JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1975, 8, 463-469 NUMBER 4 (WINTER 1975)

AN EVALUATION OF TIME-SAMPLE MEASURES OF BEHAVIOR[1]

J. POWELL, A. MARTINDALE, AND S. KULP

CALIFORNIA STATE COLLEGE, CALIFORNIA, PENNSYLVANIA

Continuous and time-sample measures of the in-seat behavior of a secretary were obtained. Measurement error, *i.e.,* the extent to which the sample measures deviated from the continuous measure, was a function of the frequency of the sample measurements and the criterion used to score an example of the behavior. If the behavior had to be exhibited throughout the observational interval (*whole-interval time sampling*), there was a consistent underestimate of the continuous measure. If the behavior had to be exhibited only briefly within the observational interval (*partial-interval time sampling*), there was a consistent overestimate of the continuous measure. And, if the behavior had to be exhibited at the end of the observational interval (*momentary time sampling*), overestimations and underestimations of the continuous measure occurred about equally often. As expected, the more frequently the sample measures were made the closer was the agreement between the sample and continuous measures. Two conclusions concerning measurement error in interval time sampling were made. The first was that the error will be a function of the mean time per response. The second is that this error will not be consistent across experimental conditions.

DESCRIPTORS: time sampling, whole interval, partial interval, momentary, measurement error, continuous *versus* time-sample recording

---

The observation and recording of overt behavior is a defining characteristic of applied behavior analysis. The purpose of this measurement process is to quantify behavioral phenomena.

Existing experimental literature reveals that two time-sampling procedures, interval and momentary, are commonly used to measure behavior. In interval time-sampling, a session is divided into equal periods and the observer views the behavior during these intervals. At the end of each interval, the observer records whether the behavior occurred within that interval. To be scored as an interval in which the behavior occurred, the investigator may require that the behavior be exhibited throughout the interval (Born and Davis, 1974; Peterson, Cox, and Bijou, 1971), or any portion of it (Hall, Lund, and Jackson, 1968; Mitchell and Stoffelmayr, 1973). In momentary time-sampling, the behavior is assessed at regular (Bushell, Wrobel, and Michaelis, 1968) or irregular (Kubany and Sloggett, 1973) periods of time. If the behavior is exhibited at the moment of observation, an instance of its occurrence is recorded.

Both present (Hutt and Hutt, 1970; Meyers and Grossen, 1974) and past (Goodenough, 1928; Arrington, 1937) usage has been to designate the above measurement procedures by the generic term, time sampling. However, as each of these procedures uses different rules for recording behavior, it would seem advisable to employ specific terms for them. The proposed solution is that the terms (1) *whole interval time-sampling,* (2) *partial interval time-sampling,* and (3) *momentary time-sampling* be employed when the behavior must occur through-

[1]We wish to thank Mrs. Karen Keys, the subject in this investigation, and Dr. Rob Hawkins who provided us with the term, momentary time-sampling. Reprints may be obtained from J. Powell, Department of Special Education, California State College, California, Pennsylvania 15419.

out the interval, any part of the interval, or at the end of the interval, respectively.

Because these measurement procedures use different rules for recording behavior, the question arises as to whether they yield different results. The way to answer that question would be to compare the results obtained through time sampling with the results obtained through a continuous measure of behavior. Continuous measures, such as total time on task (Surratt, Ulrich, and Hawkins, 1969) represent the population parameters from which time samples are drawn. The question is one of determining if time-sample measures of behavior, *i.e.,* sample statistics, are representative of a continuous measure of behavior, *i.e.,* a population statistic. The present study assessed the adequacy of time-sample recordings by comparing the results of time-sample measures of a secretary's in-seat behavior with the results of a continuous measure of the same behavior.

METHOD

Subject

A female secretary, employed in the office of the Dean of Education at California State College, performed normal secretarial duties, including typing, filing, answering the phone, and directing student workers. The subject consented to being observed but was unaware of the behavior being recorded, *i.e.,* her in-seat behavior, which was defined as posterior in contact with the seat of the chair.

Procedure

In-seat behavior was measured during two 20-min sessions each morning and each afternoon. Approximately 10 min intervened between each pair of sessions. In all sessions, a continuous measure of behavior was obtained using a videotape recorder. The video camera was suspended from the ceiling in a corner of the office and directed at the subject's desk. The video recorder and television monitor were located in an adjoining room and were not visible to the subject during the experimental sessions. Precisely 20 min of recording time was accomplished by having an electronic timer (Gra Lab Darkroom Timer, Model 300) start and stop the video camera. After taping the daily sessions the recordings were replayed and, using a stopwatch, the total time the subject was in her seat was summated. If this time was greater than 18 min, or less than 2 min, the session was voided and rerun at a later date. These limits were placed on the continuous measure so that the sample measures could deviate both above and below the continuous measure. Periodically, when reviewing these tapes, the number of times the subject got into her seat was also recorded.

Interval and momentary time-sample recordings of in-seat behavior were made by observing the television monitor. A cassette recorder, also started and stopped by the electronic timer, signalled when observations and records were to be made. For momentary time-sampling, the observer, when an observation was scheduled, glanced at the monitor and recorded if the subject was in or out of her seat at that moment. For interval time-sampling, the observer viewed the monitor and at the end of each interval recorded if the subject was in her seat for (1) the entire interval, (2) part of the interval or (3) none of the interval.

For momentary time-sampling, the variable manipulated was the time between observations. Six sessions each were conducted when observations occurred every 10, 20, 40, 80, 120, 240, 400, and 600 sec. Thus, the number of observations in a 20-min session ranged from a high of 120 to a low of two. For interval time-sampling, the variable manipulated was the length of the observational interval. Six sessions each were conducted when the interval length was 10, 20, 40, 80, and 120 sec. The number of observations per session, therefore, ranged from a high of 120 to a low of 10.

During one session at each of the above values, a second observer simultaneously and independently recorded in-seat behavior. This second observer also replayed the videotape for that session and measured the total time the subject was in-seat. Interobserver agreement as to the total time in-seat was calculated by dividing the lesser time observed by the greater time observed and multiplying by 100. For the 13 sessions subjected to this calculation, reliability was never less than 98%. Interobserver agreement for the two sample measures was calculated by dividing the number of agreements by agreements plus disagreements and multiplying by 100. This formula was used to compute the percentage of agreement for (1) occurrences of the behavior, (2) nonoccurrences of the behavior, and (3) occurrences plus nonoccurrences of the behavior. For interval time-sampling, these three figures were 100, 94, and 99%; for momentary time-sampling, they were 100, 100, and 100%.

RESULTS

The results of the continuous (closed circles) and momentary time-sample (open circles) measures of in-seat behavior, expressed as percentages, are shown in Figure 1. The continuous measure was derived by dividing the total time in-seat by the session length and multiplying by 100; the momentary time-sample measure was derived by dividing the number of observed occurrences of in-seat behavior by the total number of observations and multiplying by 100. Figure 1 shows that when the momentary time-sample measures were made each 10, 20, 40, 80, and 120 sec there was little difference between the time sample and continuous measures. At these five values, the largest discrepancy between the two was only 8% (Session 26). When the period between the momentary time samples was increased beyond 120 sec, larger differences were observed. At 240 sec, the absolute difference between the two measures averaged 15% with a range of 5 to 34%. At 400 sec, this mean difference was 12%, with a range of 1 to 27%; at 600 sec, it was 33%, with a range of 12 to 74%. This figure also shows that the in-seat behavior, as described by the continuous measure, usually occupied between 60 and 80% of the 20-min sessions. Finally, Figure 1 reveals that the momentary time-sample measure exceeded the continuous measure in 21 sessions and was less than the continuous measure in 27 sessions. The number of occurrences of in-seat behavior (not shown) averaged 2.6 per session with a range of one to seven occurrences.

Figure 2 shows the relative mean difference between the momentary time-sample and continuous measures for each observational value; this relative mean difference never exceeded 2% until the final observational value of 600 sec, where it was approximately −18%.

Figure 3 shows the continuous and interval time-sample measures of in-seat behavior. The closed circles represent the percentage of total time that in-seat behavior occurred. The closed triangles represent that percentage of intervals during which in-seat behavior occurred throughout the interval; open circles represent the sum of the percentage of intervals in which the behavior occurred during all or any part of an interval. When only those intervals are considered where in-seat behavior was exhibited throughout the interval, the mean difference and range between the two measures for the five values of the independent variable were: 3.6% (0.8 to 6.4), 4.5% (1.7 to 7.6), 6.0% (0.5 to 14.5), 16.8% (8.5 to 28.3), and 22.1% (9.2 to 40.8). Adding the partial intervals of in-seat behavior to the whole intervals produced the following mean differences and ranges: 1.4% (0.4 to 2.5), 2.9% (1.7 to 5.4), 3.3% (1.7 to 5.5), 8.7% (4.2 to 13.3), and 17.9% (10.9 to 24.3). The continuous measure in Figure 3 shows that in-seat behavior usually occupied between 40 and 90% of the 20-min sessions. Also, this figure shows that the whole-interval measure of in-seat behavior was less than the continuous measure in all 30 sessions, whereas in 27 of the 30 sessions, the whole-plus-partial interval measure exceeded the continuous measure. The number of occurrences of in-seat behavior (not shown) averaged 2.2 per session, with a range of one to six occurrences.

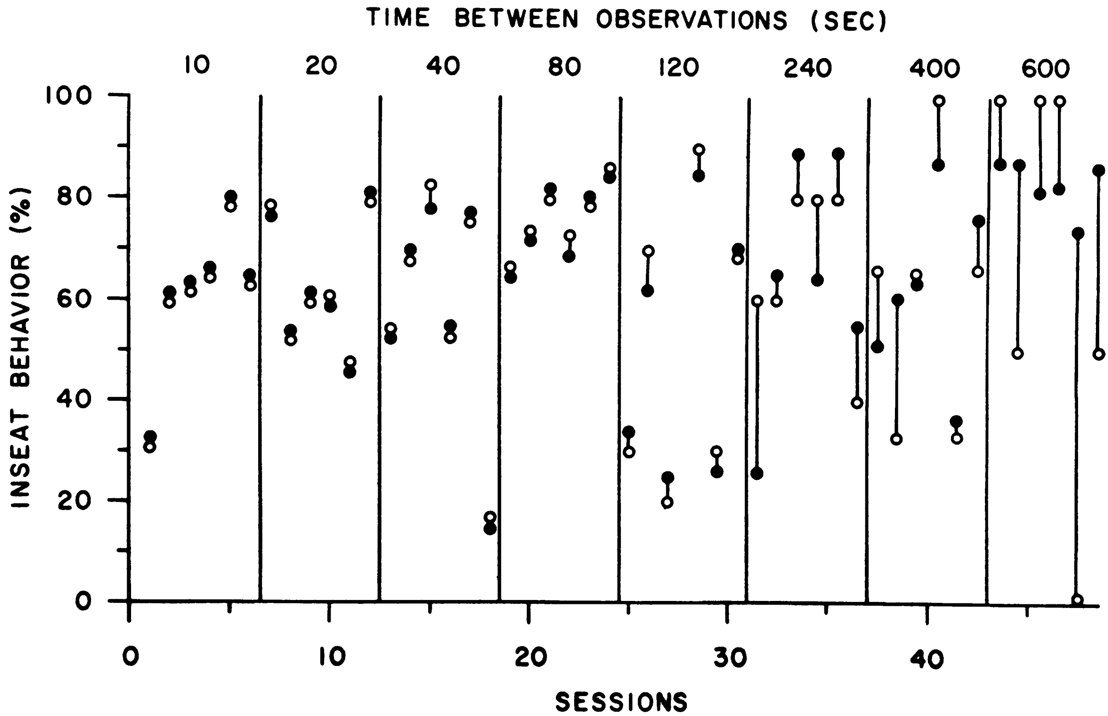

Fig. 1. Continuous and momentary time-sample measures of in-seat behavior. The closed circles represent the percentage of total time that the behavior occurred. The open circles represent the percentage of observations during which the behavior was observed. All sessions were 20 min long. The vertical lines within sessions are intended only to facilitate comparisons.

Figure 4 is a plot of the relative mean differences between the two interval measures (whole and whole-plus-partial) and the continuous measure. This figure demonstrates that both of these differences were monotonically increasing functions of the length of the observational interval, or conversely, a function of the number of observations per session; the whole-interval measure (closed triangles) was less than the continuous measure and the whole-plus partial interval measure (open circles) was greater than the continuous measure.

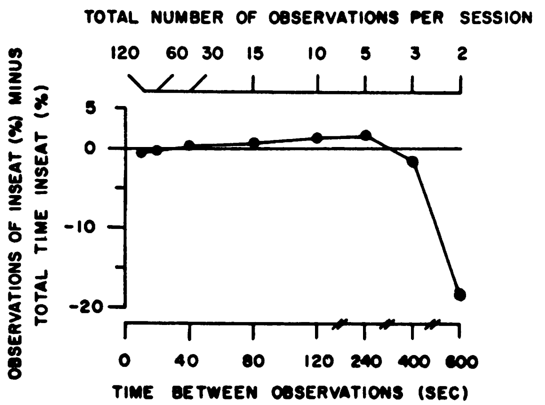

Fig. 2. The difference between the percentage of observations of in-seat behavior and the percentage of time in-seat behavior occurred. The lower abscissa shows the time between the observations and the upper abscissa the number of observations in each 20-min session. Each data point is the mean value for six sessions.

DISCUSSION

This experiment evaluated time-sample measures of behavior by assessing the degree of correspondence between the time-sample measures and a continuous measure of the behavior. A continuous measure, as it contains all examples of the behavior, represents the "true" state of nature, and to the extent that sample measures deviate from the continuous measure one has measurement error.

When momentary time-samples were conducted each 120 sec or more often, the sample and continuous measures agreed closely; discrep-

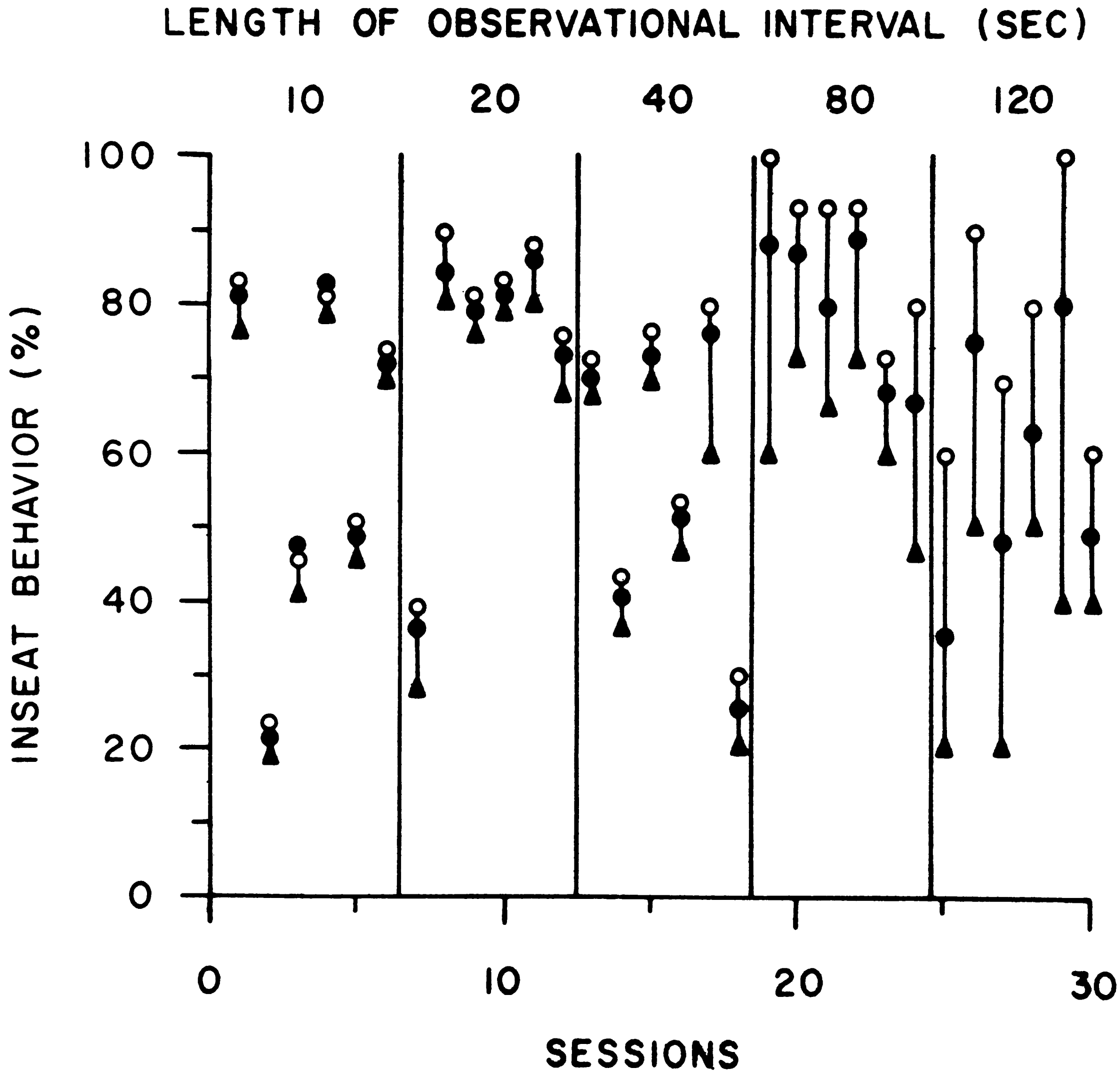

Fig. 3. Continuous and interval time-sample measures of in-seat behavior. The closed circles represent the percentage of total time that the behavior occurred. The closed triangles represent that percentage of intervals during which the behavior occurred throughout the interval. The open circles represent the sum of the percentage of intervals during which the behavior occurred throughout the interval and the percentage of intervals during which the behavior occurred any part of the interval. All sessions were 20 min long. The vertical lines within sessions are intended only to facilitate comparisons.

ancies as large as 74% were observed when the momentary time-samples occurred each 600 sec. When the time between observations is increased, the frequency of those observations is necessarily decreased and, as this study found, can result in large intersession differences between the sample and continuous measures. However, because momentary time-sampling results in both positive and negative differences, relative to a continuous measure, a summation of these differences over sessions can give a result that quite closely approaches a continuous measure (see Figure 2).

When interval time-sampling was compared with a continuous measure, the difference between the two was found to be a function of the frequency of observation. Specifically, the shorter the observational interval was, the greater the correspondence between the two measures. This is to be expected, because the more frequently

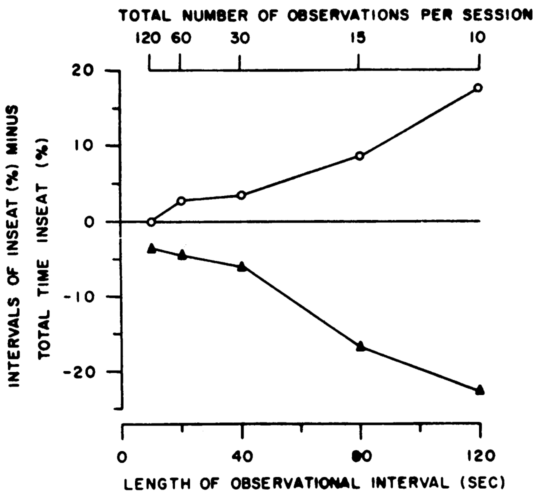

Fig. 4. The difference between the percentage of observed intervals of in-seat behavior and the percentage of time in-seat behavior occurred. The lower abscissa shows the length of the observational intervals and the upper abscissa the number of observations in each 20-min session. This difference is shown for those intervals during which the behavior occurred throughout the interval (closed triangles) and the sum of those intervals where the behavior occurred throughout the interval or any part of an interval (open circles). Each data point is the mean value for six sessions.

measures are taken, the more closely a continuous measure of the behavior is approached. Two significant findings concerning the interval time-sampling of behavior were obtained. First, if only those intervals where the behavior occurred throughout the interval are considered (whole interval time-sampling), there was always an underestimate of the continuous measure. This underestimate occurred because behavior in the partial intervals was excluded. Second, requiring only an identifiable instance of behavior within an interval to score that interval (partial interval time-sampling) resulted in an overestimate of the continuous measure (this finding assumes that the intervals identified as whole intervals of behavior in this study would also have been identified as partial intervals of behavior). This overestimate results because any instance of behavior within an interval is treated as if it occurred the entire interval.

The behavior observed in this study was a low frequency, high duration activity and no attempt was made to manipulate these response dimensions. Given the recording rules of interval time-sampling and the present results, it is possible, nevertheless, to draw several conclusions as to how changes in these two response dimensions will influence interval measures. First, as the number of responses during a given duration of behavior is increased, measurement error will increase. This is predicated on the fact that increasing the number of responses that comprise a given duration of behavior will necessarily decrease the time per response,

$$\frac{\text{total duration of responses}}{\text{number of responses}}$$

and result in, on the average: (1) fewer whole intervals of behavior and (2) less behavior in those intervals scored as partial intervals. Number one above will result in an increasing underestimate, and number two an increasing overestimate of the true state of nature. Second, it is probable that measurement error in interval time-sampling is not consistent across experimental conditions. For example, consider a baseline condition (condition A) where a continuous measure shows the behavior occurring 80% of the time, and following an experimental manipulation (condition B) the continuous measure shows the behavior occurring 20% of the time. Using partial interval time-sampling, there would be an overestimate in condition A, but this measurement error could not exceed 20%. In condition B, the overestimate would again be present, but measurement error could range up to 80%. Using whole interval time-sampling, this relationship would be reversed. That is, a large underestimation would be possible in condition A, followed by a much smaller underestimation in condition B. Indirect evidence in this study supports the above conclusion, in that the observed behavior did occur a high percentage of the time and resulted in consistently greater underestimations than overestimations (see Figure 4).

REFERENCES

Arrington, R. E. An important implication of time sampling in observational studies of behaviour. *American Journal of Sociology*, 1937, **43**, 284-295.

Born, D. and Davis, M. Amount and distribution of study in a personalized instruction course and in a lecture course. *Journal of Applied Behavior Analysis*, 1974, **7**, 365-375.

Bushell, D., Wrobel, P., and Michaelis, M. Applying group contingencies to the classroom study behavior of preschool children. *Journal of Applied Behavior Analysis*, 1968, **1**, 55-61.

Goodenough, F. Measuring behavior traits by means of repeated short samples. *Journal of Juvenile Research*, 1928, **12**, 230-235.

Hall, V., Lund, D., and Jackson, D. Effects of teacher attention on study behavior. *Journal of Applied Behavior Analysis*, 1968, **1**, 1-12.

Hutt, S. J. and Hutt, C. *Direct observation and measurement of behavior*. Springfield, Ill.: Charles C Thomas, 1970.

Kubany, E. and Sloggett, B. Coding procedures for teachers. *Journal of Applied Behavior Analysis*, 1973, **6**, 339-344.

Meyers, L. and Grossen, N. *Behavioral research: theory, procedure, and design*. San Francisco: W. H. Freeman and Company, 1974.

Mitchell, W. and Stoffelmayr, B. Application of the Premack principle to the behavioral control of extremely inactive schizophrenics. *Journal of Applied Behavior Analysis*, 1973, **6**, 419-423.

Peterson, R., Cox, M., and Bijou, S. Training children to work productively in classroom groups. *Exceptional Children*, 1971, **37**, 491-500.

Surratt, P., Ulrich, R., and Hawkins, R. An elementary student as a behavioral engineer. *Journal of Applied Behavior Analysis*, 1969, **2**, 85-92.

Received 9 October 1974.
(Final acceptance 28 April 1975.)