*Journal of Consulting and Clinical Psychology*
1972, Vol. 39, No. 2, 273–281

# SELF-MONITORING EFFECTS IN A PROGRAM FOR POTENTIAL HIGH SCHOOL DROPOUTS: A TIME-SERIES ANALYSIS[1]

JOHN M. GOTTMAN
*Indiana University*

RICHARD M. McFALL[2]
*University of Wisconsin*

Seventeen "potential high school dropouts" in a special-education program were instructed to self-monitor their frequency of oral class participation (or nonparticipation). The reactive and therapeutic effects of this self-monitoring were assessed on three dependent variables using a multiple-time-series design with two experimental groups, four experimental periods, and a crossover of experimental treatments. Self-monitoring significantly affected *Ss*' rates of oral class participation and to some extent their daily grades, but not their frequency of extraclassroom visits to the teachers' office. The *Ss* showed an increase in the particular response they were monitoring (i.e., talking or no talking). The use of a time-series methodology is discussed.

The present study served two purposes: Substantively, it was an extension of a previous study by McFall (1970) on the effects of self-monitoring in behavioral assessment and therapy. Methodologically, it was a demonstration of the use of time series in psychological research, as previously advocated by Gottman, McFall, and Barnett (1969).

It is becoming increasingly common for behavior therapists to engage patients as active collaborators in the data collection and management aspects of their own therapy programs (Kanfer & Phillips, 1970). This development, which has far-reaching clinical and theoretical implications, has stimulated research interest in two related problems: (*a*) the reliability, validity, and reactivity of self-monitoring as a data-collection procedure (McFall, 1970; Webb, Campbell, Schwartz, & Sechrest, 1966), and (*b*) the potential efficacy of self-monitoring as a treatment component in behavior-modification programs (Harris, 1969; Johnson & White, 1971; Kanfer, 1970; Kolb, Winter, & Berlew, 1968; Leitenberg, Agras, Thompson, & Wright, 1968; McFall, 1970; McFall & Hammen, 1971; Rutner & Bugle, 1969; Simkins, 1969; Stollak, 1967; Stuart, 1967).

These two research interests are reciprocally related. From the standpoint of data collection, it would be undesirable if self-monitoring proved to be a reactive procedure, that is, if it exerted an uncontrolled influence on the behaviors being observed. Thus, research in the area of data collection is concerned with assessing the extent to which self-monitoring is reactive, and with developing methods for eliminating or controlling such effects. On the other hand, from the therapeutic standpoint, any reactivity associated with self-monitoring could be considered a potential asset; that is, it would be desirable if such effects could be harnessed and used constructively to foster positive behavioral change. Research designed to uncover the basic parameters governing the operation of self-monitoring effects would bear on both of these problem areas.

Unfortunately, there has been little controlled research on self-monitoring. One of the few such studies was reported by McFall (1970). Following a base-line period, in which all students' smoking rates during class periods were recorded unobtrusively, some students were instructed to monitor and record

[1] This study was sponsored, in part, by the Madison Public Schools and the Instructional Research Laboratory, University of Wisconsin. The authors are indebted to William Marsh, Beth Loeb, Missy Taylor, Thulasi Bidappa, and Joyce Pajala for their help. The contributions of Robert Clasen to the implementation of the study, Donald Walsh to the data collection, and Peter Moulton to the data analysis are greatly appreciated. This study constitutes a portion of a more extensive research project submitted by John M. Gottman in partial fulfillment of the requirements for the PhD degree.

[2] Requests for reprints should be sent to Richard M. McFall, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

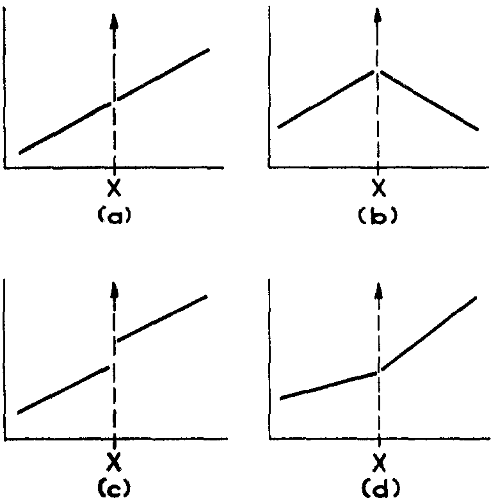


FIG. 1. Time series with (a) no experimental effect but significant difference between means, (b) an experimental effect but no significant difference between means, (c) an experimental effect with a change in level at the point of intervention but no change in slope, and (d) an experimental effect with a change in slope but no change in level at the point of intervention.

their own smoking frequency, while other students were instructed to record the number of occasions on which they considered smoking but, for whatever the reason, did not do so. Despite the fact that both groups were told that the validity of the study depended on their continuing to smoke "normally," students who monitored instances of smoking significantly increased their smoking frequency, while students who monitored instances of deciding not to smoke decreased their smoking frequency. These results were interpreted as demonstrating that self-monitoring has limitations as a data-collection procedure because of its reactivity; conversely, it is potentially useful as a behavior-modification agent for the same reason.

The present study was a logical extension of McFall's study. It involved an attempt to replicate the earlier pattern of results on the reactivity of self-monitored data collection in a different setting, while also assessing the therapeutic potential of self-monitoring procedures in a more controlled manner.

The data in this study consisted of daily observations on each of three dependent variables for each *S* over a period of nearly one semester. As such, the data represented dependent observations collected over time and provided an appropriate context within which to employ a time-series analysis, as suggested by Gottman et al. (1969).

Although time-series procedures are not yet commonly used in the behavioral sciences, they are potentially applicable in a wide variety of psychological research settings, for example, in laboratory research, in *n*-of-one research, or when monitoring progress toward objectives in an action-research setting. For example, Glass and Maguire (1968) and Campbell (1969) have illustrated the usefulness of time series for assessing the effects of social reforms. The advantages of a time-series methodology are the following: (*a*) Time-series methodology provides a continuous, descriptive record of the experimental variables over the entire course of an experiment. (*b*) Time-series data provide a source of post hoc hypotheses when coupled with a log of nonexperimental events. (*c*) Time-series and multiple-time-series designs provide varying increments of control over nonexperimental variables in settings permitting only quasi-experimental designs.

A number of statistical procedures have been suggested for analyzing time-series data. An analysis of variance, or correlated *t* test, might be used to test for shifts in means between base-line and experimental periods; however, such an analysis could be misleading and inappropriate. For example, if it were applied to a series with a constant linear trend (see Figure 1a), it might detect significant differences between base-line and experimental means, although such differences would not be due to genuine treatment effects. Conversely, if it were applied to a series where the means of two treatment periods were the same but the slopes were different (see Figure 1b), it might lead to the erroneous conclusion that there were no significant differences. To overcome some of these problems, Campbell (1969) has suggested using a regression discontinuity analysis, which tests for shifts in the slope and intercept of linear regression lines fitted separately to each experimental period. Glass and Maguire (1968) and Gottman et al. (1969) have suggested using models based on the dependency nature of the time series to test for shifts in both level and slope. These models resolve the analysis problems illustrated in Figure 1a and 1b by analyzing separately for changes in level (Figure 1c) and slope (Figure 1d).

*Statistical Analysis*

The time-series analysis in the present study was performed using a computer program developed by Glass and Maguire (1968). This program is based on a model called an "integrated moving average model with deterministic drift" previously developed by Box and Jenkins (1966, 1970). In this model the $n_1$ observations prior to intervention are described by

$$Z_1 = L + \beta_1, \quad Z_t = L + \gamma \sum_{i=1}^{t-1} \beta_i + \beta_t$$

and by

$$Z_t = L + \delta + \gamma \sum_{i=1}^{t-1} \beta_i + \beta_t$$

for the $n_1$ observations after the intervention, where

- $Z_t$ is the value of the variable at time $t$;
- L is a fixed but unknown location parameter;
- $\gamma$ is a parameter descriptive of the dependency in the time series, and takes values $0 < \gamma < 2$;
- $\delta$ is the change in level of the series caused by the intervention;
- $\beta_t$ is a random normal variable with mean $\mu$ and variance $\sigma^2$; and
- $\mu$ describes the rate of ascent or descent of the time series.

If $\beta$ is written as $\mu + \alpha$, where $\alpha$ is a normal random variable with mean 0 and variance $\sigma^2$, the equation for $Z_t$ prior to intervention becomes

$$Z_t = L + \mu\gamma(t-1) + \mu + \gamma \sum_{i=0}^{t-1} \alpha_i + \alpha_t$$

and one can see that at time $t$ the series will have drifted $\mu\gamma t$ units. In the analysis, least-squares estimates are obtained for $\delta$ and its significance is tested. When $\gamma$ is known, L, $\mu$, and $\delta$ can be estimated from the least-squares normal equations for the general linear model (see Glass & Maguire, 1968, p. 72). These least-squares estimates each have a $t$ distribution with $df = n_1 + n_2 - 3$ when divided by appropriate estimates of their standard errors. When $\gamma$ is not known, as is usually the case, Glass and Maguire say that a

> Bayesian analysis using sample information about $\gamma$ is used in making inferences about $\delta$. The posterior distribution, $h(\gamma|Z)$, of $\gamma$ given a set of $N$ observations and assuming a uniform prior distribution is known to within a constant of proportionality. The posterior distribution of $\gamma$ assuming a uniform prior (in which case the posterior distribution is equivalent to the likelihood distribution of $\gamma$) is given to within a constant of proportionality . . . [p. 73].

Practically, this means that a table is generated for various values of $\gamma$ between 0 and 2. The maximum value of the posterior distribution fixes $\gamma$ and two $t$ values, one for the change in the level of the time series and the other for the change in slope. These $t$ values test the null hypothesis that the series following the intervention can be represented by the same model as the preintervention series, with no significant shift in level or slope.

## METHOD

*Subjects*

Seventeen high school sophomores from an inner-city school (11 males, 6 females) served as *S*s. They all were participants in a larger, ongoing, special-education project designed for "potential dropouts."[3] All students involved in the larger project had been selected for inclusion as a function of having been labeled by the high school staff, at some point, as "emotionally disturbed," "culturally disadvantaged," "socially disadvantaged," "alienated," or "a potential dropout." An examination of the high school records revealed, however, that the one behavioral factor that best discriminated between project students and students in the regular academic program was their prior history of truancy. The mean rate of absenteeism among project students was approximately twice that of nonproject students (19.12 half days absent per semester as opposed to 8.51; $t = 2.77$, $p < .005$).

The 17 experimental *S*s were all members of one particular class in which a wide range of content topics, such as literature, social studies, politics, family life, drama, and English, was taught. The class met in the first period of the day, for 30 minutes, five days a week for one semester. The teacher was an attractive young female with four years of experience and a teaching certificate in special education.

[3] A more complete description of the special-education project from which *S*s were drawn is reported by Gottman (1971).

TABLE 1

SUMMARY TABLE OF TIME-SERIES ANALYSES: TALK VARIABLE

| Group | Time period | *df* | *t* for change in level | *t* for change in slope | Maximum likelihood value of $\gamma$ |
|---|---|---|---|---|---|
| 1 | B-P1 | 447 | −1.00 | 4.67* | 1.03 |
| | P1-P2 | 77 | −14.50* | −1.46 | 1.27 |
| | P2-FU | 69 | 20.20* | −2.17* | 1.35 |
| 2 | B-P1 | 447 | 20.70* | −3.89* | .85 |
| | P1-P2 | 77 | 26.00* | .07 | 1.79 |
| | P2-FU | 69 | −4.27* | .79 | 1.54 |

Note.—B = base line; P1 = Period 1; P2 = Period 2; FU = follow-ups.
* $p < .05$.

## *Dependent Variables*

The dependent variables in the present experiment were selected by means of a pilot study designed to elicit the major goals of the project, as seen by all the teachers involved. The teachers were asked to keep a daily log of classroom events they considered to be characteristic and important reflections of the project's aims. Based on these "critical incident" data (Flanagan, 1954), two *E*s generated a *Q*-sort deck (Block, 1961) of positive goal statements, along with their bipolar opposites. The teachers then sorted these goal statements on dimensions of "characteristic–uncharacteristic" and "important–unimportant." Similarly, a random sample of project students sorted the statements on these two dimensions to reflect how they thought their teachers viewed the project's goals. A principal-components analysis, using a cutoff at eigenvalues less than one, revealed that eight factors accounted for 80% of the sorting variance.

Ultimately, seven specific objectives were articulated from the factor analysis, and from these the *E*s and teachers were able to define operationally three critical dependent measures to be used for purposes of program evaluation. These, in turn, were selected as the relevant dependent variables in the present study. They were as follows:

*Talk scores.* Improved class participation emerged from the factor analysis as the primary program objective. Therefore, using a random time-sampling procedure,[4] each student's rate of oral class participation was observed for 30 minutes daily and recorded. Throughout the 30-minute observation period, every three seconds was designated as a talk unit (see Flanders, 1967). An *O* recorded a check for each *S* who spoke during each three-second interval. From these data, each *S* was given a talk score, which was equal to the total number of checks he received in each three-minute segment; thus, each *S* received 10 talk scores per day. The *O* also tape recorded all observed sessions. These recordings permitted *O* to check the accuracy of his scoring; moreover, using a table of random numbers, a simple random sample of 10 tapes was scored by a second *O* as a validity check. A satisfactory level of inter-*O* agreement was found ($r = .82$).

*Daily grade.* At the end of each class, each student was assigned a daily grade for that class by his teacher. The grades ranged from 0 to 5, and were intended to reflect task performance (0 = absent, 1 = present but no performance, 2 = minimal attempt at task performance, etc.).

*Office visits.* The teachers, who all shared one office, kept an unobtrusive record of the number of times per day that each student entered the office for any reason.

## *Procedure*

The experiment involved a multiple-time-series design, with two experimental groups, and a crossover of the experimental treatments. The experimental time line, which extended over one semester, was divided into the following periods: one week for a pre-base-line period, eight weeks for a base-line period, one week for the first self-monitoring experimental period, one week for the second self-monitoring period (where the treatments for the two groups were reversed), and approximately two weeks for a follow-up or return-to-base-line period.

The first week of the semester was devoted primarily to class organization. The pre-base-line period began in the second week of the semester, when *O* entered the class for the first time and was casually introduced by the teacher with this simple explanation: "This is Don Walsh, who is here to observe me and my teaching, and to help me be a better teacher." The *O*, who was a young adult male, sat silently in the rear of the room during this class for the remainder of the semester. He avoided interacting with the students or responding to their questions. At first his presence and reserved manner seemed to provoke some students into seeing if they could elicit a response from him; however, they seemed to habituate to his presence in about one week.

The official base-line recording of students' oral classroom participation began in the third week of the semester. In all, the base line consisted of 40 days of observation. During this period, the teacher, who still was unaware of the impending experimental manipulations, attempted in every way she could to encourage *S*s to participate more actively in class discussions, but with little success.

[4] The school day was divided into eight 30-minute time blocks, or "modules." To avoid a data overload situation and to satisfy teachers' demands that each student's performance be monitored daily, a time-sampling approach was devised by sampling from modules so that all students were observed at least once a day for one continuous module. Missing data were handled by incrementing observations.

On the forty-first day of class, the self-monitoring experiment was initiated by giving the students the following instructions:

> I would like to ask your help in an experiment to find out two things: how often you talk in the class discussion, and how often you would like to talk in the class discussion but do not. I am passing out pink and green index cards [pass out cards to predetermined individuals]. Those of you with pink cards should make a check each time you talk, whether it is a long comment or a very short one. If you say something, and someone else says something, and then you say something again, then that would call for two checks. Those of you with green cards should put down a check each time you would like to talk but, for any reason, do not. I would like you to participate just as you normally would. Don't let the fact that you are putting checks on these cards change the way you want to participate. Any questions?

Five males and three females were given pink cards; six males and three females were given green cards. At the end of the class period, the teacher collected the cards, which were 3 × 5 unlined index cards with the name of each student already typed at the top. This procedure was repeated for five class days.

The second experimental period, which covered the next five school days, was initiated by simply reversing the instructions to the two experimental groups; students who had been self-monitoring their talking now were told to monitor their nontalking, and vice versa. Otherwise, the experimental procedure was the same as in the preceding experimental period.

The return-to-base-line period was initiated simply by informing the students that the experiment had been completed, thanking them for their cooperation, and ceasing to distribute the colored cards. This period continued until the end of the semester.

Throughout the entire semester, teachers continuously kept a daily log of critical incidents, assigned daily grades to each student, and recorded the frequency of office visits.

## RESULTS

### *Talk Scores*

The effects of the self-monitoring instructions can be seen in Figure 2, which presents group means for each time period. Table 1 presents the $t$ ratios for changes in level and slope, degrees of freedom, and maximum likelihood values for $\gamma$ in the Box and Jenkins model for each pair of time periods.

Figure 2 and Table 1 clearly show the predicted crossover effect with talk-monitor instructions resulting in an increase and no-talk–monitor instructions resulting in a decrease in rate of oral class participation, as

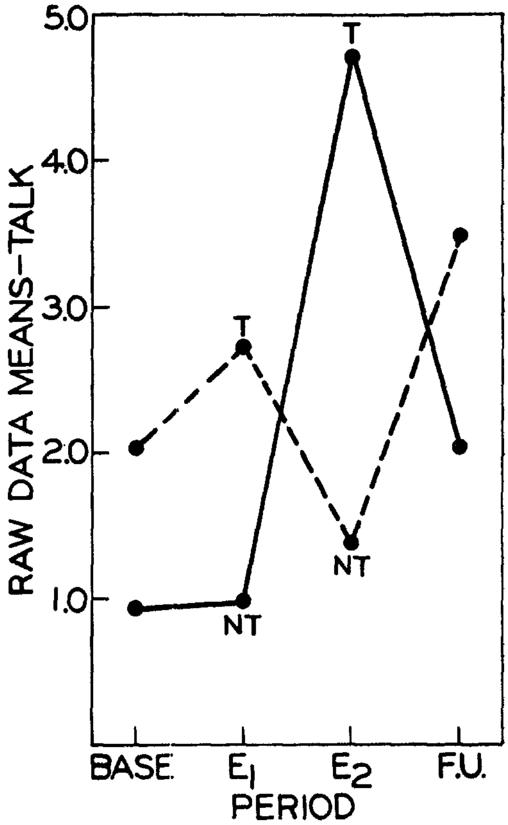


FIG. 2. Group means on talk variable.

measured by O's records. However, the nature of the effect on the within-group time series is unclear. In the first experimental period, the group receiving talk-monitor instructions showed a significant increase in the slope of their series ($t = 4.67$) but no significant shift in the level of their series following intervention ($t = -1.00$). In fact the negative $t$ value for level suggests a slight nonsignificant decline immediately following intervention. In the second experimental period, the group receiving the talk-monitor instructions showed a significant shift in level ($t = 26.00$) but not in slope ($t = .07$).

The within-group effect is similarly unclear for the no-talk–monitor instructions. For the first experimental period, the no-talk–monitor instructions resulted in a sharp immediate rise in level and a sharp decrement in slope ($t$ for level $= 20.70$; $t$ for slope $= -3.89$). In Experimental Period 2, no-talk–monitor instructions produced a significant

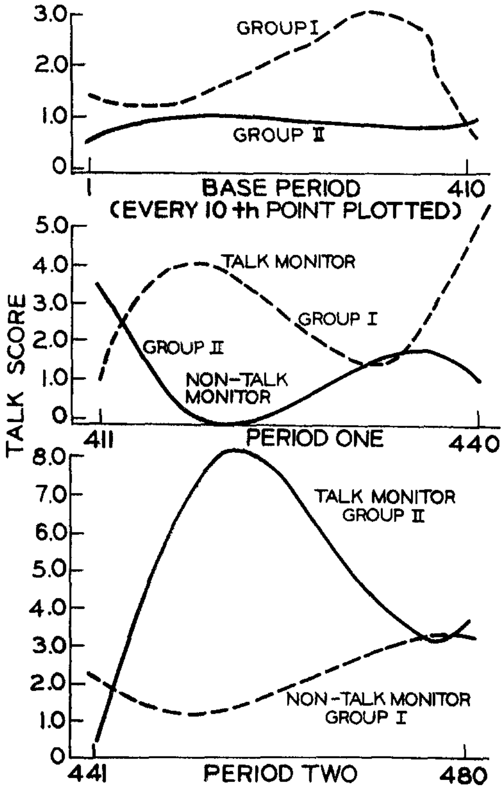


FIG. 3. Polynomial regression curves for description of within-period effects.

immediate shift downward in level ($t$ for level $= -14.50$) with a slight nonsignificant shift downward in slope ($t$ for slope $= -1.46$).

The between-group effects of self-monitoring were analyzed by using a first-order autoregressive time-series model (Watts, 1967). There were no significant differences in the base-line period ($F = .67$) or follow-up period ($F = .75$), but the groups differed significantly in the two experimental periods (for Period 1, $F = 12.00$, $df = 1/78$, $p < .05$; for Period 2, $F = 37.50$, $df = 1/78$, $p < .05$).

To obtain a clear description of within-period shifts, a third-order polynomial fit to the data was performed. The results are presented graphically in Figure 3. The talk-monitor instructions were characterized within periods by the absence of linear trend, by an immediate response component, and by a delayed decay component. The no-talk–monitor instructions were characterized within periods by an immediate decay response component and by a delayed return-to-base response component.

In summary, then, the reversal of the two self-monitoring instructions produced a defi-

TABLE 2

SUMMARY TABLE OF TIME-SERIES ANALYSES: GRADE VARIABLE

| Group | Time period | $df$ | $t$ for change in level | $t$ for change in slope | Maximum likelihood value of $\gamma$ |
|---|---|---|---|---|---|
| 1 | B-P1 | 41 | −.15 | .15 | .01 |
| | P1-P2 | 5 | .26 | −.26 | .01 |
| | P2-FU | 9 | −.25 | .24 | .01 |
| 2 | B-P1 | 41 | .11 | −.11 | .01 |
| | P1-P2 | 5 | −.03 | .03 | .01 |
| | P2-FU | 9 | −.12 | .12 | .01 |

Note.—B = base line; P1 = Period 1; P2 = Period 2; FU = follow-up.

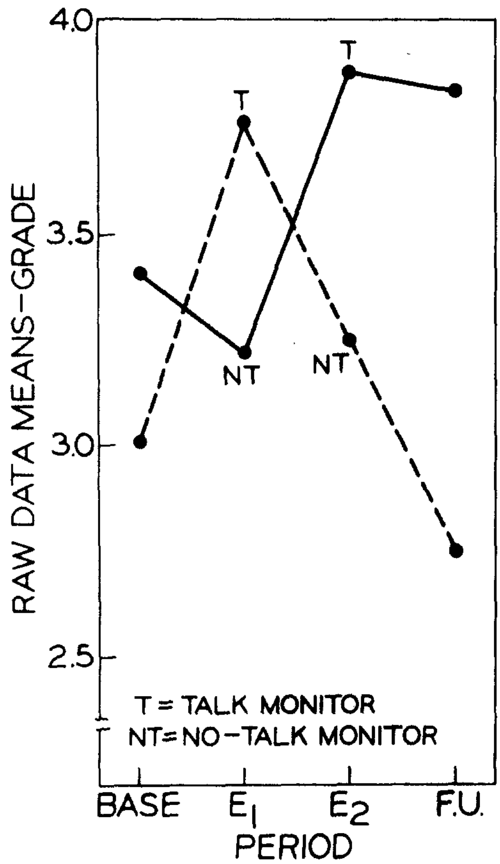


FIG. 4. Group means on grade variable.

nite crossover effect. The *S*s in both groups significantly increased their talking when instructed to monitor talking; when instructed to monitor nontalking, they tended to decrease their talking, although this effect was not as pronounced. In general, these results provide evidence for the differential reactive effects of self-monitoring instructions on oral class participation.

*Grades*

Figure 4 presents the group means for the grade variable. Using the Glass and Maguire (1968) computer program, no significant within-group differences were found (see Table 2). However, a between-groups analysis, using the autoregressive time-series model, revealed that there were no significant differences at base line ($F = 2.00$) but there were significant differences in each of the remaining periods (for Period 1, $F = 50.00$, $df = 1/6$, $p < .05$; for Period 2, $F = 150.00$, $df = 1/6$, $p < .05$; for follow-up, $F = 5.00$, $df = 1/10$, $p < .05$).

In general, the effects of self-monitoring instructions were less evident in the grade variable than they were in the talk variable. This may be due to the fact that grades were less directly related to the behavior being monitored, or it may be due simply to the fact that there were fewer observations on the grade variable (one grade per day) than on the talk variable (10 scores per day). Fewer observations would result in less power to detect any effects. The possibility that significant within-group effects would have been obtained with an increase in observations is suggested by the presence of significant between-group differences and also by the fact that there was a strong correlation between total daily talk and daily grade for the entire sample of students in the project during the base period (75% of the students' correlation coefficients exceeded .40).

*Office Visits*

Figure 5 presents group means for office visits. There were no significant within-group differences (see Table 3), but significant between-group differences were present in all periods (for the base-line period, $F = 2.30$, $df = 1/42$, $p < .05$; for Period 1, $F = 5.00$, $df = 1/6$, $p < .05$; for Period 2, $F = 14.00$, $df = 1/6$, $p < .05$; for follow-up, $F = 75.00$, $df = 1/10$, $p < .05$). In summary, the groups maintained their distance from each other

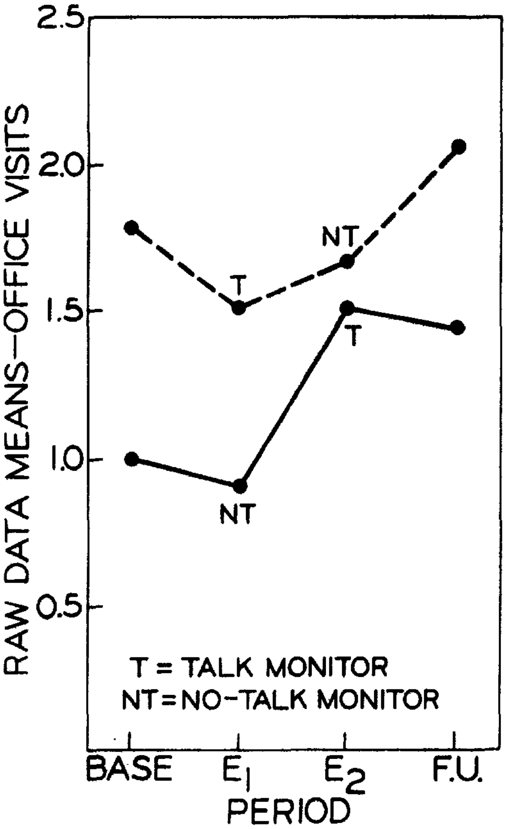


FIG. 5. Group means on office visit variable.

TABLE 3

SUMMARY TABLE OF TIME-SERIES ANALYSES: OFFICE VISITS VARIABLE

| Group | Time period | *df* | *t* for change in level | *t* for change in slope | Maximum likelihood value of $\gamma$ |
|---|---|---|---|---|---|
| 1 | B-P1 | 41 | 1.23 | −1.26 | .01 |
| | P1-P2 | 5 | −.38 | .38 | .01 |
| | P2-FU | 9 | −.44 | .42 | .01 |
| 2 | B-P1 | 41 | .82 | −.85 | .01 |
| | P1-P2 | 5 | −1.03 | 1.02 | .01 |
| | P2-FU | 9 | 1.07 | −1.10 | .01 |

Note.—B = base line; P1 = Period 1; P2 = Period 2; FU = follow-up.

throughout the experiment and failed to show the type of crossover effect which was evident in the talk variable. Office visits and talk rate were essentially uncorrelated. Only 17.7% of the students' correlation coefficients between total daily talk and office visits exceeded .40. In contrast to the grade variable, therefore, it is likely that even with more power a crossover effect would not be detected in this variable.

## Discussion

The present experiment provided evidence in support of the hypothesis that self-monitoring is a reactive data-gathering procedure. When the student-*S*s were instructed to monitor and record their frequency of oral classroom participation (or nonparticipation), significant treatment effects were obtained on the talk variable, marginal effects were found on the grade variable, and no effects were evident on the office visit variable. Thus, the only measure to show clear-cut self-monitoring effects was the one which students were monitoring.

In general, instructions to self-monitor positive instances of oral participation led to increased participation, while instructions to monitor negative instances, or nonparticipation, resulted in less participation. The *S*s showed an increase in the particular behavior they were monitoring. This directional pattern of effect is consistent with the pattern of results reported by McFall (1970). It remains to be seen, of course, whether this particular directional pattern is always obtained when *S*s self-monitor or, as seems more probable, whether the particular direction of the change is differentially affected by such variables as the response being monitored, the value *S* assigns to that response, the value *E* assigns to that response, the point in the response sequence where *S* makes his self-monitoring response, and the consequences resulting from shifts in the self-monitored behavior. Kanfer (1970) has discussed these and other variables, and has pointed to the need for more research into the specific parameters of the self-monitoring problem.

There was some indication in the results of the present experiment that the magnitude of self-monitoring effects for the talk variable was greater in the second experimental period than in the first. This observation raises the interesting possibility that there was a "rebound effect," with *S*s who initially suppressed their talking under no-talk–monitor instructions showing an increased sensitivity to the effect of talk-monitor instructions during the second experimental period. Appropriate control groups for testing this possibility were not included in the present design, but future research should explore this observation.

Orne (1970) has suggested that in order to demonstrate the therapeutic potential of self-monitoring, it would be necessary to show effects with *S*s who represented extreme cases, and to show that these effects could not easily be achieved through simple instructions alone. The present study, with its special *S* population of "negativistic" problem students and its extensive base-line efforts at achieving increased classroom performance, would seem to meet these criteria. Thus, these results provide strong evidence of the potential therapeutic utility of self-monitoring procedures.

The present study was carried out in a field setting, it dealt with a "real-life" *S* population of potential high school dropouts, its dependent variables were selected precisely because they represented the special-education goals of the teachers, and the experimental assessment procedures were built into the ongoing special-education project within which the self-monitoring experiment was performed; therefore, any improvements in classroom performance resulting from the experimental interventions represented authentic therapeutic changes within the population and setting being studied. In other words, it is difficult to question the generalizability of the self-monitoring effects because the experimental situation was essentially isomorphic with the situation to which one would wish to generalize the results. One implication of this fact is that the use of the *O* to record students' talk behavior can be considered an example of an unobtrusive measurement procedure. Because the *O* was a natural part of the environment to which one wishes to generalize, his data may be considered as archival, hence nonreactive, in much the same way that grades are considered to be archival data (see Webb et al., 1966).

Finally, the multiple-time-series methodology and analyses employed in this study made it possible to achieve a satisfactory level of quasi-experimental control in a research setting as complex and confounded as this one. Thus, the study has not only added to the general understanding of the potential reactive and therapeutic effects of self-monitoring procedures, but it has illustrated the use of specific quasi-experimental designs and data analysis techniques that should foster more innovative, controlled, and informative research on self-monitoring in the future.

## REFERENCES


BLOCK, J. *The Q-sort method in personality assessment and psychiatric research.* Springfield, Ill.: Charles C Thomas, 1961.

BOX, G. E. P., & JENKINS, G. M. *Models for prediction and control: III. Linear nonstationary models.* (Tech. Rep. No. 79) Madison: University of Wisconsin, Department of Statistics, 1966.

BOX, G. E. P., & JENKINS, G. M. *Time series analysis: Forecasting and control.* San Francisco: Holden-Day, 1970.

CAMPBELL, D. T. Reforms as experiments. *American Psychologist,* 1969, **24**, 409–429.

FLANAGAN, J. C. The critical incident technique. *Psychological Bulletin,* 1954, **51**, 327–358.

FLANDERS, N. A. Teacher influence in the classroom. In E. J. Amidon & J. B. Hough (Eds.), *Interaction analysis, theory, research, and application.* Reading, Mass.: Addison-Wesley, 1967.

GLASS, G. V., & MAGUIRE, T. O. *Analysis of time-series quasi-experiments.* (Final report, Project No. 6-8329) Boulder, Colo.: University of Colorado, Laboratory of Educational Research, 1968.

GOTTMAN, J. M. Time-series analysis in the behavioral sciences and a methodology for action research. Unpublished doctoral dissertation, University of Wisconsin, 1971.

GOTTMAN, J. M., MCFALL, R. M., & BARNETT, J. T. Design and analysis of research using time series. *Psychological Bulletin,* 1969, **72**, 299–306.

HARRIS, M. B. A self-directed program for weight control: A pilot study. *Journal of Abnormal Psychology,* 1969, **74**, 263–270.

JOHNSON, S. M., & WHITE, G. Self-observation as an agent of behavioral change. *Behavior Therapy,* 1971, **2**, 488–497.

KANFER, F. H. Self-monitoring: Methodological limitations and clinical applications. *Journal of Consulting and Clinical Psychology,* 1970, **35**, 148–152.

KANFER, F. H., & PHILLIPS, J. S. *Learning foundations of behavior therapy.* New York: Wiley, 1970.

KOLB, D. A., WINTER, S. K., & BERLEW, D. E. Self-directed change: Two studies. *Journal of Applied Behavioral Science,* 1968, **4**, 453–471.

LEITENBERG, H., AGRAS, W. S., THOMSON, L. E., & WRIGHT, D. E. Feedback in behavior modification: An experimental analysis in two phobic cases. *Journal of Applied Behavior Analysis,* 1968, **1**, 131–137.

MCFALL, R. M. Effects of self-monitoring on normal smoking behavior. *Journal of Consulting and Clinical Psychology,* 1970, **35**, 135–142.

MCFALL, R. M., & HAMMEN, C. L. Motivation, structure, and self-monitoring: Role of nonspecific factors in smoking reduction. *Journal of Consulting and Clinical Psychology,* 1971, **37**, 80–86.

ORNE, M. T. From the subject's point of view, when is behavior private and when is it public: Problems of inference. *Journal of Consulting and Clinical Psychology,* 1970, **35**, 143–147.

RUTNER, I. T., & BUGLE, C. An experimental procedure for the modification of psychotic behavior. *Journal of Consulting and Clinical Psychology,* 1969, **33**, 651–653.

SIMKINS, L. Some methodological problems in the recording of self behavior. Unpublished manuscript, University of Missouri, 1969.

STOLLAK, G. E. Weight loss obtained under different experimental procedures. *Psychotherapy: Theory, Research and Practice,* 1967, **4**, 61–64.

STUART, R. B. Behavioral control of overeating. *Behaviour Research and Therapy,* 1967, **5**, 357–365.

WATTS, D. G. *An introduction to linear modeling of stationary and nonstationary time-series.* (Tech. Rep. No. 129) Madison: University of Wisconsin, Department of Statistics, 1967.

WEBB, E. J., CAMPBELL, D. T., SCHWARTZ, R. D., & SECHREST, L. *Unobtrusive measures.* Chicago: Rand McNally, 1966.


(Received March 12, 1971)