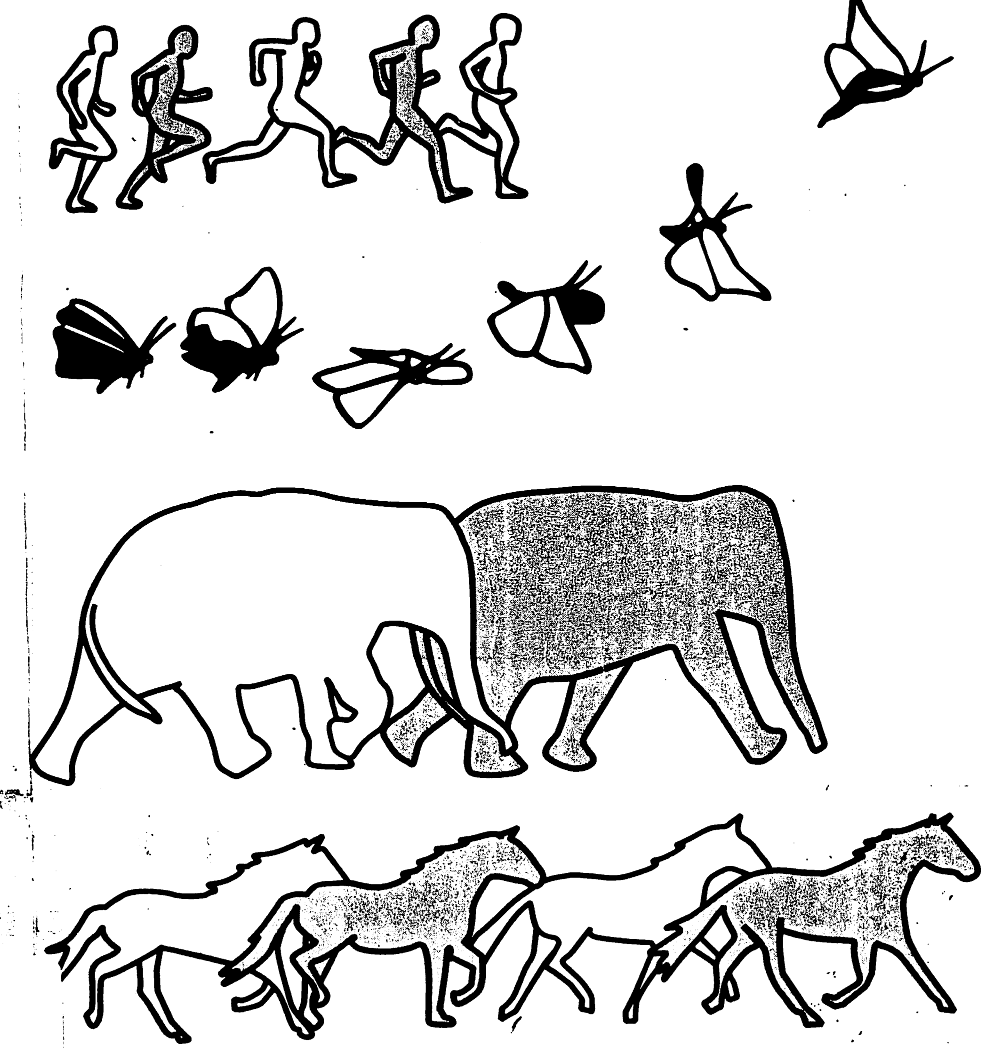
MEASURING BEHAVIOUR

An introductory guide

SECOND EDITION



Paul Martin and Patrick Bateson

6

Recording methods

A Sampling rules: *ad libitum*, focal, scan and behaviour sampling

When deciding on systematic rules for recording behaviour, two levels of decision must be made. The first, which we refer to as **sampling rules**, specifies which subjects to watch and when. This covers the distinction between *ad libitum* sampling, focal sampling, scan sampling and behaviour sampling. The second, which we refer to as **recording rules**, specifies *how* the behaviour is recorded. This covers the distinction between continuous recording and time sampling (which, in turn, is divided into instantaneous sampling and one-zero sampling; see Fig. 6.1). In this section, we consider the four different sampling rules.

1 ***Ad libitum sampling*** means that no systematic constraints are placed on what is recorded or when. The observer simply notes down whatever is visible and seems relevant at the time.

Clearly, the problem with this method is that observations will be biased towards those behaviour patterns and individuals which happen to be most conspicuous. For example, in a study of agonistic behaviour in rhesus monkeys, Bernstein (1991) found that *ad libitum* sampling tended to miss brief responses and under-estimated the involvement of younger individuals in social interactions. Provided this important limitation is borne in mind, however, *ad libitum* sampling can be useful during preliminary observations, or for recording rare but important events.

2 ***Focal sampling*** (or 'focal animal sampling') means observing one individual (or one dyad, one litter or some other unit) for a specified

amount of time and recording all instances of its behaviour – usually for several different categories of behaviour. Ideally, the choice of focal individual is determined prior to the observation session. When recording the social behaviour of the focal individual it may also be necessary to record certain aspects of other individuals' behaviour, such as who initiates interactions and to whom behaviour is directed. Focal sampling is generally the most satisfactory approach to studying groups.

Inevitably the focal individual will, on occasions, be partially obscured or move completely out of sight, in which case recording has to stop until it is visible again. Any such interruption should be recorded as 'time out', and the final measure calculated according to the time for which the focal individual was visible. Note, however, that omitting 'time out' may introduce bias if subjects systematically tend to do certain things whilst out of sight. For example, many animals seek privacy when eating or mating, so their behaviour when visible to the observer is not necessarily representative of their behaviour as a whole (e.g., Harcourt & Stewart, 1984).

Focal sampling can be particularly difficult under field conditions because the focal individual may leave the area and disappear completely. If this happens, should it be pursued and if so, for how long; or should the observations be stopped and a new focal individual chosen? Explicit rules are needed for deciding what to do if the focal individual does disappear in the middle of a recording session.

Some authors confusingly use 'focal (animal) sampling' as a synonym for continuous recording (see section 6.C), conflating a sampling rule (who is watched) with a recording rule (how their behaviour is recorded). In fact, any of the three different recording rules (continuous recording, instantaneous sampling or one-zero sampling) can be used when recording the behaviour of a single subject (focal sampling).

3 ***Scan sampling*** means that a whole group of subjects is rapidly scanned, or 'censused', at regular intervals and the behaviour of each individual at that instant is recorded. The behaviour of each individual scanned is, necessarily, recorded by *instantaneous sampling* (see section 6.D), although the two terms are not synonymous. Scan sampling usually restricts the observer to recording only one or a few simple categories of behaviour, such as whether or not a particular activity is occurring or which individuals are asleep.

The time for which each individual is watched in a scan sample should, in theory, be negligible; in practice, it may at best be short and roughly constant. A single scan may take anything from a few seconds to several minutes, depending on the size of the group and the amount of information recorded for each individual (see Altmann, 1974).

An obvious danger with scan sampling is that results will be biased because some individuals or some behaviour patterns are more conspicuous than others. For example, Harcourt & Stewart (1984) found that previously published figures for the amount of time spent feeding by gorillas in the wild were too low. They argued that this was because previous studies had used scan sampling, and gorillas are less visible to observers when they feed. Focal animal sampling with the subjects continuously in view, which was used by Harcourt & Stewart to obtain a better estimate of time spent feeding, is not subject to this source of bias.

For some purposes, though, scan sampling has distinct practical advantages. For example, de Ruiter (1986) used scan sampling to study the activity budgets of two groups of wild capuchin monkeys. The technique enabled him to obtain data that were evenly representative across all individuals, time of day and season, allowing various behavioural and ecological comparisons to be drawn between the two groups. Such a broad spread of data would not have been practicable had focal animal sampling been used.

Scan sampling can be used in addition to focal sampling during the same observation session. For example, the behaviour of a focal individual may be recorded in detail, but at fixed intervals (say, every 10 or 20 min) the whole group is scan-sampled for a single category, such as the predominant activity or proximity to each other.

If scan samples are to be used as separate data points, rather than averaged to provide a single score, they must be statistically independent of one another. This means that they must be adequately spaced out over time: clearly, scan samples taken at, say, 30-s intervals would not constitute independent measurements.

Note that some authors use 'scan sampling' to refer to instantaneous sampling (see section 6.D), again conflating a sampling rule with a recording rule.

4 Behaviour sampling means that the observer watches the whole group of subjects and records each occurrence of a particular type of behaviour, together with details of which individuals were involved.

Behaviour sampling is mainly used for recording rare but significant types of behaviour, such as fights or copulations, where it is important to record each occurrence. Rare behaviour patterns would tend to be missed by focal or scan sampling. Behaviour sampling is often used in conjunction with focal or scan sampling and is subject to the same source of bias as scan sampling, since conspicuous occurrences are more likely to be seen. Indeed, behaviour sampling is sometimes referred to as 'conspicuous behaviour recording'.

B Recording rules: continuous recording versus time sampling

Recording rules are of two basic types:

1 Continuous recording (or 'all-occurrences' recording). This method aims to provide an exact and faithful record of the behaviour, measuring true frequencies and durations and the times at which behaviour patterns stopped and started. It is described in more detail in section 6.C.

2 Time sampling. Here, the behaviour is sampled periodically, therefore less information is preserved and an exact record of the behaviour is not necessarily obtained. Time sampling can be subdivided into two principal types: **instantaneous sampling** and **one-zero sampling**. These are dealt with in sections 6.D and 6.E. The hierarchy of sampling rules and recording rules is represented in Fig. 6.1.

Time sampling is a way of condensing information, thereby making it possible to record several different categories of behaviour simultaneously. In order to do this the observation session is divided up into successive, short periods of time called **sample intervals** (see Fig. 6.2). (The choice of sample interval is discussed in section 6.F.) The instant of time at the end of each sample interval is referred to as a **sample point** (see Fig. 6.2). For example, a 30-min observation session

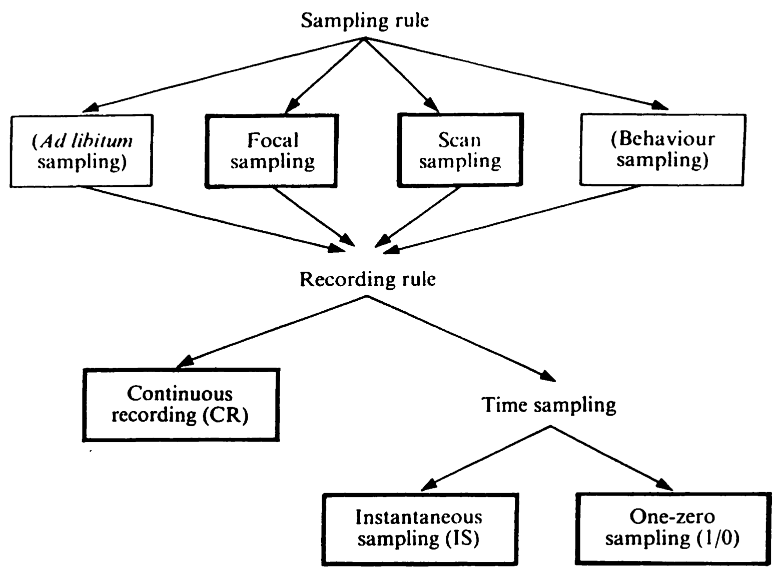


Fig. 6.1. The hierarchy of *sampling rules* (determining who is watched and when) and *recording rules* (determining how their behaviour is recorded).

might be divided up into 15-s sample intervals, giving 120 sample points. Successive sample points are denoted by a stopwatch or, more conveniently, by a **beeper** – a small electronic timer which gives the observer an audio cue through an earphone once every sample interval (see Appendix 4). The distinction between continuous recording and time sampling also applies when computers are used to record data automatically.

A final point worth emphasising is that continuous recording and time sampling can be used simultaneously for recording different categories of behaviour.

C Continuous recording

With continuous recording (or 'all-occurrences' recording) *each occurrence* of the behaviour pattern is recorded, together with information about its time of occurrence. True continuous recording aims to produce an exact record of the behaviour, with the times at which each instance of the behaviour pattern occurred (for events), or began and ended (for states).

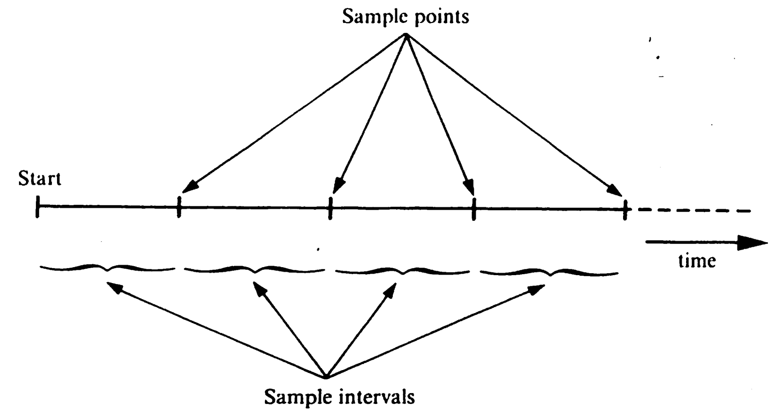


Fig. 6.2. The division of an observation session into successive, short units of time, or *sample intervals*, for the purposes of time sampling. The end of each sample interval (*sample point*) is often denoted by an audio *beeper*.

1 ***Measures obtained.*** For both events and states, continuous recording generally gives true frequencies, and true latencies and durations if an exact time base is used. However, bias can arise if a measurement of duration or latency is terminated before the bout of behaviour actually ends, either because the recording session ends or because the subject disappears from view. This is because the longer a bout of behaviour lasts, the more likely it is that its duration will be under-estimated by the termination of recording.

2 ***Applications.*** Continuous recording preserves more information about a given category of behaviour than time sampling, and should be used whenever it is necessary to measure true frequencies or durations accurately. Continuous recording is also necessary when the aim is to analyse sequences of behaviour. However, its use can be limited by practical considerations, since continuous recording is more demanding for the observer than time sampling. One consequence of this is that fewer categories can be recorded at any one time. Trying to record everything can mean that nothing is measured reliably.

In practice, continuous recording is typically used for recording the frequencies of discrete events and for recording the durations of behavioural states, particularly when it is important to preserve information about the sequence of behaviour patterns.

D Instantaneous sampling

With instantaneous sampling (or 'point sampling', or 'fixed-interval time point sampling'), the observation session is divided into short sample intervals. *On the instant of each sample point (i.e., on the 'beep'), the observer records whether or not the behaviour pattern is occurring* (see Fig. 6.3).

1 ***Measure obtained.*** The score obtained by instantaneous sampling is expressed as the *proportion of all sample points on which the behaviour pattern was occurring*. For example, if a 30-min recording session was divided into 15-s sample intervals, and a behaviour pattern occurred on 40 out of the 120 sample points, then the score would be 40/120 = 0.33. (Note that an instantaneous sampling score is a dimensionless index with no units of measurement.) Instantaneous sampling gives a single score for the whole recording session: obviously, individual sample points within a session cannot be treated as statistically independent measurements.

Instantaneous sampling does not give true frequencies or durations. However, if the sample interval is short relative to the average duration of the behaviour pattern then instantaneous sampling can produce a record that approximates to continuous recording. The shorter the sample interval, the more accurate instantaneous sampling is at estimating duration and the more closely it resembles continuous recording. If the sample interval is short then an instantaneous sampling score gives a direct estimate of the

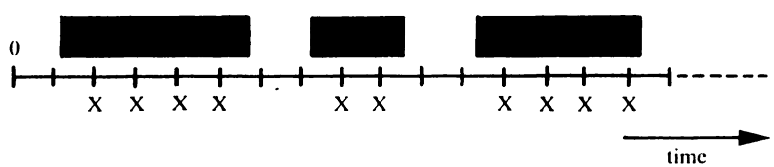


Fig. 6.3. Instantaneous sampling (or 'point sampling'). The behaviour pattern, denoted here by the black bars, is scored (by an 'X') according to whether or not it is occurring on the instant of each successive sample point ('on the beep').

proportion of time for which the behaviour occurred. For example, if the instantaneous sampling score was 0.25, the best estimate of the proportion of time spent performing the behaviour would also be 0.25.

The accuracy of instantaneous sampling depends on the length of the sample interval (which should be as short as possible), the average duration of the behaviour pattern (which should be long relative to the sample interval) and, strictly speaking, the average duration of the interval between successive bouts of the behaviour (which should also be long relative to the sample interval). Of course, if a very short sample interval is used then the practical benefits of time sampling are lost, in which case continuous recording might as well be used instead.

A number of empirical studies have shown that instantaneous sampling can give a good approximation to the proportion of time spent performing a behaviour pattern (e.g., Dunbar, 1976; Leger, 1977; Simpson & Simpson, 1977; Rhine & Flanigon, 1978; Tyler, 1979; Pöysä, 1991). Choosing the right length of sample interval is discussed in section 6.F.

2 ***Applications.*** Instantaneous sampling is used for recording behavioural states that can unequivocally be said to be occurring or not occurring at any instant in time; for example, measures of body posture, orientation, proximity, body contact, or general locomotor activity. Instantaneous sampling is *not* suitable for recording discrete events of short duration. Neither is it suitable for recording rare behaviour patterns, since a rare behaviour pattern is unlikely to be occurring at the instant of any one sample point and therefore will usually be missed.

One potential source of bias with instantaneous sampling is the observer's natural tendency to record conspicuous behaviour patterns even if they occur slightly before or after the sample point. The sample point is therefore stretched out from an instant to become a window of finite duration, making the sampling no longer 'instantaneous'. If, as is likely, this is mainly done with the more noticeable or important behaviour patterns then these will tend to be overestimated relative to less prominent behaviour patterns.

Note that instantaneous sampling is sometimes confusingly referred to as 'scan sampling' in the behavioural literature.

E One-zero sampling

In one-zero sampling, as with instantaneous sampling, the recording session is divided up into short sample intervals. *On the instant of each sample point (i.e., on the 'beep'), the observer records whether or not the behaviour pattern has occurred during the preceding sample interval.* This is done irrespective of how often, or for how long, the behaviour pattern has occurred during that sample interval (see Fig. 6.4). An equivalent procedure is to record the behaviour pattern when it first occurs, rather then waiting until the end of the sample interval.

1 ***Measure obtained.*** The score obtained by one-zero sampling is expressed as ***the proportion of all sample intervals during which the behaviour pattern occurred.*** For example, if a behaviour pattern occurred during 50 out of the 120 15-s sample intervals in a 30-min recording session, the score would be 50/120 = 0.42. Note that, as with instantaneous sampling, one-zero sampling gives a single, dimensionless score for the whole recording session. Again, individual sample points within a recording session obviously cannot be treated as statistically independent measurements.

One-zero sampling does not give true or unbiased estimates of durations or frequencies. The proportion of sample intervals in which the behaviour occurred to any extent cannot be equated either with the length of time spent performing the behaviour, or with the number of times the behaviour occurred. This point must be emphasised even more strongly than in the case of instantaneous sampling since, in addition to not necessarily being accurate, one-zero sampling also introduces systematic bias. One-zero sampling consistently *over*-estimates duration to some extent, because the behaviour is recorded as though it occurred throughout the sample interval

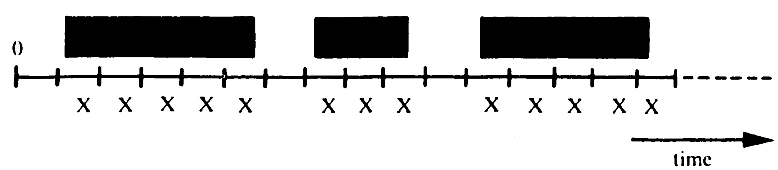


Fig. 6.4. One-zero sampling. The behaviour pattern, denoted here by the black bars, is scored according to whether or not it has occurred during the preceding sample interval.

when it need not have done so in reality. One-zero sampling also tends to *under*-estimate the number of bouts performed, because the behaviour could have occurred more than once during a sample interval. The shorter the sample interval relative to the average duration of the behaviour pattern, the more closely one-zero sampling approximates to instantaneous sampling. There are methods for estimating frequencies and durations from time-sampled records (Chow & Rosenblum, 1977).

If one-zero scores are compared – either between subjects or for different occasions – then problems can arise unless the mean bout length of the behaviour remains roughly constant. This is because the error in estimating frequency or duration depends on the ratio of mean bout length to sample interval. Thus, if the mean bout length of the behaviour varies between individuals (or, for the same individual, varies between different recording sessions) then the error in estimating frequency or duration will also vary (Dunbar, 1976).

Scores obtained using one-zero sampling are sometimes referred to as 'Hansen frequencies', but this term is best avoided because it encourages the mistaken view that one-zero scores actually are frequencies. One-zero sampling is also sometimes referred to as 'fixed-interval time span sampling'.

2 ***Applications.*** Because of the problems outlined above, one-zero sampling should be used with caution. Indeed, some authorities assert that one-zero sampling should ***never*** be used. Our own view is that one-zero sampling is useful and acceptable for recording certain types of behaviour patterns for which neither continuous recording nor instantaneous sampling is suitable. This issue is discussed in section 6.G.

F Choosing the sample interval

The size of the sample interval used in time sampling will depend on how many categories are being recorded, as well as the nature of the behaviour. The shorter the sample interval the more accurate a time-sampled record will be. However, the shorter the sample interval the more difficult it is reliably to record several categories of behaviour at once – especially if the behaviour is complicated or occurs rapidly.

In practice, it is always necessary to balance the theoretical accuracy of measurement (which requires the shortest possible sample interval) against ease and reliability of measurement (which requires an adequately long interval). If the sample interval is too short, observer errors can make recording less reliable than if a slightly longer interval were chosen. Thus, the sample interval should be the shortest possible interval that allows the observer to record reliably, after a reasonable amount of practice.

The best sample interval depends on what is being measured and is partly a matter of trial and error. To give some idea, though, many observers use a sample interval in the range from 10 s to 1 min, with sample intervals of 15, 20 or 30 s being common under laboratory conditions. Field studies, especially those in which long recording sessions are used, may need to use longer sample intervals.

Rather than relying on a combination of common sense and trial-and-error the sample interval can be chosen objectively, although this requires a considerable amount of additional work. First, a fairly large sample of the behaviour must be measured using continuous recording, in order to give a true picture of what actually happened. Scores are then calculated for each category as though the behaviour had been recorded using time sampling with various sample intervals (e.g., 10, 20, 30, . . . s). The discrepancy between the continuous record and the simulated time sampling measure can then be calculated for each sample interval. The error due to time sampling will increase as the simulated sample interval becomes larger. However, it may be possible to distinguish an obvious 'break-point', above which time sampling is too inaccurate, but below which it gives a reasonable approximation to continuous recording. This point marks the longest sample interval that can be used if the record is to be reasonably accurate for that category of behaviour.

In most studies, of course, several behavioural categories are recorded and the sample interval used must be suitable for all of them. The simplest approach here is to specify the maximum acceptable discrepancy between the continuous record and the time-sampled measure for any category (say, 10%), and plot the proportion of categories where this condition is satisfied as a function of the sample interval. An example is shown in Fig. 6.5.

One problem with this whole process is the need initially to measure the behaviour using continuous recording in order to provide a true record for comparison. In many cases time sampling is used precisely because continuous recording is simply not practicable, which would obviously rule out this procedure.

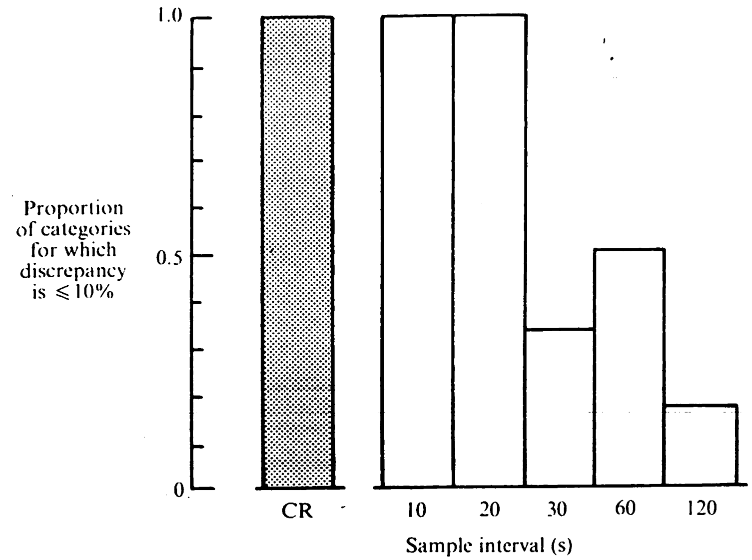


Fig. 6.5. One way of choosing the length of sample interval used in time sampling. First, a pilot sample of behaviour is measured using continuous recording (CR). This provides a true score for each category of behaviour. The score for each category is then calculated as though it had been recorded using time sampling with various different sample intervals (in this example, 10, 20, 30, 60 and 120 s). The percentage difference between the true (CR) score and the time-sampled score is then calculated for each sample interval, for each category of behaviour. The histogram shows the proportion of categories where the discrepancy between the continuous and time-sampled scores is 10% or less. In this example, a sample interval of 20 s would provide a good approximation to continuous recording, whereas longer sample intervals would introduce substantial inaccuracies for several categories. Note that a different criterion (say, 5% or 15%) might be used to define the maximum acceptable discrepancy between the continuous and time-sampled scores. (After Francis, 1966.)

G The advantages and problems of time sampling

As we have pointed out, time-sampling methods are not perfect. Neither instantaneous sampling nor one-zero sampling gives accurate estimates of frequency or duration unless the sample interval is short relative to the average duration of the behaviour pattern (see

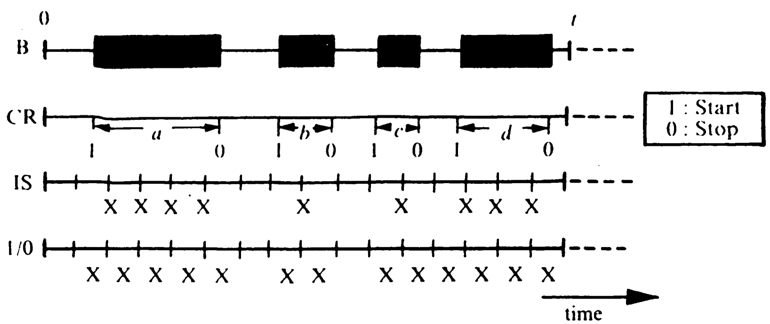


Fig. 6.6. Comparison between *continuous recording* (CR) and the two types of time sampling: *instantaneous sampling* (IS) and *one-zero sampling* (1/0). The black bars on the upper trace represent four successive occurrences of a behaviour pattern (B), during an observation period lasting t units of time and divided into 16 sample intervals. Assuming (arbitrarily) that each sample interval lasted 10 units of time (i.e., $t = 160$ units of time), then the following scores would be given by the various recording rules:
Continuous recording: Total duration $= a + b + c + d = 39 + 16 + 12 + 28 = 95$ units of time. Mean duration $= 95/4 = 23.8$ units of time. Proportion of time spent performing the behaviour pattern $= 95/160 = 0.59$. Frequency $= 4/t = 0.025$ per unit time. Total number of occurrences = 4.
Instantaneous sampling: Score $= 9/16 = 0.56$.
One-zero sampling: Score $= 13/16 = 0.81$.
Note that instantaneous sampling gives a reasonably close approximation to the proportion of time spent performing the behaviour pattern (0.56 versus 0.59) and accurately records that four separate 'bouts' occurred. One-zero sampling considerably over-estimates the proportion of time spent (0.81 versus 0.59) and records only three separate 'bouts', rather than four.

Fig. 6.6) – although analytical techniques are available for deriving unbiased estimates of frequency and duration from time-sampled measures (e.g., Quera, 1990). Moreover, time sampling is not generally suitable for recording sequences of behaviour unless the sample interval is very short. This is because with one-zero sampling two or more behaviour patterns can occur within the same sample period, while instantaneous sampling can miss changes in behaviour that occur between sample points.

Several empirical studies have, however, shown that time sampling can in practice give reasonable estimates of durations or frequencies. For example, by re-analysing data as though they had been acquired using one-zero and instantaneous sampling, Tyler (1979) found that both methods gave accurate approximations to continuous recording, depending on the particular characteristics of the behaviour. One-zero sampling is usually less satisfactory than instantaneous sampling, but actually produces better estimates of frequencies and durations for certain types of behaviour.

A major practical advantage of time sampling is that, by condensing the information recorded and thereby reducing the observer's work load, it allows a larger number of categories to be measured than is possible with continuous recording. This can be an important consideration, especially in a preliminary study where it may be necessary to record a large number of categories. Time sampling also allows a larger number of subjects to be studied, if the individuals in a group are watched cyclically (i.e., using a scan sampling procedure). For example, to record the behaviour of a group of 12 subjects the observer might look at them cyclically, noting the behaviour of each subject (using instantaneous sampling) every 15 s, thereby watching each subject once every 3 min.

Because it is simpler and less demanding than continuous recording, time sampling also tends to be reliable (e.g., Rhine & Linville, 1980). Some types of behaviour occur too rapidly for each occurrence to be recorded, making time sampling a necessity. The practical benefits of time sampling are, of course, achieved at the expense of preserving less information about the behaviour than is the case with continuous recording.

The fact that one-zero sampling does not give unbiased estimates of frequency or duration has led some authorities to argue that it should never be used. We disagree with this view, and not just because one-zero sampling is easy and reliable to use. One-zero sampling is, arguably, the only practicable method for recording a type of behaviour pattern encountered in many studies, namely intermittent behaviour. Here we are referring to behaviour patterns which start and stop repeatedly and rapidly, and last only briefly on each occasion; for example, play behaviour in some species. In such cases, continuous recording or instantaneous sampling are not practicable since it is difficult (or impossible) to record each occurrence of the behaviour or to specify at any one instant whether or not the behaviour is occurring. However, in such cases it usually is possible to state unequivocally whether or not the behaviour has occurred during the preceding sample interval. To give a specific example, in a study of social play in captive rhesus monkeys, White (1977) found that one-zero sampling was the only effective way of recording some categories of play.

A second point is that one-zero scores are, arguably, valid measures of behaviour in their own right, in so far as they provide a meaningful index of the 'amount' of behaviour. One-zero scores are often highly correlated with both frequency and duration measures of the same behaviour (Leger, 1977; Rhine & Linville, 1980), which means that they give a composite measure of 'amount' of behaviour (see Kraemer, 1979a). In contrast, frequency and duration measures of the same behaviour are not always highly correlated with one another. In addition, we see no compelling reason why behaviour patterns should necessarily be described only in terms of frequency or duration (see Rhine & Linville, 1980; Smith, 1985). Indeed, in some cases a one-zero score may actually be a more realistic index than frequency or duration (Slater, 1978; Smith, 1985).

H Identifying individuals

For many studies, being able to identify individuals is essential. Focal sampling, for example, is virtually impossible without a reliable identification method. Moreover, the theoretical interest of a study is usually enhanced when differences in the behaviour of known individuals are recorded. Only by identifying and watching individuals has it become clear that all individuals in a species do not behave in the same 'species-typical' way. On the contrary, marked intraspecific variations in behaviour have been described for many species, and many of these differences make functional sense (see Slater, 1981; Davies, 1982; Trivers, 1985; Krebs & Davies, 1987).

In the laboratory, identification of individuals by rings, tags, tattoo marks, painting the skin or fur, toe-clipping, ear-punching, collars, belts, freeze-branding, fur-clipping and so forth does not usually offer major practical difficulties (Lane-Petter, 1978; Twigg, 1978). However, it is important to realise that marking an individual may alter its behaviour or that of other individuals with which it interacts. To give one example, an experiment showed that coloured plastic leg bands worn by zebra finches affect how attractive they are to members of the opposite sex. Female zebra finches prefer males wearing red leg bands over unbanded males, while males prefer females with black leg

bands. Both males and females avoid members of the opposite sex wearing green or blue leg bands (Burley *et al.*, 1982; see also Trivers, 1985, p. 256). These results clearly show that for zebra finches, and probably many other species, identification markings can have a significant effect on behaviour.

In the field, trapping and marking can present formidable problems. Capturing animals using traps, nets or stupefying drugs can be difficult (Eltringham, 1978) and some forms of marking do not last long under field conditions. Careful thought should be given to minimising the distress caused by marking animals, for scientific as well as ethical reasons.

In addition to visual markings, modern techniques enable animals to be tracked over large distances using miniature radio transmitters or sources of low-level radioactivity attached to the subject (Linn, 1978; Amlaner & Macdonald, 1980; Mech, 1983; Macdonald & Amlaner, 1984; Kenward, 1987). Miniature radio-telemetry systems have been developed that can discriminate automatically between individuals and are suitable for use with small animals, both in the laboratory and in the field (e.g., Berdoy & Evans, 1990).

In some species, individuals have naturally distinctive markings. For example, zebras' stripes are like human finger prints, no two individuals being identical. Similarly, gorillas' noses, elephants' ears, the whisker spots of lions, cheetahs' tails and the bills of Bewick's swans, to list only a few, are highly variable, enabling experienced observers to recognise individuals (e.g., Scott, 1978). Many animals living in the wild acquire distinctive marks through injury, such as torn ears, damaged tails, scars or stiff limbs. Again, features such as these can be used to distinguish one individual from another. Identifying animals by naturally occurring features can be difficult and requires patience and practice (Pennycuick, 1978), but it is the best approach in terms of minimising suffering and disruption.

When, as is often the case, experienced observers are convinced that they can recognise individuals without recourse to written records, some public demonstration of their ability is advisable. One technique is to photograph the animals as they are simultaneously identified by the observer. The tester then removes from each photograph additional environmental cues which might be helpful in identification, and records the identity of each individual. Days or weeks later, the tester presents the observer with the pictures in random order, and asks the observer to name each individual. In one such test, an experienced observer who could individually recognise some 450 adult Bewick's swans identified the individual swan correctly in 29 out of 30 photographs (Bateson, 1977).

I Further reading

Altmann's (1974) review of observational recording methods has been influential, particularly in discouraging the use of one-zero sampling. See also Dunbar (1976), Slater (1978), Kraemer (1979a), Lehner (1987), Donát (1991) and Rogosa & Ghandour (1991). A different view of one-zero sampling is given by Tyler (1979) and Smith (1985), who defend its use on practical and theoretical grounds. Others have also concluded, for a variety of reasons, that one-zero sampling can be an acceptable method for recording certain types of behaviour (e.g., Rhine & Flanigon, 1978; Slater, 1978; Rhine & Linville, 1980; Smith & Connolly, 1980). A comprehensive explanation of recording methods is given by Lehner (1979). For details about the capture and tagging of animals, see Stonehouse (1978), Amlaner & Macdonald (1980) and Kenward (1987).