Hormones and Behavior 63 (2013) 13–19

Contents lists available at SciVerse ScienceDirect

Hormones and Behavior

journal homepage: www.elsevier.com/locate/yhbeh

Women's attractiveness changes with estradiol and progesterone across the ovulatory cycle

David A. Puts [a,*], Drew H. Bailey [b], Rodrigo A. Cárdenas [c], Robert P. Burriss [d], Lisa L.M. Welling [a], John R. Wheatley [a], Khytam Dawood [c]

[a] Department of Anthropology, Pennsylvania State University, University Park, PA 16802, USA
[b] Department of Psychological Sciences, University of Missouri, Columbia, MO 65211, USA
[c] Department of Psychology, Pennsylvania State University, University Park, PA 16802, USA
[d] Department of Psychology, University of Sterling, Sterling, FK9 4LA Scotland, UK

ARTICLE INFO

Article history:
Received 10 May 2012
Revised 6 November 2012
Accepted 8 November 2012
Available online 15 November 2012

Keywords:
Attractiveness
Estradiol
Face
Menstrual cycle
Progesterone
Voice

ABSTRACT

In many species, females are more sexually attractive to males near ovulation. Some evidence suggests a similar pattern in humans, but methodological limitations prohibit firm conclusions at present, and information on physiological mechanisms underlying any such pattern is lacking. In 202 normally-cycling women, we explored whether women's attractiveness changed over the cycle as a function of two likely candidates for mediating these changes: estradiol and progesterone. We scheduled women to attend one session during the late follicular phase and another during the mid-luteal phase. At each session, facial photographs, voice recordings and saliva samples were collected. All photographs and voice recordings were subsequently rated by men for attractiveness and by women for flirtatiousness and attractiveness to men. Saliva samples were assayed for estradiol and progesterone. We found that progesterone and its interaction with estradiol negatively predicted vocal attractiveness and overall (facial plus vocal) attractiveness to men. Progesterone also negatively predicted women's facial attractiveness to men and female-rated facial attractiveness, facial flirtatiousness and vocal attractiveness, but not female-rated vocal flirtatiousness. These results strongly suggest a pattern of increased attractiveness during peak fertility in the menstrual cycle and implicate estradiol and progesterone in driving these changes.

Introduction

Estrus is the phase of the ovulatory cycle immediately preceding ovulation and is accompanied by changes in appearance, odor and behavior across a variety of female mammals (Gangestad and Thornhill, 2008). Males respond to these cues by intensifying their mating efforts with estrous females. In *Homo sapiens*, estrus is not heralded by obvious signs, leading multiple researchers to conclude that natural selection favored the suppression of such signs in our species (Alexander and Noonan, 1979; Benshoof and Thornhill, 1979; Gangestad and Thornhill, 2008; Strassman, 1981; Symons, 1979).

Suppression of ovulatory cues by females would pose an adaptive problem for males. Males who could detect ovulatory cues, however slight, would be at an advantage in more efficiently directing their mating effort toward fertile females. Recent research indicates that some observable characteristics in women change over their cycles, including voice pitch (Bryant and Haselton, 2009) and skin color (Van den Berghe and Frost, 1986). Other research suggests that men may be capable of detecting these cues, generally preferring women's odors (Doty et al., 1975; Gildersleeve et al., 2012; Havlicek et al., 2006; Kuukasjärvi et al., 2004; Singh and Bronstad, 2001; Thornhill et al., 2003), faces (Roberts et al., 2004), and voices (Fischer et al., 2011; Pipitone and Gallup, 2008, 2011) during the late follicular (fertile) phase (see also Haselton and Gildersleeve, 2011 for a review).

Ancestral women may also have benefited from detecting the ovulatory status of other women. Not only do women appear to be more attractive at mid-cycle, but they are also more sexually attracted to men of putatively high genetic quality (e.g., Gangestad and Thornhill, 1998; Penton-Voak et al., 1999; Puts, 2005) and report more extra-pair sexual interests (Gangestad et al., 2002, 2005, 2010; Garver-Apgar et al., 2006; Haselton and Gangestad, 2006; Pillsworth and Haselton, 2006) and behavior (Bellis and Baker, 1990) at this time. At mid-cycle, women may therefore pose a greater threat to their same-sex rivals' ability to attract and retain mates. Some evidence suggests that women rate the odors (Doty et al., 1975; Kuukasjärvi et al., 2004), faces (Roberts et al., 2004) and voices (Pipitone and Gallup, 2008) of women near ovulation as being more attractive.

However, Gildersleeve et al. (2012) noted several methodological limitations of previous work in this area, including suboptimal data analysis (treating raters rather than stimulus donors as the unit of analysis) and design (between-subjects rather than within-subjects), reliance on

* Corresponding author. Fax: +1 814 863 1474.
E-mail address: dap27@psu.edu (D.A. Puts).

0018-506X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.yhbeh.2012.11.007

self-report data to establish ovulatory cycle position, and small sample sizes. Ultimately, the strongest support for cyclic changes in attractiveness will come from elucidation of the physiological mechanisms underlying them. Estradiol and progesterone are likely candidates for mediating any such changes (Haselton and Gildersleeve, 2011; Kuukasjärvi et al., 2004). For example, estradiol and progesterone receptors are expressed in laryngeal tissues (Ferguson et al., 1987; Marsigliante et al., 1996; Voelter et al., 2008), and puberty, pregnancy, menopause (Caruso et al., 2000), hormone replacement therapy (Firat et al., 2009) and hormonal contraceptive use (Amir et al., 2002) involve changes in both these hormones and vocal acoustics. Moreover, the day of ovulation can be precisely estimated using estrogen and progesterone (Baird et al., 1995), so if women's attractiveness varies with fertility, then it should also vary with these hormones. Yet, these associations remain unexplored.

We therefore investigated relationships between menstrual cycle fluctuations in women's estradiol and progesterone levels and their attractiveness to men and perceived mating threat to women (measured by attractiveness and apparent flirtatiousness). We explored these relationships in two characteristics highly salient to human mating, faces and voices (Puts et al., 2012b), using a within-subjects design and the largest sample yet collected for these purposes.

Material and methods

Participants

Two hundred and two normally-cycling women (mean age 19.6 ± 1.6 years) from 159 unique sibling groups (43 sister pairs, plus 116 singletons) participated in this research as part of a larger study involving siblings at a large Midwest U.S. university. Self-reported ethnicities were 91.6% White, 3.5% Asian, 2.0% Black or African American, 0.5% Hispanic or Latino, 0.5% Native Hawaiian or Other Pacific Islander, and 1.5% Other. Participants were scheduled for two laboratory sessions according to self-reported menstrual cycle length and date of the beginning of last menstrual bleeding. One laboratory session was scheduled within one day of expected peak estradiol production during the follicular phase, and the other session was scheduled within two days of expected peak progesterone production (mid-luteal phase), according to the methods of Puts (2006). Session order was counterbalanced across participants, and sessions occurred between 1300 h and 1600 h to minimize the influence of circadian hormonal fluctuations. Because we statistically analyze hormone levels rather than self-reported cycle phase, our use of the term "session" henceforth refers to first or second session rather than presumed follicular or luteal session. Approximately 12% of women attended only the first session.

Saliva collection and hormonal analysis

Participants collected approximately 9 ml of saliva in sodium azide-treated polystyrene test tubes during both sessions. Contamination of saliva samples was minimized by having participants not eat, drink (except plain water), smoke, chew gum, or brush their teeth for 1 h before each session. Participants rinsed their mouths with water before chewing a piece of sugar-free Trident gum (inert in salivary hormone assays) to stimulate saliva flow. The tube was capped and left upright at room temperature for 18–24 h to allow mucins to settle. Tubes were then frozen at −20 °C until analysis by the Neuroendocrinology Assay Laboratory at the University of Western Ontario, Canada.

Per previous research (e.g., Hampson et al., 2005; Oinonen and Mazmanian, 2007), progesterone was assayed using ^{125}I Coat-A-Count assay kits (Diagnostic Products Corporation, Los Angeles, CA) modified for use with saliva. Similar to previous research (e.g., Finstad et al., 2009), estradiol was assayed using ^{125}I Ultra-Sensitive E_2 RIA DSL-4800 kit (Diagnostic Systems Laboratories, Webster, TX) modified for use with saliva. Each sample was assayed twice, and average hormone levels for each sample were used in our analyses. Assay sensitivities were 0.65 pg/ml and 5 pg/ml, and intra-assay coefficients of variation were 5.1% and 10.7%, for estradiol and progesterone, respectively.

Facial photographs

Participants were provided wet wipes and instructed to remove any makeup, jewelry or spectacles and to assume a neutral expression. Facial photographs were taken with a tripod-mounted Canon PowerShot S10 digital camera at a distance of approximately 1 m, a height adjusted to the participant, and using constant lighting across participants. All face images were cropped beneath the chin, normalized on interpupillary distance, and rotated so that both pupils lay on the same horizontal plane.

Voice recording and analysis

Participants were recorded reading an excerpt from a standard voice passage (Fairbanks, 1960) in an anechoic, soundproof booth using a Shure SM58 vocal cardioid microphone. A curved wire kept the participant's mouth approximately 9.5 cm from the microphone. Voices were recorded using Goldwave software in mono at a sampling rate of 44,100 Hz and 16-bit quantization, and saved as uncompressed .WAV files.

Each recording was analyzed using Praat software (version 4.4.11). Pitch floor and ceiling were 100 Hz and 500 Hz, in accordance with the programmers' recommendations (Boersma and Weenik, 2009); otherwise, default settings were used. Across each recording, we measured mean (mean = 208.6 ± 17.6 Hz) and standard deviation (mean = 38.1 ± 9.4 Hz) of fundamental frequency (F_0, the acoustic correlate of pitch), duration (mean = 5.36 ± .90 s), number of voice breaks (mean = 14.6 ± 2.8), harmonics (mean = 15.5 ± 1.4 Hz), four measures of jitter (cycle-to-cycle variation in fundamental frequency), and five measures of shimmer (cycle-to-cycle variation in amplitude) using the 'voice report' function in Praat. All jitter (r > .90, mean r = .94) and shimmer (r > .47, mean r = .79) variables were correlated, so they were standardized (mean = 0, SD = 1) and summed (jitter: mean = −.41 ± 3.40; shimmer: mean = −.07 ± 4.32).

We also measured formant frequencies F_1 through F_4. Lower, more closely spaced formants correspond with a deeper vocal timbre. Formants were measured at each glottal pulse and averaged across measurements, as in Puts et al. (2012a). Formant measurements obtained by this method correlate highly (.93 ≤ r ≤ .98) with measurements obtained by measuring and averaging across individual vowels (Puts et al., 2012a). We then computed formant position (P_f, mean = 0.85 ± 0.40), defined as the average standardized formant value for the first four formants, using the method described in Puts et al. (2012a). The following between-sexes means and SDs were used to standardize formants: F_1 = 482.6 ± 49.8 Hz, F_2 = 1643.2 ± 145.7 Hz, mean F_3 = 2544.7 ± 173.9 Hz and mean F_4 = 3618.8 ± 266.8 Hz.

Face and voice ratings

Face photographs and voice recordings were rated by 568 men (mean age: 19.4 ± 1.8 years) and 558 women (mean age: 19.1 ± 2.4 years) from a large northeast U.S. university. Raters had a comparable ethnic distribution to the women who provided the photographs and recordings. Each rater assessed 24.9 ± 2.6 voice recordings and 24.1 ± 2.4 face photographs (including those of men and hormonally contracepting women not used here). Raters were presented a random sample of face photographs and voice recordings, except that no rater was presented with more than one photograph or recording from each participant. Using 7-point Likert scales, men rated stimuli on attractiveness for short- and long-term relationships, and women rated