) observed any relationship between either LH or estrogen levels and follicular-phase length.

Most studies have not observed any relationship between P4 and either cycle or follicular-phase length, except for a single report of higher follicular-phase P4 (note that follicular-phase P4 is of adrenal and not ovarian origin) and higher luteal-phase peak P4 in cycles with short follicular phases (Windham et al., 2002). Three studies reported that longer luteal phases were positively associated with higher luteal-phase P4 levels (Landgren et al., 1980; Smith, Lenton, & Cooke, 1985; Windham et al., 2002), but only Windham and colleagues (2002) observed higher follicular-phase P4 in cycles with long luteal phases.

In the Swedish sample, luteal phase length was also negatively correlated with peak preovulatory E2 but had no relation to postovulatory E2 levels (Landgren et al., 1980). On the other hand, Windham and colleagues (2002) observed short luteal phases to have higher total preovulatory estrogen metabolite levels *and* higher daily mean postovulatory estrogen levels. In addition, the longest luteal phases in this sample did not differ from "normal" luteal phases in estrogen levels in either phase. Displaying yet another pattern, short luteal phases in a smaller study of U.S. women did not differ from longer phases in follicular-phase estrogen levels but did have higher estrogen levels in the luteal phase (Smith et al., 1985). Finally, analyses of daily serum samples from 100 Shanghai women (aged 19 to 35 years) were in general agreement with the observations from the Swedish study, except that luteal phase length did not correlate with any hormonal index, nor did P4 correlate with any cycle variable (Liu et al., 1986).

Most of the studies did not find any relationship between any cycle or phase length and FSH (Broom et al., 1981; Landgren et al., 1980; Schipper, de Jong, & Fauser, 1998). However, Smith and colleagues (1985) observed lower FSH and LH during the luteofollicular transition (early follicular phase) in cycles with short luteal phases.

In sum, it is difficult to generate a gestalt of just how hormonal variation influences (or does not influence) the lengths of phases, and of the cycle, from these various studies. Landgren and colleagues (1980) noted that the combination of high initial E2 levels with low initial

Table 3. *Hormonal Variation versus Cycle and Phase Lengths*

	Estrogen		Progesterone		Gonadotrophins		Method	N	Source
	Follicular	Luteal	Follicular	Luteal	FSH	LH			
Cycle length	negative correlation with several indices	no relation	no relation	no relation	no relation		blood	39	Schipper et al. (1998)
Cycle length					no relation	positive correlation with several indices	blood	68	Landgren et al. (1980)
Follicular phase (days):									
length	negative correlation with mean of days 1 to 6, & with peak		no relation	no relation	no relation		blood	39	Schipper et al. (1998)
length					no relation	positive correlation with mean of days −3 to −7 before ovulation		68	Landgren et al. (1980)
length	no relation	no relation	no relation	no relation	no relation	no relation	blood	51	Broom et al. (1981)
>23 v. shorter	no difference						urine	258	Harlow et al. (2000)
>23 v. 12–23	higher total	higher daily mean	no difference	no difference			urine	411	Windham et al. (2002)
<12 v. 12–23	higher baseline & daily mean; lower total	no difference	higher	higher peak			urine	411	ibid
Luteal phase (days):									
length	negative correlation with peak	no relation	no relation	positive correlation with peak	no relation	no relation	blood	68	Landgren et al. (1980)
>14 v. 11–14	no difference	no difference	higher	higher total & peak			urine	411	Windham et al. (2002)
<11 v. 11–14	higher total	higher daily mean	no difference	lower total & peak			urine	411	ibid
<12 v. longer	no difference	higher		no difference in mid-luteal; lower in late luteal phase	no relation to mid-cycle levels; lower during luteo-follicular transition		blood	31	Smith et al. (1985)

223

LH levels was associated with relatively short cycles, and the opposite pattern with relatively long ones. Smith and colleagues (1985) concluded that a short luteal phase was not necessarily indicative of poor corpus luteum function or poor follicle development. The implications of these differences within and among women for fecundity remain uncertain. Because as many as half of all examined cycles from samples of *fertile* women have hormonal profiles whose shapes deviate markedly from "normative" profiles, several investigators question the suitability of so-called clinical "standards" of hormonal profiles (Alliende, 2002; Alvarado, Rivera, Ruiz, Arenas, & Rueda, 1988; DeCherney, Romero, & Polan, 1982; Renaud et al., 1980). Certainly the underlying hormonal profiles of cycles or phases, of any length, cannot be assumed to be comparable to any idealized standard.

Detecting the Occurrence and Timing of Ovulation

Perhaps the one clear conclusion from this discussion of cycle patterns is that the weakest analytical approach is aligning cycles by the first day of menstrual bleeding and then assuming peak hormone levels and/or peak fecundity occur during some span of days relative to that first cycle day. A somewhat more accurate approach (because luteal phase length is less variable than the follicular phase length) is aligning cycles on the first bleeding day of the subsequent cycle; in this case, for example, ovulation could be assumed to have occurred on cycle day −14 and the fertile window could be defined as days −19 to −14. Clearly, however, these defined variables are, at best, rough guesses.

A better approach is to align cycles on the day of ovulation, which would allow one to identify the fertile window (and the period of highest fecundity within it) with reasonable precision. Noninvasive methods for detecting the occurrence and timing of ovulation depend upon evaluating one or more biomarkers obtained either through a woman's self-examination or assayed in a body fluid. These include changes in ovarian steroid and/or gonadotrophin concentrations, and hormonally driven physiological changes, including an increase in basal body temperature and increased clarity and elasticity of cervical mucus. Table 4 lists the biomarkers in common use; each has advantages and drawbacks, and whenever possible, the use of multiple biomarkers is desirable (Campbell & Rockett, 2006; see O'Connor et al., 2006, for a discussion of the issues and a combined hierarchical method for detecting ovulation). It is critical to appreciate that some biomarkers indicate the likely *timing* of ovulation but cannot inform as to whether ovulation has actually occurred in a given cycle. For example, sufficiently high E2 prompts an LH surge that is followed by ovulation in about 30 hours (on average) *if* ovulation occurs. But an LH surge can and often does occur without a subsequent ovulation (i.e., the cycle is anovulatory). Therefore, the detection of an LH surge itself is *not* a marker of the occurrence of ovulation.

Covariation of Hormonal Biomarkers and Fecundity

Fecundity varies over the course of a cycle—the fertile window spans only a handful of days ending within a day after ovulation (Dunson et al., 1999, 2002)—and hence is associated with rises and falls in hormone concentrations. However, other than unusually low concentrations of either ovarian steroid, indicative of a failure to ovulate (which obviously precludes conception), it remains

Table 4. *Change in Biomarkers Relative to Day of Ovulation[a]*

Biomarker	Time Relative to Ovulation				Source
	Mean (days)	SD	Minimum	Maximum	
Rise of basal body temperature	1.1	2.0	−2	7	Flynn et al. 1988
Rise of salivary progesterone	1.5				Riad-Fahmy et al. 1987 Vitzthum et al. 2004
Rise of urinary pregnanediol-3-glucuronide (PdG)	0.2	3.0			Flynn et al. 1988
Rise of serum progesterone	−0.3	0.3	−1.3	0	Collins 1985
Peak of serum LH	−0.7	0.2	−1.9	−0.3	Collins 1985
Rise of serum LH	−1.3	0.3	−2.3	−1	Collins 1985
Peak of serum estradiol	−1.0	0.3	−2	0	Collins 1985
Peak of salivary estradiol	−1.5				Riad-Fahmy et al. 1987
Rise of serum estradiol	−3.4	0.9	−7	−2	WHO 1980
Rise of salivary estradiol	−5.0				Riad-Fahmy et al. 1987
Peak of urinary estrone-3-glucuronide (E_1G)	−1.3	1.9	−9	4	WHO 1983a
Rise of urinary estrone-3-glucuronide (E_1G)	−5.6	1.9	−11	−2	WHO 1983a
Peak of E_1G/PdG	−2.5	2.3	−10	0	WHO 1983a
Rise of E_1G/PdG	−7.2	2.8	−15	−2	WHO 1983a
Drop of E_1G/PdG	±2				Baird et al. 1991
Peak of fertile mucus	−0.4	2.2	−10	5	WHO 1983a
Peak volume of mucus	−2.0	1.0	−4.5	−1.5	Usala & Schumacher 1983
First day of fertile mucus	−5.1	2.6	−12	−1	WHO 1983a

[a]Modified from Campbell and Rockett (2006).

unclear which patterns and concentrations of hormones are necessary for full fecundity.

To address this issue, several studies have compared hormone concentrations in conception and nonconception cycles, but the findings are inconsistent. Two studies observed higher luteal-phase P4 levels in conception cycles (Baird et al., 1999; Stewart, Overstreet, Nakajima, & Lasley, 1993), but four studies did not (Li et al., 2001; Lipson & Ellison, 1996; Venners et al., 2006; Vitzthum et al., 2004). Li and colleagues (2001) reported that conception cycles had higher serum estradiol on the day that FSH peaked, but total urinary E2 metabolite levels in conception and nonconception cycles were similar. Likewise, urinary estrogen levels did not differ once other hormonal confounders were controlled (Baird et al., 1999), but other studies have reported higher estrogen levels in conception compared to nonconception cycles (Lipson & Ellison, 1996; Stewart et al., 1993; Venners et al., 2006).

As noted earlier, findings from Project REPA (Vitzthum et al., 2004) refute the hypothesis that variation in steroid levels across populations is necessarily associated with variation in fecundity (Ellison et al., 1993; Ellison, 1994; Lipson, 2001). The P4 concentrations that accompanied conception and implantation in these Bolivian women were similar to those in their own nonconception cycles and were also about 70% of those in the conception cycles of the Chicago sample (see Figure 7). Although P4 continued to be relatively lower throughout these pregnancies, the live births were of normal weight (i.e., >2.5 kg). In other words, the P4 concentrations in high altitude Bolivian women were adequate for normal conception and successful reproduction. Indicative of their high fecundability (the monthly probability of conception), between the ages of 20 and 30 years, these women averaged four live births. For each live birth, they breastfed on demand for one to two years, with a typical period of postpartum amenorrhea lasting about one year. Thus, adding time for gestation, the fertility levels of the studied Bolivian women do not suggest relatively lower fecundity. Certainly there is no evidence of a reduction in fecundity to only 70% of that of the Chicago women (if one were to assume a linear relationship between fecundity and P4 levels).

There are not as yet any other studies evaluating hormone levels at the time of conception in women living in arduous conditions. However, as would be expected given the findings from Project REPA, none of the endocrinologists or epidemiologists who reported ovarian steroid levels in Asian women to be ≤80% of the levels observed in "white" women in the United States and the United Kingdom (Bernstein et al., 1990; Dickinson et al., 1974; Key et al., 1990; MacMahon et al., 1974; Shimizu et al., 1990; Trichopoulos et al., 1984; Wang et al., 1991) suggested there was any difference in fecundity between women of Asian and European ethnicity, nor do there appear to be any other reports that such a difference might exist.

Hormones and Sexuality

It is widely assumed that endogenous hormone levels play an important role in women's sexual functioning,
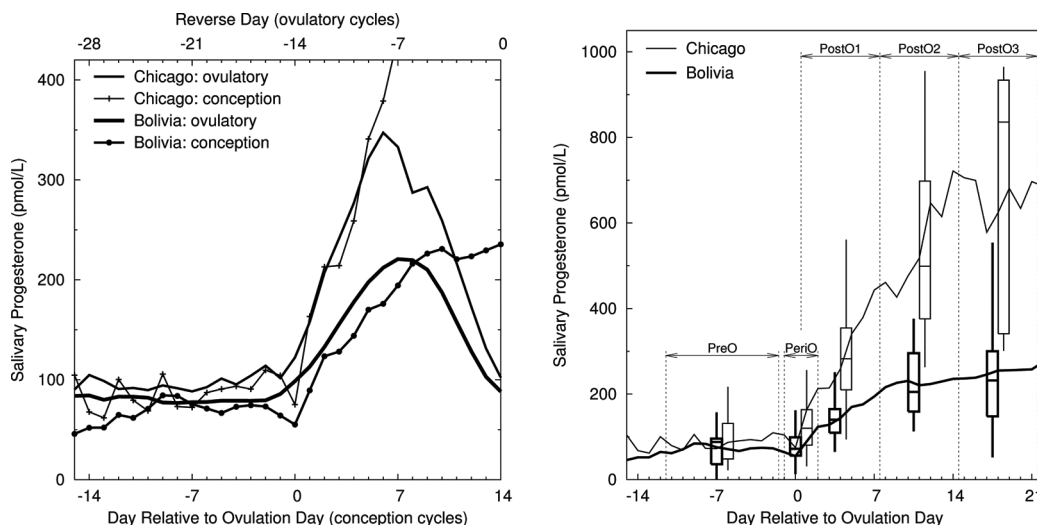
Figure 7. Left plot: Salivary progesterone [P4] profiles in ovulatory nonconception and conception cycles in women from Chicago and rural Bolivia (cycle days 1 to 28). Ovulatory cycles are aligned on the first day of the subsequent cycle (days numbered backward); conception cycles are aligned on the putative day of ovulation (day 0). P4 levels in ovulatory cycles are significantly lower in Bolivian women than in women from Chicago throughout the ovarian cycle; in each sample, conception and ovulatory cycles have comparable P4 levels. Right plot: Salivary P4 levels in conception cycles through 21 days postconception. Box plots display median, quartiles, and range of P4 indices corresponding to the range of days delimited by vertical dashed lines to the respective left and right of box plot. P4 levels do not significantly differ in women from Bolivia and Chicago during the follicular phase but are significantly different during and subsequent to ovulation and through implantation. PreO = preovulatory; PeriO = periovulatory; PostO = postovulatory (Reproduced from Vitzthum, Spielvogel, & Thornburg, 2004).

but there is surprisingly little rigorously collected evidence to support (or unequivocally refute) this expectation, in large part due to the difficulties of collecting the necessary data. Although it is patently obvious that human females can and do engage in sexual behavior (and other expressions of sexuality) on any and all days of the ovarian cycle and in the absence of any cycling (e.g., during pregnancy and lactation; after menopause), this fact does not preclude the possibility that the frequencies (and other features) of sexual functioning vary in ways that are at least partially influenced by varying hormone levels.

Extrapolating from innumerable demonstrations of a link between hormones and sexual behavior in nonhuman animals, many scholars have posited that comparable links exist in human females. More specifically, some authors have argued that natural selection favors increases in the expressions of sexuality as ovulation approaches so as to increase the probability of conception. A few evolutionary psychologists have even proposed that human females experience an estrus (a restricted period of heightened sexual activity) comparable to that which occurs in some other primates and mammals.

Before reviewing the available hormonal evidence on the posited relationships between human female sexuality and hormones, three key relevant points are worth noting. First, natural selection does *not* inexorably favor conception. As explained in the section on Evolutionary Models of Reproductive Functioning, there are always trade-offs in an organism's allocation of resources. Every opportunity to invest in reproduction carries costs, some of which may actually reduce lifetime reproductive fitness were the opportunity pursued (Vitzthum, 2008b, 2009; Williams, 1966). Thus, in some species under some conditions, possible advantages of an increase in sexual expressions near the time of ovulation may be outweighed by associated costs. This point is not simply theoretical. In fact (and this brings us to the second point), it is not the case that estrus is a feature found in virtually all primates (Dixson, 2009), nor are humans unique in having "concealed ovulation" (an absence of overt signaling accompanying ovulation).

It is beyond our scope to discuss the evidence and arguments regarding variability in the physical, psychosocial, and behavioral manifestations that do and don't accompany ovulation in nearly 400 nonhuman primate species (see Hrdy & Whitten, 1987; Dixson, 2009, 2012, for reviews). Nonetheless, it is valuable when considering the links between women's sexuality and hormones to recognize that neither estrus nor more subtle changes that could accompany ovulation are ubiquitous among primates, and that overt signaling carries costs as well as potential benefits (for models and evidence, see Alberts & Fitzpatrick, 2012; Deschner, Heistermann, Hodges, & Boesch, 2003, 2004; Hrdy & Whitten, 1987; Nunn, 1999).

The third and final point concerns the timing of the fertile window in women. Although it is the case that the fertile window is fairly narrow (about six days, ending within 24 hours after ovulation (Wilcox et al., 1995; see earlier section, *Phase durations, ovulation, and the fertile period*), it does *not* follow that the fertile window occurs during a narrow range of days during the menstrual cycle. To the contrary, because the timing of ovulation during a cycle is quite variable (see Figure 8a; based on data in Lenton, Landgren, Sexton, & Harper, 1984), women have a 10% or greater probability of being in their fertile window on *every* day from cycle days 6 through 21, and more than 70% of women are in their fertile window before cycle day 10 or after cycle day 17 (Wilcox, Dunson, & Baird, 2000). It is instructive to plot the daily probability of conception on each cycle day (see Figures 8b and 8c; based on data from Wilcox, Dunson, Weinberg, Trussell, & Baird, 2001) and to calculate the mean probability of conception (i.e., clinical pregnancy following a single act of unprotected
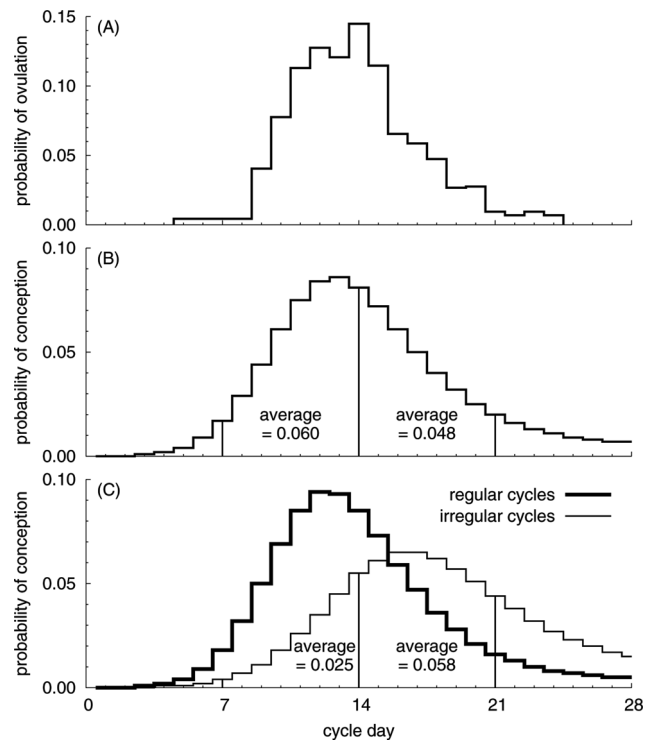
Figure 8. Panel A: The probability of ovulation by cycle day. Redrawn from Lenton, Landgren, Sexton, & Harper (1984). Normal variation in the length of the follicular phase of the menstrual cycle: effect of chronological age. *British Journal of Obstetrics and Gynaecology*, *91*, 681–684 (used with permission from John Wiley and Sons). Panel B: Daily probability of conception on each cycle day; mean probability of conception during cycle days 7 to 14 = 0.060 and during cycle days 14 to 21 = 0.048. Panel C: Daily probability of conception on each cycle day for women reporting regular cycles (thick line) and for those reporting irregular cycles (thin line); in latter sample, the average probability of conception during cycle days 7 to 14 = 0.025 and during cycle days 14 to 21 = 0.058. Panels B and C based on data from Wilcox, Dunson, Weinberg, Trussell, & Baird (2001). Likelihood of conception with a single act of intercourse: providing benchmark rates for assessment of post-coital contraceptives. *Contraception*, *63*(4), 211–215. Used with permission from Elsevier.

intercourse on a random day) during any span of cycle days. Notably, the average probability during cycle days 7 to 14 (0.060) is only 25% higher than that during cycle days 14 to 21 (0.048) (Figure 8b). The average probability during the first two weeks of the cycle (0.035) is only 16% higher than that during the next two weeks (0.030). Furthermore, in that subset of women who reported having irregular cycles, a not uncommon pattern (Burkhart et al., 1999; Creinin et al., 2004; Ferrell et al., 2005; Steiner et al., 2001; Vitzthum, Ellison, Sukalich, Caceres, & Spielvogel, 2000; Williams, 2006; see earlier section, *Cycle (segment) length and menses duration*), the average probability during cycle days 7 to 14 (0.025) is *less than half* of that during cycle days 14 to 21 (0.058) (Figure 8c).

Clearly, an investigator who assumes that ovulation occurs on cycle day 14, and then defines fecund and nonfecund periods relative to that assumed day of ovulation, is not cognizant of the substantial probability of conception during much of the ovarian cycle. For this reason, our focus in the remainder of this article is principally on those studies in which hormones were measured and/or the day of ovulation was estimated by a biomarker other than cycle day.

Hormones and Sexuality in Premenopausal Women

Studies in which endogenous hormones have been measured in healthy premenopausal women follow one (or sometimes both) of two study designs: (a) evaluating hypothesized covariation between indicators of sexual functioning and hormone levels over the course of the menstrual cycle, which addresses questions regarding normal variation, and (b) comparing hormone levels in a healthy control population to those in women diagnosed with some sexual dysfunction, which addresses sources of pathology. Because these two types of studies are addressing different types of questions, it is not surprising that their answers regarding the roles of hormones in sexual functioning are different.

Note that a fundamental challenge in ascertaining the links between sexual functioning and hormones is the construction of valid and reliable measures of "sexual functioning" (i.e., the outcome or dependent variable). Investigators have used different indicators, including sufficient lubrication, orgasmic ability and/or frequency, vaginal pulses, desire, arousal, initiation, masturbation, and responsivity. Discussion of measures of sexual functioning, an important issue not considered here, may be found in Arrington, Cofrancesco, and Wu (2004), Bancroft (2009), and Bancroft and Graham (2011).

Studies of Cyclically Varying Hormone Levels and Sexuality

Persky and his colleagues (Persky, Charney, et al., 1978; Persky, Lief, Strauss, Miller, & O'Brien, 1978)

were among the first investigators to measure hormones over the course of the ovarian cycle as part of an evaluation of putative cycle-dependent changes in sexual functioning. They followed 11 married women (aged 21 to 24 years) for three menstrual cycles, collecting twice-weekly hormonal measurements and interviews, and daily self-reports on degree of sexual gratification. There was no evidence of any significant differences in several aspects of sexual behavior (initiation, responsivity, avoidance, couple interaction, mood, gratification, frequency) across three study-defined phases of the cycle (follicular, ovulatory, luteal), nor were there any associations between changing E2 levels and these sexuality indicators (Persky, Charney, et al., 1978). In contrast, levels of various androgens (testosterone [T], dihydrotestosterone [DHT], dehydroepiandrosterone [DHEA], dehydroepiandrosterone sulfate [DHEAS], and androstenedione [A2]) were significantly associated with various measures of sexual functioning (Persky, Lief, et al., 1978; Persky et al., 1982). T and A2 were negatively correlated with sexual avoidance; T was positively correlated with initiation and with responsivity; and DHEA, A2, T, and DHT all showed significant intercorrelation scores with intercourse frequency but not with orgasmic frequency.

Another early study that collected frequent measurements of hormones (three to five times weekly during two sequential cycles in 14 healthy women) also found no relationship between concentrations of E2 or P4 and number of heterosexual activities or enjoyment of heterosexual activity (Abplanalp, Rose, Donnelly, & Livingston-Vaugh, 1979). Both the Persky and Abplanalp studies are commendable for the relatively high density of hormonal sampling and interviews. However, the sample sizes were very small, and the then-available statistical methods did not take into account either the lack of statistical independence of repeated measurements from a single woman or the increased false-positive rate that is a consequence of conducting numerous statistical tests. Multilevel modeling is a much more powerful tool for dealing with these types of data (West, Welch, & Galecki, 2006), and different results may be found if the studies were to be replicated (or improved upon) using advanced statistical approaches.

Although reports subsequent to those from Persky and his colleagues also suggested that T levels may be influencing some aspects of women's sexuality, the findings were inconclusive. In a study of premenstrual syndrome (PMS), women kept daily diaries of sexual activities, sexual interest, and other psychological variables; measurements of T were made at least thrice weekly (Bancroft, Sanders, Davidson, & Warner, 1983). In those women who masturbated ($n = 21$), levels of midcycle T and whole-cycle mean T (which are highly correlated) were each significantly correlated with frequency of masturbation. But T concentrations were *not* correlated with other indicators of sexuality. In a

study specifically designed to test the findings reported by Persky, Lief, and colleagues (1978), average intercourse frequency (over a period of about 100 days) in a sample of 43 married women was significantly correlated with midcycle indices of total T and free T but *not* with baseline or mean levels of either (Morris, Udry, Khan-Dawood, & Dawood, 1987). Elevations of hormone level were not associated with the days before, of, or after reported coitus. The authors noted, "No theoretical mechanism is self-evident to explain why a midcycle value of T predicts average frequency of intercourse in a sample of married women when baseline or average levels do not" (p. 33).

As noted, it has been repeatedly proposed that the periovulatory period should be associated with heightened female sexuality. But as of 1980, only 8 of 32 studies had reported peaks of sexual behavior in women at ovulation; peaks were also reported premenstrually (17 studies), postmenstrually (18 studies), and menstrually (4 studies) (Schreiner-Engel, Schiavi, Smith, & White, 1981). Nearly all of these studies lacked reliable biomarkers of ovarian functioning. In two later studies using such biomarkers, indices of sexuality appear to be more influenced by feelings of well-being than by ovulation. In the study by Bancroft and colleagues (1983) discussed previously, indicators of positive mood and sexuality were correlated in a subsample of 13 women lacking significant cycle-attributed problems, but there was no evidence of increased sexual interest or activities associated with ovulation. There appeared to be a midfollicular (postmenses) peak in sexual activity with a partner, but this increase was not statistically significant. Contrary to these findings, coital frequency in 25 Zimbabwe couples was significantly greater (than average coital rate) on the day of the onset of the LH surge ($p < 0.05$) but was not greater on either the day of the LH peak or during the ovulatory phase (defined in several ways) (Hedricks, Piccinino, Udry, & Chimbira, 1987). A third study produced yet another pattern (Dennerstein et al., 1994). Sexual interest [SI] in 168 Australian women was higher in the postmenses *and* periovulatory phases (the two phases did not differ in SI) than in the luteal (postovulatory through premenstrual) and menstrual phases ($p < 0.0001$). However, as nearly all the women had reported symptoms of PMS, this pattern in SI may be attributable to the effects of feeling poorly during some part of the luteal phase or to the effects of approaching ovulation or to neither. Notably, feelings of well-being and sexual interest were significantly correlated ($r = 0.29$, $p < 0.0001$), but no significant correlations were found between either mood variable and hormone (urinary estrogens and pregnanediol) indices (Dennerstein et al., 1994).

In another ambitious study (166 women completed daily checklists, questionnaires, and menstrual cycle charts for three cycles, as well as an exit questionnaire), Burleson, Gregory, and Trevathan (1995) tested several hypotheses regarding changes in sexual functioning across the menstrual cycle. Although hormones were not measured, the occurrence and timing of ovulation was determined from self-monitoring of basal body temperature and cervical mucus. They found that sexual activity was least likely to occur during the midluteal phase of an ovulatory cycle (when P4 is often highest) and that anovulatory cycles were associated with higher rates of sexual activity—patterns they attributed to a suppressive effect of P4. As intriguing as these results are, one caveat is that these women were specifically trained to record biomarkers of their own fecundity (i.e., basal body temperature and cervical mucus consistency). Thus, if a study participant recognized that she had ovulated, she may have been less inclined for a few days after ovulation to have sex for fear of pregnancy; knowledge of the very low probability of fecundity after ovulation (Wilcox et al., 1995; Dunson et al., 1999, 2002) was neither appreciated nor widespread in the scientific and lay communities at the time of the study. However, if she recognized that she was not fecund (i.e., that her cycle was anovulatory), she may have felt freer to have more sex. One of the clear evolutionary advantages of our big brains is that knowledge can trump hormones.

The most convincing evidence that ovulation is associated with a higher frequency of sexual intercourse comes from a study of 171 ovulatory cycles in 68 heterosexually active women who were using an intrauterine device or had had a tubal ligation (Wilcox et al., 2004). Intercourse and menstrual bleeding were recorded daily for up to three months, and day of ovulation was determined with hormonal biomarkers. As previously discussed, the fertile window is about six days long (principally spanning the day of ovulation and the previous five days; see Figure 6; Dunson et al., 1999, 2002; Wilcox et al., 1995). In these 68 women, the six consecutive cycle days with the highest coital frequencies corresponded exactly to the six days of the fertile window. These data are particularly compelling because during these six days the coital frequency was lowest at the beginning of the fertile window, rising monotonically over the subsequent days and peaking just before and on the day of ovulation (Wilcox et al., 2004: Figure 1). The authors pointed out, however, that these data do not permit distinguishing the causal direction: increases in intercourse may be a consequence of approaching ovulation, or intercourse may accelerate the occurrence of ovulation. In support of the latter mechanism, supplemental analyses of 867 cycles from 285 women suggested that those who preferentially had sex on weekends (Fridays through Sundays) had a significantly higher probability of ovulating on Sunday-Monday-Tuesday than would be expected by chance. However, such a pattern may also arise from weekend-associated patterns in stress, sleep, or other behaviors that might influence ovulation (Wilcox et al., 2004).

In addition, Dobbins (1980) has presented a model supporting the argument that avoidance of sex during menses (which is known to occur widely across the world; see Brewis & Meyer, 2005) can result in heaping of sexual activity at about the time of ovulation without there being any hormonal influences. It is possible that the combination of menses-associated avoidance of and weekend-associated preferences for intercourse and changes in behaviors may all intersect to create heaping independent of ovulation. Further study is needed.

Studies Comparing Sexually Functional and "Dysfunctional" Women

Based on a single measurement during the first 10 days of the menstrual cycle, Stuart, Hammond, and Pett (1987) compared T levels in 11 controls and 11 women diagnosed with inhibited sexual desire (ISD) based on multiple criteria. Both groups of women proved to have T and prolactin levels within normal ranges, and mean levels of these hormones were not significantly different.

Schreiner-Engel, Schiavi, White, and Ghizzani (1989) compared 13 demographically similar controls to 17 women diagnosed with hypoactive sexual desire disorder (HSDD) based on *DSM III-R* criteria. Assays of blood samples drawn every three to four days during a single cycle demonstrated that all of the women in both groups had E2, P4, and T levels within normal ranges. Moreover, the groups did not differ in the average levels of these hormones.

Davis, Davison, Donath, and Bell (2005) investigated the hypothesized relationship between androgen levels and self-reports of sexual functioning, determined by a self-administered questionnaire, in a sample of 343 Australian women (aged 18 to 45 years). A single blood sample for hormone assays (DHEAS, A2, and T) was drawn after a morning fast on some day between cycle day 8 and the onset of the next menses. Those women with a low domain score for sexual desire, or arousal, or responsiveness had a significantly higher risk of having a DHEAS level below the 10th percentile. Nonetheless, the majority of women in the sample with low DHEAS levels did not have low sexual function. The authors concluded their data "suggest that sex steroids influence female sexual function, but that there is no serum androgen level that defines female androgen insufficiency" (p. 96).

Basson, Brotto, Petkau, and Labrie (2010) compared 124 healthy women (>35 years of age, mean 48.3 years) to 121 women with HSDD (mean age 52.1 years) recruited from a sexual medicine center who presented with general desire and arousal disorders that were acquired after age 35. A single blood sample was collected from each woman, scheduled on cycle day 8 through day 10 for those who were not menopausal. Analyses were adjusted for age but not for menopausal status. There were no significant differences between the groups in levels of E2, estrone, estrone sulfate, or T. In contrast, levels of androstene-3β, 17β-diol and DHEAS were significantly lower in women with HSDD. However, as was the case in the samples studied by Davis and colleagues (2005), some women had low levels of DHEAS without corresponding sexual dysfunctions.

Riley and Riley (2000) compared 15 healthy controls (aged 18 to 45 years) to 15 patients classified as having "a long standing absence of sexual drive" based on several criteria. Testosterone, 5-dihydrotestosterone (5-DHT), E2, SHBG, and prolactin were measured in a single blood sample taken within 48 hours of a positive result from an in-home ovulation test kit. For those (five patients and three controls) with no positive test (i.e., cycle apparently anovulatory), the sample was taken within 48 hours of the last of five test kits. Measurements in anovulatory cycles were not excluded nor controlled for in statistical analyses. Most of the women had hormone levels that were within normal ranges, and none of the hormones differed between the two groups other than significantly lower levels of free T in the dysfunctional group. Within the control group, average daily thoughts about sex correlated to total T, free T index, and free 5-DHT index. In the patient group, coital frequency and total T level were correlated. However, none of the correlation analyses controlled for age, which might explain changes in androgens and sexual behavior independently of one other.

Turna and colleagues (2005) compared 20 premenopausal (21 to 51 years old) and 20 postmenopausal (48 to 60 years old) controls to 20 premenopausal (24 to 51 years old) and 20 postmenopausal (45 to 70 years old) patients who had decreased libido for at least six months and were in "stable relationships." Hormones (T3, T4, TSH, E2, SHBG, total T, and DHEAS) were assayed in a single blood sample drawn during days 8 through 15 of the menstrual cycle. The authors reported that significant differences were found between the patient and control groups in levels of total T, free T, and DHEAS. However, a closer examination of their data suggests that the distribution of ages within each group was dissimilar: eight (20%) of the women in the control group were 20 to 29 years old, but only three (7.5%) fell into that age range in the patient group, and the maximum age in the patient group was 10 years older than in the control group. Yet the statistical analyses of hormonal differences between the groups did not control for either age or menopausal status.

Summary

Several studies that included measurements of hormone concentrations failed to find associations between these and indicators of sexuality in cycling women (Abplanalp et al., 1979; Dennerstein et al., 1994; Persky, Charney, et al., 1978; Schreiner-Engel et al., 1981). There are perplexing and inconsistent reports that

concentrations of one androgen or another may influence female sexuality (Bancroft et al., 1983; Morris et al., 1987; Persky, Lief, et al., 1978; Persky et al., 1982). In light of the marked interindividual variation in hormone levels now known to exist (see earlier section, Hormonal Biomarkers of Fecundity), it may simply be that the sample sizes needed to detect hormone-sexuality associations are larger than those in the studies to date. There is one very convincing report of an association between ovulation and coital frequency (Wilcox et al., 2004). But the causal direction remains to be determined, and behavioral repertoires that are independent of hormonal influence could be contributing to the synchrony between intercourse and ovulation (Dobbins, 1980). There is substantial support for the argument that ovulation is, in fact, successfully concealed in humans. For example, Hadza hunter-gatherers in Tanzania believe that the period of greatest fecundity is at the end of menstrual bleeding rather than at some midpoint of the ovarian cycle (Marlowe, 2004), and analyses of demographic data from 13 countries detected reduced coital frequency during menses but no increase at ovulation (Brewis & Meyer, 2005). Nonetheless, it remains plausible that there are sexually related behavioral and/or psychological states which vary in concert with hormonal changes during the ovarian cycle. But such hormonal influences are subtle at most, as evidenced by the difficulties in finding them.

The majority of studies evaluating women with sexual dysfunctions (however defined) found no compelling evidence of clear hormonal differences between these women and healthy women. Interpretations of the findings from the two studies reporting some hormonal differences between healthy and nonhealthy women are hampered by small sample sizes and statistical limitations. Moreover, the findings from Project REPA and other studies (discussed previously) suggest that fecundity does not necessarily covary with absolute levels of progesterone and estrogens but rather covaries with changes in hormone levels (as is the case with seasonally varying workloads). Likewise, if hormones and pathologies in sexual functioning are linked, changes in hormone levels may be far more important than absolute hormone levels. If so, the study designs to address this question will need to be very different from the approaches that have been used thus far.

Menopause, Sex, and Hormones

As populations age in industrialized countries, increasing attention has been paid to the sexual lives of peri- and postmenopausal women (to avoid wordiness, the present discussion hereafter will use *menopause* to refer to both peri- and postmenopause). Although most researchers have assumed that sexual activity declines with age, evidence of this is inconsistent, and many menopausal women report that sex remains an important part of their lives (Avis et al., 2009; Huang et al., 2009; Winterich, 2003). Likewise, findings are contradictory as to whether those older women who do experience declines in sexual functioning might be more susceptible to the hormonal changes that occur during menopause (Bancroft, 2009). Perhaps the only conclusion agreed upon by most investigators is that female sexuality, even in (or perhaps, especially in) menopausal women, is the outcome of a complex set of psychosocial *and* biological factors.

Debate is ongoing as to whether androgens, particularly T, or estrogens have the greatest influence on female sexual functioning (Persky et al., 1982; Riley & Riley, 2000; Turna et al., 2005). Although the currently available evidence is far from conclusive, T may influence libido in menopausal women (Bachman & Leiblum, 1991; McCoy & Davidson, 1985), whereas estrogen is necessary for genital maintenance (Traish, Botchevar, & Kim, 2010) and for normal blood flow and lubrication, both of which are important for comfort and satisfaction during sexual intercourse (Bachman & Leiblum, 2004).

Some of the barriers to achieving a clearer understanding of the links between hormones and sexual functioning in menopausal women arise from methodological challenges. These include cross-sectional designs that are unable to capture the hormonal changes occurring with menopause (Bancroft, 2009; Dennerstein, Randolph, Taffe, Dudley, & Burger, 2002; Hayes & Dennerstein, 2005); the difficulties in measuring low levels of hormones, particularly androgens, with current technologies (Bancroft, 2009; Wierman et al., 2010); the use of a single hormone measurement as a proxy for the entirety of hormonal production over the course of an ovarian cycle (Vitzthum, 2009); the inclusion of study participants who are taking hormonal supplements, in whom supraphysiological hormone levels may not accurately represent the effects of endogenous hormone levels (Bancroft, 2009; Bancroft & Graham, 2011; Bachman & Leiblum, 2004; Cawood & Bancroft, 1996); and selection bias in the study samples, arguably among the most difficult issues to address.

Probably the largest selection biases arise from two sources. First, because hormone replacement therapy (HRT) is widespread in many industrialized countries (where nearly all studies have taken place), fewer menopausal women in these populations are available for studies of endogenous hormones. More important, women with the most extreme hormonal changes may experience more severe menopausal symptoms and thus may be the most likely to use HRT (McCoy & Davidson, 1985). If so, study participants not taking HRT may have fewer menopausal symptoms and perhaps lower absolute and/or fluctuating changes in hormone levels compared to the larger population of menopausal women. Second, most studies have primarily recruited

Caucasian women who are at least middle class and can afford health care; many studies have recruited from physicians' or counselors' offices. Women in different cultures likely have different perceptions of menopause, and it is also now well established that women with different sociocultural backgrounds and/or lifestyles can have dramatically different natural hormone levels (Briggs & Briggs, 1972; Haiman et al., 2002; Key et al., 1990; Vitzthum, 2008a, 2008b, 2009; Vitzthum et al., 2002; Windham et al., 2002). It stands to reason that this hormonal variation would continue through the menopausal transition.

The task of operationalizing different aspects of sexuality and designing suitable data collection instruments appears to be even more daunting than recruiting more representative samples or adequately measuring hormones. Clearly the use of different instruments, and different definitions of the same term, makes cross-study comparability very difficult (Hayes & Dennerstein, 2005). Waters may become even murkier if the reliability and validity of some indicator of sexual functioning is uncertain (Gracia et al., 2004). Reducing complex behavioral and emotional repertoires to a single score is problematic (Geirhart, 2006), but so are multiple analyses of individual but highly correlated indicators of sexual functioning. The use of sophisticated multivariate and multilevel analyses is probably called for, along with the larger samples necessary for these methods to be valid.

Given these challenges, it is not surprising that there are but a handful of longitudinal studies of menopausal women who are not using HRT. Although their findings are inconsistent, these studies serve as a guide for what needs to be undertaken to better understand how endogenous hormones may influence sexual functioning in menopausal women.

Analyses from perhaps the first longitudinal study focusing solely on perimenopausal women not using HRT suggested that declines in coital frequency occurring over the course of the menopausal transition were better correlated with declines in T than in E2 concentrations (McCoy & Davidson, 1985). However, the somewhat weaker correlations with E2 may be attributable to the use of a single measurement during the very early follicular phase as a proxy for the entirety of the E2 concentrations during an ovarian cycle. More frequent measurements of E2, or even a single measurement more likely to represent peak E2 levels (e.g., periovulatory or midluteal collection), may have yielded different conclusions. Although quarterly interviews and hormone measurements, continued for each woman until several months after her last menstrual period, were a significant advance over cross-sectional studies, the small sample (16 women) and unavailability of sophisticated statistical techniques better suited to the structure of the data (e.g., multilevel modeling) makes it difficult to extrapolate the findings to a larger population or to select one hormone as more influential in sexual functioning than another.

A larger ($n = 141$), albeit shorter (four weekly interviews), repeated measures study (Bancroft & Cawood, 1996; Cawood & Bancroft, 1996) found no relationship between androgens or estrogen and sexuality or well-being, having controlled for potential confounders (age, menopausal status, BMI, smoking). However, DHEA significantly predicted measures of well-being.

Dennerstein et al. (2002) recruited 226 Australian women in the early stages of the menopausal transition and followed them with annual visits for eight years, by which time all were postmenopausal. In addition to interviews and questionnaires, blood for hormone measurements was drawn between days 4 and 8 of the menstrual cycle (based on participant menstrual diaries). Relying on responses to the Personal Experiences Questionnaire (SPEQ), a woman was classified as sexually dysfunctional if her SPEQ total score ≤7. In the first year of the study, 42% of the sample was classified as sexually dysfunctional; by the eighth year, this percentage had more than doubled to 88%. Based on their analyses, the authors concluded that this increase in sexual dysfunction was related to declining estrogen levels but not to changes in androgens. One is left pondering, however, on the suitability of a measurement that classifies nearly all otherwise healthy women as "dysfunctional" (see Geirhart, 2006).

Gracia and colleagues (2004) has conducted perhaps the largest and longest study yet on sexual functioning and endogenous hormones of women in the years leading to menopause ($n = 326$ women, 30 to 47 years old; every 8 months, participants were visited twice (one month apart and intended to coincide each time with the first six cycle days), over the course of 4 years = 12 visits). Absolute concentrations of T, DHEAS, E2, LH, and FSH were not associated with sexual dysfunction. However, women in whom T levels were relatively stable over the course of the study were significantly less likely to report a decline in sexual interest compared to those with substantial fluctuations in T (but note the authors' own caveat that "the reliability and validity of the question assessing libido has not been tested" [p. 149]). Their finding that declines in libido were associated with the magnitude of fluctuations, rather than the absolute levels of T, is unique (and is also consistent with the argument made earlier in the present paper that change in hormonal concentrations may be more biologically salient than absolute magnitude).

Are We Asking the Right Questions about Menopausal Women's Sexuality?

Intriguingly, regardless of which (or whether any) hormone was a significant factor in explaining variation in sexuality in menopausal women, the studies discussed also found that several other variables were either as important as or more important than these biomarkers. Cawood and Bancroft (1996) reported that "tiredness"

was positively and significantly predictive of depression and was the primary negative predictor of well-being; the most important predictors of sexual well-being in their sample were variables related to the "quality of the relationship with the partner" and their "state of well-being." Vaginal dryness, depression, and children living at home (Gracia et al., 2004), and aging (Dennerstein, Dudley, & Burger, 2001) were also found to be important predictors of sexual dysfunction. In a recent review of relevant studies, Bancroft (2009) concluded that mental health and wellness are more important than either age or menopausal status in predicting sexual functioning.

It is critical to keep in mind that hormones are not produced in a biological vacuum, nor is there a unidirectional relationship between hormones and behavioral or emotional states. Both external and internal signals alter the production of and response to hormones (Christiansen, 2001). Thus it should come as no surprise that the quality of her relationship has as much to do with a menopausal woman's sexuality as does any level of any hormone. Moreover, the hormone level itself may be a consequence of relationship quality. For example, experiencing fewer daily stressors, or better defenses against those stressors, typically translates into lower cortisol levels, which can lead to higher T levels (Christiansen, 2001).

It's also worth reiterating that human behaviors, emotions, and relationships are as much a product of cultural values as they are of biological substrates. Although still far too scant, current evidence from the studies of mostly middle-class Caucasian women suggests that relationship and general health issues may be among the most important factors in these women's sexual functioning. It may be that these or perhaps other factors are significant for women from other socioeconomic or cultural groups (Avis et al., 2009; Cain et al., 2003; Huang et al., 2009).

Premenstrual Syndrome and Ovarian Hormones

The very existence of PMS is a persistent debate in the medical and sex research fields. The belief that women are subject to cyclical changes in well-being as a consequence of their normal bodily functions can be traced from antiquity to today. For example, women's perceived deficiencies and problems have long been tied to their uteruses (hysteria, a "psychological condition" long considered unique to women, is derived from the Greek word for uterus, *hystera*). In contrast, men's emotional and physical states are rarely, if ever, seriously blamed on their reproductive organs (Marks, 2001; Martin, 1987). Although such cultural biases regarding women's bodies and functioning do not themselves negate women's experiences of premenstrual symptoms, it is instructive to recognize that the

complaints most often cited (irritability and hostility) are the antithesis of stereotypical feminine virtues (Gottlieb, 1988). The frequent linking of women's physical and emotional/mental states in medical literature serves to reinforce the assumption that women's minds as well as their bodies are tightly bound to their hormonal states. Given the historical tensions between women and the (mostly) male physicians who have attempted to mold their bodies and minds to suit these virtues, the literature on PMS is best read cautiously.

The hypothesized relationship between the premenstrual portion of the ovarian cycle and premenstrual "syndrome" assumes that there are links between normal changes in physiological processes and changes in well-being. More explicitly, the hormone etiology hypothesis for PMS predicts a correlation between changes in hormone levels and changes in symptoms of illness. Here we are leaving aside a consideration of how to define PMS (see Kadian & O'Brien, 2012 for one impressively complex classification scheme), even though the extensive debates on this issue raise fundamentally important questions regarding study design and measurement error. Rather, our focus is on the work that sought to measure the hormonal deficiencies, excesses, and/or fluctuations which are still widely thought to cause PMS. As is discussed in the next section, there is no unequivocal evidence for such a relationship.

The Evidence

Most studies of premenstrual symptoms rely on imprecise estimates of the timing of ovulation and/or phases of the menstrual cycle rather than on measurements of hormone concentrations themselves. As previously discussed, such nonhormonal biomarkers are poor proxies for evaluating the hormonal correlates of PMS because of substantial individual variation in hormone patterns, cycle length, and phase lengths. Such studies are not further considered here.

Frank's 1931 report is widely acknowledged as one of the earliest efforts to investigate a hormonal etiology for PMS. However groundbreaking and ambitious for its time, this work suffers from serious methodological errors. Despite Frank's claims, too little was known about the menstrual cycle to accurately identify an underlying hormonal relationship with the PMS symptoms reported in the paper (which, incidentally, spanned from "tension" to "almost crazy" to "impossible to live with"). Frank's evaluation of hormonal patterns during the ovarian cycle was based on injecting dried human blood into mice, then examining the changes in cells in vaginal smears from those mice (Frank & Goldberger, 1928). Not surprisingly, most, if not all, of the conclusions regarding female reproductive hormones reported in Frank and Goldberger (1928) are now known (based on modern hormone sampling and assaying technologies) to be grossly oversimplified or even erroneous.

Advancements in technologies have not, however, led to a clear understanding of what roles, if any, cyclically varying hormone levels play in PMS. Two influential early studies used sensitive assays to compare hormone concentrations in women with cycle-attributed symptoms to those in asymptomatic controls, and reported differences in estrogens, P4, FSH, and SHBG on some, but not all, of the days preceding menses (Bäckström & Carstensen, 1974; Bäckström, Wide, Södergård, & Carstensen, 1976). In the final six days of the luteal phase, P4 was lower on three of six days and estrogens were higher on four of six days in a sample of 10 women with cycle-attributed anxiety, compared to eight controls (Bäckström & Carstensen, 1974). The second study, of the final 10 days of the luteal phase, compared 15 women experiencing slight to moderate cycle-attributed symptoms and 17 asymptomatic controls (Bäckström et al., 1976). In the symptomatic group, P4 was lower on six of 10 days and the estrogen peak and profile appeared to be shifted to later in the luteal phase with the result that estrogen levels were significantly lower on two days in the early luteal phase and higher on the five days preceding menses. LH and albumin did not differ between the groups, but FSH was significantly higher on four of 10 days and SHBG was higher on one day in the symptomatic group. Although provocative, these studies have significant methodological and statistical limitations that could not be appreciated 40 years ago, including age differences in study samples, possible inclusion of anovulatory cycles, autocorrelation across cycle days, and repeated statistical testing without suitable adjustments. As noted in the following section, later studies have not, in fact, confirmed these early reports regarding links between concentrations of P4 and/or estrogens and PMS.

The first study to evaluate daily-reported mood and enjoyment of activities in conjunction with frequent measurements of P4 and E2 (three to five times weekly during two sequential cycles) found no evidence of any cycle-phase associated or hormone-level associated variation in emotional states in a sample of 14 healthy U.S. women (Abplanalp et al., 1979). The authors were cautious in their claims and specifically noted the small and carefully screened study sample that was absent of any women with significant premenstrual difficulties. Despite the apparent homogeneity (with respect to cycle length and regularity, physical and psychological health, and other potential confounders) of their study participants, there was considerable interindividual variability in hormonal and psychological measures during each defined cycle phase.

Findings from a larger study (Sanders, Warner, Bäckström, & Bancroft, 1983; Bäckström et al., 1983) of European women who reported that they did or did not experience significant cycle-associated problems are consistent with and expand upon the findings from Abplanalp and colleagues (1979). Hormone measurements were made at least thrice weekly, and women provided daily self-ratings of psychological variables. As might be expected, a composite indicator of "well-being" did not vary significantly with cycle phase in asymptomatic women ($n = 16$). In contrast, compared to other phases of the cycle, "well-being" was lowest in the late-luteal (premenstrual) and early follicular (menses) phases in women who had reported cycle-associated symptoms ($n = 18$, $p < 0.0009$) and was lowest during the late luteal phase in clients at a clinic for PMS symptoms ($n = 19$, $p < 0.0001$). All three samples experienced statistically significant phase-associated variation in an indicator of "physical distress" (being greatest during the late luteal phase), but none of the samples experienced phase-associated variation in indicators of "sexuality" or "outwardly directed emotions." The authors suggested that phase-associated changes in physical symptoms may occur in many cycling women and in some women may be interacting with other (nonhormonal and/or hormonal) factors to produce phase-associated changes in "well-being." Interestingly, compared to the two other samples, women in the clinic sample had had more pregnancies and miscarriages, were more likely to have small children and be working at home, and had a higher rate of adverse reactions to oral contraceptives. However, there were no significant differences across the three samples in the mean concentrations of E2, P4, A2, or T, nor were there significant hormonal differences related to the degree of mood change in these three samples or in analyses of selected subsamples (Bäckström et al., 1983). Noting that women with the most severe symptoms ($n = 12$) had similar temporal patterns for changes in E2, P4, and some mood and physical variables, the authors suggested that variation in sensitivity to changes in hormone concentrations may be more important than absolute hormone concentrations per se (Bäckström et al., 1983).

Subsequent studies produced findings comparable to those from the study reported in Abplanalp and colleagues (1979) and those reported in Sanders and colleagues (1983) and Bäckström and colleagues (1983). Rubinow and colleagues (1988) carefully screened participants for meeting DSM-III-R diagnostic criteria, controlled for the large interindividual variation in baseline hormone levels, and used suitable statistical methods. They reported that ovulatory cycles from 17 women diagnosed with "peri-luteal phase dysphoric disorder" and nine control women did not differ in levels of E2, P4, T, DHT, DHEAS, cortisol, prolactin, SHBG, FSH, or LH. Likewise, a comparison of 11 women who met diagnostic criteria for PMS did not differ in levels of P4, total-T, A2, DHEAS, or SHBG (measured at three points during a cycle) compared to 11 age-matched controls with no premenstrual symptoms (Eriksson, Sundblad, Lisjo, Modigh, & Andersch, 1992).

233

However, this was the first report that levels of free T were consistently and significantly higher in those with PMS than in the control samples. Also, DHEAS was higher around ovulation ($p < 0.05$) and 17-OHP levels were higher during the luteal phase ($p < 0.05$) in the PMS sample. Noting that there was no difference between the samples in either SHBG levels or the total-T/SHBG ratio, and that other factors (e.g., stress associated with PMS) may underlie the elevated free-T levels, the authors cautiously suggested that their findings pointed to a role for androgens in PMS.

In a sample of 168 women (of whom 65 were classified as having PMS and 51 as having "menstrual distress"), well-being varied across cycle phases ($p < 0.0001$, lowest in the late luteal/early follicular phases). However, there were *no* significant correlations between well-being and hormonal variables (estrogens and PdG measured in daily 24-hour urine samples; Dennerstein et al., 1994). Likewise, there was *no* temporal correspondence between an index of adverse mood and levels of E2 or P4 in a carefully screened sample of 19 healthy women (Van Goozen, Wiegant, Endert, Helmond, & Van de Poll, 1997). There were also *no* differences in levels of P4, A2, T, DHEAS, SHBG, or cortisol in subsamples of those who reported some cycle-attributed symptoms (but had not been diagnostically evaluated for PMS, $n = 11$) and those who did not ($n = 8$). However, those who did *not* report symptoms had significantly higher E2 ($p < 0.01$) and higher E2/P4 ($p < 0.01$), findings that contradict hypotheses that PMS is a consequence of elevated E2 and/or elevated E2/P4. Other investigators have also failed to find significant synchronous relationships between P4 or E2 concentrations and cycle-attributed symptoms but have suggested that it is the magnitude and/or rate of change in the levels of these hormones some days before the appearance of symptoms that is etiologically salient (Halbreich, Endicott, Goldstein, & Nee, 1986; Redei & Freeman, 1995). Although there may be merit in this hypothesis, small sample sizes and analytical challenges have prevented any definitive conclusions.

In sum, of those studies in which investigators measured hormone concentrations throughout the course of the ovarian cycle in samples of women with and/or without cycle-attributed symptoms, all failed to find unequivocal evidence of any relationship between baseline and/or fluctuations in hormone levels and changes in mood or physical indicators. Given the substantial variability in hormonal, psychological, and physical measures reported in these and other studies, it could be that the sample sizes were too small to have sufficient statistical power to discern any phase- or hormone-associated variation in these indicators. Several analyses also neglected to account for autocorrelation among observations, repeated hypothesis testing, and potential confounders including age. These challenges can now be better addressed with multilevel modeling (West et al.,

2006), a statistical approach not readily available until relatively recently.

However, even with the use of appropriate statistical tests, there are reasons to be skeptical that more advanced analytical methods might reveal robust and consistent hormone-symptom associations in women with cycle-attributed changes in psychological and physical indicators. For example, ovarian suppression with leuprolide (an agonist analogue of GnRH) reduced symptoms in 10 of 18 women diagnosed with PMS (Schmidt, Nieman, Danaceua, Adams, & Rubinow, 1998). Symptoms returned in those 10 upon replacement administration of either E2 or P4 while continuing on leuprolide. Mood in a control sample (15 women with "minimal" cycle-attributed symptoms) did not change during either experimental regime. The authors emphasized the different response to the experimental protocol in women with and without PMS, concluding that "normal plasma concentrations of gonadal steroids can trigger an abnormal response—deterioration in mood state—in susceptible women" (Schmidt et al., 1998, p. 216). On the other hand, we find it particularly notable that 44% (8 of 18) of the women who had met the same strict diagnostic criteria for PMS had *no* change in cycle-attributed symptoms despite ovarian suppression.

Limited cross-cultural evidence suggests that women throughout the world may experience changes in well-being during the premenstrual phase, but the prevalence and reported symptoms vary widely (Chau, Phil, Chang, & Chang, 1998; Sadler et al., 2010; Takeda, Tasaka, Skata, & Murata, 2006; van den Akker, Eves, Service, & Lennon, 1995). Although little is understood about the reasons for interpopulational variation in cycle-attributed symptoms, they appear to be influenced by race and/or ethnicity (Takeda et al., 2006), education and stress levels (Sadler et al., 2010), susceptibility to major depression (Treloar, Heath, & Martin, 2002), and, tellingly, culturally transmitted fears about menstruation (van den Akker et al., 1995).

Evolutionary Hypotheses

Despite the lack of clear evidence that normal endogenous hormone levels influence PMS, suggestions that there may be evolutionary explanations for PMS persist (Reiber, 2008, 2009; Rosseinsky & Hall, 1974; Yonkers, O'Brien, & Eriksson, 2008). However, fundamental flaws with such models render them implausible. The most significant of these flaws is that extended periods of monthly cycling is a very new phenomenon in human history. Most (if not all) evolutionary hypotheses regarding PMS assume frequent ovulatory cycles in order for natural selection to act on the corresponding behavioral and emotional phenotypes. But as already noted, models that assume nonstop machine-like ovarian functioning lack biological validity. For

234

most of our history, human females have spent most of their lives not cycling, either as a result of pregnancy or lactational amenorrhea; even when cycling does occur, ovulation is not guaranteed (Strassmann, 1997; Vitzthum, 2009). Thus, there is little logic in positing a significant fitness advantage for a suite of unpleasant physiological and emotional states that vary with ovarian cycling in light of the fact that ovarian cycling for most premodern women would have been uncommon and erratic.

Recognizing this limitation, there have been claims that PMS is a by-product of selection for some other trait that does have fitness advantages. A recent example of such a model proposes PMS exists as a by-product of evolutionary forces that "maximize chances of mating and fertilization" by heightening a woman's sexual attractiveness during the ovulatory period (Reiber, 2008, 2009). Rather than actually explaining the mechanism behind PMS, Reiber's proposal attempts to explain why some women may experience negative PMS symptoms but others clearly do not. She proposes that a woman who has sufficient nutritional, social, and financial resources (i.e., whose condition is propitious for reproduction), and is approaching menopause (i.e., she has little time left to reproduce) will experience a positive upswing in well-being around the time of ovulation, which will increase her ability to attract potential mates, and that this upswing in well-being will serve as a buffer against the negative experiences associated with declining hormone levels in the premenstrum (Reiber, 2008, 2009). On the other hand, women who are not experiencing conditions favorable to immediate reproduction (women without resources who are not nearing menopausal status) will not experience positive upswings in mood, as they are not in a position to attract mates and become impregnated (Reiber, 2008). As a result, these women will have no hormonal buffer when they reach the premenstrual phase of their cycles and so will experience only the negative symptoms associated with PMS (Reiber, 2008). Symptoms of PMS, then, are simply the result of the diminishing heightened sense of well-being during the ovulatory phase. Although Reiber (2009) suggests there may be some evidence for her hypothesis, in fact there are other equally (or more) plausible explanations for her observations.

Alternative nonevolutionary explanations for PMS focus on the medicalization of women's bodies and emotions, and often draw attention to the cultural scripts of passive femininity that women are expected to follow (Gottlieb, 1988; Martin, 1987; Zita, 1988). This perspective points out the ways in which assumptions about women's bodies, their natures, and their roles in society are built into research regarding PMS. This bias "often constitutes . . . a collection of negative facts about women's nature, a nature which in turn is seen as requiring medical surveillance and management, along with a 'protective' secondary citizenship" (Zita, 1988, p. 79).

This is not to belittle women's negative experiences or to recommend their symptoms go untreated. It is, however, to suggest that the treatment for chronic, severe PMS may require more than a prescription for pills or hormone supplements.

Summary

As a clinically useful concept, is PMS as "unsatisfactory" today as it was 20 years ago (Bancroft, Williamson, Warner, Rennie, & Smith, 1993)? Recent reviews of PMS state, "The symptoms can begin in the early, mid or late luteal phase and are not associated with defined concentrations of any specific gonadal or non-gonadal hormone" (Rapkin & Akopians, 2012, p. 52). Premenstrual symptoms (more than 200 have been reported in the literature) may be physical, behavioral, and/or psychological; may be associated with ovulatory or nonovulatory cycles; and may continue through menses or may be present in the absence of menstruation (Kadian & O'Brien, 2012). Many women undoubtedly suffer from cycle-related maladies. Yet faced with this current "understanding" of PMS, one is prompted to wonder if there are many illnesses that wouldn't fit the description listed, and how many women go untreated for one serious malady or another because it is assumed to be attributable to "their periods"?

All evidence to date leads to the conclusion that it is very unlikely a better understanding of PMS is to be found through observational studies measuring basal and/or fluctuating levels of hormones and attempting to correlate these with any indices of symptoms, a point made by Rubinow and colleagues in 1988. Others have stressed the problems arising from questionnaire bias (Meaden, Hartlage, & Cook-Karr, 2005) and self-reporting of symptoms (Halbreich & Endicott, 1985).

Several authors (e.g., Sanders et al., 1983; Rubinow & Schmidt, 1995) have suggested that etiology is a matter of *sensitivity* to normal hormonal states and/or fluctuations rather than a result of hormonal abnormalities, although the mechanisms remain unclear. Perhaps because sensitivity and other hypotheses are also not amenable to testing with observational protocols, experimental studies are becoming more common. The report from Schmidt and colleagues (1998) of a reduction in symptoms in 60% of the study participants following artificial suppression of ovarian function is intriguing, if inconclusive. But those protocols involving supraphysiological levels of hormones (reviewed in Poromaa, Smith, & Gullinello, 2003) may having little bearing on the lived experiences of PMS, particularly if (as many have argued) PMS is not a strictly biological phenomenon. For example, the experiences and expectations of symptoms may reinforce each other, making it impossible to disentangle cause and effect (Anson, 1999; van den Akker et al., 1995).

The current absence of knowledge regarding PMS simultaneously perpetuates the medicalization of women's bodies and has hampered the development of safe, effective treatments. Hormonal contraceptives may alleviate symptoms in some (for a brief review of treatments, see Cunningham, Yonkers, O'Brien, and Eriksson, 2009). However, contraceptive therapies must be used cautiously because of (ironically) PMS-like side effects including a range of emotional and physical symptoms (Graham, Ramos, Bancroft, Maglaya, & Farley, 1995; Sabatini & Cagiano, 2006; Sulak, Scow, Preece, & Riggs, 2000). If, in fact, some women are more vulnerable to the effects of hormones than are others, it is likely that individuals will require different treatments, depending on the underlying source of the vulnerability, be it biological, social, or some combination. Models that account for variation in endogenous hormone levels, as well as social factors and expectations of cycle-attributed symptoms, could potentially help clarify the confusion and debate surrounding the etiology and experience of PMS. At the least, it does appear unlikely that a universal causative factor will suffice to explain PMS.

Menstrual Synchrony or Menstrual Myth?

Menstrual synchrony (MS) is widely understood to be a phenomenon whereby the duration of two or more women's (or animals') ovarian cycles shorten or lengthen so as to bring the timing of the onset of their menses into mutual alignment. In the literature on MS, it is typically thought to occur as a consequence of a pheromonal or social signal between the menstruating women rather than as the result of an alignment of each woman's cycle to some environmental signal external to the dyad or group. Confusion can arise in debates if this distinction regarding the trigger of any apparent MS is not clarified. Some authors consider any co-occurrence of menses to be MS, independent of the possible mechanism, but others see demonstration of an external trigger to be a refutation of MS *sensu stricto*. This sort of confusion regarding the definition of and possible mechanisms behind MS is common in both the scientific and more general literature.

MS is often assumed to be a well-documented feature of women's biology but, in fact, there is surprisingly little (if any) undisputed evidence to support the existence of any mechanism that functions to create synchrony among women's cycles. Rather, apparent synchronization is readily attributable to chance convergence arising from the finite and variable length of menstrual cycles and the rules of probability. Thus, given an average cycle length of 28 days, the maximum number of days by which two women can differ in menstrual onset is 14 days, and the average difference is only seven days. In light of the evidence presented earlier on the natural "irregularity" of women's cycles (i.e., about half or more of women who claim to have "regular" cycles, in fact, have a range in segment length of at least six days and about a quarter have a range greater than two weeks), it is hardly surprising that menses onset is coincident at some time or another in a pair of women. Nevertheless, it is worth reviewing the history of research on MS and considering why the hypothesized phenomenon has garnered so much attention and persists in spite of much evidence to the contrary.

Origins of the Menstrual Synchrony Hypothesis (MSH)

Martha McClintock (1971) first proposed and tested the MSH in a sample of 135 female college dormitory residents, aged 17 to 22 years old. Three times during the school year, these women were asked for the timing of their last and penultimate menstrual cycles, and to list those women with whom they had spent the most time. Menses dates were compared between different relationship dyads: there appeared to be a significant decrease in the time between menstrual onset dates of those women who spent more time together over the study period, with roommates becoming more closely aligned than did friends. McClintock (1971) suggested that these patterns were consistent with the hypothesis that synchrony was established and maintained by pheromonal signals.

Some subsequent studies have lent support to the MSH. Graham and McGrew (1980) recruited 79 university women who were classified as 18 pairs of close friends, 18 pairs of neighbors, 18 random pairs, and 15 groups based on closeness of living spaces. Contraceptive users (35 of 79 participants) were included in the expectation that these women might influence the cycling of the nonusers. Over the course of four cycles, the time between menstrual cycle onset dates for pairs of close friends significantly decreased, but this was not the case for pairs of neighbors, leading the authors to conclude that a shared environment is not sufficient for MS. Rather, social factors (perhaps including pheromonal communication) might be the important determining factor in MS (Graham & McGrew, 1980). In an effort to test the pheromone hypothesis, Russell, Switz, and Thompson (1980) treated the upper lips of five women with a mixture of alcohol and underarm sweat three times a week for four cycles. They reported that four of the five women so treated synchronized with the sweat donor, with the mean deviation of onset dates dropping from 9.3 days to 3.4 days.

Critiques of MSH Studies

Wilson (1992) and Yang and Schank (2006), among others, have criticized the study design, methods, and statistics used by McClintock (1971) and others who have claimed evidence of MS. For example, McClintock

appears to have incorrectly used the Page test for ordered hypotheses with multiple treatments (she used the same groups of women repeatedly instead of independent treatments), making it impossible to evaluate the true level of significance of her reported findings. Likewise, reports of greater estrous synchrony in chimpanzees caged together than in those caged apart (Wallis, 1985) and synchronization of estrogen peaks in a sample of five golden lion tamarins (French & Stribley, 1987) are rendered moot by the use of unsuitable statistical tests (Schank, 2001). Furthermore, computer simulations suggested that the null hypothesis of no synchronization could not be rejected in either the chimpanzee or tamarin samples (Schank, 2001).

Classification of subject dyads in various studies is also problematic. McClintock (1971) assumed in her analyses that "roommates" and "closest friends" were mutually exclusive groups, which they may not have been, an assumption that undermines the validity of her statistical tests (Yang & Schank, 2006). A perusal of the relevant literature quickly reveals the inconsistencies regarding the claims of which dyads are expected to synchronize because they are supposedly "closer" than other dyads (e.g., roommates with whom one might share the exact same environment for at least eight hours of a 24-hour-day, or close friends with whom one may have greater physical or emotional contact, or coworkers with whom one spends eight or more waking hours). Collectively, studies have reported MS in any of these groups at least some of the time (Graham, 1991), a pattern of findings better explained by chance coincidence of menses onset than by MS.

McClintock (1971) also appears to have miscalculated the menses onset dates for the study subjects, which artificially inflated the calculation of the initial divergence among subjects (Wilson, 1992). As a consequence, the study appeared to show significant decreases in the difference between the timing of menses onset, when in fact any convergence was reasonably attributable to chance. Wilson (1992) also argued that the initial menses timing in the sample studied by Graham and McGrew (1980) is skewed toward asynchrony, and that the probability that 14 out of 18 pairs will show decreasing onset differences ("convergence" of cycle timing) is reasonably attributable to chance alone. Likewise, the observations reported by Russell et al. (1980) are attributable to chance alone (Wilson, 1992).

Empirical Refutations of MSH

In a study specifically designed to address the criticisms and obstacles that other studies had faced, Trevathan, Burleson, and Gregory (1993) tested the MSH in a sample of 29 cohabiting lesbian couples, none of whom was having sex with men, and who were older than the college-aged samples recruited in most studies. The investigators argued that, more so than in

any other study, this population met the requirements necessary for MS to occur (i.e., the subjects had regular cycles, spent a great deal of intimate time together, and were the least likely to experience possible male "interference" in signaling).

Over the course of three menstrual cycles, women kept daily diaries to report changes in menstrual status, sexual activities, stressors, illnesses, and other details about their health and well-being. Although some cycle convergence did occur, cycles were more likely to diverge (Trevathan et al., 1993). The investigators also tested for ovulatory synchrony, which has been implicitly assumed to occur if MS occurs (the evidence presented earlier in Table 1 on variation in follicular phase lengths refutes this assumption). Using basal body temperature as a biomarker for ovulation, they did not find any evidence of ovulatory synchrony. Trevathan and colleagues (1993) concluded that menstrual synchrony "is not a real phenomenon" but suggested that study of the "sociosexual regulation of ovarian function is warranted."

Strassmann (1997, 1999) has conducted the only test of the MSH in a natural fertility population (i.e., no use of any parity-specific methods to control fertility), specifically, the Dogon in Mali. This is a particularly valuable evaluation of the hypothesis because such populations are the best contemporary model for the reproductive and menstrual patterns that would have been present during much of the evolution of *Homo sapiens*. If MS had evolved for any adaptive reason (see discussion that follows), then MS should be evident among Dogon women.

Notably, as would be expected in married women not using contraception, Dogon women do not experience a lengthy series of regular ovulatory cycles as is common for women in industrialized populations. Rather, after a handful of cycles, some of which are anovulatory, married Dogon women become pregnant. If the conception is not lost, they eventually give birth and breastfeed for an extended period (which suppresses menstrual cycling), then experience another handful of cycles before becoming pregnant again. Similar ovarian cycle patterns have been documented in other natural fertility populations (Wood, 1994). These patterns alone are reason enough to prompt skepticism that MS might have evolved for any specific purpose. If a phenomenon is uncommon and occurs erratically, then selection favoring it is likely to be weak at most.

Strassmann (1997) analyzed the Dogon women at three levels of relatedness based on the amount of time they were thought to spend together: all women in the village, all women living in a particular lineage of related males, and all women who regularly ate and worked together. She used Cox regression to determine whether, for any given cycle, one woman's risk of menstruating was influenced by the number of other women who were menstruating. In brief, there was no evidence for MS on

237

any of the three levels of social interaction (Strassmann, 1997, 1999).

Yang and Schank (2006) studied 186 Chinese women living in dormitories. Variation in the cycles from these women was comparable to the cycle variation observed by McClintock (1971). Statistical analyses indicated that any apparent synchronization of menses onset in these Chinese women was attributable to chance. Yang and Schank (2006) concluded that the "common perception of synchrony" is merely the result of chance occurrences that are inevitable due to cycle variability (p. 434). Unavoidably, but deceptively, cycle variability leads to oscillating periods of menstrual onset date clustering *and* divergence.

Studies of reproductive functioning in nonhuman primates have also failed to support the MSH *sensu stricto* (Dixson, 2012). In a population of wild chimpanzees, Matsumoto-Oda and colleagues (2007) determined that, if anything, females were asynchronous as compared to chance ($p < 0.001$). Likewise, there is no evidence for synchrony in baboons (Tobler, Pledger, & Linklater, 2010), mandrills (Setchell, Kendal, & Tyniec, 2011), macaques (Fürtbauer, Munry, Heistermann, Schülke, & Ostner, 2011), golden lion tamarins (Monfort, Bush, & Wildt, 1996), or ring-tailed lemurs (Pereira, 1991). The absence of estrous synchrony in ring-tailed lemurs is, perhaps, the most unexpected because they are highly seasonal breeders with conceptions typically occurring each year within a period of less than two months (Pereira, 1991). Despite the current lack of evidence that female dyads (or larger groups) within any primate species are synchronizing to one another's cycles, it may still be that the estrus cycle in some primates is sufficiently influenced by climatic or other environmental changes such that individuals are all roughly synchronized to that external factor (e.g., see Wallis, 1995, on chimpanzees and Dunbar, 1980, on baboons).

Because discriminating between these two putative routes to synchronization is challenging in free-living humans and animals, there have been several efforts to discover the specific mechanism by which MS purportedly occurs. Environmental factors, pheromonal signals, and social factors have all been proposed and debated (e.g., McClintock, 1971, 1981; Meredith, 2001; Stern & McClintock, 1998; Trotier et al., 2000; Weller & Weller, 1993). Although a review of these studies is beyond the scope of the present work, it is notable that as yet no compelling evidence supports any of these hypotheses. Of course, absence of evidence regarding a mechanism for MS is not proof that MS does not occur. But such absence of evidence is cause for healthy skepticism.

Is MS an Evolutionary Adaptation?

McClintock (1981) did not consider MS to have an adaptive function, but others have advanced evolutionary explanations for the hypothesized phenomenon.

Recall that for a phenotype to be considered an adaptation it must increase the lifetime reproductive success of the individuals having the phenotype relative to other individuals who do not have the phenotype. This requirement is sorely lacking in most (if not all) of the hypothesized evolutionary explanations.

For example, Knight (1991) proposed that the coordination of cycles in a group of females encourages a male to invest in resource acquisition activities (and subsequently in provisioning females) instead of expending energy on guarding females during infecund periods. However, Foley and Fitzgerald (1996) countered that if all females are in their peak fertile periods concurrently, this would increase female competition. Therefore, the selection for "cheaters" (i.e., females who are not responsive to the hypothetical synchronizing pheromones) would be strong because these females would have much less competition during the times when synchronizing females are not fecund. The advantages of cheating create an unstable condition in which there would not be directional selection for MS. Foley and Fitzgerald (1996) evaluated these arguments with computer simulations that tracked the hypothetical costs and benefits of synchronizers and cheaters. The simulations showed that synchronizers enjoyed no reproductive benefits unless synchronization somehow led to a drastic reduction of infant mortality or to a significant reduction in the interbirth interval. They concluded that given the demographic conditions of human evolution including high infant mortality rates, there was little likelihood that MS could have evolved (Foley & Fitzgerald, 1996).

Strassmann (1997) has argued forcefully that there is simply no reason to expect MS. Many factors influence the onset of a menstrual cycle including early pregnancy loss, breastfeeding behaviors, and energetic and psychosocial stressors (Strassmann, 1997; Vitzthum, 2009; Wasser & Barash, 1983). These factors, as well as variability between the cycles of different women and within a woman's lifetime of cycles, are serious obstacles to the establishment and maintenance of MS.

Summary

An appreciation of the likely patterns of ovarian cycling throughout much of human evolutionary history (until the 20th century) coupled with data on the extraordinary variation within and among contemporary women in cycle length quickly leads to a nagging doubt regarding the likelihood of MS *sensu stricto*. Add a good dose of probability theory and the fact that reasonably well designed studies have failed to support the MSH, and one is left wondering why so much attention has been given to searching for elusive mechanisms and constructing convoluted evolutionary scenarios.

In light of the lack of empirical evidence for MS *sensu stricto*, it seems there should be more widespread doubt

than acceptance of this hypothesis. That this has not occurred is intriguing. Although the question is understudied, there are several plausible explanations for the persistence of a belief in MS.

A large part of the answer may simply be unfamiliarity with the extent to which the menstrual cycle in healthy women is "irregular" and that the laws of probability reveal how very likely it is that the appearance of synchrony is merely coincidence. More subtle influences may include Western cultural constructs of healthy humans as smooth functioning machines (hence "regular" cycles are normal) interwoven with cultural ideals of women as particularly empathetic and linked to one another and to "nature" (e.g., lunar cycles). Or it may be that most women in industrialized populations do not realize the infrequency of ovarian cycling in natural fertility populations (i.e., during nearly all of human history). It may also be that those more familiar with estrus cycles and seasonal breeding in other animals expect to see something similar in human biology. But then one must ask how and why synchrony could co-occur in a species that lacks estrous cycles and has hidden ovulation, as do human females. The seven billion of us on this planet attest to the evolutionary success of these (and other) reproductive adaptations. Menstrual synchrony in women would serve only to counter the adaptive advantages of these well-documented reproductive phenotypes.

Concluding Remarks

Reproductive ecologists and sex researchers have much to offer each other and would mutually benefit from a greater exchange of perspectives and methods. Although not the focus of the present article, human evolutionary biologists should certainly incorporate measures of human sexual behaviors more nuanced than, for example, "copulation frequency" into their studies of mating effort. Likewise, sex researchers would gain much from replacing nonhormonal biomarkers of ovarian function with hormonal biomarkers and using these with due acknowledgment of the fact that hormone levels vary greatly within as well as among women. It is also essential to recognize that the causes and significance of this variation are as yet far from clear.

Currently available data do not convincingly support a role for hormone-associated variation in sexually related behaviors or psychological states over the course of the ovarian cycle, nor for any association between absolute hormone levels and sexual dysfunction. Changes in hormone levels during the perimenopausal transition may influence sexual functioning, but it appears that other factors, especially those related to relationship quality, may be far more important determinants of sexual functioning. There is also a dearth of clear evidence to support a hormonal foundation for premenstrual syndrome. In addition, after four decades of studies and debates, the balance of evidence strongly supports the position that any apparent "menstrual synchrony" *sensu stricto* is attributable to random co-occurrences reflecting the finite but variable length of an ovarian cycle.

Before undertaking any additional studies on these hypothesized relationships between hormones and sexually related phenotypes, it is useful to consider whether there are any arguments supporting these hypotheses that are consistent with the evolutionary history of human females and with the known variation in the associated suite of reproductive phenotypes. Specifically, throughout much of the world today (and throughout all human populations until very recently) premenopausal adult women are typically in heterosexual relationships, not using contraception, spending several years pregnant and lactating, and only occasionally experiencing a few sequential ovarian cycles. Women in such natural fertility populations average only about 40 ovarian cycles in a lifetime. In addition, we now know that within a woman, there is only a modest correlation between sequential cycles in cycle length, phase lengths, timing of ovulation, and hormone levels, and that these biomarkers of ovarian functioning vary widely among women and populations without necessarily affecting fecundity. Moreover, ovulation is substantially (and arguably wholly) concealed, and sexual behavior occurs throughout an ovarian cycle regardless of the timing of the fertile period (although it may increase modestly prior to ovulation, the reasons for this are uncertain), and also occurs when women are not fecund (i.e., during pregnancy, lactation, and postmenopausal). Undeniably, sex in humans (and in other species; see Dixson, 2009, 2012) is about reproduction *and* many other aspects of sociality and psychology (e.g., pleasure, affection, power, identity). In light of these well-documented reproductive phenotypes in humans, there is little reason to expect a simple covariation of hormones and sexually related behaviors or psychological states over the course of an ovarian cycle and even less reason to expect menstrual synchrony. Likewise, given the extraordinary variation in hormone levels, there is little reason to think there will be a clear association between sexual functioning and absolute hormone levels, although it remains to be adequately investigated whether a substantial fluctuation in a woman's typical hormone level acts as a biologically salient signal that influences sexual functioning.

Although we have expressed caution throughout this review, we remain optimistic that coupling the tools and models of reproductive ecologists and sex researchers is a highly productive union for expanding our understanding of the evolutionary history of, and contemporary variation in, human sexuality.

References

Abplanalp, J. M., Rose, R. M., Donnelly, A. F., & Livingston-Vaugh, L. (1979). Psychoendocrinology of the menstrual cycle, II: The relationship between enjoyment of activities, moods, and reproductive hormones. *Psychosomatic Medicine*, 41, 605–615.

Alberts, S. C., & Fitzpatrick, C. L. (2012). Paternal care and the evolution of exaggerated sexual swellings in primates. *Behavioral Ecology*, 23, 699–706.

Alliende, M. E. (2002). Mean versus individual hormonal profiles in the menstrual cycle. *Fertility and Sterility*, 78, 90–95.

Alvarado, G., Rivera, R., Ruiz, R., Arenas, M., & Rueda, D. (1988). The characteristics of the menstrual bleeding pattern in a group of normal women from Durango. *Ginecologia y Obstetricia de México*, 56, 127–131.

Anson, O. (1999). Exploring the bio-psycho-social approach to premenstrual experiences. *Social Science and Medicine*, 49, 67–80.

Apter, D. (1997). Development of the hypothalamic-pituitary-ovarian axis. *Annals of the New York Academy of Sciences*, 816, 9–21.

Arrington, R., Cofrancesco, J., & Wu, A. W. (2004). Questionnaires to measure sexual quality of life. *Quality of Life Research*, 13, 1643–1658.

Avis, N. E., Brockwell, S., Randolph, J. F., Shunhua, S., Cain, V. S., Ory, M., et al. (2009). Longitudinal changes in sexual functioning as women transition through menopause: Results from the study of women's health across the nation (SWAN). *Menopause*, 16, 442–452.

Bachman, G. A., & Leiblum, S. R. (1991). Sexuality in sexagenarian women. *Maturitas*, 13, 43–50.

Bachman, G. A., & Leiblum, S. R. (2004). The impact of hormones on menopausal sexuality: A literature review. *Menopause*, 11, 120–130.

Bäckström, T., & Carstensen, H. (1974). Estrogen and progesterone in plasma in relation to premenstrual tension. *Journal of Steroid Biochemistry*, 5, 257–60.

Bäckström, T., Sanders, D., Leask, R., Davison, D., Warner, P., & Bancroft, J. (1983). Mood, sexuality, hormones, and the menstrual cycle, II: Hormone levels and their relationship to the premenstrual syndrome. *Psychosomatic Medicine*, 45, 503–507.

Bäckström, T., Wide, L., Södergård, R., & Carstensen, H. (1976). FSH, LH, TeBG-capacity, estrogen, and progesterone in women with premenstrual tension during the luteal phase. *Journal of Steroid Biochemistry*, 7, 473–476.

Bailey, J., & Marshall, J. (1970). The relationship of the post-ovulatory phase of the menstrual cycle to total cycle length. *Journal of Biosocial Sciences*, 2, 123–132.

Baird, D. D., Weinberg, C. R., Wilcox, A. J., & McConnaughey, D. R. (1991). Using the ratio of urinary oestrogen and progesterone metabolites to estimate day of ovulation. *Statistics in Medicine*, 10, 255–266.

Baird, D. D., Weinberg, C. R., Zhou, H., Kamel, F., McConnaughey, D. R., & Kesner, J. S. (1999). Preimplantation urinary hormone profiles and the probability of conception in healthy women. *Fertility and Sterility*, 71, 40–49.

Bancroft, J. (2009). *Human sexuality and its problems*. New York: Churchill Livingstone.

Bancroft, J., & Cawood, H. H. (1996). Androgens and the menopause: A study of 40–60-year-old women. *Clinical Endocrinology*, 45, 577–587.

Bancroft, J., & Graham, C. A. (2011). The varied nature of women's sexuality: Unresolved issues and a theoretical approach. *Hormones and Behavior*, 59, 717–729.

Bancroft, J., Sanders, D., Davidson, D., & Warner, P. (1983). Mood, sexuality, hormones, and the menstrual cycle, III: Sexuality and the role of androgens. *Psychosomatic Medicine*, 45, 509–516.

Bancroft, J., Williamson, L., Warner, P., Rennie, D., & Smith, S. (1993). Perimenstrual complaints in women complaining of PMS, menorrhagia, and dysmenorrhea: Toward a dismantling of the premenstrual syndrome. *Psychosomatic Medicine*, 55, 133–145.

Basson, R., Brotto, L. A., Petkau, A. J., & Labrie, F. (2010). Role of androgens in women's sexual dysfunction. *Menopause: Journal of the North American Menopause Society*, 17, 962–971.

Belsey, E. M., & Peregoudov, S. (1988). Determinants of menstrual bleeding patterns among women using natural and hormonal methods of contraception, I: Regional variations. *Contraception*, 38, 227–242.

Belsey, E. M., & Pinol, A. P. (1997). Menstrual bleeding patterns in untreated women. *Contraception*, 55, 57–65.

Bernstein, L., Yuan, J. M., Ross, R. K., Pike, M. C., Hanisch, R., Lobo, R., et al. (1990). Serum hormone levels in pre-menopausal Chinese women in Shanghai and White women in Los Angeles: Results from two breast cancer case-control studies. *Cancer Causes and Control*, 1, 51–58.

Bongaarts, J. (1980). Does malnutrition affect fecundity? A summary of evidence. *Science*, 208, 564–569.

Borgerhoff Mulder, M. (1992). Reproductive decisions. In E. A. Smith & B. Winterhalder (Eds.), *Evolutionary ecology and human behavior* (pp. 339–374). New York: Aldine de Gruyter.

Borgerhoff Mulder, M. (2000). Optimizing offspring: The quantity-quality tradeoff in agropastoral Kipsigis. *Evolution and Human Behavior*, 21, 391–410.

Brewis, A., & Meyer, M. (2005). Demographic evidence that human ovulation is undetectable (at least in pair bonds). *Current Anthropology*, 46, 465–471.

Bribiescas, R. G. (2001). Reproductive ecology and life history of the human male. *American Journal of Physical Anthropology*, 116, 148–176.

Briggs, M. H., & Briggs, M. (1972). Steroid hormone concentrations in blood plasma from residents of Zambia, belonging to different ethnical groups. *Acta Endocrinologica*, 70, 619–624.

Broom, T. J., Matthews, C. D., Cooke, I. D., Ralph, M. M., Seamark, R. F., & Cox, L. W. (1981). Endocrine profiles and fertility status of human menstrual cycles of varying follicular phase length. *Fertility and Sterility*, 36, 194–200.

Burkhart, M. C, de Souza, M. L., Salazar, S., & Hess, T. (1999). Incidence of irregular cycles among Mayan women who reported having regular cycles: Implications for fertility awareness methods. *Contraception*, 59, 271–275.

Burleson, M. H., Gregory, W. L., & Trevathan, W. R. (1995). Heterosexual activity: Relationship with the ovarian function. *Psychoneuroendocrinology*, 20, 405–421.

Cain, V. S., Johannes, C. B., Avis, N. E., Mohr, B., Schocken, M., Skurnick, J., et al. (2003). Sexual functioning and practices in a multi-ethnic study of midlife women: Baseline results from SWAN. *Journal of Sex Research*, 40, 266–276.

Cameron, N. (1976). Weight and skinfold variation at menarche and the critical weight hypothesis. *Annals of Human Biology*, 76, 279–282.

Campbell, K. L., & Rockett, J. C. (2006). Biomarkers of ovulation, endometrial receptivity, fertilisation, implantation and early pregnancy progression. *Paediatric and Perinatal Epidemiology*, 20(Suppl. 1), 13–25.

Campbell, K. L., & Wood, J. W. (1994). Human reproductive ecology: Interactions of environment, fertility, and behavior. *Annals of the New York Academy of Sciences*, 709, 1–429.

Caro, T. M., & Borgerhoff Mulder, M. (1987). The problem of adaptation in the study of human behavior. *Ethology and Sociobiology*, 8, 61–72.

Cawood, H. H., & Bancroft, J. (1996). Steroid hormones, the menopause, sexuality, and well-being of women. *Psychological Medicine*, 26, 925–936.

Charnov, E. L. (1993). *Life history invariants*. Oxford, England: Oxford University Press.

Charnov, E. L., & Ernest, S. K. (2006). The offspring-size/clutch-size trade-off in mammals. *American Naturalist*, *167*, 578–582.

Chau, J. P. C., Phil, M., Chang, A. M., & Chang, M. Z. (1998). Relationship between premenstrual tension syndrome and anxiety in Chinese adolescents. *Journal of Adolescent Health*, *22*, 247–249.

Chiazze, L., Brayer, F. T., Macisco, J. J., Parker, M. P., & Duffy, B. J. (1968). The length and variability of the human menstrual cycle. *Journal of the American Medical Association*, *203*, 377–380.

Chisholm, J. S. (1993). Death, hope, and sex: Life-history theory and the development of reproductive strategies. *Current Anthropology*, *34*, 1–24.

Chisholm, J. S., & Coall, D. A. (2008). Not by bread alone: The role of psychosocial stress in age at first reproduction and health inequalities. In W. R. Trevathan, E. O. Smith & J. J. McKenna (Eds.), *Evolutionary medicine and health: New perspectives* (pp. 134–148). New York: Oxford University Press.

Christiansen, K. (2001). Behavioral effects of androgen in men and women. *Journal of Endocrinology*, *170*, 39–48.

Clutton-Brock, T. H., & Harvey, P. H. (1979). Comparison and adaptation. *Proceedings of the Royal Society London B*, *205*, 547–565.

Collet, M. E., Wertenberger, G. E., & Fiske, V. M. (1954). The effect of age upon the pattern of the menstrual cycle. *Fertility and Sterility*, *5*, 437–448.

Collins, W. P. (1985). Hormonal indices of ovulation and the fertile period. *Advances in Contraception*, *1*, 279–294.

Cooney, K. A., Nyirabukeye, T., Labbok, M. H., Hoser, P. H., & Ballard, E. (1996). An assessment of the nine-month lactational amenorrhea method (MAMA-9) in Rwanda. *Studies in Family Planning*, *27*, 102–171.

Creinin, M. D., Keverline, S., & Meyn, L. A. (2004). How regular is regular? An analysis of menstrual cycle regularity. *Contraception*, *70*, 289–292.

Cunningham, J., Yonkers, K. A., O'Brien, S., & Eriksson, E. (2009). Update on research and treatment of premenstrual dsyphoric disorder. *Harvard Review of Psychiatry*, *17*, 120–137.

Davis, R., Davison, S. L., Donath, S., & Bell, R. J. (2005). Circulating androgen levels and self-reported sexual function in women. *Journal of the American Medical Association*, *294*, 91–96.

DeCherney, A. H., Romero, R., & Polan, M. L. (1982). Ultrasound in reproductive endocrinology. *Fertility and Sterility*, *37*, 323–333.

Dennerstein, L., Dudley, E., & Burger, H. (2001). Are changes in sexual functioning during midlife due to aging or menopause? *Fertility and Sterility*, *76*, 456–460.

Dennerstein, L., Gotts, G., Brown, J. B., Morse, C. A., Farely, T. M. M., & Pinol, A. (1994). The relationship between the menstrual cycle and female sexual interest in women with PMS complaints and volunteers. *Psychoneuroendocrinology*, *19*, 293–304.

Dennerstein, L., Randolph, J., Taffe, J., Dudley, E., & Burger, H. (2002). Hormones, mood, sexuality, and the menopausal transition. *Fertility and Sterility*, *7*, S42–S48.

Deschner, T., Heistermann, M., Hodges, K., & Boesch, C. (2003). Timing and probability of ovulation in relation to sex skin swelling in wild West African chimpanzees, *Pan troglodytes verus*. *Animal Behavior*, *66*, 551–560.

Deschner, T., Heistermann, M., Hodges, K., & Boesch, C. (2004). Female sexual swelling size, timing of ovulation, and male behavior in wild West African chimpanzees. *Hormones and Behavior*, *46*, 204–215.

Dickinson, L. E., MacMahon, B., Cole, P., & Brown, J. B. (1974). Estrogen profiles of Oriental and Caucasian women in Hawaii. *New England Journal of Medicine*, *291*, 1211–1213.

Dixson, A. F. (2009). The oestrus that never was. In A. F. Dixson (Ed.), *Sexual selection and the origins of human mating systems* (pp. 106–123). New York: Oxford University Press.

Dixson, A. F. (2012). *Primate sexuality: Comparative studies of the Prosimians, monkeys, apes, and humans*. Oxford, England: Oxford University Press.

Dobbins, J. (1980). Implication of a time-dependent model of sexual intercourse with the menstrual cycle. *Journal of Biosocial Science*, *12*, 133–140.

Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *American Biology Teacher*, *35*, 125–129.

Döring, G. K. (1969). The incidence of anovular cycles in women. *Journal of Reproduction and Fertility*, *6*, 77–81.

Dunbar, R. I. M. (1980). Demographic and life history variables of a population of gelada baboons (*Theropithecus gelada*). *Journal of Animal Ecology*, *49*, 485–506.

Dunson, D. B., Baird, D. D., Wilcox, A. J., & Weinberg, C. R. (1999). Day-specific probabilities of clinical pregnancy based on two studies with imperfect measures of ovulation. *Human Reproduction*, *14*, 1835–1839.

Dunson, D. B., Columbo, B., & Baird, D. D. (2002). Changes with age in the level and duration of fertility in the menstrual cycle. *Human Reproduction*, *17*, 1399–1403.

Ecochard, R., Boehringer, H., Rabilloud, M., & Marret, H. (2001). Chronological aspects of ultrasonic, hormonal, and other indirect indicates of ovulation. *BJOG*, *108*, 822–829.

Ecochard, R., & Gougeon, A. (2000). Side of ovulation and cycle characteristics in normally fertile women. *Human Reproduction*, *15*, 752–755.

Ellison, P. T. (1994). Advances in human reproductive ecology. *Annual Review of Anthropology*, *23*, 255–275.

Ellison, P. T. (2001). Introduction: Ecology, reproduction, and human evolution. In P. T. Ellison (Ed.), *Reproductive ecology and human evolution* (pp. 1–11). New York: Aldine.

Ellison, P. T., & Lager, C. (1986). Moderate recreational running is associated with lowered salivary progesterone profiles in women. *American Journal of Obstetrics and Gynecology*, *154*, 1000–1003.

Ellison, P. T., Panter-Brick, C., Lipson, L. F., & O'Rourke, M. (1993). The ecological context of human ovarian function. *Human Reproduction*, *8*, 2248–2258.

Ellison, P. T., Peacock, N. R., & Lager, C. (1989). Ecology and ovarian function among Lese women of the Ituri Forest, Zaire. *American Journal of Physical Anthropology*, *78*, 519–526.

Eriksson, E., Sundblad, C., Lisjo, P., Modigh, K., & Andersch, B. (1992). Serum levels of androgens are higher in women with premenstrual irritability and dysphoria than in controls. *Psychoneuroendocrinology*, *17*, 195–204.

Fehring, R. J., Schneider, M., & Raviele, K. (2006). Variability in the phases of the menstrual cycle. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*, *35*, 376–384.

Ferrell, R. J., O'Connor, K. A., Rodriguez, G., Gorrindo, T., Holman, D. J., Brindle, E., et al. (2005). Monitoring reproductive aging in a five year prospective study: Aggregate and individual changes in steroid hormones and menstrual cycle lengths with age. *Menopause*, *12*, 567–577.

Fisher, D. C. (1985). Evolutionary morphology: Beyond the analogous, the anecdotal, and the ad hoc. *Paleobiology*, *11*, 120–138.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford, England: Clarendon.

Flynn, A. M., Docker, M. F., McCarthy, N., & Royston, J. P. (1988). Detection of the fertile phase from changes in cervico-vaginal fluid volume. *International Journal of Fertility*, *33*, S17–S23.

Foley, R. A., & Fitzgerald, C. M. (1996). Is reproductive synchrony an evolutionarily stable strategy for hunger-gatherers? *Current Anthropology*, *37*, 539–545.

Fraenkel, L. (1926). Zeit und Ursächlichkeitsverhältnis zwischen Ovulation und Menstruation [Time and the causal relationship between ovulation and menstruation]. *Handbuch der Normalen und der Pathologischen Physiologie* [Handbook of Normal and Pathological Physiology], xiv/1:454.

France, J. T., Graham, F. M., Gosling, L., Hair, P., & Knox, B. S. (1992). Characteristics of natural conceptual cycles occurring in a prospective study of sex preselection: Fertility awareness

symptoms, hormone levels, sperm survival, and pregnancy outcome. *International Journal of Fertility*, 37, 244–255.

Frank, R. T. (1931). The hormonal causes of premenstrual tension. *Archives of Neurology and Psychiatry*, 26, 1053–1057.

Frank, R. T., & Goldberger, M. A. (1928). Clinical data obtained with the female sex hormone blood test. *Journal of the American Medical Association*, 90, 106–110.

French, J. A., & Stribley, J. A. (1987). Synchronization of ovarian cycles within and between social groups in golden lion tamarins (*Leontopithecus rosalia*). *American Journal of Primatology*, 12, 469–478.

Frisch, R. E. (1978). Population, food intake, and fertility. *Science*, 199, 22–30.

Frisch, R. E., & McArthur, J. W. (1974). Menstrual cycles: Fatness as a determinant of minimum weight for height necessary for their maintenance or onset. *Science*, 185, 949–951.

Frisch, R. E., & Revelle, R. (1970). Height and weight at menarche and a hypothesis of critical body weights and adolescent events. *Science*, 169, 397–399.

Frisch, R. E., & Revelle, R. (1971a). Height and weight at menarche and a hypothesis of menarche. *Archives of Disease in Childhood*, 46, 695–701.

Frisch, R. E., & Revelle, R. (1971b). The height and weight of girls and boys at the time of initiation of the adolescent growth spurt in height and weight and the relationship to menarche. *Human Biology*, 43, 140–159.

Fürtbauer, I., Munry, R., Heistermann, M., Schülke, O., & Ostner, J. (2011). You mate, I mate; Macaque females synchronize sex not cycles. *PLoS ONE*, 6(e26144), 1–6.

Gadgil, M., & Bossert, W. H. (1970). Life historical consequences of natural selection. *American Naturalist*, 104, 1–24.

Geirhart, B. S. (2006). When does a ''less than perfect'' sex life become female sexual dysfunction? *Obstetrics and Gynecology*, 107, 750–751.

Geronimus, A. T., Bound, J., & Waidmann, T. A. (1999). Health inequality and population variation in fertility-timing. *Social Science and Medicine*, 49, 1623–1636.

Ginsburg, F., & Rapp, R. (1991). The politics of reproduction. *Annual Review of Anthropology*, 30, 311–343.

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society London B*, 205, 581–598.

Gould, S. J., & Vrba, E. S. (1982). Exaptation—A missing term in the science of form. *Paleobiology*, 8, 4–15.

Gottlieb, A. (1988). American premenstrual syndrome: A mute voice. *Anthropology Today*, 4, 10–13.

Gracia, C. R., Sammuel, M. D., Freeman, E. W., Liu, L., Hollander, L., & Nelson, D. B. (2004). Predictors of decreased libido in women during the late reproductive years. *Menopause*, 11, 144–150.

Graham, C. A. (1991). Menstrual synchrony: An update and review. *Human Nature*, 2, 293–311.

Graham, C. A., & McGrew, W. C. (1980). Menstrual synchrony in female undergraduates living on a coeducational campus. *Psychoneurendocrinology*, 5, 245–252.

Graham, C. A., Ramos, R., Bancroft, J., Maglaya, C., & Farley, T. M. M. (1995). The effects of steroidal contraceptives on the wellbeing and sexuality of women: A double blind, placebo-controlled, two-centre study of combined and progestogen-only methods. *Contraception*, 52, 363–369.

Gray, R. H., Campbell, O. M., Zacur, H. A., Labbok, M. H., & MacRae, S. L. (1987). Postpartum return of ovarian activity in nonbreastfeeding women monitored by urinary assays. *Journal of Clinical Endocrinology and Metabolism*, 64, 645–650.

Haiman, C. A., Pike, M. C., Bernstein, L., Jaque, S. V., Stancyzk, F. Z., Afghani, A., et al. (2002). Ethnic differences in ovulatory function in nulliparous women. *British Journal of Cancer*, 86, 367–371.

Halbreich, U., & Endicott, J. (1985). Methodological issues in studies of premenstrual changes. *Psychoneuroendocrinology*, 10, 15–32.

Halbreich, U., Endicott, J., Goldstein, S., & Nee, J. (1986). Premenstrual changes and changes in gonadal hormones. *Acta Psychiatrica Scandinavica*, 74, 576–586.

Harlow, S. D., Baird, D. D., Weinberg, C. R., & Wilcox, A. J. (2000). Urinary oestrogen patterns in long follicular phases. *Human Reproduction*, 15, 11–16.

Harris, A. L., & Vitzthum, V. J. (2012). Project REPA (Reproduction and Ecology in Provincía Aroma): The predictability of women's progesterone concentrations from cycle to cycle. *American Journal of Human Biology*, 24, 226.

Hayes, R., & Dennerstein, L. (2005). The impact of aging on sexual function and sexual dysfunction in women: A review of population-based studies. *Journal of Sexual Medicine*, 2, 317–330.

Hedricks, C., Piccinino, L. J., Udry, J. R., & Chimbira, T. H. (1987). Peak coital rate coincides with onset of luteinizing hormone surge. *Fertility and Sterility*, 48, 234.

Hill, K., & Hurtado, A. M. (1996). *Ache life history: The ecology and demography of a foraging people*. New York: Aldine de Gruyter.

Hill, K., & Kaplan, H. (1999). Life history traits in humans: Theory and empirical studies. *Annual Review of Anthropology*, 28, 397–430.

Holman, D. J. (1996). *Total fecundability and pregnancy loss in Bangladesh*. PhD dissertation, Pennsylvania State University, University Park, PA.

Hrdy, S. B., & Whitten, P. L. (1987). Patterning of sexual activity. In B. B. Smuts, D. L. Cheney, R. M. Seyfarth, W. R. Wrangham & T. Struthsacker (Eds.), *Primate societies* (pp. 370–384). Chicago: University of Chicago Press.

Huang, A. J., Subak, L. L., Thom, D. H., Van Den Eeden, S. K., Ragins, A. I., Kupperman, M., et al. (2009). Sexual function and aging in racially and ethnically diverse women. *Journal of the American Geriatrics Society*, 57, 1362–1368.

Jacob, F. (1977). Evolution and tinkering. *Science*, 196, 1161–1166.

Jasienska, G. (2001). Why energy expenditure causes reproductive suppression in woman. In P. T. Ellison (Ed.), *Reproductive ecology and human evolution* (pp. 59–84). New York: Aldine.

Jeyaseelan, L., & Rao, P. S. (1993). Correlates of menstrual cycle length in south Indian women: A prospective study. *Human Biology*, 65, 627–634.

Kadian, S., & O'Brien, S. (2012). Classification of premenstrual disorders as proposed by the International Society for Premenstrual Disorders. *Menopause International*, 18, 43–47.

Key, T. J., Chen, J., Wang, D. Y., Pike, M. C., & Boreham, J. (1990). Sex hormones in women in rural China and Britain. *British Journal of Cancer*, 62, 631–636.

Knight, C. (1991). *Blood relations: Menstruation and the origins of culture*. New Haven, CT: Yale University Press.

Konner, M., & Worthman, C. (1980). Nursing frequency, gonadal function, and birth spacing among !Kung hunter-gatherers. *Science*, 207, 788–791.

Landgren, B. M., & Diczfalusy, E. (1980). Hormonal effects of the 300 microgram norethisterone (NET) minipill, I: Daily steroid levels in 43 subjects during a pretreatment cycle and during the second month of NET administration. *Contraception*, 21, 87–113.

Landgren, B. M., Undén, A. L., & Diczfalusy, E. (1980). Hormonal profile of the cycle in 68 normally menstruating women. *Acta Endocrinology*, 94, 89–98.

Lasker, G. W. (1969). Human biological adaptability. *Science*, 166, 1480–1486.

Lauder, G. V. (1996). The argument from design. In M. R. Rose & G. V. Lauder (Eds.), *Adaptation* (pp. 55–92). San Diego: Academic Press.

Lenton, E. A., Landgren, B., Sexton, L., & Harper, R. (1984). Normal variation in the length of the follicular phase of the menstrual cycle: Effect of chronological age. *British Journal of Obstetrics and Gynaecology*, 91, 681–684.

Leroi, A. M., Rose, M. R., & Lauder, G. V. (1994). What does the comparative method reveal about adaptation? *American Naturalist*, *143*, 381–402.

Leslie, P. W., & Fry, P. H. (1989). Extreme seasonality of births among nomadic Turkana pastoralists. *American Journal of Physical Anthropology*, *79*, 103–115.

Leslie, P. W., & Little, M. A. (2003). Human biology and ecology: Variation in nature and the nature of variation. *American Anthropologist*, *105*, 28–37.

Lessells, C. M. (1991). The evolution of life histories. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 32–68). Oxford, England: Blackwell.

Li, H., Nakajima, S. T., Chen, J., Todd, H. E., Overstreet, J. W., & Lasley, B. L. (2001). Differences in hormonal characteristics of conceptive versus nonconceptive menstrual cycles. *Fertility and Sterility*, *75*, 549–53.

Lipson, S. F. (2001). Metabolism, maturation, and ovarian function. In P. T. Ellison (Ed.), *Reproductive ecology and human evolution* (pp. 235–248). New York: Aldine.

Lipson, S. F., & Ellison, P. T. (1992). Normative study of age variation in salivary progesterone profiles. *Journal of Biosocial Sciences*, *24*, 233–244.

Lipson, S. F., & Ellison, P. T. (1996). Comparison of salivary steroid profiles in naturally occurring conception and non-conception cycles. *Human Reproduction*, *11*, 2090–2096.

Liu, S. F., Wang, Z. X., Yuan, Y. E., Bing, S. M., Zhang, B. Z., Wu, J. Z., et al. (1986). Hormone changes during the menstrual cycle of Chinese women. *Journal of Reproduction and Fertility*, *76*, 43–52.

Livingstone, F. B. (1958). Anthropological implications of sickle cell gene distribution in West Africa. *American Anthropologist*, *60*, 533–562.

Lock, M. (2001). The tempering of medical anthropology: Troubling natural categories. *Medical Anthropology Quarterly*, *15*, 478–492.

MacMahon, B., Cole, P., Brown, J. B., Aoki, K., Lin, T. M., Morgan, R. W., et al. (1974). Urine oestrogen profiles of Asian and North American women. *International Journal of Cancer*, *14*, 161–167.

Malina, R. M. (1983). Menarche in athletes: A synthesis and hypothesis. *Annals of Human Biology*, *10*, 1–24.

Malthus, T. R. (1798). *An essay on the principle of population*. London: Reeves and Turner.

Mansfield, M. J., & Emans, S. J. (1984). Adolescent menstrual irregularity. *Journal of Reproductive Medicine*, *29*, 399–410.

Marks, L. (2001). ''Andromeda freed from her chains'': Attitudes towards women and the contraceptive pill, 1950–1970. *Clio Medica*, *61*, 217–244.

Marlowe, F. W. (2004). Is human ovulation concealed? Evidence from conception beliefs in a hunter-gatherer society. *Archives of Sexual Behavior*, *33*, 427–432.

Martin, E. (1987). *The woman in the body: A cultural analysis of reproduction*. Boston: Beacon Press.

Matsumoto, S., Nogami, Y., & Ohkuri, S. (1962). Statistical studies on menstruation: A criticism on the definition of normal menstruation. *Gunma Journal of Medical Sciences*, *11*, 294–318.

Matsumoto-Oda, A., Hamai, M., Hayaki, H., Hosaka, K., Hunt, K. D., Kasuya, E., et al. (2007). Estrus cycle asynchrony in wild female chimpanzees, *Pan troglodytes schweinfurthii*. *Behavioral Ecology and Sociobiology*, *61*, 661–668.

Mazess, R. B. (1975). Biological adaptation: Aptitudes and acclimatization. In E. S. Watts, F. E. Johnston & G. W. Lasker (Eds.), *Biosocial interrelationships in population adaptation* (pp. 9–18). The Hague, Netherlands: Mouton.

McClintock, M. K. (1971). Menstrual synchrony and suppression. *Nature*, *229*, 244–245.

McClintock, M. K. (1981). Social control of the ovarian cycle and the function of estrous synchrony. *American Zoologist*, *21*, 243–256.

McCoy, N. L., & Davidson, J. M. (1985). A longitudinal study of the effects of menopause on sexuality. *Matuitas*, *7*, 203–210.

McDade, T. W., & Worthman, C. M. (1998). The weanling's dilemma reconsidered: A biocultural analysis of breastfeeding ecology. *Journal of Development and Behavioral Pediatrics*, *19*, 286–299.

Meaden, P. M., Hartlage, S. A., & Cook-Karr, J. (2005). Timing and severity of symptoms associated with the menstrual cycle in a community-based sample in the Midwestern United States. *Psychiatry Research*, *134*, 27–36.

Menken, J., Trussell, J., & Watkins, S. (1981). The nutrition fertility link: An evaluation of the evidence. *Journal of Interdisciplinary History*, *11*, 425–441.

Meredith, M. (2001). Human vomeronasal organ function: A critical review of best and worst cases. *Chemical Senses*, *26*, 433–445.

Messinis, I. E. (2006). Ovarian feedback, mechanism of action and possible clinical implications. *Human Reproduction Update*, *12*, 557–571.

Metcalf, M. G. (1983). Incidence of ovulation from the menarche to the menopause: Observations of 622 New Zealand women. *New Zealand Medical Journal*, *96*, 645–648.

Metcalf, M. G., & Mackenzie, J. A. (1980). Incidence of ovulation in young women. *Journal of Biosocial Sciences*, *12*, 345–352.

Meyer, V. F. (2001). The medicalization of menopause: Critique and consequences. *International Journal of Health Services*, *31*, 769–792.

Migliano, A. B., Vinicius, L., & Lahr, M. M. (2007). Life history trade-offs explain the evolution of human pygmies. *Proceedings of the National Academy of Sciences*, *104*, 20216–20219.

Monari, P., & Montanari, A. (1998). Length of menstrual cycles and their variability. *Genus*, *54*, 95–118.

Monfort, S. L., Bush, M., & Wildt, D. E. (1996). Natural and induced ovarian synchrony in golden lion tamarins (*leontopithecus rosalia*). *Biology of Reproduction*, *55*, 875–882.

Morbeck, M. E., Galloway, A., & Zihlman, A. (Eds). (1997). *The evolving female: A life history perspective*. Princeton, NJ: Princeton University Press.

Morris, N. M., Udry, J. R., Khan-Dawood, F., & Dawood, M. Y. (1987). Marital sex frequency and midcycle female testosterone. *Archives of Sexual Behavior*, *16*, 27–37.

Nepomnaschy, P. A., Welch, K., McConnell, D., Strassmann, B. I., & England, B. G. (2004). Stress and female reproductive functioning: A study of daily variations in cortisol, gonadotrophins, and gonadal steroids in a rural Mayan population. *American Journal of Human Biology*, *16*, 523–532.

Nunn, C. L. (1999). The evolution of exaggerated sexual swellings in primates and the graded-signal hypothesis. *Animal Behavior*, *58*, 229–246.

O'Connor, K. A., Brindle, E., Miller, R. C., Shofer, J. B., Ferrell, R. J., Klein, N. A., et al. (2006). Ovulation detection methods for urinary hormones: Precision, daily and intermittent sampling and a combined hierarchical method. *Human Reproduction*, *21*, 1442–1452.

Panter-Brick, C., Lotstein, D. S., & Ellison, P. T. (1993). Seasonality of reproductive function and weight loss in rural Nepali women. *Human Reproduction*, *8*, 684–690.

Pereira, M. E. (1991). Asynchrony within estrous synchrony among ringtailed lemurs (Primates: Lemuridae). *Physiology and Behavior*, *49*, 47–52.

Persky, H., Charney, N., Lief, H. I., O'Brien, C. P., Miller, W. R., & Strauss, D. (1978). The relationship of plasma estradiol to sexual behavior in young women. *Psychosomatic Medicine*, *40*, 523–535.

Persky, H., Dreisbach, L., Miller, W. R., O'Brien, C. P., Khan, M. A., & Lief, H. I. (1982). The relation of plasma androgen levels to sexual behaviors and attitudes of women. *Psychosomatic Medicine*, *44*, 305–319.

Persky, H., Lief, H. I., Strauss, D., Miller, W. R., & O'Brien, C. P. (1978). Plasma testosterone level and sexual behavior of couples. *Archives of Sexual Behavior*, *7*, 157–173.

Pinheiro, S. P., Holmes, M. D., Pollak, M. N., Barbieri, R. L., & Hankinson, S. E. (2005). Racial differences in premenopausal

endogenous hormones. *Cancer Epidemiology Biomarkers and Prevention, 14,* 2147–2153.

Poromaa, I. S., Smith, S., & Gullinello, M. (2003). GABA receptors, progesterone, and premenstrual dsyphoric disorder. *Archives of Women's Mental Health, 6,* 23–41.

Prior, J. C. (1985a). Hormonal mechanisms of reproductive function and hypothalamic adaptation to endurance training. In J. Puhl (Ed.), *The menstrual cycle and physical activity* (pp. 63–75). Champaign, IL: Human Kinetics.

Prior, J. C. (1985b). Luteal phase defects and anovulation: Adaptive alterations occurring with conditioning exercise. *Seminars in Reproductive Endocrinology, 3,* 27–33.

Prior, J. C. (1987). Physical exercise and the neuroendocrine control of reproduction. *Baillieres Clinical Endocrinology and Metabolism, 1,* 299–317.

Promislow, D. E. L., & Harvey, P. H. (1990). Living fast and dying young: A comparative analysis of life-history variation among mammals. *Journal of Zoology, 220,* 417–437.

Promislow, J. H., Baird, D. D., Wilcox, A. J., & Weinberg, C. R. (2007). Bleeding following pregnancy loss before 6 weeks' gestation. *Human Reproduction, 22,* 853–857.

Quandt, S. A. (1984). Nutritional thriftiness and human reproduction: Beyond the critical body composition hypothesis. *Social Science and Medicine, 19,* 177–182.

Rapkin, A. J., & Akopians, A. L. (2012). Pathophysiology of premenstrual syndrome and premenstrual dysphoric disorder. *Menopause International, 18,* 52–59.

Redei, E., & Freeman, E. W. (1995). Daily plasma estradiol and progesterone levels over the menstrual cycle and their relation to premenstrual symptoms. *Psychoneuroendocrinology, 20,* 259–267.

Reiber, C. (2008). An evolutionary model of premenstrual syndrome. *Medical Hypotheses, 70,* 1058–1065.

Reiber, C. (2009). Empirical support for an evolutionary model of premenstrual syndrome. *Journal of Social, Evolutionary, and Cultural Psychology, 3,* 9–28.

Renaud, R. L., Macler, J., Dervain, I., Ehret, M. C., Aron, C., Plas-Roser, S., et al. (1980). Echographic study of follicular maturation and ovulation during the normal menstrual cycle. *Fertility and Sterility, 33,* 272–276.

Riad-Fahmy, D., Read, G. F., Walker, R. F., Walker, S. M., & Griffiths, K. (1987). Determination of ovarian steroid hormone levels in saliva: An overview. *Journal of Reproductive Medicine, 32,* 254–272.

Riley, A., & Riley, E. (2000). Controlled studies on women presenting with sexual drive disorder, I: Endocrine status. *Journal of Sex and Marital Therapy, 26,* 269–283.

Rose, M. R., & Lauder, G. V. (1996). *Adaptation.* San Diego: Academic Press.

Rosseinsky, D. R., & Hall, P. G. (1974). An evolutionary theory of premenstrual tension. *Lancet, 304,* 1024.

Rubinow, D. R., Hoban, M. C., Grover, G. N., Galloway, D. S., Roy-Byrne, P., Andersen, R., et al. (1988). Changes in plasma hormones across the menstrual cycle in patients with menstrually related mood disorder and in control subjects. *American Journal of Obstetrics and Gynecology, 158,* 5–11.

Rubinow, D. R., & Schmidt, P. J. (1995). The neuroendocrinology of menstrual cycle mood disorders. *Annals New York Academy of Sciences, 771,* 648–659.

Russell, M. J., Switz, G. M., & Thompson, K. (1980). Olfactory influences on the human menstrual cycle. *Pharmacology Biochemistry and Behavior, 13,* 737–738.

Sabatini, R., & Cagiano, R. (2006). Comparison profiles of cycle control, side effects, and sexual satisfaction of three hormonal contraceptives. *Contraception, 74,* 220–223.

Sadler, C., Smith, H., Hammond, J., Bayly, R., Borland, S., Panay, N., et al. (2010). Lifestyle factors, hormonal contraception, and premenstrual symptoms: The United Kingdom Southampton Women's Survey. *Journal of Women's Health, 19,* 391–396.

Sanders, D., Warner, P., Bäckström, T., & Bancroft, J. (1983). Mood, sexuality, hormones, and the menstrual cycle, 1: Changes in mood and physical state: Description of subjects and method. *Psychosomatic Medicine, 45,* 487–501.

Santos, J. R. M. D. (1997). Medicalization of menopause and public health. *Journal of Psychosomatic Obstetrics & Gynecology, 18,* 175–180.

Schank, J. C. (2001). Measurement and cycle variability: Reexamining the case for ovarian-cycle synchrony in primates. *Behavioural Processes, 56,* 131–146.

Schipper, I., de Jong, F. H., & Fauser, B. C. (1998). Lack of correlation between maximum early follicular phase serum follicle stimulating hormone concentrations and menstrual cycle characteristics in women under the age of 35 years. *Human Reproduction, 13,* 1442–1448.

Schmidt, P. J., Nieman, L. K., Danaceua, M. A., Adams, L. F., & Rubinow, D. R. (1998). Differential behavioral effects of gonadal steroids in women with and in those without premenstrual syndrome. *New England Journal of Medicine, 338,* 209–216.

Schreiner-Engel, P., Schiavi, R. C., Smith, H., & White, D. (1981). Sexual arousability and the menstrual cycle. *Psychosomatic Medicine, 43,* 199–214.

Schreiner-Engel, P., Schiavi, R. C., White, D., & Ghizzani, A. (1989). Low sexual desire in women: The role of reproductive hormones. *Hormones and Behavior, 23,* 221–234.

Scott, E. C., & Johnston, F. E. (1985). Science, nutrition, fat, and policy: Tests of the critical fat hypothesis. *Current Anthropology, 26,* 463–473.

Sellen, D. W., & Smay, D. B. (2001). Relationship between subsistence and age at weaning in "preindustrial" societies. *Human Nature, 12,* 47–87.

Setchell, J. M., Kendal, J., & Tyniec, P. (2011). Do non-human primates synchronise their menstrual cycles? A test in mandrills. *Psychoneuroendocrinology, 36,* 51–59.

Shimizu, H., Ross, R. K., Bernstein, L., Pike, M. C., & Henderson, B. E. (1990). Serum oestrogen levels in postmenopausal women: Comparison of American Whites and Japanese in Japan. *British Journal of Cancer, 62,* 451–453.

Smith, S. K., Lenton, E. A., & Cooke, I. D. (1985). Plasma gonadotrophin and ovarian steroid concentrations in women with menstrual cycles with a short luteal phase. *Journal of Reproduction and Fertility, 75,* 363–368.

Snowden, R., & Christian, B. (1983). *Patterns and perceptions of menstruation.* New York: St. Martin's Press.

Stearns, S. C. (1992). *The evolution of life histories.* Oxford, England: Oxford University Press.

Steiner, M. J., Hertz-Picciotto, I., Taylor, D., Schoenbach, V., & Wheeless, A. (2001). Retrospective vs. prospective coital frequency and menstrual cycle length in a contraceptive effectiveness trial. *Annals of Epidemiology, 11,* 428–433.

Stern, K., & McClintock, M. A. (1998). Regulation of ovulation by human pheromones. *Nature, 392,* 177–179.

Stewart, D. R., Overstreet, J. W., Nakajima, S. T., & Lasley, B. L. (1993). Enhanced ovarian steroid secretion before implantation in early human pregnancy. *Journal of Clinical Endocrinology and Metabolism, 76,* 1470–1476.

Strassmann, B. (1997). The biology of menstruation in *Homo sapiens*: Total lifetime menses, fertility, and nonsynchrony in a natural-fertility population. *Current Anthropology, 38,* 123–129.

Strassmann, B. (1999). Menstrual synchrony pheromones: Cause for doubt. *Human Reproduction, 14,* 579–580.

Strassmann, B. I., & Gillespie, B. (2002). Life–history theory, fertility, and reproductive success in humans. *Proceedings Biological Science, 269,* 553–562.

Stuart, F. M., Hammond, D. C., & Pett, M. A. (1987). Inhibited sexual desire in women. *Archives of Sexual Behavior, 16,* 91–106.

Sulak, P. J., Scow, R. D., Preece, C., & Riggs, M. W. (2000). Hormone withdrawal symptoms in oral contraceptive users. *Obstetrics & Gynecology, 95,* 261–266.

Takeda, T., Tasaka, K., Skata, M., & Murata, Y. (2006). Prevalence of premenstrual syndrome and premenstrual dysphoric disorder in Japanese women. *Archives of Women's Mental Health*, 9, 209–212.

Thomas, R. B., Gage, T. B., & Little, M. A. (1989). Reflections on adaptive and ecological models. In M. A Little & J. D. Haas (Eds.), *Human population biology* (pp. 296–319). New York: Oxford University Press.

Tobler, R., Pledger, S., & Linklater, W. (2010). No evidence for ovarian synchrony or asynchrony in hamadryas baboons. *Animal Behavior*, 80, 829–837.

Traish, A. M., Botchevar, F., & Kim, N. N. (2010). Biochemical factors modulating female genital sexual arousal physiology. *Journal of Sexual Medicine*, 7, 2925–2946.

Treloar, A. E., Boynton, R. E., Behn, B. G., & Brown, B. W. (1967). Variation of the human menstrual cycle through reproductive life. *International Journal of Fertility*, 12, 77–126.

Treloar, S. A., Heath, A. C., & Martin, N. G. (2002). Genetic and environmental influences on premenstrual symptoms in an Australian twin sample. *Psychological Medicine*, 32, 25–38.

Trevathan, W. R., Burleson, M. H., & Gregory, W. L. (1993). No evidence for menstrual synchrony in lesbian couples. *Psychoneuroendocrinology*, 18, 425–435.

Trichopoulos, D., Yen, S., Brown, J., Cole, P., & MacMahon, B. (1984). The effect of westernization on urine estrogens, frequency of ovulation, and breast cancer risk. A study of ethnic Chinese women in the Orient and the USA. *Cancer*, 53, 187–192.

Trotier, D., Eliot, C., Wassef, M., Talmain, G., Bensimon, J. L., Doving, K. B., et al. (2000). The vomeronasal cavity in adult humans. *Chemical Senses*, 25, 369–380.

Trussell, J. (1980). Statistical flaws in evidence for the Frisch hypothesis that fatness triggers menarche. *Human Biology*, 52, 711–720.

Turna, B., Apaydin, E., Semerci, B., Atlay, B., Cikili, N., & Nazli, O. (2005). Women with low libido: Correlation of decreased androgen levels with female sexual function index. *International Journal of Impotence Research*, 17, 148–153.

Usala, S. J., & Schumacher, G. F. (1983). Volumetric self-sampling of cervico vaginal fluid: A new approach to ovulation timing. *Fertility and Sterility*, 39, 304–309.

van den Akker, O. B. A., Eves, F. F., Service, S., & Lennon, B. (1995). Menstrual cycle reporting in three British ethnic groups. *Social Science & Medicine*, 40, 1417–1423.

van der Walt, L. A., Wilmsen, E. N., & Jenkins, T. (1978). Unusual sex hormone patterns among desert-dwelling hunter-gatherers. *Journal of Clinical Endocrinology and Metabolism*, 46, 658–663.

van der Wijden, C., Brown, J., & Kleijnen, J. (2003). Lactational amenorrhea for family planning. *Cochrane Database of Systematic Review*, 4, CD001329.

Van Goozen, S. H. M., Wiegant, V. M., Endert, E., Helmond, F. A., & Van de Poll, N. E. (1997). Psychoendocrinological assessment of the menstrual cycle: The relationship between hormones, sexuality, and mood. *Archives of Sexual Behavior*, 26, 359–382.

Venners, S. A., Liu, X., Perry, M. J., Korrick, S. A., Li, Z., & Yang, F. (2006). Urinary estrogen and progesterone metabolite concentrations in menstrual cycles of fertile women with non- conception, early pregnancy loss or clinical pregnancy. *Human Reproduction*, 21, 2272–2280.

Vitzthum, V. J. (1989). Nursing behaviour and its relation to duration of post-partum amenorrhoea in an Andean community. *Journal of Biosocial Science*, 21, 145–160.

Vitzthum, V. J. (1990). *An adaptational model of ovarian function*. Research Report No. 90–200. Ann Arbor: University of Michigan, Population Studies Center.

Vitzthum, V. J. (1992). Infant nutrition and the consequences of differential market access in Nuñoa, Peru. *Ecology of Food and Nutrition*, 28, 45–63.

Vitzthum, V. J. (1994a). Causes and consequences of heterogeneity in infant feeding practices among indigenous Andean women. *Annals of New York Academy of Sciences*, 709, 221–224.

Vitzthum, V. J. (1994b). The comparative study of breastfeeding structure and its relation to human reproductive ecology. *Yearbook of Physical Anthropology*, 37, 307–349.

Vitzthum, V. J. (1997). Flexibility and paradox: The nature of adaptation in human reproduction. In M. E. Morbeck, A. Galloway & A. Zihlman (Eds.), *The evolving female: A life history perspective* (pp. 242–258). Princeton, NJ: Princeton University Press.

Vitzthum, V. J. (2001). Why not so great is still good enough: Flexible responsiveness in human reproductive functioning. In P. T. Ellison (Ed.), *Reproductive ecology and human evolution* (pp. 179–202). New York: Aldine.

Vitzthum, V. J. (2008a). Evolution and endocrinology: The regulation of pregnancy outcomes. In S. Elton & P. O'Higgens (Eds.), *Medicine and evolution: current applications, future prospects* (pp. 99–126). Boca Raton, FL: CRC Press.

Vitzthum, V. J. (2008b). Evolutionary models of women's reproductive functioning. *Annual Review of Anthropology*, 37, 53–73.

Vitzthum, V. J. (2009). The ecology and evolutionary endocrinology of reproduction in the human female. *Yearbook of Physical Anthropology*, 52, 95–136.

Vitzthum, V. J., Bentley, G. R., Spielvogel, H., Caceres, E., Thornburg, J., Jones, L., et al. (2002). Salivary progesterone levels and rate of ovulation are significantly lower in poorer than in better-off urban- dwelling Bolivian women. *Human Reproduction*, 17, 1906–1913.

Vitzthum, V. J., Ellison, P. T., Sukalich, S., Caceres, E., & Spielvogel, H. (2000). Does hypoxia impair ovarian function in Bolivian women indigenous to high altitude? *High Altitude Medicine and Biology*, 1, 39–49.

Vitzthum, V. J., & Smith, S. L. (1989). Evaluation of data on menstrual status and activity: An evolutionary perspective. *American Journal of Physical Anthropology*, 78, 318.

Vitzthum, V. J., Spielvogel, H., Caceres, E., & Gaines, J. (2000). Menstrual patterns and fecundity among non-lactating and lactating cycling women in rural highland Bolivia: Implications for contraceptive choice. *Contraception*, 62, 181–187.

Vitzthum, V. J., Spielvogel, H., Caceres, E., & Miller, A. (2001). Vaginal bleeding patterns among rural highland Bolivian women: Relationship to fecundity and fetal loss. *Contraception*, 6, 319–325.

Vitzthum, V. J., Spielvogel, H., & Thornburg, J. (2004). Interpopulational differences in progesterone levels during conception and implantation in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 1443–1448.

Vitzthum, V. J., Spielvogel, H., Thornburg, J., & West, B. (2006). A prospective study of early pregnancy loss in humans. *Fertility and Sterility*, 86, 373–379.

Vitzthum, V. J., Thornburg, J., Schaebs, F., Deimel, C., & Deschner, T. (2011). East meets West: Evaluations of the potential role of early life conditions on adult ovarian steroid levels in Mongolian and German women. *American Journal of Human Biology*, 23, 282–283.

Vitzthum, V. J., Thornburg, J., & Spielvogel, H. (2009). Seasonal modulation of reproductive effort during early pregnancy in humans. *American Journal of Human Biology*, 21, 548–558.

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

Vitzthum, V. J., Worthman, C. M., Spielvogel, H., & Thornburg, J. (2007). Adequacy of steroid levels for the regulation of gonadotrophin secretion during the ovarian cycles of Bolivian women. *American Journal of Human Biology*, 19, 286.

Voland, E. (1998). Evolutionary ecology of human reproduction. *Annual Review of Anthropology*, 27, 347–374.

Vollman, R. F. (1977). The menstrual cycle. *Major Problems in Obstetrics and Gynecology*, 7, 1–193.

Waller, K., Swan, S. H., Windham, G. C., Fenster, L., Elkin, E. P., & Lasley, B. L. (1998). Use of urine biomarkers to evaluate menstrual function in healthy premenopausal women. *American Journal of Epidemiology*, 147, 1071–1080.

Wallis, J. (1985). Synchrony of estrous swelling in captive group-living chimpanzees (*Pan troglodytes*). *International Journal of Primatology*, 6, 335–350.

Wallis, J. (1995). Seasonal influence on reproduction in chimpanzees of Gombe National Park. *International Journal of Primatology*, 16, 435–451.

Wang, D. Y., Key, T. J., Pike, M. C., Boreham, J., & Chen, J. (1991). Serum hormone levels in British and rural Chinese females. *Breast Cancer Research and Treatments*, 18(Suppl. 1), S41–S45.

Wasser, S. K., & Barash, D. P. (1983). Reproductive suppression among female mammals: Implications for biomedicine and sexual selection theory. *Quarterly Review of Biology*, 58, 513–538.

Weller, L., & Weller, A. (1993). Human menstrual synchrony: A critical assessment. *Neuroscience and Biobehavioral Reviews*, 17, 427–439.

West, B. T., Welch, K. B., & Galecki, A. T. (2006). *Linear mixed models: A practical guide using statistical software*. Boca Raton, FL: Chapman & Hall.

Wierman, M. E., Nappi, R. E., Avis, N., Davis, S. R., Labrie, F., Rosner, W., et al. (2010). Endocrine aspects of women's sexual function. *Journal of Sexual Medicine*, 7, 561–585.

Wilcox, A. J., Baird, D. D., Dunson, D. B., McConnaughey, R., Kesner, J. S., & Weinberg, C. R. (2004). On the frequency of intercourse around ovulation: Evidence for biological influences. *Human Reproduction*, 19, 1539–1543.

Wilcox, A. J., Dunson, D., & Baird, D. D. (2000). The timing of the "fertile window" in the menstrual cycle: Day specific estimates from a prospective study. *British Medical Journal*, 312, 1259–1262.

Wilcox, A. J., Dunson, D. B., Weinberg, C. R., Trussell, J., & Baird, D. D. (2001). Likelihood of conception with a single act of intercourse: Providing benchmark rates for assessment of post-coital contraceptives. *Contraception*, 63, 211–216.

Wilcox, A. J., Weinberg, C. R., & Baird, D. D. (1995). Timing of sexual intercourse in relation to ovulation: Effects on the probability of conception, survival of the pregnancy, and sex of the baby. *New England Journal of Medicine*, 333, 1517–1521.

Wiley, A. S. (1992). Adaptation and the biocultural paradigm in medical anthropology: A critical review. *Medical Anthropology Quarterly*, 6, 213–236.

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton, NJ: Princeton University Press.

Williams, S. R. (2006). Menstrual cycle characteristics and predictability of ovulation of Bhutia women in Sikkim, India. *Journal of Physiological Anthropology and Applied Human Sciences*, 25, 85–90.

Wilson, C. H. (1992). A critical review of menstrual synchrony research. *Psychoneuroendocrinology*, 17, 565–591.

Windham, G. C., Elkin, E., Fenster, L., Waller, K., Anderson, M., Mitchell, P. R., et al. (2002). Ovarian hormones in premenopausal women: Variation by demographic, reproductive and menstrual cycle characteristics. *Epidemiology*, 13, 675–684.

Winterhalder, B., & Smith, E. A. (2000). Analyzing adaptive processes: Human behavioral ecology at 25. *Evolutionary Anthropology*, 9, 51–72.

Winterich, J. A. (2003). Sex, menopause, and culture: Sexual orientation and the meaning of menopause for women's sex lives. *Gender and Society*, 17, 627–642.

Wood, J. W. (1994). *Dynamics of human reproduction: Biology, biometry, demography*. New York: Aldine de Gruyter.

Wood, J. W., Johnson, P. L., & Campbell, K. L. (1985). Demographic and endocrinological aspects of low natural fertility in Highland New Guinea. *Journal of Biosocial Sciences*, 17, 57–79.

World Health Organization. (1980). Temporal relationships between ovulation and defined changes in the concentrations of plasma estradiol-17 β, luteinizing hormone, follicle stimulating hormone and progesterone, I: Probit analysis. *American Journal of Obstetrics and Gynecology*, 138, 383–390.

World Health Organization. (1983a). A prospective multi-centre trial of the ovulation method of natural family planning, III: Characteristics of the menstrual cycle and of the fertile phase. *Fertility and Sterility*, 40, 773–778.

World Health Organization. (1983b). Temporal relationships between indices of the fertile period. *Fertility and Sterility*, 39, 647–655.

Worthman, C. M. (2003). Energetics, sociality, and human reproduction: Life history theory in real life. In K. W. Wachter & R. A. Bulatao (Eds.), *Offspring: Human fertility behavior in biodemographic perspective* (pp. 289–321). Washington, DC: National Academics Press, NRC.

Yang, Z., & Schank, J. C. (2006). Women do not synchronize their menstrual cycles. *Human Nature*, 17, 433–447.

Yonkers, K. A., O'Brien, P. M., & Eriksson, E. (2008). Premenstrual syndrome. *Lancet*, 371, 1200–1210.

Zera, A. J., & Harshman, L. G. (2001). The physiology of life history trade-offs in animals. *Annual Review of Ecology, Evolution, and Systematics*, 32, 95–126.

Zita, J. N. (1988). The pre-menstrual syndrome: "Dis-easing" the female cycle. *Hypatia*, 3, 77–99.