In: Handbook of Methods in Cultural Anthropology, H.R. Bernard, ed. 1998. Altamira Press. pp: 301-331

ALLEN JOHNSON
ROSS SACKETT

Nine

Direct Systematic Observation of Behavior

Introduction

Behavior—what people *do*—is a fundamental dimension of cultural diversity Whether we think of behavior as customs, habits, practices, lifeways, or activity patterns, an important task of all ethnographic inquiry is to illuminate the patterns of action and interaction of the people we study. The central thesis of this chapter is that direct systematic observation is our best approach to developing trustworthy accounts of people's behavior. As such, it deserves a more prominent place in the ethnographic toolkit than most anthropologists seem to appreciate.

Systematic observation is structured by explicit rules about who we observe, when and where we observe them, what we observe, and how we record our observations. These entail selection among options, each with associated tradeoffs. Here we guide the newcomer to systematic observation through the maze of choices from project conceptualization to the sampling and recording of behavior patterns.

Why Do Direct Systematic Observation?

We distinguish three broad methods of ethnographic field research: interview, participant observation, and direct systematic observation. Interview research, an eclectic category that includes a large and diverse array of specific methods, relies entirely on research subjects as sources of ethnographic knowledge. Participant observation places the ethnographer at the scene, where a combination of direct

observation and interview provide the evidence from which rich ethnographic accounts may be constructed. By direct systematic observation, we mean those ethnographic methods that—in contrast to interviews—rely primarily on the researcher's first-hand observations and that—in contrast to participant observation—are seriously attentive to problems of sampling and measurement.

Rather than study behavior for its own sake, anthropologists commonly undertake behavioral description as one component in more holistic ethnographic investigations that include people's attitudes, discourse, and organization. In these holistic efforts, interviewing and participant observation—the traditional mainstays of ethnography—are often the methods used for describing behavior as well as for describing what people think, how they are organized, and so on. We endorse multiple methods and recognize the complementarity between them. We argue here, however, in the strongest terms that interviews and participant observation are, by themselves, inadequate to the task of constructing trustworthy accounts of activity patterns. There is an irreducible need for other, more rigorous observational methods.

Interviews Versus Direct Systematic Observation

The methodological distinction between interviews and direct systematic observation is related (but not identical) to the emic/etic distinction in the sense that to obtain emic data—which requires an interpretation by a research subject—it is usually necessary to conduct some sort of interview, whereas the data of direct observation are usually classified as etic. What confounds this neat dichotomy is that interview data may also be classified as etic to the degree that our research subjects are regarded as reporters of events they have witnessed first hand (Harris 1990:53). The general distinction between interview and direct observation, therefore, is most properly that between interview methods that emphasize meaning, interpretation, and subjective experience as against research methods (including some interviews) that emphasize accurate reporting of observed scenes and activities—ideally, without interpretation. This last caveat is fundamental and raises methodological problems that are too often swept under the rug by researchers unwilling to face the painful implications.

The fact is that humans (including both trained field workers and untrained research subjects) are surprisingly incapable of accurately describing scenes they have observed with their own eyes (and ears and other senses). Abundant evidence shows that when research subjects are asked to report on their own behavior, and these reports are compared to researchers' records of the subjects' behavior based on direct observation, the research subjects' accounts of their own behavior are substantially "wrong"—that is, they show errors of from 50%–80% when compared to the observational data (Bernard et al. 1986:388; Engle and Lumpkin 1991). This means, to put it bluntly, that anthropologists who rely uncritically on their research

subjects for descriptions of behavior are more likely to be wrong than right. Certainly, this is a methodological issue worthy of the most serious attention.

Where does the problem lie? It is most serious for researchers who treat their informants' descriptions of the past as accurate descriptions of *behavior*. For example, for years the University of California asked faculty to report annually on their use of time for the previous year, the information to be used in helping set public policy on such issues as how much time professors were expected to spend in teaching and supervising students as opposed to research, conferences, and outside consulting. Routinely, professors reported their average work week to be more than 60 hours, with abundant time devoted to teaching. The question skeptics are entitled to ask is, "Are these self-reports to be believed?" Or, more generically, are our research subjects the "videotape-like creatures with near-perfect retrieval systems" that our research methods sometimes assume (D'Andrade 1974:124)?

The methodological issue is clarified by the distinction between short- and long-term memory (D'Andrade 1995:42–44). Roughly speaking, directly observed events are first stored in short-term memory, which appears to be relatively objective and accurate about recording what actually transpired, but which has a very limited storage capacity. Quickly, therefore, knowledge stored in short-term memory is transferred elsewhere in the brain—to long-term memory. How quickly this happens is debated, but it is generally thought to take a matter of seconds.

Long-term memory has the advantage over short-term memory in that its storage capacity is far greater, but it has a serious defect from the standpoint of research methods: The correlation between records made immediately on observation (depending on short-term memory) and records based on later recall (long-term memory) is extremely low—generally around $r = .25$ (Shweder 1982). What happens to information in long-term memory? The answer is not surprising to cultural anthropologists: The direct observations that were stored in short-term memory are systematically distorted to conform to cultural expectations. In particular:

> What people remember as going together are the kinds of behavior they judge to be similar. Humans show a systematic distortion in their memories. They falsely recall "what goes with what" based on "what is like what." This effect has been demonstrated across a wide range of kinds of materials, not just behavior frequencies. . . Overall, these results throw doubt on a broad class of retrospectively based research data.
>
> . . . The results [however] are much happier for the study of the organization of *culture*. . . Cognitively shared salient features are an interesting part of a society's culture. (D'Andrade 1995:84)

We can expect, then, that when anyone is asked to report on his or her own behavior in the past, or on the observed behavior of others, systematic distortions—selective remembering and forgetting—will shape the memory so strongly as to make it largely irrelevant as a description of the detailed behavior in question.

The silver lining of this cloud is that, if in fact people's cultural models of behavior more or less accurately reflect *average* behavior in their communities, then their long-term memories will not be far off as reports of what people usually do (Romney et al. 1986). What the reports of UC professors tell us, therefore, is most likely some combination of their implicit, shared views of how professors typically behave, with the addition of some more-or-less deliberate distortion based on their perceived self-interest: After all, professors know quite well that the activity questionnaires play a role in the political process of university funding. The compiled data almost certainly exaggerate the length of the faculty work week and likely also the proportions of time spent in teaching. They are, of course, entirely useless as records of such potentially interesting behaviors as socializing in the hallways or staring out the window.

Anthropologists and other behavioral scientists should be more distressed by this problem than they appear to be. Although there are certainly times during ethnographic research when long-term memory must be relied on for behavioral descriptions, field workers should refrain from doing so casually. Although direct observation is sometimes difficult or awkward, informant recall—or even the researcher's write-up of fieldnotes at day's end—is no substitute. Recall data, based on long-term memory, are about cultural pattern (the informant's or the researcher's), not about observed behavior.

Participant Observation Versus Direct Systematic Observation

The problem of erroneous descriptions arising from cultural distortion also weakens participant observation as a method of behavioral research. One problem is that anthropologists don't naturally or automatically gather *representative* data on behavior. On the contrary, like all people, they are driven by their interests and cultural models to privilege certain behaviors over others as "relevant" or "interesting." To demonstrate this, Sackett (1996) examined the spectrum of activities described in a number of anthropological accounts drawn from textbooks, ethnographies, and five years of articles from the *American Anthropologist*. He took a random sample of descriptions of behavior found in these publications, which resulted in 750 activity descriptions divided equally from these three kinds of anthropological writings. He then compared these activity descriptions with the actual frequency of those activities in people's behavior as reported in time allocation studies from around the world (Table 1).

Two features of Table 1 are especially noteworthy. First, the relative conspicuousness of activities reported is quite similar across the three modes of anthropological writing (for example, the frequency of discussion of political and religious behaviors hovers around 36%). This suggests that writers in the various

anthropological modes share a view of which activities are important and worth attention and which can safely be ignored.

TABLE 1
Comparison of Anthropological Activity Descriptions with Actual Time Allocation

	% of Behavior Descriptions				
Activity	Textbooks (n = 250)	Ethnographies (n = 250)	*Amer. Anth.* (n = 250)	Overall (n = 750)	Global Adult Time Allocation
political/religious	38.4	36.4	34.0	36.3	2.8
food production	22.0	18.4	26.4	22.3	11.0
eating, rest, sleep	14.0	18.4	12.8	15.1	62.0
commercial	12.0	9.6	10.0	10.5	8.8
family care	9.2	11.2	8.8	9.7	12.0
making/fixing things	4.4	6.0	8.0	6.1	3.4

SOURCE: Sackett 1996.

Second, there is *no consistent relationship* between the conspicuousness of a behavior in our professional discourse and the amount of time people around the world actually spend engaged in that behavior. In fact, political and religious practice—mainly, public ceremonials like rituals and rallies—commands the greatest proportion of anthropological attention (36%) yet actually occupies the least amount of people's time (<3%). Self-absorbed activities like eating, sleeping, and relaxing occupy by far the greatest proportion of people's lives (over 60% of their time), yet constitute only 15% of published descriptions.

Sackett's research also reveals a strong gender bias in activity descriptions (Table 2). Far more anthropological attention is being paid, at least in professional writings, to activities cross-culturally associated with men rather than women. Men's activities are *eight times more likely* to be described in these writings than are women's activities, although the actual frequencies with which these activities are performed by the men and women of the world are not greatly different (12.3% for predominantly women's activities, 18.6% for predominantly men's).

Tables 1 and 2 provide evidence of strong shared biases in how anthropologists describe behavior. Our point here does not concern whether or not these biases are appropriate, only that they exist and carry important implications for behavioral measurement. As long as participant observation alone stands as the main method of behavioral description, we can expect these biases to shape the way observations are recorded during fieldwork and later reported in scholarly publications. The same kinds of biases that shape long-term memory also direct the field worker's attention and colleagues' and editors' judgments regarding what is deserving of description, analysis, and publication.

The methodological solution to this problem is to take the researcher's attention off auto-pilot, so to speak, by introducing rigorous procedures for sampling and

TABLE 2
Comparison of Anthropological Attention to Women's and Men's Activities

Cross-cultural division of labor	% of Behavior Descriptions				
	Textbooks (n = 250)	Ethnographies (n = 250)	*Amer. Anth.* (n = 250)	Overall (n = 750)	Global adult time allocation
Activities performed predominantly by women	6.8	8.8	6.4	7.3	12.3
Activities showing little cross-cultural gender bias	29.6	34.4	32.8	32.3	69.1
Activities performed predominantly by men	63.9	56.8	60.8	60.4	18.6

SOURCE: Sackett 1996.

recording behavior—in effect requiring field workers to observe and report behaviors they might otherwise neglect. Such neglect is virtually inevitable unless addressed. Over time, field workers become so familiar with their subjects' behavior that they begin to stop noticing quotidian commonplaces. Field workers fall into comfortable field routines that make some scenes and locations much more likely to be observed than others; and they find it easier to remember cases that confirm their own understandings of what is going on and to forget the negative cases that defy their understandings.

Hence, there is much in anthropological method—indeed, in human nature—to make it highly improbable that routine descriptions of behavior based on participant observation research will validly describe actual behavior. For example, we can't make sense of an ethnographer's report that child care is predominantly women's work unless we are told precisely what child care is: Does it include producing the food the child eats, or having socially recognized responsibility for its well-being, or is it limited to directly handling the child? And what is the quantitative basis for the report: Is it the relative frequency of care, total time spent, the amount of physical exertion, or some combination of all of these?

The current state of anthropological fieldwork is such that ethnographers are generally inconsistent in the ways they describe activities, differing substantially in their definitions of such basic categories as labor, housework, and child care. It is now a commonplace in activity studies that how much "work" people do depends sensitively on just which activities we consider work. Even a small shift in definitions can lead to radically different conclusions (Johnson 1975). Some differences between ethnographies are inevitable given the difference in social and cultural

contexts in which research is done, but lack of explicit attention to definitions makes it impossible to distinguish real cultural diversity from differences in ethnographers' arbitrary and idiosyncratic conventions. Although the solutions to these difficulties come from scientific methodology, the broad goal, of value to scientists and humanists alike, is *verisimilitude*—a recognizable similarity to the actual lives of the people we study.

The Challenge of Behavior Measurement

A commitment to include behavioral measurement in fieldwork immediately brings us to a dilemma: Measuring the ongoing stream of behavior in its natural detail, complexity, and context is so daunting as to be a practical impossibility. Behavior —somebody doing something—refers in phenomenological terms to observable changes in location, posture, expression, and vocalization. Simply describing a subject's location requires at least six pieces of information (latitude, longitude, altitude, compass azimuth, orientation to the horizon, and time). Posture is far more complicated: Given the human body's 206 bones and their articulation at joints, we would have to recreate 218 joint angles precisely in order to reproduce a subject's posture at any given instant (Alexander 1992). Add to this a minimum of 58 facial muscle groups involved in expression, and we would have to record a total of 282 pieces of information (6 location measures, plus 218 degrees of postural freedom, plus 58 facial muscle groups) simply to describe someone's behavior at a moment in time.

And these 282 pieces of information allow us to dip only once into the stream of behavior. To animate our description with the verisimilitude of a motion picture (which itself is only a sampling of the stream of behavior), we would have to update all these variables 24 times each second. At this frame speed, we would have to make 406,080 measurements per minute of observation, amounting to over 24 million measurements to describe just one hour of spontaneous behavior. And we haven't even addressed the problem of describing a subject's speech or the setting in which the behavior is happening.

This dilemma is real for every behavioral researcher and has long attracted serious methodological discussion (for example, Chapple and Arensberg 1940; Harris 1964). A large part of the solution is beyond the scope of this chapter: In the process of research design, clarifying the purposes of the research allows the field worker to select and distill from all the possible ways of describing behavior those that most efficiently answer the key questions of the research project (Martin and Bateson 1993; Bernard 1994). Hence, flexing the shoulder, extending the elbow, pronating the hand, and flexing the fingers around a piece of fruit become "grasping an orange," and that, in turn, becomes part of an act of food procurement, food preparation, gift exchange, theft, or eating (as the case may be). Any of these

higher-order categories of activity may or may not be of interest to any given research project.

Once we have a clear sense of the subject and goals of the research, we need to address several questions that have profound implications for how we conduct our behavioral study: Is observation an appropriate technique? How does systematic observation integrate with other activities? What information do we need to collect?

Is Observation Appropriate?

Whiting and Whiting (1973) have argued that behavior observation is more expensive in field and analysis time than other ethnographic activities and should only be used when other techniques are less effective. To justify this investment, first and foremost the study must be about behavior. That is, activity descriptions must be able to help answer the key questions posed by the research proposal. When those questions ask what activities are like, who performs them, and in what contexts they occur, observation is appropriate. Questions about the consequences of activities for individual well-being, alteration of the environment, genetic fitness, and so on, can also be answered by observations carefully supported by other ethnographic evidence.

Practical concerns limit the appropriateness of behavioral observations. Will research subjects tolerate the presence of the observer? Will they act in the observer's presence as they would if the observer were not there? Are some kinds of observation less acceptable to subjects than others? Can the behaviors of interest be observed without violating subjects' sense of decency and privacy? Is the observer likely to witness illegal or stigmatizing behaviors and, if so, are the consequences of reporting or not reporting them morally and ethically acceptable?

Experienced field workers will agree, however, that these practical concerns arise in any participant observation research. Perhaps the methodological rigor of systematic observation highlights the problems, for the option of quietly putting away one's notebook, turning away, or leaving the scene is available to the participant observer more so than to the systematic observer. And, given the serious problems associated with using informant recall to answer behavioral questions, we must be prepared to make the investment in direct systematic observation where it is the only appropriate method for answering key research questions.

How Does Systematic Observation Integrate with Other Ethnographic Activities?

Participant observation, interview, and systematic observation compete for our limited field time and attention. The inflexible scheduling often required of systematic observation highlights the conflict: Most sampling strategies require us to

be present at particular times and places, when we might prefer to pursue interesting events elsewhere in the community—to interview an unexpectedly cooperative informant or to escape the community in private contemplation. Furthermore, systematic observation is sometimes a less comfortable role for the researcher, owing to its requirements of detachment and objectivity. It creates a distance, in contrast to the essentially friendly, mutually attentive, and empathetic relationship fostered by interviews and participant observation.

But it's best to recognize and embrace the synergy between systematic observation and both interviewing and participant observation. On the one hand, systematic observation keeps participant observation honest by explicitly confronting the implicit biases of long-term memory, variations in contemporary research fashions, and inconsistent definitions. On the other hand, it addresses the potential biases and ambiguities of an interview, helping identify conscious and unconscious deceptions and distortions. Reciprocally, both interview and participant observation increase our confidence in the validity of the results of systematic observation by helping us understand the consequences of behavior for subjects' well-being, motivations, and emic conceptions of activity.

What Information Do We Need to Collect?

Actors, Actions, and Settings

We find it helpful to distinguish three broad categories of variables common in activity studies: actors (including both the focal subject(s) and the social others with whom she or he interacts), actions (the behaviors we want to study, whether specific acts and activities, the content of speech, or the consequences of behavior) and settings (including the location of action, details of the physical space in which the actors act and interact, and the "props" which they manipulate and use, such as furniture, implements, foods, etc.).

For ethnographic research, the most common and most useful focus is on actors, a sample of whom are selected for observation such that their actions and the settings in which they act are the content of the observations. More rarely, it is an action (for example, political speeches) or a setting (for example, a public speaking space) that is the focus.

Level of Behavior Measurement

Different research questions require different levels of behavior measurement. Since higher levels of measurement usually cost more (in various ways), we should be clear about the minimal data quality needed to answer a question before settling on a suitable method (Martin and Bateson 1993). In general, actions can be measured as nominal variables or as quantitative variables.

1. Nominal variables. Some issues can be resolved with nominal-level data—that is, as presence or absence of particular behaviors. We might, for example, ask whether a particular behavior is in the repertoire of the population we study (for example, does anyone in the community fish, or make pottery, or observe Hindu food restrictions?). Alternatively, we may want to know whether a certain behavior occurred during a given time interval (for example, does anyone fish in winter, or make pottery on Sunday?). Such questions are relatively easy to answer and may even be reliably addressed through informant recall.

   The chief disadvantage of nominal data is that they are poor indicators of how much of a behavior people do or of a behavior's relative importance in their lives. For example, two subjects could have fished on the same day, but the actual time they spent fishing and the number of fish they caught could have differed greatly.
2. Quantitative variables. Some questions require data on the actual amount of behavior. Measuring the actual amount of behavior performed increases our capacity to evaluate the importance of the behavior. Behavior researchers commonly distinguish four quantitative measures of behavior:

   Frequency refers to the number of occurrences of a behavior during a given time interval, most conveniently expressed as a rate (instances per unit time).

   Duration is the length of time for which a single occurrence of a behavior lasts, measured in time units.

   Total duration is the total amount of time spent performing the activity during a particular interval of time, numerically equivalent to the frequency of the behavior multiplied by the average duration. It is the measure most commonly used in time allocation studies, expressed either in time units (for example, hours per day) or as a percentage of time.

   Intensity refers to the pace at which the behavior is performed. It may be measured by a local rate such as elemental acts per unit of time (ax swings per minute) or rate of production (meters of cloth woven per day), or by a more generally comparative measure like energy expenditure (Montoye et al. 1996).

Behavior Flow

In addition to measuring the presence or amount of a behavior, we might also be interested in behavior flow. The flow of behavior is the sequence of behaviors from one activity to another. We may be interested in the flow of behavior in the short term, such as the component actions in weaving a basket, or in longer term sequences like the daily round or annual cycle.

Measuring the amount and flow of activities increases the power of our data but it generally entails more fieldwork and greater risk of influencing the behavior of our subjects (reactivity). Measuring behavior also places greater constraints on the kinds of sampling and recording methods we use. Below, we examine further the tradeoffs among alternative methods.

Temporal Resolution

The focus of behavioral observation may be on short episodes (events) or long ones (states), each with different methodological implications. An example of an event is a schoolyard fight: A glare is met by a challenge, prompting a shove in response, erupting in a brief skirmish that is quickly stopped by schoolyard monitors; within seconds the event is over.

The interesting feature of behavioral events tends to be their *frequency*: How often do schoolyard fights occur and are they frequent in relation to other behaviors on the schoolyard or elsewhere? Since events last such a short time, they require a deliberate sampling strategy, such as *continuous monitoring* for periods of time.

By contrast, the daily commute to work has the qualities of a behavioral state: It is extended in duration, often rather repetitive, and possibly monotonous. The salient features of states are usually their *total duration* over some period of time and the *duration* of each occurrence.

For example, the duration of the daily commute between home and workplace, and the total time spent commuting as a part of the working day are much more revealing about how we spend our lives than the fact that we do it twice a day. Because states are of longer duration than events, they are easier to encounter during research and can be observed either by continuous monitoring, or by sampling the behavior stream at discrete moments, a strategy called *instantaneous time-sampling*.

Descriptive Resolution

We have discussed the overwhelming amount of detail possible in behavioral observation. A simple activity like grasping an orange could potentially require thousands of points of measurement. The more detailed the resolution of behavioral study, the more costly the research in time, technology, and, probably, the good will of the research subjects. Detailed description of the discrete *acts* that go into making up the generalized sequences we call *activities*, therefore, must be justified by the theoretical and practical goals of the research project.

In our experience, even those observers who collect data on specific acts tend, in the long run, to abstract and generalize these into descriptions of activities anyway. For example, the act of "walking down the trail while carrying a rifle" is detailed and specific, but the most useful form of description usually ends up being "hunting deer." Why not save a step and record the activity "hunting deer" in the first place? If we collect too much detail in the field, we waste field time and then require additional time later—probably after the field site is long behind us and far away—trying to infer meaningful activity from data sheets full of incremental acts.

Sampling

All field workers interested in fair and accurate descriptions of what people do have wished at one time or another to be a fly on the wall or an omnipresent (and invisible) observer—gifted with the oft-mentioned "god's-eye view." The only realistic option is, of course, a compromise in the form of sampling. Sampling for systematic behavioral observation requires several steps:

Step 1: Establishing the Limits of the Study

We first need to establish the *sampling universe*, the envelope of people, places and times from which we will select entities for observation.

Social Boundaries

Although it is typical in ethnographic research for many practical issues to influence the selection of a population for study—such as access to the population, the investigator's language facility, the research budget—it is still the case that the best research population is one that is most appropriate for answering the key research questions identified in the research proposal. The size of the targeted social universe may vary greatly, from a single Yanomami family (Lhermillier and Lhermillier 1983) to entire national (Niemi, Kiiski, and Liikkanan 1981) and multinational (Szalai 1972) communities.

In large-scale studies (and even some smaller ones), the temptation is great to reduce the scope of the research by selecting specific categories of individuals for focus, using criteria like age, sex, social class, or ethnicity. While attractive as a way of conserving resources, we urge researchers to resist this temptation. In our experience, it leads to no end of lamentation once the research is over, primarily because the most important "context" for individuals' activities is the activities of the others around them. Researchers who study only the women in a community regret the absence of information on men; those who study only children wish they had included adults; those who study only the Hindus of some village wish they had data on neighboring Muslims for comparison; and so on. It is far more satisfying to develop sampling methods that include the whole range of kinds of people in research populations.

Geographical Boundaries

Once a research population has been identified and a sample drawn, it often happens that individuals in that sample leave the field site and are beyond the limits to which the researcher can travel to observe them. This is true when subjects leave to work in distant cities or go on trek. Field workers must impose a distance rule

on observations, such as the *day-range* of common activities, limited by the time it takes to travel to a site and back in a single day (Carlstein 1982). A common strategy in such cases is to define a category of activity coded "Away," to indicate that subjects are temporarily resident elsewhere (see, for example, Paolisso and Sackett 1988). It is possible to add information to this code by spelling out what the purpose of the trip was stated to be, such as working for wages, shopping, or visiting relatives. But these data are not comparable to the data of direct observation and should be treated separately in later analyses.

Temporal Boundaries

The major issues of temporal sampling concern how long a period of study is required to get a good sample of ordinary behavior and how to handle the problem of nighttime activity. The convention in ethnographic research is to study a community for at least a full year, to observe the variations of the seasons and the full annual cycle: This rule applies as well to behavioral observation. Arguments for shorter time frames can be made in some cases, where seasonal variations can be demonstrated not to be very important (perhaps in some cities or with respect to key activities like infant care).

Regarding the daily "observation window," the ideal is to make observations throughout the 24-hour cycle, but such research is rare because it is intrusive and making night visits is often culturally prohibited and dangerous. Therefore, "dawn-to-dusk" studies are often a practical necessity, but they are problematic because they systematically overestimate time use in some activities and underestimate it in others (Gross 1984; Scaglion 1986). When systematic observations are limited to daytime, some other technique such as participant observation or interview should be used to establish in a general way what goes on at night.

Step 2: Choosing the Units to Observe

Having chosen a sample for observation, we must decide what to target for observation. As we have said, most behavior studies concern some combination of three variables: actors, activities, and settings. Usually, we focus on one of these and observe how the others vary in relation to it.

Targeting People

Most commonly, research focuses on individuals, who should be selected for their *representativeness*. In a small community, it may be possible to include all individuals for observation, as is often done in time allocation studies of neighborhoods or villages. In a larger community, samples of individuals may be drawn for observation. The best sample in such cases is always a random sample,

drawn using a lottery or table of random numbers. It may be tempting to build an opportunity sample in such cases, using people with whom the researcher is already familiar (and, by implication, avoiding those who are unfamiliar or with whom the researcher doesn't get along). This should be avoided because it will compromise the reliability of the sample as being representative of the population and thus threaten the study's overall validity. By the same token, all aspects of the observation routine should be randomized as far as possible. For example, if the schedule of observations is such that individual X is always observed on Tuesday, even though X was chosen at random for inclusion in the study, the observations on X will not be representative of her or his general activity patterns.

Targeting Activities or Settings

Occasionally, the research questions call for targeting certain activities. For example, Lee (1979) observed the time his !Kung (Botswana) informants took to make ostrich eggshell canteens (360 minutes), arrow poison (300 minutes), and shelters (900 minutes). By determining who made each type of artifact, its unit labor cost, its useful lifetime, and the number of such objects maintained by a family, Lee could estimate the total time men and women devoted to manufacturing and repair each day (men, 64 minutes; women, 45 minutes).

Another possibility is to target settings, like houses, meeting places, schoolyards, etc. In the 1970s, psychologists Bornstein and Bornstein (1976) examined spontaneous walking speed in 15 communities ranging in size from 365 people (Psychro, Crete) to over 2,600,000 (Brooklyn, New York). In the downtown of each community they marked a 50-foot stretch of walkway along a main street and recorded the time passers-by took to cover the distance. They found a remarkably strong correlation ($r = 0.91$) between the size of the community and walking speed: The larger the community, the faster people walked. Later research established this as a general and robust cross-cultural pattern and even related it to the incidence of heart disease (Levine 1990).

Scans or Follows?

Whatever the focus on actor, activity, or setting, researchers generally have the choice of concentrating on a single individual (*follows*) or tracking the activities of members of a group (*scans*). In *focal-individual follows*, we select one subject and monitor her or his behavior over time, to the exclusion of others, as the Bornsteins did in their research. In *group scans*, we record either sequentially or synchronously the behaviors of a number of associating individuals (Altmann 1974; Dunbar 1976). A common strategy in time allocation studies is *household scan sampling*, which targets whole households and records the activities of all the members present at the moment of contact (Johnson 1975; Baksh 1989–90). In some ethnographic situations

it may be practical to scan an entire community sequentially by walking a circuit from house to house (for example, Hames 1979; Betzig and Turke 1985; Flinn 1988).

Step 3: Scheduling Observations

A strategy for scheduling observations should be developed at the beginning of behavioral research to ensure that the study period is sampled evenly. The main issues are: when to begin each observing session, how long the session should last, and how to sample time within the session (Gross 1984; Bernard 1994).

Scheduling Observing Sessions

Behavior patterns are usually structured by the time of day, day of the week and month, weather patterns, and season. Unless we use a strategy to ensure representative sampling, these habitual activity rhythms will confound the patterns we think we see in the data. For example, in contemporary society if behavior observations were made only on weekdays, the absence of weekend behaviors would strongly distort our data as a description of the full range of behavior patterns. The best solution is to schedule observing sessions at random, since with a large enough sample the habitual activity rhythms will be fairly captured in the data, rather than distorting it. Under some circumstances (not holistic ethnographic research), it may be acceptable to limit observations to a single time or circumstance: For example, Borenstein and Borenstein (1976) limited their observations of walking speed in all cases to sunny days of moderate temperature, to eliminate weather as a confounding factor in their study.

Sampling Strategies

How long do we keep the target individual or group under observation? The answer depends on which of four main sampling strategies we choose: continuous monitoring, fixed-interval instantaneous sampling, random-interval instantaneous sampling, and one-zero sampling,

Continuous Monitoring

For the fullest account we may choose to target a single individual and follow her or him throughout the day (for example, Dufour 1984; Hill et al. 1985) or for several days in succession (Szalai 1972). Continuous monitoring is Bernard's (1994) term, synonymous with focal-animal sampling (Altmann 1974), continuous recording (Martin and Bateson 1993), systematic observation (Whiting and Whiting 1973), continuous scan/focal follow (Hames 1992, for groups and individuals), and continuous observation (Gross 1984). With its goal of producing a detailed description

of the subject's stream of behavior over the observation session, this is the most intuitive of the sampling strategies, a kind of ethnographic narrative structured by operational rules for sampling and description.

The most common method is to note the onset of each activity and the time at which it began. Whenever the subject changes activities, therefore, a new notation is made. The resulting record is a *list of behavioral events* that can be analyzed for duration and sequence. Of all the systematic observation strategies, continuous monitoring results in the most comprehensive and flexible descriptions. It provides narratives of scenes and processes that can provide illustrative anecdotes; it chronicles both brief events and protracted states; it can provide levels of measurement from nominal lists of behaviors to quantitative behavior frequencies, durations, and even intensities. One of its great attractions is maintaining a record of the sequencing of behavior ("flow"), including the sequences of verbal or nonverbal exchange between subjects that are indispensable in studies of interaction.

> Example: Panter-Brick (1989), Tamang (Nepal). In farming communities, pregnant women commonly work through much of their pregnancy and nursing periods, yet little in detail is known about how pregnancy and lactation affect women's behavior. In a remote Tamang farming community, Panter-Brick found she could not use spot checks because people dispersed too much during daily work, so she selected focal-subject follows. For a whole annual cycle (1982–83) and all daylight hours (11 in winter, 13.5 in summer) she used local assistant pairs to keep two people each day under continual direct observation, recording their activities minute-by-minute, resulting in 7,678 hours of observations on 297 woman days, with 24 women who were either pregnant or lactating, and 19 who were neither. Among her key findings: pregnant and lactating women work significantly less than others during slack seasons (1.5 to 2.5 hours less per day), but work nearly as much as others during heavy work seasons (under 0.5 hours less per day), when their labor input is decisive for their household food supply.

As with all systematic observation methods, there are tradeoffs to be considered with continuous monitoring. If its chief advantages are its detail, faithful rendering of sequences, and an intuitive storylike record, its costs are nonetheless significant. For one thing, it imposes heavy demands on observers, making errors and reduced levels of detail more likely the longer the session goes on. Whiting and Whiting (1975) found in their comparative study of child socialization that after only a few minutes even trained observers became increasingly telegraphic, selective, and unreliable in their descriptions.

For another, it raises the likelihood of subject reactivity, since it is difficult to act normally in the presence of a foreigner (or even a fellow community member working as a field assistant) following closely behind making continuous notations on a clipboard. Continuous monitoring is subject to statistical problems that result from the lack of independence of observations. It is subject to difficulties in determining exactly where activities begin and end (Gross 1984:539), and to

problems associated with diverting too much field time away from other important ethnographic activities.

For these reasons, continuous monitoring is most appropriate when behavior measurement is the primary goal of the research and the powerful data that the method provides are essential to answering key research questions. Otherwise, it is a useful adjunct to other ethnographic methods as a way to document in great detail specific cultural practices. Other strategies are available that produce the same kinds of data at lower cost and often at greater reliability.

Fixed-Interval Instantaneous Sampling

This is also known as time-point sampling (Martin and Bateson 1993), instantaneous and scan sampling (Altmann 1974), and fixed-interval sampling (Gross 1984), and is but one level of abstraction from continuous monitoring. Instead of recording the continuous flow of behavior, fixed-interval sampling divides the observing session into discrete equal-length time intervals and makes observations on the instant marking the transition from one interval to the next. Typically, observations are made once a minute on the minute, but intervals from once every 15 seconds to every 5 minutes are common. Less demanding on observers, and less precise, it results in a *list of intervals*, each characterized by a particular behavior. The resulting record is much like a motion picture, with each observation representing a brief frame in a sequence of data points over the observing session. In longer intervals, such as the 5-minute intervals in Whiting and Whiting (1975), behavior in the interval is described in narrative fashion, a kind of minicontinous monitoring.

The primary use of fixed-interval sampling is to measure the total duration of activities. If we are making one observation per minute for a 60-minute session, for example, and find that, say, "resting" has been described for 22 of those observations, then we may calculate that 22/60, or 37% of that subject's time during that hour was spent resting. By scheduling observations over a full day (either "daylight" or around the clock), total daily time spent in each activity can be estimated.

Example: Bailey, R. C. and Peacock, N. (1988), Efe (Zaire). The Efe pygmies are tropical forest foragers who live in close symbiosis with settled Lese horticulturalists, a relationship of considerable theoretical interest in cultural anthropology. Bailey following the men and Peacock the women, they used 1-hour focal follows, wherein every minute on the minute (as indicated by a digital watch with count-down functions) the observer recorded the subject's activity using a verb-object code, their posture, location, and the nearest neighbor within 10 meters. Where appropriate, they also recorded the type of food being eaten, the weight of the load carried, responsibility for tending a cooking pot, type of weapon carried, and social interactions (who interacted, what type of interaction, verbal exchanges, child care, etc.). Subjects were always aware that they were being observed but rarely interacted with the observers; when the observer's presence had an obvious impact on their behaviors, this

was noted and the observation was omitted from the data in the sample. Results showed these "foragers" spending surprising amounts of time tending Lese gardens (Efe women spent 111 minutes per day in swiddens as compared to only 40 minutes per day for Efe men), and much of their "foraging" activity is spent obtaining honey and meat for trade to Lese villagers.

The advantages of fixed-interval sampling over continuous monitoring are that field workers experience less fatigue and can collect much more information at each sampling interval than they could on a continuous observation routine. When subjects remain in one place, the field worker can also monitor the activities of more than one person, but this advantage evaporates when individuals separate during the observation, forcing the observer to do a focal-follow on an individual or subgroup that remains together.

Some of the disadvantages of continuous monitoring remain true for fixed-interval sampling: the continuous presence of the field worker increases the risk of subject reactivity, and the successive observations within a session are not statistically independent. In contrast to continuous monitoring, it is debatable whether estimates of frequency, duration, and sequence are entirely trustworthy when derived solely from fixed-interval sampling. As a rule, such estimates are reliable only if the sampling interval is shorter than the duration of each instance of a given behavior. *But we cannot know this unless we already have information on activity duration.* Since behaviors differ in their typical duration (for example, events versus states), a given sampling interval may give stable estimates of frequency for some behaviors but not others. Techniques for choosing the appropriate sample interval are discussed by Martin and Bateson (1993:93).

Perhaps because of these potential drawbacks, fixed-interval sampling is less often used than other techniques of systematic observation. If the goal is data on activity frequencies, durations and sequences, then continuous monitoring is more trustworthy and has greater face validity. If all we need are total durations of activities, then random-interval sampling is cheaper, more compatible with other ethnographic activities, and perhaps more reliable.

Random-Interval Instantaneous Sampling

Also known as spot observations, spot checks, or simply as the time allocation technique, this is a highly preferred method for certain purposes. Behavior, like time, flows seamlessly from moment to moment. Continuous monitoring records the flow, abstracted only by restricting observation to sessions with clearly marked start and end points. Fixed-interval observation abstracts further from the behavior flow by sampling the observation session itself at a series of evenly spaced discrete moments. Random-interval sampling takes this process of abstraction to its logical extreme: The *entire observing session* is reduced to a single instantaneous observation. If the record of fixed-interval sampling is analogous to a motion picture, spot

observation is analogous to an album of random photographs: A few might seem scattershot and unrepresentative, but a large number of truly random snapshots provide an *overall* summary of how people spend their time. But also like such a photo album, and unlike either continuous monitoring or fixed-interval sampling, the order of images carries little or no information about the sequence of activities that precede and follow each observation.

Spot observation begins with an observing schedule that specifies the day and time of each observation and the target individual(s) whose behavior is to be described. Target individuals and observation times should be selected at random, although compromises for practical reasons—such as targeting whole households at random rather than individuals, or following a circuit through the community—are often acceptable. Ideally, the observer finds the target individual at the selected time and records the activity at the instant the target was first seen (or, if the target saw the observer first, the activity she or he was doing *just before* she or he paused to greet the observer). Context data (location, weather, tools being used, etc.) can also be recorded at this moment. If the research subject is absent, interviews can be used to fill in the description, although this should be kept as a separate kind of data to be distinguished in later analysis.

Since each observing session is only a moment, the field worker is left with ample time for detailed descriptions of setting, informal *ad libitum* observations, and ethnographic interviews. This almost seamless integration with other ethnographic activities makes the method appealing even in research where behavioral measurement is secondary to other ethnographic goals. In a sense, spot observation is only a formalized (and randomized) version of common ethnographic visiting practices.

Example: Johnson, A. and O. Johnson (1988), Machiguenga (Peru). Wanting a description of how tropical forest forager-horticulturalists spend time, but not wanting to commit large amounts of field time to continuous monitoring, the researchers applied the spot observation technique (Johnson 1975). Between July 1972 and August 1973, they visited all households in a scattered settlement of 13 households at random during daylight hours (6 am to 7 pm) only, "because travel after dark is hazardous and visiting at night is not encouraged." Households and visiting times were specified in advance using a table of random numbers. Visits were not made every day, because of other demands on the researchers' time, but over the research year visits were made on 135 days totaling 3,495 spot observations of individuals. Activities were coded into a hierarchical coding scheme based on eleven major activity categories (with subcategories in each): eating, food preparation, child rearing, manufacture, wild food collection, garden labor, idleness, hygiene, visiting, school, and wage labor. "The brief time spent in recording activities took only a small fraction of the total field time. In fact, the visits brought us into contact with community members who could be interviewed for other purposes" (Johnson 1975:304). Among the findings of the study was the sensitivity of conclusions about "work" to differences in definition: If "work" is garden labor, then men work 16% of the daytime, women only 7%; if work includes all food production, the percentage climbs to 34% for men, 13% for women; add food

preparation and manufacture, 46% for men, 47% for women; add child care and the men's figure remains the same but women's rises to 56%. Not only does the absolute amount of work change, but also our evaluation of whether men or women work longer hours (see Baksh 1990).

Gross (1984:540–541) writes, "Random spot checks are in fact very economical of observer time. . . . [Compared to other techniques] in terms of sampling validity and level of detail . . . there is little question that random spot checks is the method of choice in time allocation studies." The method also improves the ethnographic enterprise by requiring researchers to be present at places and times they might not ordinarily choose, broadening exposure to local scenes and bringing serendipitous insights. Because each observation is statistically independent, some statistical problems disappear.

On the other hand, tradeoffs exist here as well. Researchers have noted the difficulty of observing an instant of behavior. In reality, field workers tend to watch during a brief "observation window" before they come to a judgment about which behavior the subject is performing: even here, without the benefit of long-term memory, the researcher is still making a choice about which aspects of the stream of behavior to code into "activity" (Altmann 1974; Martin and Bateson 1993). Cross-cultural psychologists have found evidence that individuals in a variety of cultures integrate successive events over a period of about three seconds to construct a subjectively experienced "present" (Feldhutter et al. 1990). Perhaps we should acknowledge that "instantaneous" behavior sampling actually involves a three-second observation window during which the observer constructs the observed activity in ways that have not yet been explored.

Furthermore, if spot observers did consistently record the behavior that occurred the instant they first spotted the target, the most prominent activity would likely be "greeting anthropologist" (Scaglion 1986). Hence, the rule: Record the activity in progress before the presence of the observer became known to the target. But, to the degree to which we use context to decide what the person was doing, we collapse context and behavior where ideally they would be separate data points that we could compare later, confounding efforts to use observational data to understand how environment (context) influences behavior.

Another problem with spot checks is that as a practical reality we end up relying on informant reports and recall to describe a significant amount of behavior, namely, that of individuals not locatable during the spot check. On the one hand, this opens us up to the same problems of memory distortion and biased attention as any other interview data. On the other hand, it introduces other biases, since activities conducted away from the observing eye of the field worker are often certain kinds of activities, including activities at a distance (hunting, wage labor) and those done in private (sex, defecation, violence). Spot checks of readily visible activities will underestimate these activities, and are also liable to overestimate

social activities (Hawkes et al. 1987), and household activities (Borgerhoff Mulder and Caro 1985).

In common with other methods of direct observation, another bias of spot observations concerns nighttime activities (Bernard 1994:325–327). Confining the observations to daytime, typical of spot observations, has the obvious result of overrepresenting daytime activities. Scaglion's (1986) study of 24-hour time allocation in an Abelam community (New Guinea) makes clear that many interesting activities occur at night, including ritual, socializing, food preparation, and even such food production as hunting and gardening.

Spot observations will give a rough picture of time allocation within a small community with a surprisingly small number of observations, perhaps as low as 1,000. That is, in a community of 200, as few as five observations on each individual, if conducted strictly at random, would allow for a broadly accurate picture of how people spend time (Baksh 1989–90). But the more detailed a description needs to be, the more observations must be made. Rare activities, for example, will be missed or inaccurately estimated by small numbers of spot observations. Bernard and Killworth (1993) have estimated the number of observations needed to achieve varying levels of accuracy depending on how rare or common a given activity is (see Bernard 1994:325–326).

One-Zero or Activity Presence Sampling

One-zero sampling is useful when all we need to know is whether a particular behavior of interest—say, hunting or wage labor—occurred during some specified interval of time, but we don't care how much of the activity was performed. Several techniques are commonly used in one-zero sampling, including person-day records, leave-and-return logs, and the in-out diary. The person-day diary is a running account of who performed a particular activity each day of the study, based either on direct observation or interview. The leave-and-return log is a refinement that records the time the subject leaves to perform an activity and the time she or he returns. It is analogous to an industrial worker's time card in that it gives a list of the tasks performed and the window of time within which they occurred, but does not tell just how much of the time window was actually spent in the specified activity as opposed to rest, socializing, eating, and so forth.

In practice, activity-presence recording is most useful when applied to general activity categories of considerable duration, such as labor patterns. Since each subject can be surveyed quickly, it allows the researcher to keep track of a large number of subjects each day. It gives a semiquantitative assessment of activity by measuring the *number of intervals* in which the target activity occurs, sometimes called a "Hansen frequency" (Martin and Bateson 1993). The label "frequency" is misleading, however, since the method doesn't provide true frequencies and systematically underestimates activity frequencies (since some activites are missed)

while overestimating durations (since uncounted activities also helped fill the time window).

Example: Hurtado and Hill (1987), Cuiva (Venezuela). Scholars had disputed whether Cuiva still lived as foragers or had now become settled agriculturalists. Fieldwork in 1985 used a battery of methods to document subsistence ecology. During one 25-day field visit, the authors used an "in-out notebook" to record departure and return times of all individuals leaving the settlement, noting food resources brought back and the composition of foraging parties (n=921 person-days, 517 for men and 404 for women). Unlike other foragers, this study shows that Cuiva men produce more food than the women do, and do so more efficiently; it also established that the bulk of their food still comes from wild game and roots, confirming their status as foragers.

The chief advantages of activity-presence recording are that it is cheap in observer time, nontechnical, objective, and reliable within the limits of what it seeks to measure. Minimal observer judgment is required to score behavior, and recall data can be used with perhaps less error than in other forms of behavioral observation. It gives valuable data about labor variation among individuals, age-sex groups, days of the week, and seasons. Because the observer doesn't follow or encounter the subject, the risk of reactivity is minimal. This can be very valuable in activities like hunting, where the observer's presence is highly likely to lead the hunters to modify their route, pace, and perhaps the game being sought. The method is also highly compatible with other ethnographic activities.

The real limitation of the method is the ambiguity of interpreting the results. Just what do the data indicate in levels of time and effort? Two individuals may be reported to spend the same number of person-days at work, but for one this may represent a few hours of labor whereas for another full days at work. This problem can be addressed by using participant observation or systematic behavior observation to detail daily patterns of activity among a subset of individuals that can then be used to interpret person-day records.

Recording Strategies

Descriptions and Codes

At the moment of recording observations, we have two approaches open to us: Describe the activities we observe or code them. At one end, behavioral descriptions are open-ended textual accounts of behavior; at the other end, codes are brief, usually symbolic representations with a selective, explicit, and preferably unambiguous meaning. This leaves us with three options: (1) avoid coding entirely by keeping textual descriptions and treating our data as a corpus of qualitative narratives; (2) collect textual descriptions to be coded later using categories

developed once the research is complete; or (3) code activities at the moment of observation using categories developed prior to or during early phases of the research.

Text descriptions have the advantage of being like conventional ethnographic description and may be used as such during analysis. They can be as detailed and topically broad as the situation allows and as finely contextualized as necessary. They have a potential richness that can allow reanalysis at a later date or by different researchers. The recording style is free to adapt as the research progresses and the goals of the research change or become refined. On the other hand, use of text descriptions of behavior allows conceptual confusion to lay buried in ways that would be immediately exposed by requiring explicit rules of description. It allows for descriptions of similar events to vary greatly from one instance to the next, and places great demands on the observer.

Coding has its own advantages. It calls for a degree of selectivity: The goal of coding is not to capture reality in all its complexity but to highlight those aspects we consider significant in the context of the research. Codes are ideally explicitly defined and objective, typically in a detailed codebook that spells out how codes are to be applied so that different coders can achieve a high degree of interobserver reliability. With codes, it is ideally possible to have a complete set of codes to cover all activities, with discrete codes for discrete activities, minimizing overlap (mutually exclusive and exhaustive; Robson 1993). These strengths of coding are also, in some contexts, limitations: codes reduce the complexity and ambiguity of real behavior and to that degree represent a loss of information in exchange for order and clarity.

The Behavior Record

Each observation in a behavior record should have a unique identification code for record keeping. Each subject, too, will have a unique identification code. As part of a general ethical requirement to protect the anonymity of research subjects, these codes will not, as a rule, identify the subject to a casual outside observer. The subject should be on a census list that contains relevant information like age, gender, household, position in the family, and perhaps other information like occupation, ethnicity, and health status. Each observation should also pinpoint the time of observation on a 24-hour clock

Structural and Functional Descriptions

Researchers have distinguished two sorts of behavioral descriptions: structural descriptions of an actor's physical movements and functional descriptions of the presumed objectives or consequences of the behavior (Borgerhoff Mulder and Caro 1985; Hames 1992). Structural descriptions document the actor's posture (sitting,

walking, climbing, etc.), manipulatory movements (carrying, handling, touching, etc.), and social interactions (speaking, striking, receiving an object, etc.). Structural descriptions may be conveniently recorded in grammatical data "sentences"—"individual X walks rapidly" (subject-verb-adverb) or "individual Y peels potato" (subject-verb-object). Since structural codes are straightforward descriptions of observable actions, they can be recorded reliably and objectively, with minimal inference on the part of observers.

However, the same physical act can have very different consequences depending on the context. For example, the action of repeatedly raising and lowering a hoe while standing may have different outcomes depending on whether it is performed while weeding a garden, clearing an irrigation ditch, or killing a snake that wandered into the yard. The goal of functional description is to characterize—at least in a general way—the purposes and likely outcomes of the actor's movements.

Thus, rather than recording in detail all the component movements of the actor's behavior, the observer describes the action using broader functional categories such as food production, child care, cooking, hygiene, and so forth. The specific functional activity classification scheme—which is best worked out in advance of making observations—depends on the nature of the research problem and the habitual activity patterns of the study population. A period of informal observation at the beginning of fieldwork can be invaluable for testing and refining activity classification schemes.

The field worker need not choose between structural and functional description: There are good reasons for collecting both simultaneously. Recording both physical actions and their presumed consequences increases confidence that measurement was not biased by preconceived notions of function, improves the richness of our descriptions of general (functional) activity categories such as "production" and "leisure," and preserves information on both what the observer saw and how she or he interpreted it (Borgerhoff Mulder and Caro 1985).

Coding

To facilitate cross-cultural comparisons of activity patterns, Gross (1984:542) calls on investigators to "standardize their coding of behavior into a set of broad descriptive categories of behavior. This would," he points out, "make the results of different investigators collected for different purposes usable for comparative purposes." With Gross (1984), we stress that any standardized coding system can be modified, with sub-categories added that are appropriate to particular cases. Table 3 shows a hierarchical coding scheme developed for use across cultures in small-scale and peasant societies (Johnson and Johnson 1988). Analogous coding schemes have been used by sociologists working in urban societies since the 1960s (Szalai 1972).

TABLE 3
Standardized Hierarchical Activity Codes Designed for Cross-Cultural Comparisons of Time Use Patterns.

F Food production
Includes foraging, agriculture, and husbandry producing food for household consumption, including travel to and from site of food production, cultivation of nonfoods or cash crops if these are undifferentiated from subsistence production, and soil preparation.
- FA Agriculture
- FC Collecting wild plant foods
- FF Fishing
- FH Hunting and fowling
- FL Tending food and draft animals
- FX Other food production
- FU Food production, type unknown

C Commercial activities
Activities oriented toward the production and exchange of money and trade goods.
- CA Cash cropping, raising livestock for sale
- CC Collecting wild/natural products for sale, including mining
- CM Manufacturing articles for sale
- CS Shopping, buying and selling, and bartering
- CW Wage laboring, selling labor to others, service for money
- CX Other commercial activity
- CU Commercial activity, type unknown

M Manufacture
Making and repairing the household, furnishings, implements, clothing (all for household use); nonpaid work on facilities such as roads, fences, irrigation channels, traps, etc.
- MA Making or repairing portable artifacts
- MC Making or repairing clothing
- MF Building or repairing immobile facilities
- MM Acquiring materials for manufactures
- MX Other manufacture
- MU Manufacture, type unknown

P Food preparation
Processing, storing, and serving food.
- PC Cooking food
- PG Handling or processing food for storage
- PH Handling or processing food for consumption in the near future
- PS Serving or transporting food
- PX Other food preparation
- PU Food preparation, type unknown

H Housework
Household tidying and cleaning chores, cleaning related to meals, managing household water and fuel, pet care, cleaning clothing.
- HH Housekeeping, tidying, cleaning
- HM Fetching and managing household water and fuel

HX Other housework
HU Housework, type unknown

E Eating
Eating and other meal-time activities.
EE Eating, drinking, and ingestion of nonfoods
ES Suckling
EX Other food consumption
EU Food consumption, type unknown

S Social
Includes a broad range of activities not better classified into the above categories, which are distinguished by social exchanges.
SC Child care, actively tending child
SE Acquiring or giving education, information, socialization in a face-to-face context
SO Care for another (nonchild), receiving care
SP Group ceremony, ritual, political activity (including both spectating and participating)
SR Group recreational activity or public entertainment (participating or spectating)
SS Socializing, chatting, visiting
SX Other social activity
SU Social activity, type unknown

I Individual
Self-involved activities that do not fit well into other categories.
IE Acquiring education or information, alone
IG Self-grooming, dressing, hygiene
II Idle due to illness
IN Idle "doing nothing"
IP Participating in individual religious observance
IR Participating in individual recreation or entertainment
IS Sleeping, napping
IX Other individual activity
IU Individual activity, type unknown

U Away from community unobserved
This category applies only to activities that take place outside of the specified geographical range of the behavior study (for example, on an extended trip visiting another community). Secondary codes refer to the primary activity best describing the purpose of the trip, giving priority to those activities earlier in the list. Note: this code is not to be used to distinguish unobserved activities ("time-outs") when the subject is temporarily out of sight.
UF Away to do food-producing activity
UC Away to do commercial activity
UM Away to do manufacturing activity
UP Away to do food preparation (possible, but probably unlikely)
UH Away to do housework activity (also unlikely)
UE Away to eat (unlikely)
US Away to do social activity
UI Away to do individual activity
UX Away to do activity classified as "Other"
UU Away for unknown purpose

X Other
Activity that does not fit into any of the above categories. No secondary codes suggested, although field workers using this activity list undoubtedly will want to subdivide and specify this residual category.

SOURCE: Based on Standard Activity Codes, UCLA Time Allocation Project (see Johnson and Johnson 1988).

Three problems that confront the researcher during coding are *simultaneity*, *reliability*, and *context*.

1. *The simultaneity problem.* What do we record when an actor is performing what could be coded as two different activities? A common example is a woman nursing a baby while tending a cooking pot: is her behavior "child care" or "food preparation"? A small but significant proportion of activities are not mutually exclusive in this sense; such doubling-up of activities is actually an interesting empirical question: For example, it is often argued that women's activities are constrained by their responsibility for child care, so that their other activities must be compatible.

Strategies for handling the simultaneity problem:

1. Use a structural coding system and describe physical acts, perhaps along with a functional activity code (Borgerhoff Mulder and Caro 1985), thus preserving information about simultaneity.
2. Code both activities as a full observation. This preserves information about simultaneity but raises the problem of double accounting (inflated *n*).
3. Code both activities, but treat each as a half-observation. This avoids the problem of inflated *n* but discounts activities performed simultaneously: nursing a baby while cooking only counts half as much as nursing a baby alone.
4. Establish priority rules, such as coding food preparation over child care: In this case, the woman would only be counted as cooking. This solves the problem of inflated *n* but loses information about simultaneity.
5. Code the different activities as separate dimensions of behavior. In this case, there might be an activity code (cooking) and a social code (child care).
6. Develop combination codes such as cooking/child care to cover the most common cases of simultaneity.

Each of these solutions has its strengths and shortcomings. Whichever is chosen must be consistently adhered to and made explicit in publishing the data.

2. *The reliability problem.* There are two questions that frame this problem: Does a single observer consistently record the same behavior in the same way (intracoder reliability)? Do two observers consistently code the same behavior in the same way (intercoder reliability)? The reliability problem must be addressed early in the research —even prior to it —so that a detailed codebook can be constructed. The codebook should do more than just list the complete set of codes. It should spell out how the codes will be applied. Codes must be applied in practice and evaluated for usefulness; coders must practice also and

learn how to increase their consistency. A steep learning curve means that a few weeks of practice may be needed in order to stabilize recording or coding and achieve a high degree of intercoder reliability (Robson 1993).

3. *The context problem.* We use context both to help us interpret behavior and to look for connections between environment and behavior. As noted above, this is only possible to the degree that our descriptions of behavior are separate from our descriptions of context. Some aspects of context are *constant*, in the sense that they change little or not at all during the research (for example, political institutions, roads, markets). A census, map, and a general ethnographic description often suffice to supply this information. Other aspects of context are *variable* and must be recorded on each observation. Common context variables include date and time, weather, location, social interaction, and technology used. It is possible, however, to go overboard with context descriptions: in our experience few anthropologists do much with the context data they collect—people tend to do agriculture in their fields, housework in their homes, etc. As with behavioral descriptions, effort put into context descriptions should be constrained by the question: What are the data for?

Understanding of context can also be enhanced by attention to: (1) *methodological details* needed to establish the reliability of the observation, such as whether the behavior was directly observed or established by hearsay; and (2) *provision for open-ended comments* on the circumstances of the observation, amplification of the data recorded, or general ethnographic observations.

Conclusion

In ethnography, what people *do* matters. Much of ethnographic description is intended to tell us how members of a group behave: how they make a living, raise children, resolve conflicts, worship, celebrate, mourn. The accuracy, fairness, completeness—in a word, verisimilitude—of ethnographic descriptions is of central importance to us, yet we as ethnographers are astonishingly cavalier about how we develop our descriptions of behavior. We do not take seriously enough the severe limitations that the cultural construction of long-term memory place on the ability of both research subjects and fieldworkers to develop accurate descriptions of ongoing behavior.

Accurate descriptions of behavior matter because, from the most theoretically abstract to the most immediately applied concerns, our understanding of the causes and consequences of human behavior depends on them. We would ordinarily have little respect for theories or policies based on data that had error rates ranging from 50%–80%. Yet participant observation typically produces descriptions of behavior with such error rates, unless the observations are carefully recorded immediately as the behavior occurs.

Once it is accepted that accurate descriptions of the behavior of members of a community matter, the methodological issues we have addressed in this chapter become paramount. What does it mean to describe an activity? Whom should we observe, when, and where? At what level of detail? How can we sample from the stream of behavior in a community to generate a fair representation of people of differing age, gender, and status, in all the relevant contexts of their lives?

These are issues that we should address as far as possible in the design phase of research. But, because research in real communities is always unpredictable to some degree, we must remain flexible and ready to change methods to meet contingencies. The best way to do this is to develop skills in a variety of methods of direct systematic observation and an understanding of the rationales behind them. Researchers who do so will be able to make the inevitable compromises fieldwork entails while achieving the most accurate and representative descriptions possible.

REFERENCES

Alexander, R. M. 1992. *The Human Machine*. New York: Columbia University Press.

Altmann, J. 1974. The Observational Study of Behavior. *Behavior 48*:1–41.

Bailey, R. C., and N. Peacock. 1988. *Time Allocation of Efe Pygmy Men and Women of the Ituri Forest, Zaire*. Vol. 3 in *Cross-Cultural Studies in Time Allocation*. New Haven: Human Relations Area Files.

Baksh, M. 1989–90. The Spot Observations Technique in Time Allocation Research. *Cultural Anthropology Methods 1*:3; 2:6–7; 3:4–5.

Baksh, M. 1990. *Time Allocation among the Machiguenga of Camana*. Vol. 7 in *Cross-Cultural Studies in Time Allocation*. New Haven: Human Relations Area Files.

Bernard, H. R., P. J. Pelto, O. Werner, J. Boster, A. K. Romney, A. Johnson, C. R. Ember, and A. Kasakoff. 1986. The Construction of Primary Data in Cultural Anthropology. *Current Anthropology 27*:382–396.

Bernard, H. R., and P. D. Killworth. 1993. Sampling in Time Allocation Research. *Ethnology 32*:207–215.

Bernard, H. R. 1994. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*, 2d ed. Walnut Creek, CA: AltaMira Press.

Betzig, L. L., and P. Turke. 1985. Measuring Time Allocation: Observation and Intention. *Current Anthropology 26*:647–650.

Borgerhoff Mulder, M. B., and T. M. Caro. 1985. The Use of Quantitative Observation Techniques in Anthropology. *Current Anthropology 26*:323–336.

Bornstein M. H., and H. G. Bornstein. 1976. The Pace of Life. *Nature 259*:557–559.

Carlstein, T. 1982. *Time Resources, Society, and Ecology*. Vol. 1, *Preindustrial Societies*. London: George Allen and Unwin.

Chapple, E. D., and C. M. Arensberg. 1940. Measuring Human Relations: An Introduction to the Study of the Interaction of Individuals. *Genetic Psychology Monographs 22*:3–147.

D'Andrade, R. 1974. Memory and the Assessment of Behavior. In *Measurement in the Social Sciences*. H. Blalock, ed. Pp. 159–186. Chicago: Aldine.

D'Andrade, R. 1995. *The Development of Cognitive Anthropology*. Cambridge: Cambridge University Press.

Dufour, D. L. 1984. The Time and Energy Expenditure of Indigenous Women Horticulturalists in the Northwest Amazon. *American Journal of Physical Anthropology 65*:37–46.

Dunbar, R.I.M. 1976. Some Aspects of Research Design and Their Implications in the Observational Study of Behavior. *Behavior 58*:78–98.

Engle, P., and J. Lumpkin. 1991. How Accurate Are Time Use Estimates? Effects of Cognitive Enhancement and Cultural Differences on Recall Accuracy. *Applied Cognitive Psychology 6*:141–159.

Feldhutter, I., M. Schleidt, and I. Eibl-Eibesfeldt. 1990. Moving in the Beat of Seconds. *Ethology and Sociobiology 11*:511–520.

Flinn, M. 1988. Parent-Offspring Interactions in a Caribbean Village. In *Human Reproductive Behavior: A Darwinian Perspective*. L. L. Betzig, M. Borgerhoff Mulder, and P. Turke, eds. Pp. 189–200. Cambridge: Cambridge University Press.

Gross, D. R. 1984. Time Allocation: A Tool for the Study of Cultural Behavior. *Annual Reviews in Anthropology 13*:519–558.

Hames, R. 1979. A Comparison of the Efficiencies of the Shotgun and Bow in Neotropical Forest Hunting. *Human Ecology 7*:219–252.

Hames, R. 1992. Time Allocation. In *Evolutionary Ecology and Human Behavior* E. A. Smith and B. Winterhalder, eds. Pp. 203–235. New York: Aldine.

Harris, M. 1964. *The Nature of Cultural Things*. New York: Random House.

Harris, M. 1990. Emics and Etics Revisited. In *Emics and Etics: The Insider/Outsider Debate*. T M. Headland, K. L. Pike, and M. Harris, eds. Pp. 48–61 Thousand Oaks, CA. Sage Publications.

Hawkes, K., H. Kaplan, K. Hill, and A. M. Hurtado. 1987. A Problem of Bias in Scan Sampling. *Journal of Anthropological Research 43*:239–246.

Hill, K., H. Kaplan, K. Hawkes, and A. M. Hurtado. 1985. Men's Time Allocation to Subsistence Work among the Ache of Eastern Paraguay. *Human Ecology 13*(1):29–47.

Hurtado, A. M., and K. R. Hill. 1987. Early Dry Season Subsistence Ecology of Cuiva (Hiwi) Foragers of Venezuela. *Human Ecology 15*:163–187.

Johnson, A. 1975. Time Allocation in a Machiguenga Community. *Ethnology 14*: 301–310.

Johnson, A., and O. R. Johnson. 1988. *Time Allocation among the Machiguenga of Shimaa*. Vol. 1 in *Cross-Cultural Studies in Time Allocation*. New Haven: Human Relations Area Files.

Lee, R. B. 1979. *The !Kung San*. Cambridge: Cambridge University Press.

Levine, R. V. 1990. The Pace of Life. *American Scientist 78*:450–459.

Lhermillier, A., and N. Lhermillier. 1983. Vida economica y social de un nucleo familiar Yanomami. *Boletin Antropologico 3*:39–65.

Martin, P., and P Bateson. 1993. *Measuring Behavior· An Introductory Guide*, 2d ed. Cambridge: Cambridge University Press.

Montoye, H. J., H.C.G. Kemper, W.H.M. Saris, and R. A. Washburn. 1996. *Measuring Physical Activity and Energy Expenditure*. Champaign, IL: Human Kinetics.

Niemi, I., S. Kiiski, and M. Liikkanen. 1981. *Use of Time in Finland, 1979*. Helsinki: Central Statistical Office of Finland.

Panter-Brick, C. 1989. Motherhood and Subsistence Work. The Tamang of Rural Nepal. *Human Ecology 17*:205–228.

Paolisso, M. J., and R. Sackett. 1988. *Time Allocation among the Yukpa of Venezuela*. Vol. 2 in *Cross-Cultural Studies in Time Allocation*. New Haven: Human Relations Area Files.

Robson, C. 1993 *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Oxford: Blackwell.

Romney, A. K., S. Weller, and W. Batchelder. 1986. Culture as Consensus: A Theory of Culture and Informant Accuracy. *American Anthropologist 88*:313–338.

Sackett, R. 1996. Time, Energy, and the Indolent Savage: A Cross-Cultural Test of the Primitive Affluence Hypothesis. Ph.D. diss., Department of Anthropology. Los Angeles: University of California.

Scaglion, R. 1986. The Importance of Nighttime Observations in Time Allocation Studies *American Ethnologist 13*:537–545.

Shweder, R. A. 1982. *Fact and Artifact in Trait Perception. The Systematic Distortion Hypothesis*. Vol. 11 in *Normal Personality Processes: Progress in Experimental Psychology Research*. B. A. Maher and W B. Maher, eds. Pp. 65–100. New York. Academic Press.

Szalai, A., ed. 1972. *The Use of Time. Daily Activities of Urban and Suburban Populations in Twelve Countries*. The Hague and Paris: Mouton.

Whiting, B., and J.W.M Whiting. 1973. Methods for Observing and Recording Behavior In *A Handbook of Method in Cultural Anthropology*. R. Naroll and R. Cohen, eds. Pp. 282–315. New York: Columbia University Press.

Whiting, B., and J.W.M. Whiting. 1975. *Children of Six Cultures: A Psycho-Cultural Analysis*. Cambridge: Harvard University Press.