A Post Hoc Correction Procedure for Systematic Errors in Time-Sampling Duration Estimates

Hoi K. Suen1,2 and Donald Ary1

Accepted: November 19, 1985

Numerous previous studies have shown that partial-interval sampling in direct observation systematically overestimates duration and underestimates frequency. Whole-interval sampling systematically underestimates both duration and frequency. This paper presents a post hoc method through which the systematic errors in duration estimates in partial-interval sampling and whole-interval sampling can be minimized.

KEY WORDS: systematic error; behavioral observation; observer accuracy; time sampling; duration estimate.

INTRODUCTION

For over half a century direct observation techniques have been employed to study pathological and other behaviors and to assess the effect of treatment on behavior. Much of our present knowledge concerning human and animal behavior has resulted from such observations. Many of these, such as the classic Suomi and Harlow (1972) studies, employed time-sampling techniques and these techniques are still widely used.

In time sampling, an observation session is divided into a number of equal time intervals and observations are made interval by interval. There are three primary scoring methods in time sampling. In partial (or one-zero) sampling, an interval is scored 1 if the behavior of interest occurs any time

1Northern Illinois University, DeKalb, Illinois 60115.
2To whom correspondence should be addressed at Special Projects, Northern Illinois University, DeKalb, Illinois 60115.
within an interval and is scored 0 otherwise. In momentary (or instantaneous) sampling, an interval is scored 1 if the behavior occurs at the last instant of an interval and is scored 0 otherwise. In whole sampling, an interval is scored 1 if the behavior occurs in and lasts the entire interval and is scored 0 otherwise. The product of time sampling is a chain of 1's and 0's representing the occurrence/nonoccurrence of a behavior during the observation session. The behavior can be calibrated in terms of duration, which is the proportion of time the behavior occurs, or frequency, which is the number of times the behavior is initiated, or both.

Numerous studies (e.g., Ary, 1984; Dunbar, 1976; Green & Alverson, 1978; Legar, 1977; McDowell, 1973; Milar & Hawkins, 1976; Murphy & Goodall, 1980; Powell, Martindale, & Kulp, 1975; Powell, Martindale, Kulp, Martindale, & Bauman, 1977; Sanson-Fisher, Poole, & Dunn, 1980; Simpson & Simpson, 1977) have found that partial scores systematically overestimate duration and underestimate frequency, whole scores systematically underestimate both duration and frequency, and momentary scores produce unbiased duration estimates but do not yield useful frequency information. These outcomes have led many to recommend momentary sampling as the mode of choice. Rhine and Linville (1980), on the other hand, recommended partial-interval sampling as the mode of choice since partial-interval scores are a function of both frequency and duration. They contended that if actual frequency and actual duration are both valid measures of social behavior, then partial-interval scores, which are an excellent combination of the two, can hardly be dubbed inexact or invalid.

THE ESTIMATION OF DURATION AND FREQUENCY

Suen and Ary (1984) have shown that with momentary sampling, if the duration of the smallest bout and the smallest interresponse time (IRT; the time between two consecutive occurrences of the target behavior) is greater than the interval length, every bout will yield at least one score of 1 and every IRT will yield at least one score of 0. Therefore, the count of 0-1 sequences will give the correct frequency. Under the same condition, momentary sampling will also yield a reasonable approximation of duration, with a maximum random error of ±(frequency/number of intervals). In order for the 0-1 sequence to yield a correct frequency count with partial-interval sampling, not only must the shortest bout be longer than the interval length, but also the shortest IRT must be longer than twice the interval length. When the two conditions necessary for a correct frequency count are met, an unbiased estimate of duration can be obtained by

\[ \text{duration} = \frac{[(\text{partial sampling score}) - \text{frequency}]}{N}, \]  

(1)