
Jeffrey C. Schank
University of California, Davis

This article replies to commentaries by A. Weller and L. Weller (2002) and by C. A. Graham (2002). The author of this reply argues that A. Weller and L. Weller merely defined away the problem of cycle variability for synchrony by assuming either that all cycles are 28-days long or that the expected difference between 2 cycles is 1/4 the mean of the cycles of 2 rhythms. In her commentary, C. A. Graham stated that A. Weller and L. Weller’s later research did not use recall data, but the author of this reply shows that this is not true. Menstrual-synchrony research taken as a whole is plagued by a multitude of systematic errors that lead inevitably to the conclusion that there is no evidence for menstrual synchrony among women.

Ten years ago, Wilson (1992) published a comprehensive critique of the menstrual-synchrony literature. Since then, no researchers, other than Weller and Weller, have published studies reporting synchrony. Moreover, it has been 17 years since the last publication reporting synchrony in another mammalian species (French & Stribley, 1985; Handelmann, Ravizza, & Ray, 1980; McClintock, 1978; Wallis, 1985); only studies finding no significant effects of synchrony have been published (e.g., Erb, Edwards, Jenkins, Mucklow, & Wynne-Edwards, 1993; Monfort, Bush, & Wildt, 1996; Schank, 2000, 2001a, 2001c; Strassmann, 1997). In Schank (2000, 2001b), I critically reexamined the methods referred to by A. Weller and Weller (2002). In their research over the last decade, they have used two (perhaps three; H. C. Wilson, personal communication, January 28, 2002) basic approaches. In this reply article, I refer to these as their early method (Schank, 2000) and later method (Schank, 2001b). In this reply, it is their later method that I am primarily concerned with, and I respond to A. Weller and Weller’s (2002) and Graham’s (2002) commentaries by topic.

A. Weller and Weller (2002)

Cycle Variability

At the center of the issues that I (Schank, 2000, 2001b) and others before me (Arden & Dye, 1998; Strassmann, 1997, 1999; Wilson, 1992) have raised is the problem of menstrual-cycle variability. The basic problem is that for two rhythms to be synchronized, their cycles must be integer multiple lengths of each other (Winfree, 1980). For example, 28-day rhythms can be synchronized but 28- and 29-day rhythms cannot. Similarly, 20- and 40-day rhythms can be synchronized in a ratio of 1:2 (e.g., one 40-day onset to every other 20-day onset). A. Weller and Weller (2002) and L. Weller and Weller (1997) have never addressed the problems of inter- and intrawoman cycle variability. Weller and Weller’s methods simply define these problems away. In their early method (Schank, 2000), A. Weller and Weller (1993) assumed for the purpose of measurement and statistical analysis that menstrual cycles were exactly 28 (or 29) days long:

We first calculated synchrony on the basis of an expected 28 day [sic] cycle. However, since the results may be biased to synchrony when the expected cycle length is based on less than the actual cycle length, we also calculated our results against an expected 29 day cycle, which would be biased against synchrony. (A. Weller & Weller, 1993, p. 945)

In their later method (Schank, 2001b), L. Weller and Weller (1997) assumed that the expected onset difference between two rhythms is one quarter of the mean of their cycle lengths. Onset differences that are less than the expected difference are considered synchronous. For example, suppose woman A has exactly 27-day cycles and woman B has 35-day cycles. The expected difference according to A. Weller and Weller (2002) is 31/4 days (the mean for A and B), which is 7.75 days. Suppose B has a menses onset on March 2nd and A has a menses onset on March 7th. The difference for the 1st month is 5 days, which is synchronous (i.e., 5.00 < 7.75). The next onset for A is 7 + 27 = 34 days later and for B is 2 + 35 = 37 days later, a 3-day difference, which is again synchronous (i.e., 3.00 < 7.75). The third onset for A is 34 + 27 = 61 days later and for B is 37 + 35 = 72 days later, an 11-day difference, which is not synchronous (i.e., 11.00 > 7.75). According to L. Weller and Weller (1997), having 2 out of 3 months of synchrony is sufficient to say two menstrual cycles are synchronized. However, it is mathematically impossible for a 27-day rhythm to be synchronized to a 35-day rhythm: onsets repeatedly diverge and converge over time (see Schank, 2001b, for more detail). Thus, A. Weller and Weller (2002) and L. Weller and Weller (1997) introduced methods that are mathematically inconsistent with the basic concept of synchrony (Winfree, 1980).
The Expected Differences Between Cycle Onsets

As I just discussed, A. Weller and Weller (2002) assumed that the expected onset difference between two rhythms is one quarter of their mean cycle length. However, what can be assumed and what actually is assumed by a researcher’s methods can be radically different. This is why a substantial proportion of statistical research today is devoted to Monte Carlo simulation studies that investigate the robustness of new and even well-established statistical approaches. In all of Weller and Weller’s published research on synchrony, they have never investigated the validity of their key statistical assumptions.

Consider an example discussed by L. Weller and Weller (1997) regarding two women, one with 40-day cycles and the other with 20-day cycles. The mean of these two cycle lengths is 30 days, and one quarter of their length is 7.5 days. This is the same value as would be correctly calculated for two women with exactly 30-day cycles (Schank, 2001b). However, consider how far a 40-day cycle onset can be from the onsets of a 20-day rhythm. At most, it can only be 10 days from the nearest 20-day onset. If one 20-day onset is 40 days from a given 40-day onset, then there must be another 20-day onset that exactly matches the 40-day onset. The rhythms in this case are exactly synchronized in a ratio of 1:2, and the expected difference due to chance is 5 days, not 7.5 days. However, the critical question is what the expected onset differences are when using L. Weller and Weller’s (1997) three-way comparison method (Schank, 2001b).

If \( A_1, A_2, B_1, \) and \( B_2 \) are onsets of two women temporally ordered by their indices, then the initial onset according to L. Weller and Weller (1997) is the smaller of three absolute differences: \( |A_1 - B_1|, |A_1 - B_2|, \) or \( |A_2 - B_1| \). Suppose \( A \) has 20-day rhythms and \( B \) has 40-day rhythms and we randomly shuffle these two rhythms 1,000,000 times and apply the three-way comparison method (Schank, 2001b). What is the expected initial onset difference? It turns out to be 5.83 (which can also be analytically derived from the three-way comparison method), a meaningless number due to an insufficient number of onset comparisons, but nonetheless, the expected difference generated by their three-way comparison method (Schank, 2001b). Nevertheless, 5.83 is less than 7.50 and indicates an intrinsic bias toward the detection of synchrony when it does not exist, a Type I error.

Cycle Variation

In A. Weller and Weller (2002), they claimed that I used too high an estimate of cycle variation in their studies. What I did was to simulate a range of variation from no variation to just below \( SD = 5.3 \) the greatest cycle variation Weller and Weller reported (i.e., \( SD = 5.5 \); L. Weller, Weller, & Roizman, 1999). In their Bedouin families study, A. Weller and Weller (1997) did not report standard deviations, but in A. Weller and Weller (2002), they recomputed the means and found a mean cycle length of 30.26 days \( (SD = 1.52) \) for the 1st month. This standard deviation is indeed much lower than what they have typically reported, but was it computed correctly? The mean for the 1st month reported by A. Weller and Weller (1997) was 30.97 days for the same data (as extrapolated from Table 1, A. Weller & Weller, 1997, p. 147). This is a substantial discrepancy. If the 1st month’s mean cycle length was 30.26 days with an expected difference of 7.57 days, then why do all the expected scores range from 7.70 to 7.76 days in their original publication (A. Weller & Weller, 1997)?

Fifty Percent Replicability of Weller and Weller’s Synchrony Studies

In A. Weller and Weller (2002), they argued that if there are errors and biases in their methods, the replicability rate of their studies should be much higher than 50%. Frankly, I do not understand this argument. If synchrony is a real phenomenon, then their replicability should be higher than 50%. Indeed, 50% replicability suggests they have introduced a high Type I error rate.

Increasing Asynchrony Over Time

Arden and Dye (1998) first pointed out that cycle variability decreases synchrony over successive months as an artifact of A. Weller and Weller’s (1997) methods. In Table 1 of Schank’s (2001b) article, I summarized the results of all the studies using Weller and Weller’s later method. Applying a sign test, the binomial probability of synchrony decreasing over 3 months in all of their later studies was \( p = .5^3 \) and \( p = .004 \). A. Weller and Weller (2002) objected that not all decreases were significant. However, this was not the question. I was concerned with whether the artifact described by Arden and Dye occurred in their study, and it does. Synchrony scores increase over time. Indeed, A. Weller and Weller (2002) appeared unconcerned with sequences of synchrony scores (e.g., 6.32, 6.24, and 7.40; roommates—sisters, Table 1, Schank, 2001b) in which there is an unexplained large decrease in synchrony (i.e., increased difference score). Why did the synchrony score increase from 6.24 to 7.40 on the 3rd month? Why, in all of their studies (see Table 1, Schank, 2001b), do scores become more asynchronous by the third cycle when they are supposed to be stably synchronized?

Random Control Groups

McClintock (1971) introduced the method of constructing a random control group from the actual sample by randomly repairing up women and calculating their synchrony scores. The actual and constructed samples are then compared using a \( t \) test or a Mann–Whitney \( U \) test. Fortunately, this invalid statistical maneuver has only been used by a few menstrual-synchrony researchers. This procedure is invalid because the actual sample and the constructed sample are not independent. For example, if all participants are randomly repaired, then the mean cycle length and variance will be identical for the two groups.

There are also questions concerning how the random controls were constructed and whether random group controls were constructed until a significant difference was found. For example, A. Weller and Weller (1993) applied the random groups method to their dormitory samples. The mean synchrony score of the real data was approximately 7.0 days, and the mean synchrony score for the randomly constructed control group was 7.3 days (see A. Weller & Weller, 1993, Figure 2, p. 947). A. Weller and Weller reported \( p < .01 \) using the Mann–Whitney \( U \) test. However, because 7.0 days was the expected onset difference, they found a significant difference between two nonsynchronous samples! Wil-
Most Studies Rely on Participant Recall

In her commentary, Graham (2002) stated that the majority of menstrual-synchrony studies have not required women to recall their menses onsets but have asked women to fill out calendars instead. This is not true. L. Weller and Weller (1997), in their later method, began using calendars in conjunction with asking women to recall their last and next to last menses onsets. Graham stated only part of A. Weller and Weller’s (1997) method. In that study, women also filled out a questionnaire that asked them to recall “. . . dates of their last and next to last period” (A. Weller & Weller, 1997, p. 146). The reason A. Weller and Weller needed these recalled dates is because typically four onset dates are required to use the three-way comparison methods to calculate onset data differences for 3 months. Suppose $|A_1 - B_2|$ is the smallest difference and so this is the initial onset comparison for Month 1; then for Month 2, the onset difference is $|A_2 - B_1|$, and for Month 3, it is $|A_3 - B_4|$. Thus, in two thirds of the cases when either $|A_1 - B_2|$ or $|A_2 - B_3|$ were the smallest of the three comparisons, A. Weller and Weller had to use recalled onsets to get 3 months of onset comparisons.

Failure to Control for the Mutual Exchange of Information in Filling Out Calendars

According to her commentary, Graham (2002) assumes that filling out menstrual calendars is a reliable and accurate method for determining menses onsets. It is no defense to cite the use of menstrual calendars in other areas such as fertility research (Graham, 2002) because there are no studies that have quantified how accurately and reliably participants fill out calendars. The accuracy and reliability in filling out calendars may vary with the nature of the study and with the importance of the study to the participant. Participants in a fertility study may feel a greater personal obligation to reliably fill out calendars than college students fulfilling a research participation requirement in their introductory courses. Most important, menstrual-synchrony studies have differed from other areas of research in requiring friends and roommates or mothers and daughters to fill out separate calendars, and yet none of A. Weller and Weller’s (2002) studies have implemented controls for the mutual exchange of information. In a previous study (Schank, 2001b), I showed that even very small systematic errors due to the mutual exchange of information or to recall could produce the synchrony effects reported by A. Weller and Weller (e.g., 1997). It has long been known that participant recall or self-report of fact is problematic: “On average, about half of what informants report is probably incorrect in some way” (Bernard, Killworth, Kronenfeld, & Sailer, 1984, p. 503).

Discussion

Graham (2002) concluded her commentary stating “Menstrual synchrony is a difficult phenomenon to demonstrate, its occurrence among women remains uncertain, or at least unpredictable, and if it does exist, its functional significance is unclear” (p. 314). I agree with this statement, but it is a concession that 30 years of research have produced no evidence for menstrual synchrony. Indeed, A. Weller and Weller (2002) have found synchrony when it is impossible (e.g., the synchrony of 27- and 35-day rhythms). I have demonstrated the errors and biases in both their early and their later methods that allowed them to find synchrony when it does not exist. Graham’s commentary rather than undermining my original critique has highlighted how little is known about the accuracy and reliability of menstrual calendars, especially in menstrual-synchrony research. Indeed, in science, as researchers, we often take many things for granted that we should not.

References


Received February 6, 2002
Revision received February 28, 2002
Accepted March 2, 2002

---

**New Editors Appointed, 2004–2009**

The Publications and Communications Board of the American Psychological Association announces the appointment of five new editors for 6-year terms beginning in 2004.

As of January 1, 2003, manuscripts should be directed to the following individuals. Authors are strongly encouraged to submit manuscripts electronically as of January 1, 2003, through the journal’s Manuscript Submission Portal (see the Web site listed with each journal title).

- **For *Psychology and Aging* (http://www.apa.org/journals/pag.html)**, submit manuscripts to **Rose T. Zacks, PhD**, Department of Psychology, Michigan State University, East Lansing, MI 48824-1117.

- **For *Psychological Assessment* (http://www.apa.org/journals/pas.html)**, submit manuscripts to **Milton E. Strauss, PhD**, Department of Psychology, Case Western Reserve University, Cleveland, OH 44106-7123.

- **For *Journal of Family Psychology* (http://www.apa.org/journals/fam.html)**, submit manuscripts to **Anne Kazak, PhD**, ABPP, Oncology Psychosocial Research, The Children’s Hospital of Philadelphia, Room 1486 (Market Street), 34th and Civic Center Boulevard, Philadelphia, PA 19104. For overnight couriers: Room 1486, 3535 Market Street, Philadelphia, PA 19104.


- **For *Journal of Personality and Social Psychology: Personality Processes and Individual Differences* section (http://www.apa.org/journals/psp.html)**, submit manuscripts to **Charles S. Carver, PhD**, Department of Psychology, University of Miami, P.O. Box 248185, Coral Gables, FL 33124-2070.

Manuscript submission patterns make the precise date of completion of the 2003 volumes uncertain. Current editors Leah L. Light, PhD, Stephen N. Haynes, PhD, Ross D. Parke, PhD, Mark E. Bouton, PhD, and Ed Diener, PhD, respectively, will receive and consider manuscripts through December 31, 2002. Should 2003 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2004 volumes.