Women Do Not Synchronize Their Menstrual Cycles

Zhengwei Yang
North Sichuan Medical College, Nanchong, People’s Republic of China

Jeffrey C. Schank
University of California, Davis

It is widely believed that women who live together or who are close friends synchronize their menstrual cycles. We reexamined this phenomenon in two ways. First, we collected data on menstrual cycles from 186 Chinese women living in dorms for over a year. We found that women living in groups did not synchronize their cycles. Second, we reviewed the first study reporting menstrual synchrony. We found that group synchrony in that study was at the level of chance. We then show that cycle variability produces convergences and subsequent divergences of cycle onsets and may explain perceptions of synchrony.

KEY WORDS: Estrous synchrony; Human pheromones; Menstrual cycles; Menstrual synchrony; Reproductive synchrony

Menstrual synchrony was first reported among college students more than three decades ago (McClintock 1971). The belief that menstrual synchrony occurs is quite common, with one study reporting that 80% of women believe synchrony occurs and 70% believe synchrony is a pleasant experience (Arden et al. 1999). Hypotheses have been proposed for menstrual synchrony as an adaptation in human populations (Burley 1979; Turke 1984), but these hypotheses have been questioned (Kiltie 1982; Schank 2001d, 2004). Menstrual synchrony has also been cited as the basis for human pheromone research (Kohl et al. 2001), but reports of human pheromones have been disputed (Strassmann 1999; Whitten 1999; Wilson 1987, 1988; Winman 2004).

Since the publication of the first study (McClintock 1971), a number of studies have been conducted, with some apparently replicating the finding of synchrony.

In this article, we report new results that we believe resolve this controversy. First, we report results from the largest and longest prospective study ever undertaken on female college students using state-of-the-art statistical techniques. Second, since the first study (McClintock 1971) is controversial (Wilson 1992; Strassmann 1999) and conflicts with the results presented here, we review it to determine whether the level of synchrony reported differed from chance. Third, we show that although synchrony does not occur, cycle variability may lead to the common perception of synchrony.

METHODS

To test whether or not women synchronize their menstrual cycles when living in close proximity, we collected data from 29 groups of women living in dorms at a Chinese university for more than a year. Four to eight women lived in each room, which should be ideal conditions for synchronization if pheromones are the mechanism of synchrony (Stern and McClintock 1998).

Our investigation was conducted from April 1, 2000 (denoted as Day 1), to July 4, 2001 (Day 460). The school vacations during the period were July 15 to August 25 (Days 106–147), and January 13 to February 16 (Days 288–322). Chinese schools have two semesters each year with two vacations (i.e., a six-week summer vacation during July and August and a four-week winter vacation during January and February) between semesters. Most students stay off campus during the vacations, and the school year starts after summer vacation.

A total of 186 women participated in the study. Ninety-three women began college in September 1997 and 1998 (Grade 1997–1998) and resided in 13 rooms, each room housing five to eight students. The other 93 women (Grade 2000) began college on August 26, 2000 (Day 148 of the study), and resided in 16 rooms, each housing four to eight students. Grade 1997–1998 women were 19–23 years old, while Grade 2000 women were 17–21 years old. All students were from different high schools in Sichuan Province, and few had been roommates prior to college.

All Grade 1997–1998 women and half of Grade 2000 women resided in the same building; the other half of Grade 2000 women resided in another dormitory. Both are five-story buildings with rooms along two sides of a common corridor;
there are 28 rooms in each story, 14 rooms on each side of the common corridor. Rooms in the first dormitory are $5.0 \times 3.5 \times 3.3$ m, and $3.7 \times 3.1 \times 3.0$ m in the second dormitory.

Women were given a notebook and asked to keep a record of all of their menstrual onset dates during the period of investigation. During the school year, six investigators from the Grade 1997–1998 women went to the dormitory every one to two weeks to collect the recorded results. Data recorded during vacation were collected at beginning of the school term.

**Statistical Analysis**

Upon inspecting the data (Figure 1), we found cycle variability typical of women in the age range of 17 to 23 years (Harlow et al. 2000). Cycle variability mathematically rules out phase-locking synchrony (Winfree 1980), but it may be that there are subtle statistical clusterings of cycle onsets. To analyze this possibility and to avoid problems in determining initial onset dates for statistical analysis (Wilson 1992), we developed several bootstrap techniques to detect clustering of cycle onsets. The first technique used the Kuiper-Stephens non-parametric test for deviations from non-uniform distributions on a circle (Stephens 1965). The assumption that all cycles have the same period did not hold for these data, but using the mean cycle length as the cycle period for a group allowed us to detect clustered onsets for further graphical analysis. Groups in which clustered onsets were detected were then plotted to determine whether clustering persisted or dissipated over time. Dissipation of onset clustering is predicted when there is cycle variability.

To analyze the data for all groups, we used bootstrap-resampling algorithms that randomly shuffled the relationships among cycle onsets to create synthetic data sets. To do this, we wrote a program in C, which randomly shuffled the order of the cycles within women (A1) or shuffled the onset dates by a specified number of days (A2) (Schank 2001b, 2001c). For A2, onset dates were shifted by up to $\bar{m}_j/4$ days, where $\bar{m}_j$ is the mean cycle length for each group, $j$. By doing this repeatedly for each group, two sets of 100,000 synthetic data sets were created using A1 and A2.

Three measures of onset clustering were then applied to each of the synthetic data sets: (1) the number of clustered onsets detected using the Kuiper-Stephens test; (2) McClintock’s mean onset-date deviation (MD) measure (McClintock 1971):

$$\sum_{i=1}^{n} \left| x_i - \bar{x} \right|$$

where $n$ is the number of women in a group, $x_i$ is the menstrual-cycle onset date for the $i$th woman, and $\bar{x}$ is the mean of the onset dates for the group; and (3) the pairwise onset-date mean deviation (PD) measure in a group. PD is the absolute difference between each unique $(i \neq j)$ pair of onsets divided by the total number of comparisons (Schank 1997):
$\frac{(n-2)! \cdot 2}{n!} \sum_{i=1}^{n} \sum_{j=i+1}^{n} |x_i - x_j|$  

(PD)

where $x_i$ and $x_j$ are the onset dates of women $i$ and $j$, respectively. In general, the more clustered the onsets are for a group, the smaller MD and PD will be, but see Schank (1997) for problems with these types of measures.

RESULTS

Using the Kuiper-Stephens test as a detector of clustering, we found 9 of 29 groups that appeared to have at least one set of clustered onset dates (Figure 2). The first thing to notice about Figure 2 is that, for all but two of the groups, clustered onsets do not persist owing to cycle variability (groups 97–521, 97–525, 98–508, 00–408, 00–S506, 00–S509, 00–S511). Group 98–510 exhibits apparent convergence of cycle onsets after more than a year, but as the study ended, they began to diverge again. Women in group 00–S503, after 9 months of living together, exhibit apparent onset convergence for the very last onsets recorded. However, given the cycle variability in this group, we would expect that cycle onsets would have subsequently diverged. Thus, cycle variability continually changes the cycle onset relationships among women in a group, and Figure 2 illustrates that stable, convergent relationships among cycles do not persist through time.

This analysis does not, however, tell whether 9 of 29 groups with clustered onsets are more than expected by chance. To answer this question, we applied the Kuiper-Stephens test to two synthetic data sets we created with our two shuffling algorithms. We found that the observed number of 9 deviations did not differ from chance (Figure 3). Both algorithms indicated that the expected number of deviations by chance is about 10 out of 29 groups.
Figure 2. Cycle onsets for 9 of 29 groups where sets of nearest onset dates (indicated by open circles) deviated from a uniform distribution around a circle. Each group is designated as Grade-Room number. Note that cycle variability causes convergence and divergence of onsets.
In addition, we also used the two most common measures of synchrony in the literature (Schank 1997, 2001b): MD and PD. Neither MD nor PD is a good measure of deviations from a uniform distribution, but both, though not ideal, heuristically indicate group clustering (Schank 1997). Thus, for historical completeness, we included analyses using MD and PD. Neither measure differed significantly from chance (MD: $p > .5$ and PD: $p > .43$ using algorithm A1, and MD: $p > .26$ and PD: $p > .28$ using algorithm A2). Thus, there was no onset convergence and no synchronization of cycles.

**DID THE FIRST STUDY FIND MENSTRUAL SYNCHRONY?**

Our results conflict with the results from the first (McClintock 1971) and subsequent studies reporting synchrony (Goldman and Schneider 1987; Graham and McGrew 1980, 1992; Little et al. 1989; Matteo 1987; Quadagno et al. 1981; Skandhan et al. 1979; A. Weller and L. Weller 1992, 1993, 1995a, 1997; L. Weller and A. Weller 1993; Weller, Weller, and Roizman 1999; Weller et al. 1999). All of these studies have been shown to contain serious methodological errors (Arden and Dye 1998; Schank 2000b, 2001b, 2002; Strassmann 1999; Wilson 1992). However, we focus on the first study—the most cited, and the study all others aimed to replicate—to show that the level of synchrony reported in the first study was in fact at the level of chance.

In the first study (McClintock 1971), data were collected during the late 1960s from 135 women 17–22 years old living in single or double dormitory rooms be-
between late September and early April. Three times during the academic year each woman was asked to recall “when her last and second to last menstrual periods had begun . . . and listed by room number the girls \( N \leq 10 \) with whom they spent the most time, indicating which two of these they saw most often” (McClintock 1971:244). Women were then paired into 33 pairs of closest friends and 33 pairs of roommates (Wilson 1992). Two women were paired as closest friends “only if both had indicated that they saw each other most often” (McClintock 1971:244). The 66 pairs of roommates and close friends apparently synchronized their cycles slightly from October to April (McClintock 1971).

Two critical problems have been raised concerning the first study. First, Wilson (1992) showed that the method used to choose initial menstrual-cycle onset dates for women was biased toward asynchrony. Thus, the apparent increase in synchrony from October to April may have been due to an error in determining initial onset dates, which were too asynchronous, rather than a process of synchronization.

Second, were the two sets of closest friends and roommates mutually exclusive (Wilson 1992)? At least 48 women lived in single rooms (i.e., “Six smaller living areas, separated from the main corridor by at least one door, each house approximately eight girls in single rooms” [McClintock 1971:244]). Some of the other dorm rooms also housed women without roommates. If we assume 66 women lived as roommates, then there were 69 women living in single rooms. This implies that no woman listed her roommate as a closest friend and no single woman listed a woman who was a roommate as a closest friend. This seems extraordinarily unlikely. We can conclude with near certainty that women could not have been uniquely paired into closest friends or closest friends excluding roommates. This suggests that the only way to organize these data is as groups of close friends.

All women were also arranged into 15 groups of close friends (ranging from 5 to 10 women) (McClintock 1971). Group synchrony was measured using MD (see above), and mean difference scores for the 15 groups were reported to have decreased (indicating increased synchrony) by about 1.8 days from October to April (McClintock 1971; Wilson 1992). However, it is not likely that all 135 women were grouped in the range of 5 to 10 women per group (Wilson 1992). If we assume that at least one group each contained 5, 6, 7, 8, and 9 women, then at least 10 groups had 10 women in order for the groups to sum to 135. Women were included in a group only if they mutually listed each other as close friends, but women only listed up to 10 women as close friends (McClintock 1971). This implies that, in each of group of 10 women, each woman in a group listed exactly the same 10 close friends. The probability that, in 10 groups of 10 women, each woman in a group listed exactly the same 10 close friends is again extraordinarily unlikely. Wilson recognized this problem and estimated that perhaps a maximum of 105 women could have been grouped into an average of 7 women per group (Wilson 1992).

In whatever way the women were grouped, Page’s (1964) test for ordered hypotheses with multiple treatments was used to determine if the groups synchronized their cycles from October to April. Five treatment conditions were used in
this test, but Page’s test requires independent treatment conditions and not repeated
conditions (i.e., the same groups of women were used in the five treatment condi-
tions), so the test was not validly applied. Confidence intervals were also calculated
and used to analyze these data (see Figure 1 in McClintock 1971), but the confi-
dence intervals were again not valid for comparison because the treatment condi-
tions were not independent, and the synchrony scores were from different group
sizes. Thus, because inappropriate statistical procedures were used, we cannot de-
termine from the reported analysis whether or not group synchrony occurred.

Even though invalid statistical techniques were used, we can still assess whether
or not synchronization may have occurred by finding the expected value of MD for
groups when cycle onsets among women are related by chance alone. If the level of
synchrony expected by chance is approximately 6.4 days (the value reported for
October: McClintock 1971; Wilson 1992), then a drop of 1.8 days to 4.6, though
small, may have been statistically significant.

When a group consists of only two women with menstrual cycles $m$ days long
and cycle onsets are randomly related, the expected value of MD can be solved
mathematically. We found that MD has an approximate value of $m/8$.\(^1\) The mean
menstrual-cycle length for the groups in first study (McClintock 1971) was proba-
bly in the range of 28 to 30 days (Wilson 1992). For two women with menstrual
cycles that are 29 days long, MD has an expected value of 3.63 days, well below
both the 6.4 days reported at the beginning (October) and the 4.6 days (April) re-
ported at the end (McClintock 1971), suggesting that the original study’s results
were more asynchronous than expected by chance both at the beginning and the
end. However, as group size increases, the expected value of MD gradually and
asymptotically increases from approximately $m/8$ to approximately $m/4$.\(^2\) Thus, the
expected level of synchrony occurring by chance for 29-day cycles ranges from
3.63 days for groups of $n = 2$ women to 7.24 days as $n$ becomes large.

For small groups greater than two and when cycles are variable, the expected
value of MD cannot be solved analytically. However, Monte Carlo simulation can
be used to find the expected values. To calculate the expected values, we wrote a
program in C to create two types of synthetic data sets. The first consisted of cycles
with intervals of exactly $m$ days; the first cycle date was an integer randomly se-
lected from the range $[1, m]$. The second set consisted of cycle lengths drawn ran-
domly from our empirical distribution (as illustrated in Figure 1) whose first cycle
date for each individual was a random integer in the range 14 to 40 days. The mean
and standard deviation of the cycle lengths in McClintock’s (1971) study were 29.7
days, s.d. = 5.7. We truncated the extremes of our distribution so that cycles ranged
from 14 days to 40 days in length. The resulting mean was 29.7 days (s.d. = 4.2)
matching McClintock’s (1971) study mean but with less cycle variation. We found
that for group sizes in the range of 5 to 10 women, the expected value of MD is
approximately 5 days when cycle onsets are randomly related to each other (see
Figure 4). Thus, the reported level of synchrony in the first study (McClintock
1971) was too asynchronous in October (i.e., beginning of the study) and subsequently decreased by April to a level expected by chance (Figure 4). Graham and McGrew (1980, 1992) also used MD as a group synchrony measure, and although they reported that women synchronized their cycles, the beginning and ending values of MD were actually more asynchronous than expected by chance, which was likely due to errors in determining closest onset dates as described by Wilson (1992). In short, the method used to determine initial onset dates for comparison in the first study (McClintock 1971) was biased toward asynchrony (Wilson 1992). The small drop of 1.8 days from October to April can be explained as an error in calculating initial synchrony scores that were too high (i.e., too asynchronous; see Wilson 1992), which subsequently dropped to chance levels. Thus, the first menstrual-synchrony study (McClintock 1971) did not show that women synchronize their menstrual cycles.

EXPLAINING THE PERCEPTION OF SYNCHRONY

Women often have the perception that their menstrual cycles are synchronized with other women (Arden et al. 1999). With the help of Figure 2, we can explain how this perception of onset convergence is real and yet synchrony does not occur. For example, group 97–525 exhibits convergence for two cycles around day 250 of the
study. Notice that the onsets just prior to these onsets appear to be converging, and after two sets of convergent onsets, they diverge again because of cycle variability.

Figure 2 also illustrates other examples of group-onset convergence and subgroup convergence within groups. That is, although an entire group may not converge, subgroups within groups may temporarily converge (e.g., 00–408 and 00-S506). Indeed, inspection of Figure 2 (and other plots not shown) reveals that there are nearly always two or more women in a group with converging cycle onsets that subsequently diverge over time. Thus, unless a group or subgroup of roommates or friends have cycles of exactly the same length, they may perceive cycle-onset convergence if they monitor each other’s cycles over time.

DISCUSSION

In the largest study ever undertaken, we show here that synchrony did not occur when women lived together for a year or more, and that menstrual cycle onset clusterings were neither stable (Figure 2) nor occurred more than expected by chance (Figure 3). This is because rhythms with cycles of variable length cannot synchronize unless rhythms converge on the same frequency (i.e., cycles must converge on the same cycle length) (Wilson 1992; Winfree 1980). We next demonstrated that the original results (McClintock 1971) were flawed and at the level of chance. Figure 1 illustrates that cycle variability in the first study (McClintock 1971) is comparable to the cycle variability in our study, which also explains why we found that the level of synchrony reported in the first study was actually at the level of chance. We also showed that cycle variability leads to repeated cycle convergences and subsequent divergences, which may explain the perception of synchrony (Arden et al. 1999). Thus, because of menstrual cycle variability (Arden and Dye 1998; Schank 2000b, 2001b; Wilson 1992) and its mathematical implications for stable convergence of cycles (Winfree 1980), we are able to draw the strong conclusion that women do not synchronize their menstrual cycles.

In the 1980s, several studies reported estrous synchrony in groups of Norway rats (McClintock 1978), golden hamsters (Handelmann et al. 1980), chimpanzees (Wallis 1985), and golden lion tamarins (French and Stribley 1985). In the nearly 20 years since these studies were published, no other mammals have been reported to synchronize their cycles when living in groups. Subsequent methodological studies have shown that these early results were flawed, the results were due to chance (Schank 2000a, 2001a, 2001d), and they failed to replicate (Gattermann et al. 2002; Monfort et al. 1996; Schank 2001a, 2001c). Indeed, in the case of golden hamsters, if females are presynchronized (i.e., females are matched for being in the same estrous state) and then grouped, they significantly desynchronize their cycles (Gattermann et al. 2002).

There are also theoretical reasons to think that mutual synchronization of cycles may be a costly mating strategy for females in some contexts. Schank (2004) cre-
ated an agent-based simulation model of Calhoun’s (1962) study of wild Norway rats, which are promiscuous in mating. Since Calhoun (1962) created detailed maps of trails, burrows, trees, and so on, an exact reorientation of the habitat could be simulated. Calhoun kept track of the number of reproductive males and females and described in detail the scramble of males after females in heat. Schank (2004) simulated this system and found that if males differed in quality, then females who were not in heat on the same night experienced on average more matings with higher-quality males and more mating opportunities. By avoiding synchrony, competition among females for high-quality males was reduced and fitness increased. Schank (2004) also found that variable estrous cycles were all that was needed to avoid synchrony and its costs. Thus, the evidence (empirical and theoretical) to date supports the conclusion that mutual synchronization of estrous or menstrual cycles does not occur in groups of female mammals.

This study was supported by the Sichuan Youth Foundation of Science and Technology (Chuan Ke Ji [2001] 2) and NIH Grant No. 5R01MH065555–02. We thank Shuhuan Yang, Xia Xu, Jun Qu, Jing Zhu, Meijia Yu, & Chang Zhou for all their work in collecting the data. Without them, this study could never have been done.

Zheng-Wei Yang is director and professor of the Morphometric Research Laboratory, North Sichuan Medical College, China. His main research interests are in spermatogenesis, morphometry, and stereology.

Jeff Schank is associate professor of psychology at the University of California, Davis. His main research interests are in computational and biorobotic modeling of group behavior and the development of sensorimotor behavior in animals.

NOTES

1. McClintock (1971) used equation MD as a measure of group synchrony in women. The expected value of MD can be calculated when (a) either the number of women in a group is 2 or \( n \) is large (i.e., \( n \to \infty \)), (b) cycles are of the same length \( m \), and (c) cycle onsets among women are randomly related. Under these assumptions, the expected value of MD is \( m/8 \) (or approximately \( m/8 \)) for \( n = 2 \). This is intuitively correct because a mean is a center of gravity, so if the expected difference between cycle onsets is approximately \( m/4 \), the center of gravity is approximately \( m/8 \). For example, suppose two women with 28-day cycles have onsets on February 14 and 21. The mean of these two dates is 17.5. The absolute difference between the onsets is 7 days and the absolute difference from the mean for each woman is 3.5 days.

However, the exact values depend on whether the cycles are odd or even. More precisely, suppose individuals A and B have cycles with integer periods \( m > 1 \) and that the cycle onsets between women are uniformly and randomly distributed around a circle. For any onset date \( i \) of A and the closest onset date \( j \) of B, there are four solutions for the expected value of MD depending on the even and odd properties of \( m \).

If \( m \) is even and

1. if \( m/2 \) is even, \( E(MD) = \frac{m}{8} \)

2. if \( m/2 \) is odd, \( E(MD) = \frac{m}{8} + \frac{1}{2m} \)
If \( m \) is odd and

3. if \((m - 1)/2\) is even, \( E(MD) = \frac{m}{8} + \frac{1}{8m} - \frac{1}{4m^2} \)

4. if \((m - 1)/2\) is odd, \( E(MD) = \frac{m}{8} + \frac{1}{8m} + \frac{1}{4m^2} \)

When \( m \) is in the range of the typical cycle lengths for women (e.g., 28 to 31 days), then cases 2, 3, and 4 are approximately \( m/8 \) (contact Schank for the mathematical details of the solutions).

2. The expected value of MD can also be solved as \( n \to \infty \), which is approximately \( m/4 \). Suppose group size approaches infinity (i.e., \( n \to \infty \)) with cycles of integer period \( m > 1 \). If cycle onsets are in a uniform random distribution around a circle, then there are two solutions to MD depending on whether \( m \) is even or odd.

1. If \( m \) is even, \( \lim_{n \to \infty} E(MD) = \frac{m}{4} \)

2. If \( m \) is odd, \( \lim_{n \to \infty} E(MD) = \frac{m}{4} - \frac{1}{4m} \)

When \( m \) is in the range of the typical cycle lengths for women (e.g., 28 to 31 days), then case 2 is approximately \( m/4 \). This is also the expected value for a uniform distribution and is the maximum asynchrony possible for a group of women with cycles of length \( m \).

REFERENCES

Arden, M. A., and L. Dye

Arden, M. A., L. Dye, and A. Walker

Burley, N.

Calhoun, J.

Cepicky, P., F. Mandy, L. Hlavicka, and K. Sosnova

French, J. A., and J. A. Stroble

Gattermann, R. K. Ulbrich, and R. Weinandy

Goldman, S. E., and H. G. Schneider

Graham, C. A., and W. D. McGrew
Testing Menstrual Synchrony in China


Handelmann, G., R. Ravizza, and W. J. Ray

Harlow, S. D., X. Lin, and M. J. Ho

Jarett, L. R.

Kiltie, R. A.

Kohl J. V., M. Atzmueller, B. Fink, and K. Grammer

Little, B. B., D. S. Guzick, R. M. Malina, and M. D. Rocha Ferreira

Matteo, S.

McClintock, M. K.


Monfort, S. L., M. Bush, and D. E. Wildt

Page, E. B.

Quadagno, D. M., H. E. Shubeita, J. Deck, and D. Francoer

Schank, J. C.


Skandhan, K. P., A. K. Pandya, S. Skandhan, and Y. B. Mehta

Stephens, M. A.

Stern, K., and M. K. McClintock

Strassmann, B. L.


Trevathan, W. R., M. H. Burleson, and W. L. Gregory

Turke, P. W.

Wallis, J.

Weller, A., and L. Weller

Weller, L., and A. Weller

Weller, L., A. Weller, and S. Roizman

Whitten, W.

Wilson, H. C.

Wilson, H. C., S. H. Kiefhaber, and V. Gravel
Winfree, A. T.

Winman, A.