Abstract

The accurate collection of unbiased behavioral data is an important component of ethnographic description and serves as a valuable method for the evaluation of hypotheses to account for cultural variation. In this paper we present two approaches for the collection of behavioral data: behavior reports and behavior observations. Behavior reports rely on subjects recalling and recording their behavior at specific time intervals during the day whereas behavior observations rely on a researcher observing and recording the behavior of a subject. Each approach has specific strengths and weakness for the generation of candid and unbiased data. We review the recent methodological literature on both approaches, note particular problems in both approaches, and contrast their respective strengths and weakness in terms of investigating ethnographic issues.

**Key Words:** behavior research, time allocation, time diary method, instantaneous sampling

Introduction

The systematic study of human behavior produces invaluable societal and cultural insights. In our field of anthropology, researchers use systematic behavior research to accomplish a number of research objectives. We study behavior to refine our
ethnographic observations (Johnson and Sackett 1998), contribute to theory development
(Gurven and Kaplan 2006), generate cross cultural comparisons (cf. Minge-Klevana
1980; Hames 1989), and engage anthropology in applied work (cf. Paolisso et al. 1989;
2002).

Despite the centrality of human behavioral research to anthropological (and other
social science) research, discussions of the strengths and limitations of different
behavioral research approaches has been intermittent and not always very specific. In the
1980s, a small corpus of reviews and assessments of anthropology and behavioral
research was available to anthropologists interested in different approaches to behavioral
research, particularly time allocation research (Gross 1984). Early articles by Johnson
(1976) and Baksh (1989, 1990) introduced the basics of instantaneous sampling (spot
checks), and Johnson and Sackett (1998) later provided an overview of direct systematic
observation methods. More recently, others have discussed observation in the context of
data recording software (Ice 2004), sample size (Bernard and Killworth 1993) and data
entry programs and personal data assistants (PDAs) (Koster 2006; Gravlee et al., 2006).
None of these works have focused comparatively, and in detail, on the strengths and
limitations of different methods for collecting information on human behavior in the
field. Moreover, over the last 10 to 15 years, there has been considerable refinement in
the specific methods used to collect behavioral data (e.g., Koster 2006) This detailed
information may be available to researchers familiar with a subset of behavior research
methods, but there has been almost no comparative discussion of the methodological pros
and cons across behavior study approaches.
In this article, we provide very specific methodological guidance for field researchers considering the collection of behavioral data. The impetus for the writing of this article is three fold: First, we recently have been teaching a course on behavior research methods. This teaching is part of the Short Courses on Research Methods in Cultural Anthropology (SCRM), a program supported by the National Science Foundation (NSF) (http://www.qualquant.net/training/scrm.htm). The course provides Ph.D. anthropologists and other social scientists with opportunities to learn about and practice implementing different types of behavior research methods, with a focus on application in naturalistic settings. Teaching this course has convinced us that there is strong interest among a diverse group of teachers, researchers and practitioners in learning the specifics of how to undertake behavioral research for a wide range of problems and situations. We have also learned that many of the specifics of different methodological approaches are not readily accessible, outside of small circles of researchers actively working with the approaches.

Second, we both have undertaken behavioral research on numerous occasions in different field and sociocultural settings. Hames has completed behaviorally oriented field research among the Ye’kwana and Yanomamô (Hames 1987, 1996) and Paolisso has completed similar work among the Yukpa of Venezuela (Paolisso and Sackett 1988), Embu of Kenya (Paolisso et al. 1989; Baksh et al. 1994) and peasant groups in Nepal (Paolisso et al. 2002) and Honduras (Paolisso et al. 1999). In all of these research undertakings, we both grappled with a host of design and implementation challenges. Finally, it has become clear to us, based on our own behavioral research, that often
solutions to one or two key methodological challenges can make the difference between successful or unsuccessful implementation of a behavioral research project.

In one article it is not possible to compare and contrast all approaches to behavioral research. Based on our teaching and field research experiences, it seems clear to us that a fundamental methodological question that researchers must first confront is whether to collect reports of behavior or observations of behavior. Since most existing methodological reviews focus on either one or the other, there is limited comparative information available that would help researchers developing behavioral research projects evaluate the strengths and limitations of these two major approaches. In this article, we compare the methodological strengths and weaknesses of collecting reports of behavior versus collecting actual observations of behavior. More specifically, we compare two very frequently-used methods to collect reports and observations: time diary research (reports) and instantaneous sampling (observation). Our goal is to provide researchers with information that will help them decide on which approach is better suited to their research needs.

In the past, a comparative review of the time diary and instantaneous sampling may not have been necessary. Time dairies were used mainly by sociologists and demographers working in literate, Western populations where respondents could be contacted through mail or telephone and could either verbally report or complete the time diary themselves. Often trained interviewers implemented the time diary studies. The result was large, representative national sample (Stinson 1999).

In contrast, instantaneous sampling was used more by anthropologists working alone in small communities in non-Western cultural settings. Here there were no
national level census data from which to develop a sampling frame, but local communities or groups that could be censused and households randomly visited. Coding was often developed inductively based on ethnographic insights (Gross 1984; Johnson and Sackett 1998).

Today, however, these historical field and implementation differences have been greatly reduced. In technologically advanced countries, researchers from different disciplines are using cell phones, pagers and handheld computers to “beep” respondents at randomly selected times to collect immediate recall data on behavior as well as respondent’s emotional and psychological states (cf. Chick 1994; Csikszentmihalyi and Larson 1987; Gravelee et al. 2006; Weisner et al. 2001). In non-Western settings, anthropologists have trained field staff to undertake instantaneous sampling, thus increasing the sample size and opening up greater flexibility in randomizing the sample (Paolisso et al. 1991, 1999, 2002). Many of the communities anthropologists use to study are now more connected to mainstream societies and have higher levels of literacy, and thus more amenable to time diary methods.

Nonetheless, most systematic behavior research in the foreseeable future will, in all likelihood, continue be either recall or direct observation, with time diary and instantaneous sampling as two very specific implementations of these approaches. To help us compare the time diary and instantaneous sampling, we begin with a description of the history, use and methodological steps involved in each approach. In this background discussion, we also include information on other recall and observation approaches, to help situate time diary and instantaneous sampling within a larger research methods context. We hope that this background information will by itself help many
future behavior researchers help decide which method best fits their research situation.

We next move beyond background description to a discussion of some key methodological similarities and differences between the two approaches. Specifically, we discuss how time diary and instantaneous sampling vary in terms of 1) the type of samples they produce and how those samples capture variability in individual behavior; 2) how each method collects information on the activities of other individuals interacting with the target individual, and 3) how each method handles the recording of the frequent situation when the target individual is engaged in multiple activities at once. We then conclude with a comparative overview of the most significant strengths and limitations of the time diary and instantaneous sampling.

**Behavior Recall and Direct Observations of Behavior**

Behavior researchers are interested in identifying the range and diversity of activities undertaken, the temporal dimensions of the activities (e.g., time allocated to each activity and when and for how long the activity occurred), and context (e.g., location of activity and if the activity was undertaken alone or in conjunctions with other activities) (Robinson and Gobey 1999). We begin with a description of origins, methodological steps, and types of results produced by time diary and instantaneous sampling. This description provides the background for the more detailed methodological comparison of time diary and instantaneous sampling presented below.

**Behavior Recall using Time Diaries**

Perhaps the most widely used and known approach for collecting systematic information on human behavior is recall. In behavior recall, individuals are asked to report their activities for a specified period of time in the recent past. Methods to collect
behavior recall data can be grouped into two large categories. The first, which we discuss only briefly, consists of open-ended questions, included for example in a survey, that ask respondents to estimate the time they spent in specific activities. Individuals might be asked ‘How much time did you spend in “x” activity, for example “working,” “sleeping,” or in “sporting events?” (Robinson and Godbey 1999). The respondent needs to remember 1) if they did engage in activity “x” during the study period and 2) if so, for how long. While easy to administer, and acceptable for qualitative descriptions of behavior, this recall method is less useful for quantitative analysis. It has been criticized because often respondents 1) do not remember all their behaviors, 2) overestimate the time spent in activities, resulting in more minutes allocated to activities than actually available in the reference period, 3) use varying understanding of activities (watching TV while cooking gets reported as TV by one respondent and cooking by another, and 4) under-report socially undesirable or unacceptable behavior (Robinson and Godbey 1999).

The second approach to collecting recalls of behavior is the use of time diaries. “The time diary is a micro-behavioral technique for collecting self reports of an individual’s behavior in an open-ended fashion on an activity-by-activity basis” (Robinson and Godbey 1999: 66). In a time diary, “a verbatim description of the day’s activities is collected along with an assignment of the approximate starting and stopping times for each activity, recorded either in free format or in fixed 5-to-10-minute intervals” (Stinson 1999:14). In time diary studies, individuals are asked to recall in order of occurrence the activities they undertook during a specified period of time in the very recent past. A previous 24 hour period is often used as the reference frame for the recall of activities.
Time diaries are an established approach within the social sciences, and one for which there has been extensive methodological discussions (cf. Robinson and Godbey 1999). In the United States, the use of time diaries can be traced to work by the United States Department of Agriculture in the 1920s and 1930s to create daily time records for homemakers (Stinson 1999). In the early 20th century, the anthropologist Audrey Richards collected time use diaries among the Bemba of Zimbabwe (then Rhodesia) (Richards 1939). There is an equally long history of time diary studies in Europe, Canada and Australia (Stinson 1999). An extensive and well-known time diary study is the “Multinational Time Budget Study, undertaken by Szalai (1972). The time diary method has been used widely in surveys of behaviors and time allocation among Americans (Robinson 1976; Robinson and Godbey 1999). An exemplary and ongoing time diary study is the U.S. Government’s Bureau of Labor Statistics American Time Use Study (ATUS) (www.bls.gov/tus/; Horrigan and Herz 2004).

Data from time diary studies have been used to study a wide range of behaviors, including trends and gender differentials in housework (Bianchi et al. 2002), parental time with children (Sayer et al. 2004; Sandberg and Hofferth 2000), and overall leisure (Godbey and Robinson 1999; Schor 1991; Jacobs and Gerson 2004). Other time diary studies have investigated trends in TV viewing, internet use, and specific types of leisure activities (Robinson and Godbey 1999), civic involvement (Putman 2000; Sayer 2001), and religious participation (Presser and Stinson 1998).

Eliciting a valid and reliable recall of activities over a specified period of time is a fundamental requirement of the time diary methods. The goal is to collect from respondents a listing of their activities, including starting and stopping times. The first
methodological step is to establish the period of time during which respondents will be asked to recall their activities. The most commonly used recall period is the previous 24 hours. For example, respondents can be asked to recall their activities from 4:00am yesterday to 4:00am today. More distant recall periods are also possible, such as last week or even month, although they are used much less often than the 24 hour recall. The general rule is that the more distant the recall period, the more general will be the recall.

Once the recall period is determined, the next methodological step is to use a chronological framework to assist respondents in remembering their activities for that period. In the USDA studies mentioned earlier, a 12 hour clock was used. Participants were instructed to draw lines on the clock diagram to mark the beginning and ending times of their activities and to describe the activity inside the intervening spaces (See Exhibit 1 in Stinson 1999).

Today, most studies do not use a clock but a chronological listing of activities. There are two formats for these listing. First, beginning at a specified time (e.g., 4:00am or “the time you woke up”), the respondent is asked to list the activities she or he engaged in for specified increments of time during the day, for example every 15 minutes. For example, the respondent is asked “what did you do from 4:00am to 4:15,” and next “what did you do from 4:15 to 4:30,” etc. Questions about behavior for 15 minute segments are asked for the entire 24 hour recall period. Alternatively, past activities can be elicited using a less structured chronological listing. Rather than prompting behavior recall for small intervals of time, respondents are asked to list the activities they undertook in the order they completed them, starting from a specified time (e.g., midnight or 4:00am). In this approach the beginning and ending times are dictated
by the respondent’s reported length of time they spent in each behavior. There is less respondent burden in this approach, although the interviewer has less control over the recall process.

The asking of behavior for specified periods of time produces fine-grained data, if interviewers and respondents can manage the cognitive burden of recalling behavior in such small segments. Piloting is critical to determine the optimal time interval, which should be the smallest possible that guides respondents through the day’s activities without creating mental fatigue and loss of recall accuracy. The target individual provides the verbatim report of his or her activity, which is then coded.

Most time diary studies use coding schemes based on the structure developed by Alexander Szalai for the Multinational Time-Use Project of the 1960s (Szalai 1972). These activity codes are typically arranged into mutually exclusive behavior groups that cover all aspects of human activity. These primary divisions of behavior generally include personal care activities, employment-related activities, education activities, domestic activities, child care activities, purchasing goods and services, voluntary work and care activities, social and community activities, recreation and leisure, and travel time. For example, the ATUS coding lexicon uses a hierarchical structure, classifying reported activities into 17 major categories, with two additional levels of detail in each category. ATUS coders assign a six digit classification code to each diary activity (rather than the three-digit code commonly used in other time-use surveys). The first two digits represent the major activity categories; the next two digits represent the second-tier level of detail; the final two digits represent the third—the most detailed level of activity (Shelley 2005). For example, the ATUS code for “making the bed” is 020101. “Making
the bed” appears in the coding application as an example under the third-tier category, interior cleaning, which is part of the second tier category, housework, which falls under the household activities major category (Shelley 2005). The website http://www.bls.gov/tus/lexicons.htm contains the ATUS codes and modifications and instructions.

The ATUS collects information on time spent in each of more than 400 detailed activities. A data extract builder (ATUS-X) has been designed to make it easy for users to create data files that contain the time use, personal characteristic and household characteristic variables they want, thus making the data more accessible to a broader audience. The output from the ATUS-X is a data file consisting of person-level records that contain the variables a user has requested and formatted direct input into SAS, Stata or SPSS statistical programs.

Behavior Observation using Instantaneous Sampling

There is a robust literature on behavior observations methods in psychology, animal behavior, and anthropology. General reviews can be found in Altman’s classic paper (1974), Martin and Bateson’s (1993) textbook, and in anthropology by Gross (1984), Johnson and Sackett (1998), Borgerhoff Mulder and Caro (1985), and Hames (1992). By direct observation we mean observations collected by a researcher in contrast to time diary or other recall studies where the subject reports or records his or her behavior. As with behavior recall, the goal of direct observation is to collect a wide variety of quantitative data on behavior that can be used to statistically test hypotheses or to more precisely describe patterns of behavior.
Any behavior observation is a combination of who and what is observed and whether the behavior is recorded continuously or instantaneously. Following Martin and Bateson (1993: 84-86) we distinguish between sampling rules (who or what is observed) and recording rules (whether behavior is recorded continuously or instantaneously). A simplified picture of how sampling and recording rules are combined is presented in Figure 1 (see Martin and Bateson 1993: 88, for a more elaborate scheme).

Figure 1

Although ethnographers use the full range of behavior observation techniques in Figure 1, we will focus only instantaneous sampling because of its dominance in anthropological research (for a more comprehensive review see Hames, in preparation).

Instantaneous sampling, as the name suggests, simply records the behavior of the individual the instant he or she is observed. Commonly called “spot checks” (after Johnson, 1976), “scan sampling” or “instantaneous scan sampling” (Borgerhoff Mulder and Caro, 1985; Hames, 1992), instantaneous sampling is by far the most commonly used method in ethnography. The procedure consists of recording a subject’s behavior the moment the subject is observed. In community-based ethnographic studies it usually consists of serially visiting households in a community or section of a community and recording the behavior of everyone present at the moment (instant) that the individual is viewed by the researcher. After the recording is done the researcher proceeds to the next house and repeats the procedure until the entire village is sampled.

Although the goal of this method is to record behavior observed by the researcher this requirement is variably met. If the person to be recorded is not present during the sampling period the solution is for the researcher to ask someone present where the
absent subject is and what he or she is doing. Consequently, the “observation” is really a report making the data point equivalent to the recall or report approach described above. When this occurs, the researcher needs to note that the observation is really a report and must, as practicable, ensure that the report is accurate (Borgerhoff and Caro 1985; Hames 1992).

An important goal of instantaneous sampling is to generate a random sample of naturally occurring behaviors. Bias in sampling can be produced in a variety of ways that the researcher may be blind to. To avoid bias researchers randomize the time of the start of their observations, where they start, and the route they take through a settlement. These choices are made in advance of the observational round and are typically generated by a table of random numbers before the observations are to be made. These directives are followed regardless of weather conditions or the likelihood of interesting events that may occur in the future. Adherence to a rigid set of protocols is necessary to avoid biasing observations toward behaviors that are easily visible or behaviors that the researcher believes are interesting, important, extraordinary, or rare. Borgerhoff Mulder and Caro (1985), Hawkes et al. (1987), and Hames (1992) describe sources of bias when observational protocols are not followed.

Instantaneous sampling is a “dimensionless” measure since it has no duration. The only statistics that can be compiled are counts of the various behaviors recorded but such counts can be legitimately transformed into real time measures. For example, if one samples behavior during waking hours, say a fourteen hour day, and one knows that 15 percent of observations were in food preparation activities, then one could reasonably
conclude that 2.1 hours per day were spent in this activity (for examples see Hames 1992; Gurven and Kaplan 2006)

Instantaneous recording has a number of advantages. First, compared to focal follows or continuous monitoring (e.g., Bock 2005) it is very economical in terms of an ethnographer’s research time. An outcome of this economy is that it permits a large number of different individuals to be sampled. In some cases, over the course of a year, ethnographers working alone have averaged more than 300 observations per person in a village of more than 100 (Hames, 1987; Flinn, 1988; Paolisso et al., 1989). Finally, it is less obtrusive to subjects such that they less likely to modify their behavior compared to the constant scrutiny of continuous observation.

Researchers using instantaneous sampling have used both functional and structural descriptors in codes (Hames, 1992; Borgerhoff Mulder and Caro, 1985). Structural descriptions of behavior describe the bodily actions, stances, orientations, etc of the observed and can be quite detailed since one may be describing a very complex pattern of behavior in a sentence-like form. Functional descriptions focus on the purpose or design of the behavior, are simple, and conform to our intuitive understanding of behavior. In a structural code one might describe a behavior as squatting on the ground while striking plants with a machete at ground level, and occasionally tossing plants aside. Such a description might require at least three variables (body position, tool use, and body movement). Or one can functionally and simply describe the behavior as weeding. However, our intuitions may be highly inaccurate in novel cultural environments. Accurate functional descriptions presuppose that the researcher has an excellent grasp of local behavioral intentions and variability. An excellent discussion of
this problem can be found in Borgerhoff Mulder and Caro (1985:327-328) and should be read by anyone planning to observe behavior. Finally, Johnson and Sackett (1998) present a set of standardized cross-cultural behavior codes employed by researchers engaged in observation research.

**Methodological Comparison: Time Diary and Instantaneous Sampling**

Methodological critiques of the time diary method (cf. Robinson and Godbey 1999; Stinson 1999; Szalai 1972) and instantaneous sampling (cf. Johnson and Sackett 1998 and Borgerhoff Mulder and Caro 1985) are available. These articles provide detailed suggestions for addressing methodological challenges for time diary and instantaneous sampling, but do not offer any comparative critique.

In this section, we compare and contrast time diary and instantaneous sampling in terms of 1) the type of samples they produce and how those samples capture variability in individual behavior; 2) how each method collects information on the activities of other individuals interacting with the target individual, and 3) how each method handles the recording of the frequent situation when the target individual is engaged in multiple activities at once.

**Sample Differences: Average Day versus Average Individual**

Both time diary and instantaneous sampling have as their goal the production of a representative sample of individual behaviors large enough for statistical analysis. Quantitative analysis depends on a reasonably large sample of individuals in different contexts and at different times of the day and who have an adequate range of demographic qualities (old, young, female, male). Although a large sample size is always desirable, the key requirement is the production of an unbiased, random sample.
Time diary and instantaneous sampling can both produce large samples of behavior. However, sample differences between the two methods result from implicit assumptions researchers have about the sources of variability in individual behavior. In time diary, there is a sampling bias toward routine, daily behaviors, while in instantaneous there is a sampling bias toward individual behaviors over a longer period of time than a 24 hour day.

The sample unit for time diary is the 24 hour day. Typically, individuals are asked to recall activities for the previous 24 hours, as described above. Individuals provide verbatim recalls of their activities, along with some time estimates. While the individuals report specific activities, these individual activities form part of a larger sequence of continuous and linked activities. To varying degrees, preceding and proceeding activities influence the likelihood of an activity occurring, the time of day when it occurs, and its duration. For example, if an individual is ill and in bed, then many of the rest of the day’s activities will be constrained by these early behaviors or individual conditions. Also of importance is the case that most time diary studies collect only one 24 recall period per individual. The resulting sample is cross-sectional in nature, comparing [n] number of individuals’ 24 hour day of activities. There is no repeat of any individual’s behavior for another time period. Finally, the number of activities recalled per individual for his or her 24 hour day depends on how active or inactive they were on the selected study day. If, for example, an individual is sick and in bed with the flu, then there will most likely be fewer activities reported with greater amounts of time spent in each activity, on average. The 24 hour period needs to be fully accounted.
Given the above sampling strategy, time diaries produce valid and reliable
descriptions of daily, repeated, and routine behaviors. The behaviors that individuals
must or often do with great regularity are reported more frequently and with greater
accuracy than rare behaviors. The result is a good assessment of the typical day of
activities for the study individual. However, time diary studies do not capture so well the
infrequent behaviors, the spikes of activities that fall outside of the normal, daily routine.
For example, parental activities with dependent children are fairly well-captured in time
diaries because parental time investment occurs on a daily basis. However, time diaries
do not capture as well infrequent or irregular life-cycle transitions, for example birth,
graduation, church confirmation, marriage, change in job and retirement.

Also, time diaries may not be able to capture important household, family or
community transfers, such as time devoted to the care of older parents. Care giving to
elders can be lumpy and crisis-driven, and thus hard to capture adequately in the 24-hour
time diary or even a two day diary format. Such research may require that parent-child
pairs are sampled and the diaries recorded when the mother is laid up.

In contrast, instantaneous sampling produces estimates not for the average day,
but for the average individual. Typically, in instantaneous sampling, the sampling period
is much longer than one or two days and often covers an entire year, to account for
seasonal variation in behavior. Instead of disaggregating a stream of continuous and
often linked behavior into individual activities (codes), instantaneous sampling involves
completing multiple snapshots of individual behaviors over the study period. Because
significant time periods can elapse between “snapshots,” often days, there is no temporal
link between each observed activity. Each observation is a behavior for the target
individual on different days or months. The resulting individual sample is longitudinal: individual behavior can be compared at different points. Aggregating the individual observations produces data on the typical individual’s behavior over time, although often the data is analyzed as cross-sectional. Finally, all individuals have the same number of observations, which is determined by how often the behavior snapshots are taken.

The sampling strategy in instantaneous sampling is useful for capturing the variability in behavior that occurs over longer periods than an average day. Still, since the observations are made during the day (see below), the behavior recorded is used to estimate an individual’s average daily behavior, as described above. This is true even though the observation is dimensionless, with no reported or measured time spent by any individual in any activity. Instantaneous sampling does produce descriptions of the average individual’s daily activity pattern, including those activities that are infrequent and rare, and out of the routine, but important. The result is a description of the relative importance of different behaviors to an individual expressed in a comparison of the amount of time they spent in each activity. It produces less of what the average day is like, but more what is the relative importance of different activities to an individual, measured by the amount of time that individual spends in those activities, on average. (Remember the base is not an average day). Compared to time diary, its strength is that it does produce descriptions of daytime behavior and time allocation in capturing the infrequent and atypical behaviors: these behaviors will show up as a small percentage of the total observations. The result is a good proxy for the average day.

A weakness of the sampling strategy of instantaneous sampling arises from one difficulty in implementing the approach. If people move around, then instantaneous
sampling may be inefficient because they are absent when the researchers visit to observe. However, as noted above, many researchers work around this problem by asking those present to report on the whereabouts and activities of the missing person, thus allowing the researcher to complete the visit/observation of all members of the group (e.g., household).

The use of reports instead of direct observation creates problems of its own. The person reporting what an absent person is doing may classify the behavior differently than the researcher and/or the report may be what the informant thought the absent individual intended to do. There are solutions to this problem such as contacting the absent person to ask what they were doing while absent. Regardless, it is important that the researcher notes whether the observation was a true observation or a report (Borgerhoff Mulder and Caro 1985).

It is also clear than no ethnographer has free and instantaneous access to all individuals in his sample. Consequently, purely random observations are difficult. Nevertheless, there are a variety of tips and procedures that one can use to ensure as close to a random sample as possible. A researcher should always strive to observe behaviors candidly, develop techniques that insure that all relevant behaviors have an equal opportunity to be observed, and that the presence of the observer and the methods used do not affect the behavior of the observed. Of course, a culture’s preferences for privacy, openness, and observability may require the researcher to modify these requirements tactically to achieve an unbiased sample of observations.

Finally, instantaneous sampling is typically made during daytime hours and sometimes is extended into early evening or morning. In many places sampling during
nighttime hours is either dangerous or unwelcome. In a pioneering piece of research, Scaglion (1986) sampled behavior during night-time hours (7:00 PM to 6:00 AM) and discovered that in 26% of observations his New Guinea subjects were awake, and in approximately 75% of these instances they were engaged in ritual activities. The importance of sampling during nighttime hours likely varies from culture to culture. In industrial settings it is likely to be required to gain an unbiased set of observations. If nighttime is important, then the researcher is advised to do some time diary for the evening period.

To summarize, with time diary research, the sample is for the typical day, 24 hour period. With instantaneous sampling, the sample is for the typical individual. The choice of which sampling frame better captures behavior variability is an ethnographic decision: how patterned is behavior into either routine days versus seasonal activities, and how important are non-routine, infrequent life events to the research. Is the day the base unit from which to compare individual behavior, or is there a need for a more open and longer time frame from which to sample?

**Context Information: Location, Behavior of Others and Multiple Activities**

Often the meaning or significance of a specific behavior changes depending on the spatial, demographic or sociocultural contexts. Both time diary and instantaneous observation methodologies include the collection of some forms of location, demographic or sociocultural context information, such as age, sex, family membership, occupation, education, etc. How time diary and instantaneous methods allow the collection of context information is an important consideration for researchers planning to use one or the other method.
Information on the location of the activity can be useful for a variety of questions. Hames (1987) used the location to determine whose garden an individual was working in order to create measures of garden labor exchange among the Ye’kwana. In another study on the Ye’kwana, the location variable was used to measure meal sharing patterns (Hames and McCabe 2007). In both cases, when someone was observed to work in a garden not his or her own or eat a meal in another household, the owner of the field or household was known, and these cases were scored as measures of labor and food exchange, respectively. Sugawara (1988) also used the location variable to describe gender differences inter-camp visiting among San foragers and Ohtsuka et al. (2004) were able to use the location variable to understand gender based differences in exposure to environmental toxins.

Location is an example of context information that “references” back to the target individual. A second key context variable is the behavior of other individuals present. Here, there is room for much more variation, and many more possible values and information. For reach recalled or observed behavior there maybe different behaviors for others who are interacting with the target individuals.

Time diary researchers recognize that individuals do not act alone much of the time. They also understand that many important behaviors are the product of joint time allocation among two or more individuals (e.g., husband and wife, parent and child, groups of siblings). Yet, the focus of the time diary approach is collection and analysis of data for one individual, for one representative 24 hour period. While time diary approaches may collect information on other individuals present, the primary reason to collect that additional information is to code accurately the behavior of the target
individual. Generally, insufficient information is collected to code and analyze the behavior of the other individual(s), at a level comparable to the target individual.

The reluctance of time diary researchers to collect information on other individuals is explained, in part, by the difficulties in data collection and analysis. For example, most interviewing approaches (e.g., phone interview) preclude the possibility of collecting time diaries from an entire household, since trying to make contact with all household members on the same day is nearly impossible (Stinson 1999). Also, some statisticians argue that collecting clusters of inter-related, individual behaviors increases survey standard errors, due to the endogenous effects of these inter-individual activities. Others researchers argue the opposite: that the social dynamism produced by the intertwining of household members’ activities demands that households be studied as a group (Stinson 1999). There is further concern that if data are collected for more than one member of a household or social group, response rates could suffer overall, due to the difficulty of scheduling and completing the time-diary interview (Schwartz et al. 2002).

Three data collection options are available to time diary researchers interested in collecting information on more than one individual. First, for each recalled activity, the interviewer can ask if anyone else was present or involved. The data recording form used by the Australian Time Use Survey has a column for recording the informant’s response to the question “Who was with you at home, or with you away from home (e.g. no-one, family, friends) (Stinson 1999)? A short, verbatim response is entered, and later coded. The information is useful predominantly for understanding the behavior of the target individual, and can be implemented in telephone interviews.
A second approach is to develop specific sub-modules of data collection, focused, for example, on specific behaviors that are defined by a strong interactive component, such as child care. In these instances, a separate interviewing guideline is developed that seeks detailed information on who was present, did what, and for how long. In the example of child care, a target individual (a mother), is asked to recall not only her child care behavior but others who provided care to her child, for a specified period of time. Although limited to a sub-domain of behavior, this approach allows the inclusion of more and diverse individuals, undertaking a wide range of behaviors. Still, the focus is on one target individual, either the mother or the child, with the mother reporting the child’s behavior and interactions.

Finally, a third approach is to use open-ended questions. In the 1992 and 1998 Canadian General Social Survey, diary information was collected from one respondent per household, using a retrospective telephone interview. The Canadian survey augments the data from the diary by also asking respondents direct, stylized questions about their own and their partners’ use of time in several unpaid activities (Winkler 2002). Respondents began a telephone interview by reporting on their activities during the previous day in a diary format. Later, they were asked stylized direct questions about their own unpaid activities and about those of their spouses; in effect, the respondents were thereby serving as proxy reporters. Notably, the survey also asked about the activities of opposite-sex cohabiting partners. For instance, in the 1992 version of the survey, respondents were asked the following questions: (1) “Last week, did you spend any time doing housework, including cooking, cleaning, grocery shopping and laundry for your household?” and, if so, “for how many hours?” (2) Last week, did you do any
unpaid work to maintain or improve your house, yard, or automobile?” and, if so, “for how many hours?” and (3) “Last week, how many hours did you spend looking after children who live in your household?” Next, respondents were asked identical questions about their partners (with “he/she” replacing “you” in the questions).

Instantaneous sampling has also been used to chart human interaction but most of the research has been limited to the study of infant-caretaker interactions (see Hewlett, 1992 and Hewlett and Lamb, 2005). The reasons for this focus are obvious. Caretaking is a common activity, it is easily observed in a variety of settings, and it is a crucial in studies of developmental psychology, socialization, and parental investment. Nevertheless, studies of interaction should be extended beyond parent-child interactions as it is perhaps the most crucial dimension of human social interaction. Despite its potential utility for elucidating patterns of social organization very few researchers use it to study social interaction behaviorally. Exceptions include Flinn’s study (1988) of conflict between fathers and daughters and their prospective suitors, under the heading of daughter guarding; Sugawara’s (1988) study of inter-camp visiting among the San; Kimura’s (1992) analysis of association patterns among the Bongando of Zaire; and Johnson and Johnson’s (1976) study of husband and wife interactions among the Machiguenga.

In instantaneous sampling there are several different ways to collect information on interaction. In studies of child care the child is the focal subject and those providing care are added to the record in terms of their name or identification number followed by the kind of care or behavior they are directing toward the child and the caregiver has his or her behavior coded as care giving following by the name of the subject receiving care
followed by the kind of care being received (Hewlett 1989). The same basic procedure may be followed for any other kind of interaction (playing, conversation, or jointly carrying a heavy object). If the number of people interacting in a group is large then data entry can become quite tedious. A solution to this problem is to classify the behavior as an interaction and the interactants will be defined having the same date, time, and location when their entries are recorded. Finally, interactions are often asymmetric (talking-listening or feeding-being fed) and codes may be modified to note such asymmetries.

**Recording Simultaneous Activities**

A third major methodological challenge for time diary and instantaneous sampling researchers is whether and how to collect information on more than one activity for the target individual. Time diary research suggests that individuals may spend three to four hours per day engaged in multiple activities (Stinson 1999). As Szalai has remarked, “… any time-budget study which does not grapple in some way with the problem of recording secondary or parallel activities is essentially unable to give a balanced account of the great variety of activities which fill up everyday life (Quoted from Stinson 1999:18).

The importance of capturing more than one activity of the target individual is particularly pronounced in the area of child care. Both time diary and instantaneous sampling methods have included information on the multiple activities of care providers. Typically, time diary studies allow respondents to report at least one “simultaneous” or “secondary” activity to their perceived main activity.
For instantaneous sampling one can easily add secondary and tertiary activities to the record. For example, in observing the Ye’kwana it was not uncommon for Hames to come upon a woman sitting on the lever of a manioc press (to express the juice from the pulp) while nursing a child and conversing with an adjacent woman. Does the researcher code for food preparation, child care, or social (talking), or for all three activities? Johnson and Sackett (1998: 327) call this coding challenge the simultaneity problem and describe the strengths and weaknesses of six possible solutions. All their solutions are reasonable but the one we favor is to preserve the richness of the observation and code the behaviors as primary, secondary, and tertiary. This creates another problem: which behavior is primary? Information on context is essential in deciding which activities are primary, secondary or tertiary. In the above example, Hames coded the primary activity as food preparation: the woman had gone to the press to express the juice. She also needed to take care of her child, who was brought along, resulting in child care being a secondary activity. By chance, another woman happened to be there, and conversation ensured, which became the tertiary activity. The choices Hames made in assigning primary, secondary and tertiary codes clearly reflect his overall research interest in work patterns; another researcher interested in gender differences in adult interaction and communication patterns might have code conversation as the primary activity, with location and others present being codes as secondary and tertiary. The above Ye’kwana example illustrates the complexity of coding simultaneous behaviors, in deciding which behavior is primary, etc. However, we believe it is worth the effort to record simultaneous activities.

Conclusions
There are two broad options available to social scientists interested in the systematic collection of information on human behavior: reports (usually recalls) and observation. Researchers are trained in disciplinary traditions that tend to use, often with good reasons but often by simple default, one approach over the other. Methodological guidance is available for both recall and observation approaches, but there has not been enough discussion of the fundamental differences in the approaches, particularly in terms of methodological strengths and weaknesses. Such a discussion would be helpful to researchers who are planning systematic behavioral research for the first time and for experienced behavior researchers in need of combined methodological approaches to enhance ongoing data collection.

In this article, we reviewed two well-known and used methods for the systematic study of human behavior: time diary and instantaneous sampling. Our goal has been to describe each approach in sufficient detail so that future behavioral researchers would be able to understand the basics of each approach, along with some of their methodological strengths and weaknesses. We also compared both approaches in terms of their sampling biases, and how each method handles the challenges of recording multiple behaviors of and context information for the individual under study. In terms of context information, we focused on how time diary and instantaneous observation capture the behaviors of other individuals as it affects the behavior of the target respondent.

In Table 1 below we summarize some of the methodological strengths and weaknesses of time diary and instantaneous observation methods. The research attributes included have largely been discussed above. We rate each attribute as good, poor, or
variable. These are rough qualitative indicators, with “good” representing a consistent strength of a method, “poor” indicating that a method does accommodate the attribute very easily, and “variable” is used for attributes where the method can accommodate depending on specific data collection circumstances.

As shown in Table 1, the strengths of instantaneous observations include accuracy and high resolution behavior descriptions. Since the observer is recording behavior the observation is more likely to be accurate and it can have as much detail as needed to meet the researcher’s study goals. The context of the behavior can be as fully characterized as desired and simultaneous behaviors can be recorded in full detail. Finally, given individuals are repeatedly sampled one can easily produce profiles of how time is allocated over a long period of time and how it varies with social characteristics.

A particular weakness of instantaneous behavior observations is the inability to create measures of duration of behaviors or behavior sequences. (These two weaknesses can be overcome by using continuous observations of behavior.) The method can be time consuming since the researcher and not the subject must record the behavior. Finally, it is intrusive and can be challenging to implement in the field, though most who have used the approach have developed successful approaches well-suited to their field research conditions. Finally, the presence of an observer has the potential to alter the behavior of the subject and subject’s absence may force the research to rely on reports instead of observations.

The time diary method has a number of methodological strengths. First, there is a well-established methodological literature available as well as on-going discussion of the
methodological issues of time diary research. With a moderate amount of effort, a novice behavior researcher can access and even participate in current methodological discussions, all of which should be of great guidance in developing and adapting time diary methods to any particular survey or ethnographic situation. There are also extensive coding schemes and databases of coded behavior available that with little or no modification can be used in new research (cf. the American Time Use Study (ATUS)).

A second strength of the time diary approach is that, if feasible, it is a very efficient method for collecting information on daily, routine behaviors, including their sequence and duration. If feasible is a critical consideration here, since study participants must be willing to work with the investigator to produce a time diary that is detailed, accurate and covers the entire specified period of study. This contrasts sharply with the feasibility requirements of instantaneous observations, where, after initial data collection conditions are satisfied, the researcher controls many of the implementation steps. However, if feasible the time diary method is a very efficient method at collecting reliable and comparable data on behaviors, including most importantly information on duration and sequence. This efficiency enables the method to produce large, representative samples.

In terms of weaknesses, the time diary method is less well suited to collect information on multiple activities and context information, particularly, in the latter case, information on the behaviors of others who are interacting with study individual. For anthropologists and other social scientists interested in group dynamics and more holistic, descriptive accounts of behavior, this limitation can be significant. Time diary researchers have developed additional data collection modules that add insights on key
areas of behavior interaction, but these are often partial and cover only a limited range of behaviors and interactions (e.g., child care). Also, as mentioned above, time diary has sampling biases that result in better information on routine, daily behaviors, and less information on the infrequent or irregular behavior. To us, this bias is a key consideration in deciding whether to use time diary versus instantaneous observation.

To conclude, ultimately the approach that one selects will depend on the questions asked and the nature of the population sampled. Behavior observations have been indispensable in traditional ethnographic settings where literacy is absent and/or subjects are unaccustomed to handling survey forms. It is also indispensable for subjects who are young and where the specifics of the behavior being studied can only be accurately identified by trained experts. Time diaries have worked best in literate populations and when the collection of a large sample of daily activities is desired. It importantly produces data on behavior sequences and duration which are difficult to obtain using direct observation. Unlike direct observations the accuracy of time diaries depends on the training, reliability, and motivation of subjects to make accurate, candid, and timely reports of their behaviors.

In the future, the use of PDA’s and other electronic recording devices hold considerable promise because of their ability to combine the strengths of observations and reports. Like behavior observations, recording of behavior is done in real time (whenever the device requests input) thus enhancing accuracy by avoiding recall error, it is relatively unobtrusive, and this protocol can collect instantaneous or continuous observations. Finally it collects data in a digital form allowing permitting rapid statistical analysis to track trends and potential errors in research protocols. For future behavior
researchers, the methodological issue will increasingly become not whether to use time
diary or instantaneous observation, for example, but how to combine the bests of both
approaches to improve validity, reliability and relevance of behavior measurements to
increasingly complex and diverse social science questions.
## Basic Observational Methods

<table>
<thead>
<tr>
<th>Recording rules</th>
<th>Sampling rules</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Instantaneous</td>
<td>Group</td>
<td>Instantaneous scan</td>
<td>Individual</td>
</tr>
<tr>
<td>Continuous</td>
<td>Continuous scan</td>
<td>Continuous Focal</td>
<td></td>
</tr>
</tbody>
</table>

_After Hames (1992) p. 211, fig. 7.3_
<table>
<thead>
<tr>
<th>Research Attribute</th>
<th>Time Diary</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrusiveness</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Variable</td>
<td>Good</td>
</tr>
<tr>
<td>Interaction</td>
<td>Variable</td>
<td>Good</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Sequence</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Duration</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Contextual information</td>
<td>Poor</td>
<td>Good</td>
</tr>
<tr>
<td>Night-time data</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Multiple activities</td>
<td>Variable</td>
<td>Good</td>
</tr>
<tr>
<td>Subject literacy</td>
<td>Required</td>
<td>Not required</td>
</tr>
<tr>
<td>Large sample size</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Seasonal variation</td>
<td>Poor</td>
<td>Good</td>
</tr>
<tr>
<td>Rare behavior</td>
<td>Poor</td>
<td>Good</td>
</tr>
</tbody>
</table>
References


1993


36


1 We would like to thank Dr. Russ Bernard for his valuable guidance and feedback in developing this course, and to the course’s students for their many insightful comments that have helped us continue to improve the course.

2 We do not discuss *ad libitum* sampling, a kind of behavior sampling, which involves the unsystematic recording of “interesting” behaviors and is only useful for initial investigation or, perhaps, rare but important events.