Menstrual Cycle Changes and Intellectual Performance

Barbara Sommer, MA

Test scores of a large number of female college students were studied in an attempt to determine whether there is an association between intellectual performance and phases of the menstrual cycle. No systematic relationships were found. However, subjects using oral contraceptives showed a higher mean performance on all tests than those not using oral contraceptives.

The effect of the menstrual cycle on the behavior of women has long been a subject of much conjecture and little systematic investigation. The fluctuation of hormones associated with the menstrual cycle has been held responsible for morbid fears and sexual fantasies (1), criminal behavior (2), suicide (3, 4), accident rate increases (5), aggravated mental illness (6) and other aberrant behaviors of women.

The precise nature of the relationship between the sex hormones (estrogen and progesterone) and behavior is not clear. The following graph (Figure 1) shows the changes in levels of circulating estrogen and progesterone through the normal menstrual cycle.*

The general conception is that the rapid drop in progesterone level and concomitant discomfort immediately preceding and during menstruation is the source of irritability, depression, anxiety and other symptoms generally referred to as "that difficult time of the month" (10, 11). Wickham (12) speaks of the "physical disturbance of menstruation . . . ." and states that "towards the end of the cycle, the destructive forces are in control and they culminate in the menstrual period."

Although considerable research has been done on mood changes (13, 14), little data is available concerning possible effects of hormonal variation on intellectual function. Affective states often have a bearing on cognitive states. Thus, if one finds systematic and predictable mood changes in women, it is not illogical to infer the existence of cognitive changes as well. This view of woman as an unfortunate victim of her cyclic hormonal variation has been proposed by Tiger (15) who suggests that a woman is at a disadvantage when required to perform an intellectual task during her "low-performance time". Tiger relies on a study done by Dalton (16) in which English schoolgirls, ages 11-17, reportedly showed a decline in performance in both the premenstrual (5 days preceding menstruation) and menstrual (5 days of menstruation) phases of the cycle. No statistical tests were made of Dalton's data and there is some question as to whether the differences reported are statistically reliable. For example, she reported that 27% of the girls showed a drop in marks from the preceding week during the premenstrual phase of their cycle. However, 17% showed an increase and the remaining 56% showed no change at all.

Wickham (12) surveyed the aptitude test

From the Department of Psychology, University of California, Davis, Calif.

Received for publication May 4, 1971; revision received Oct 29, 1971.

Address for reprint requests: Barbara Sommer, Dept. of Psychology, University of California, Davis, Calif 95616.

* Total estrogen excretion ranges from below 5 g/24 hr to about 90 g/24 hr (8, 9). Pregnanediol (a metabolite of progesterone) excretion ranges from 1 mg/24 hrs in the follicular phase (8) to 30 mg/24 hr in the luteal phase (10).

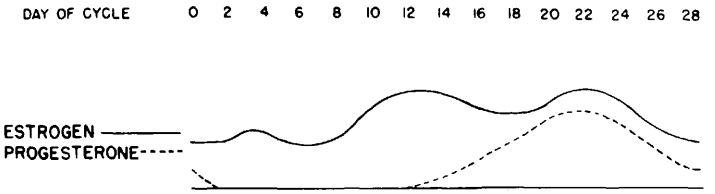

Fig 1. Relative changes in estrogen and progesterone levels throughout the menstrual cycle. Day 0 refers to onset of menstrual flow (From Kupperman [7] p 322).

performance of 1525 young women in the British service. Using two phases, Period and Non-period (Period = 3 days before and after onset of menstruation for a short cycle, 5 days for long cycles and 4 days before and after for 28-day cycle subjects), she found no significant differences on any of the individual tests. Using *progressive matrices* she retested and analyzed the results on the basis of grouping subjects by test day (day 1 to 28). Differences between observed and expected mean retest scores were irregular and "not significantly associated with a particular day or phase within the menstrual cycle." She concludes that "no advantage would be gained by taking into account the day within the menstrual cycle on which a test was given for predictive purposes." She studied other possible relations using a total of 4000 subjects and did not find any significant relationships.

In her book *The Premenstrual Syndrome,* Dalton (11) documents a sufficient number of cases to indicate that there is a genuine syndrome with often devastating effects which has at least a partial etiology in the hormonal changes associated with the menstrual cycle. It is necessary to keep in mind, however, that she is referring to a very particular group of women. The frequency of occurrence of this syndrome in the general population of women is not known. In Wickham's (12) sample of 4000 women, 24% stated they experienced some pain immediately before or during menstruation. Moos (17) administered a Menstrual Distress questionnaire to 839 women ages 20 to 30. Four hundred and twenty of them were taking oral contraceptives and 298 were not. He reports that 30% of the non-oral-contraceptive group complained of moderate, strong or acute cramps in the menstrual phase and 36% of moderate, strong or acute irritability in the premenstrual phase. Other symptoms showed lower percentages. Thus, if these samples are representative, approximately one third of the population reports unpleasant symptoms and a smaller number would be expected to suffer symptoms as debilitating as those described by Dalton.

The following studies represent successive attempts to document whether or not a decrement in intellectual performance occurs during the premenstrual and menstrual phases of the cycle. In two of these studies the Watson-Glaser Critical Thinking Appraisal (18)* was used, while other comparisons employ the students' scores on regular class examinations.

MATERIALS AND METHODS

Study 1a

Subjects. Tests were administered to *all students* in an upper division psychology class at a University of California campus. Thirty-two males and 57 females took all four tests. Their ages ranged from 18 to 22 years.

* The Watson-Glaser test was developed in 1942 and revised in 1956. It is comprised of two parallel forms containing five subtests each: a) inference, b) recognition of assumptions, c) deduction, d) interpretation, and e) evaluation of arguments.

Table 1. Mean Scores by Cycle Phase (Raw Scores) in 11 Subjects

	Sample	
	Nonpill (N = 11)	Pill (N = 11)
P	34.2	36.1
M	30.5	35.6
F	33.4	35.5
L	30.8	36.5

Table 2. Mean Scores by Cycle Phase (Raw Scores)

	Sample			
Cycle phase	No. of subjects	Nonpill	No. of subjects	Pill
P	25	32.08	17	32.12
M	40	31.78	22	33.95
F	34	31.85	15	35.47
L	36	33.61	18	33.44

Test. Two matched forms of the Watson-Glaser Critical Thinking Appraisal (18) were split-halved, resulting in four equivalent tests of high level judgmental and reasoning performance.

Procedure. The tests were given weekly in the middle of the morning class period. The subjects were told that the tests were to measure performance changes over time.

After completion of the fourth test, a questionnaire designed to obtain menstrual cycle data was given to the women; following that, a more detailed explanation of the study was given. The questionnaire included questions about cycle length and regularity and specific menstrual dates for the test period. The female subjects were separated into two groups based on whether or not oral contraceptives were used, and a note was made of those using sequential rather than combination-type pills. A subsample was selected from each of these groups comprised of women with very regular cycles and for whom each of the four tests fell into a different quarter of her cycle. These quarters were designated as premenstrual (P), menstrual (M), follicular (F) and luteal (L). The intellectual tests were not scored until after tabulation of the menstrual cycle data and the division into the comparison groups were completed.

Results. *Selected samples, pill and nonpill.* The differences in mean test performance between the various cycle phases were not significant (Table 1). Nor was there any association between cycle phase and the test session when the student did most poorly. The predicted deficit did not occur. Nor did any one phase appear particularly conducive to improved performance. In the nonpill sample an ANOVA, repeated measures, failed to show a significant test order interaction—ie the tests appeared to be sufficiently equivalent. Four girls took test A, 6 test B, 6 test C and 6 test D in their premenstrual phase. Thus, the test order is fairly random with respect to cycle phase.

Study 1b

Method. A replication of Study 1a was made the succeeding quarter with an upper-division psychology class. A total of 79 subjects met the criteria for use (regular cycles with each of the three tests taken falling clearly within a cycle quarter). The only modification was that three rather than four tests were used, as the previous class had shown signs of boredom on the fourth test. The test order was also reversed. On the menstrual questionnaire given after the administration of the tests, only six of the 106 subjects who were asked reported prior knowledge of the actual nature of the study.

Results. The differences in mean test performance among the various cycle phase groups again were not significant (Table 2), nor was there any correlation between cycle phase and poorest test session.

Study 2

Method. Rather than rely solely on performance on the Watson-Glaser test, it seemed desirable to examine the students' scores on the actual examinations in the course in which they were enrolled. Again the subjects' test scores were not examined until after they had been classified into groups according to menstrual cycle phase.

Subjects were 207 students in two lower-division psychology classes. In both classes multiple choice tests were given and were computer scored, yielding z scores which enabled cross-test comparisons. Some subjects were removed from the sample because of cycle irregularity, failure to put name on questionnaire or failure to take the regularly scheduled exams. Other subjects could not be used when the

Table 3. Obtained Mean Performance for Subjects Grouped by Menstrual Cycle Phase at the Time of the Test (z Scores)

	Tests			
	No. of subjects	1	No. of subjects	2
P.	29	-.259	30	-.152
M	35	-.517	31	-.172
F	19	-.092	23	-.171
L	23	-.155	33	-.230

examination date occurred on the borderline between two cycle phases. For subjects who did not use oral contraceptives, three different analyses of the data were presented: a) comparison of mean test performance of subjects in the premenstrual-menstrual phase (PM) and the follicular-luteal (FL) phase for the first and second examination; b) subjects' performance on the first and third examinations relative to the second examination; c) determination of the phase in which the lowest of four examination scores occurred.

Results. 1. The phase group means again fail to show an association between menstrual cycle phase and test performance (Table 3). While on Test 1 there was the expected decrement in performance particularly in the M group, on the second examination two weeks later, the L group had the lowest average. Twenty-eight of the 33 in the L group were in the M phase on the preceding test. The obtained decrement in the M phase on Test 1 appeared to result from the lower performance in general of those individuals rather than from their cycle phase.

2. The more satisfactory way of analyzing the data was to investigate the performance of the same individuals over time. Using Test 2 as a standard, a pre-test (Test 1 two weeks earlier) and a post-test (Test 3 two weeks later) performance were compared (Table 4). If a performance decrement was associated with the premenstrual and menstrual phases, the PM group should have shown improved performance on both pre-test (1) and post-test (3) in relation to Test 2, while the FL group should have shown lower scores on Test 1 and Test 3 in relation to Test 2. In only one of these instances did the results support such a prediction. The FL group did significantly better on Test 2 than on Test 1. However, they did not show the predicted decline on Test 3. Two of the mean differences were in a direction contrary to the expected one but were not significant.

3. The phase in which each subject's lowest test score occurred was tabulated. That information was available for 101 of the 144 questionnaires obtained from class A. Table 5 shows that the subjects' lowest scores did not occur significantly more often in one phase than in another.

In addition to the analyses made on the basis of cycle phase, a comparison was made between the performance of the pill (combination-type) and nonpill users. Table 6 indicates that in all three studies, the pill group showed

Table 4. Obtained Mean Performance on Tests 1, 2 and 3 for Subjects Grouped by Menstrual Cycle Phase at the Time of Test 2

Phase	No. of subjects	1	2	3	t	P
PM	70	-.0411	-.0187		.527	ns
FL	77*	-.3062	-.0778		2.145	<.02
PM	54		-.0265	-.1222	.695	ns
FL	63†		-.0481	-.0995	.458	ns

*Class A and B
†Class A only

Table 5. Number of Lowest Scores Obtained in Each Phase

Phase	No. of subjects		%
PM	56		52
FL	45	$\chi^2 = 1.2$ ns	48

a higher mean performance than the nonpill group. The mean score for the men fell between the pill and the nonpill groups. Since these are independent samples, an overall probability statement can be made [20] (Table 7).

DISCUSSION

The findings of these studies do not support the hypothesis that a decline in intellectual performance is associated with a particular phase of the menstrual cycle. However, it should be made clear that the Watson-Glaser Critical Thinking Appraisal was not designed to measure fluctuation, systematic or otherwise, over time. It was devised to assess a person's critical ability at one point in time and to detect whether or not an increase occurs after treatment, generally of an educational nature. It might also be used as a predictor for use in selection and classification (19). Its utilization in the present study was based on two points: first, it seemed to parallel some of the intellectual tasks college students are required to perform; second, it is one of the very few tests that can be given to college level students and which is amenable to multiple presentation. Essentially the same rationale was used in the choice of the Psychology 1 examinations. Large numbers of college students, possibly a majority, at one time or another in their college careers take a Psychology 1 course. Use of these test scores has the advantage of sampling actual task performance.

A more serious problem is one of the reliability or sensitivity of these tests. Most tests relating to intellectual function are highly reliable—ie, very marked changes may occur before performance scores are affected. They have generally been designed to avoid bias introduced by day-to-day fluctuation in mood, health, or alertness. As a result, such tests may not be sensitive enough to detect menstrual cycle phase effects. On the other hand, if examinations given in college classes and a test like the Watson-Glaser Critical Thinking Appraisal are believed to reflect some legitimate aspect of intellectual function and permit prediction of future performance, the absence of menstrual cycle phase fluctuation is worthy of note.

Table 6. Mean Scores for Each Group

	Group			
Study	No. of subjects	Nonpill	No. of subjects	Pill
Ia	33	32.5	24	35.0
Ib	60	31.8	27	34.1
II*	142	-.030	32	.158

*Class A only

The purpose of the present study was not to identify subtle changes which correlate with the menstrual cycle but, assuming such changes occur, to determine whether these possible changes actually have a demonstrable effect on intellectual performance on two sorts of tasks. However, the study does not rule out the possibility of associated changes which may affect other spheres of intellectual function. Such

Table 7. Combined Probabilities—Comparison Between Nonpill and Pill Groups

Experiment	t_{obs}	df	P*	
Ia	2.543	55	.018	
Ib	2.590	85	.015	$\chi^2 = 19.65$
II	1.305	172	.200	$df = 6$

*Overall significance $P = 0.01$ [20]

functions are very difficult to measure. For example, how can one reliably and repeatedly appraise "goodness of judgment" in a judge, lawyer, or administrator in such a way as to detect possible systematic fluctuation? At present, one must rely on those measures which are available—ie, standardized tests and perceptual-motor tasks, as well as self-report and observational study.

The only feasible approach is to collect samples of behavior from a sufficiently broad spectrum of situations and to determine in each case if variations in performance parallel menstrual cycle phases. This is a task beyond the powers of any single experimenter or study. Only when many studies have been made (and there is valid social justification for such a research program) can the case for cycle fluctuation affecting intellectual performance be confirmed or disproved.

The finding of the superior performance of the oral contraceptive group was surprising and consistent. It seems likely that women who are highly motivated about their college careers would show a higher performance level and would take measures to insure against unwanted pregnancy. The variability in intelligence and socioeconomic background is so small in this particular population that motivational rather than general intelligence or background factors are probably involved. On the other hand, it is possible that many of the students using oral contraceptives are married or have fairly stable relationships and thus suffer less disruptive anxiety. A third possibility, less likely than the others (at least to the writer), is that the menstrual cycle does indeed affect performance in some still unspecified way and that the pill difference is a result of the stabilizing of hormonal levels which occurs with the use of combination-type oral contraceptives.

SUMMARY

In order to ascertain whether an intellectual decrement is likely to be associated with phases of the menstrual cycle, three groups of female college students were tested. Two groups were given equivalent forms of the Watson-Glaser Critical Thinking Appraisal at weekly intervals. No association was found between test performance and the cycle quarter in which the test was taken.

Class examination scores were evaluated for the third group. Again, no association was found between menstrual cycle phase and test performance.

A comparison of average performance of students in all three groups indicated that oral contraceptive (combination-type) users performed at a significantly higher level than non-oral-contraceptive users. The reason for this difference requires further investigation. Neither age, general intelligence differences nor socioeconomic factors appear responsible for the difference.

ACKNOWLEDGMENTS

Particular appreciation is due to Dr. Karen Paige and Dr. Gary Mitchell for their assistance and encouragement, and especially to Dr. Robert Sommer.

REFERENCES

1. Benedek T: Sexual functions in women and their disturbance, American Handbook of Psychiatry. Vol 1. Edited by S Arieti. New York, Basic Books, 1959, pp 727–748
2. Dalton K: Menstruation and crime. Br Med J 2:1752–1753, 1961
3. Tonks CM, Rack PH, Rose MJ: Attempted suicide and the menstrual cycle. J Psychosom Res 11:319–323, 1968
4. Mandell A, Mandell M: Suicide and the menstrual cycle. JAMA 200:792, 1967
5. Dalton K: Menstruation and accidents. Br Med J 2:1425–1426, 1960
6. Dalton K: Menstruation and acute psychiatric illness. Br Med J 1:148–149, 1959
7. Kupperman HS: Human Endocrinology. Philadelphia, FA Davis Company, 1963
8. Loraine JA, Bell ET: Fertility and Conception in the Human Female. Baltimore, The Williams & Wilkins Company, 1968
9. Eechaute W, Demeester C: Fluorimetric de-

termination of total estrogens after gel filtration of enzymatically hydrolyzed urine. J Clin Endocrinol Metab 25:480–490, 1965
10. Hamburg DA: Effects of progesterone on behavior. Res Publ Assoc Res Nerv Ment Dis 43:251–265, 1966
11. Dalton K: The Premenstrual Syndrome. Springfield, Ill, Charles C Thomas Publisher, 1964
12. Wickham M: The effects of the menstrual cycle on test performance. Br J Psychiatry 49:34–41, 1958
13. Ivey M, Bardwick J: Patterns of affective fluctuation in the menstrual cycle. Psychosom Med 30:336–345, 1968
14. Paige K: The effects of oral contraceptives on affective fluctuations associated with the menstrual cycle. Doctoral dissertation, University of Michigan, 1969
15. Tiger L: The possible biological origins of sexual discrimination. Impact of Sc on Soc 20:29–44, 1970
16. Dalton K: Effect of menstruation on schoolgirls' weekly work. Br Med J 1:326–328, 1960
17. Moos RH: Psychological aspects of oral contraceptives. Arch Gen Psychiatry 19:87–94, 1968
18. Watson G, Glaser MG: Watson-Glaser Critical Thinking Appraisal. New York, World Book Company, 1952
19. Watson G, Glaser MG: Watson Glaser Critical Thinking Appraisal. New York, Harcourt, Brace and World, Inc, 1964
20. Winer BJ: Statistical Principles in Experimental Design. New York, McGraw-Hill Book Company, 1962

THE ANNUAL MEETING OF THE SOCIETY FOR PSYCHOPHYSIOLOGICAL RESEARCH

The next annual meeting of the Society for Psychophysiological Research will take place on November 9–12, 1972 at the Sheraton-Plaza Hotel in Boston, Mass. For information write to: B.M. SHMAVONIAN, Department of Behavioral Science, Temple University Medical School, c/o EPPI, Henry Avenue, Philadelphia, Pa 19129.