We shall argue that an economical description of the flow of behaviour in time can be represented as a sequence of decisions, a term which we shall define operationally. Underlying this definition, and more important than the decision concept itself, is the notion that Uncertainty or Predictability is a measurable variable which is fundamental both to the description and the understanding of behaviour. Detailed description of behaviour is a pre-requisite for the analysis of its mechanism. By specifying precisely what the animal does and does not do, we can cut down the number of tenable hypotheses about underlying mechanism. What we are calling the uncertainty or decision structure of behaviour is not only the key to making descriptions economic and manageable, it is also a major feature of behaviour organisation which needs explaining in its own right.

Behaviour itself is a temporal sequence of muscular contractions. A complete account of it as such would be ludicrously lengthy and cumbersome. It would be impossible to publish an account of every small movement made by an animal during an hour's observation period. Therefore, in describing what occurs, the ethologist must detect any repeated patterns in the stream of observed events so that the volume of description can be reduced by induction, that is by the elimination of redundancy from the record. (It is interesting that sensory systems apparently do the same thing and for the same reason—H. Barlow, 1961.) For example, if an observer notices that a posture Q is always followed by another posture U then it will be unnecessary for him to
record every occurrence of U after Q. He need only report on when Q occurred and can subsequently infer from his abbreviated record when U happened. On the other hand, where there is no such pattern or redundancy, for example if behaviour Q is known to be likely to be followed by either B or C or U, then it will be essential for him to report each occurrence of U since it is unpredictable from knowledge of what has occurred before. Thus considerable economy of description can be accomplished by detecting pattern or redundancy in the account of behaviour and not reporting the exact course of frequently repeated patterns every time they occur.

If one is interested in the colour preference of chicks one might record all the pecks at spots of different colours. But pecks occur in rapid bursts to the same colour. Once one peck has been given to, say, the red spot, it is highly likely to be followed by several more. If the bursts were of constant length there would be no need to record pecks after the first one in each burst; they would be redundant. We might say that the 'unit of decision' is not the peck but the burst, and we should do better to count bursts not pecks. However, if bursts are not of constant length, we might also have to record the end of each burst as another 'decision-point'. In fact the situation is more complex still (L. Machlis, pers. comm.) but it illustrates the point that not all instants in the temporal stream of behaviour are of equal importance, since some are to a greater or lesser extent statistically dependent on other instants.

Behaviour, then, is described compactly in terms of a sequence of 'decision points', moments of high information content from which the rest of the behaviour, being redundant, can potentially be reconstructed. This link with the language of information theory (Shannon & Weaver, 1949) gives us the opportunity to make an operational definition of a decision. Information is the reduction in uncertainty. A decision situation may be said to arise when there are alternatives, i.e. there is uncertainty about future behaviour. After a decision has been taken the alternatives are reduced, ideally to one. A decision then is an event which leads to a sudden decrease in the uncertainty of future behaviour. When there is high uncertainty as to what the animal will do next it is said to be in a decision-making state. If this period is followed by a period in which the successive movements are highly predictable, the animal is said to have made its decision and to be carrying it out. The 'Bit' of Information should be a suitable measure of the uncertainty of behaviour, and we have used it in this study.

When a word such as 'decision' which is much used in everyday
language is taken over for an operationally defined scientific usage it is necessary to justify borrowing it. Sometimes only part of the original usage is captured in the more precise definition. Some psychologists have taken over those aspects of the word decision which relate to purpose, goal, or 'pay-off' (e.g. Logan, 1965). We (and Dawkins, 1969) are making operational a different, but equally important aspect of decisions, which is related to predictability and uncertainty.

When we describe a man as 'decisive', we are really saying something about the predictability of his behaviour, rather than about the wisdom or quality of his decisions or what he bases them on. He does not vacillate, starting to do one thing then breaking off and doing something else. Once he has 'decided' what to do, he does it. An observer analysing his behaviour statistically would note long sequences in which the behaviour followed a predictable pattern once it had become clear which of several alternative patterns he had decided upon. In between these predictable sequences there would be moments of high uncertainty when the observer would have no idea which way the man was going to commit himself. These moments or periods of uncertainty would be called times of decision-making.

By contrast, an 'indecisive' man spends all his time in an intermediate state of predictability. He will start one pattern of behaviour, then change his mind half way through and switch to another. The statistically minded observer will not in this case see any clear-cut periods of high predictability interspersed with clear decisions. Instead he will detect a continuous intermediate level of 'decisioniness'.

Similarly if we say "The Chancellor's decision to devalue was not his own, it was economically inevitable", we mean that no uncertainty was reduced by the Chancellor's announcement; any competent economist could have forecast it. If we admit that the precise timing of the announcement was not predictable, so much is the limit of 'decision' accorded to the Chancellor. "He has just decided to marry her" means that the probability of the marriage has just increased markedly, and also presupposes that the decision is his to make, i.e. her behaviour in this matter is wholly predictable. "He never knew exactly when he decided to emigrate" means that the probability of his emigrating increased gradually, and did not dramatically rise at any time.

This usage of the term decision will not appeal to everyone, and we do not wish to overstate it. In particular we most emphatically do not want to be taken to imply that there is an all or none distinction between redundant predictable points in time, and information-rich decision-
points. There will doubtless be a continuous gradation. The quantitative extent to which any particular item of behaviour should be regarded as a ‘decision-point’ must be measured empirically.

Methods for doing this are already available in the form of the well developed statistical techniques which ethologists have used for detecting sequential and temporal patterns in behaviour (e.g. BAERENDS et al., 1955; NELSON, 1964; DELIUS, 1969; FENTRESS, 1972). These methods may have other defects, for instance, they may be as unable to cope with the complexity of some behaviour as with that of some aspects of language (MARSHALL; VOWLES, 1970), but for our present purposes we would like to rectify two omissions.

The first omission is a neglect of short term or moment to moment patterns of posture. Statistics are normally applied to the relations between already identified action patterns, given names like ‘tail-beating’, ‘wing-raising’ or ‘post-copulatory steaming’, over a relatively large time scale. The recognition of these action patterns has usually been subjective (DANE et al., 1959; G. BARLOW, 1968). But what they really are of course is frequently recurring sequences of movements on a finer time scale, which should be picked out by the very same statistical techniques as are used to detect patterns between them. In this case the raw data for the analysis would be single muscular contractions, or measurements from single frames of film. (For a stimulating discussion of this in relation to the Fixed Action Pattern concept, see BARLOW, 1968). The end product of this analysis would then be the raw data for the type of higher order analysis which is conventionally done. When the ethologist subjectively recognises his animal’s action patterns he must be performing a similar statistical analysis, but using the ‘computer in the head’. This is probably because the human brain is relatively good at short-term temporal pattern recognition, a fact exploited by HASS and EIBL-EIBESFELDT (HASS, 1970) who speeded up film of the movements of people and were then able to use the naked eye to detect longer-term patterns totally missed when the film was played at normal speed. However, it might also be that there really is more pattern at the short term level. Whatever the reason, it would seem desirable to use the same methods to detect patterns on both time scales, providing a comprehensive picture of any regularities that might exist. Slowing down film or videotape makes it possible to apply objective publicly verifiable statistical methods both to the long term relations between behaviour patterns and to the recognition of the internal relations between successive postures making up these patterns.
We should not stop at distinguishing only two different time scales of decision analysis. If short-term patterns tend to cluster in time with other, perhaps functionally related, short-term patterns (e.g. WIEPKEMA, 1961), the initiation of a new cluster would involve a higher order of 'decision' than the initiation of a new short-term pattern within a cluster which is already under way. The subsidiary within-cluster decisions would be partly predictable from prior knowledge of what cluster the animal is already in. We must be prepared for a multi-level hierarchy of decisions corresponding to the hierarchical schemes of KÖRTLANDT (1955), BAERENDS and TINBERGEN (TINBERGEN, 1951) and MARSHALL and VOWLES (VOWLES, 1970).

However, although the decision-structure concept is equally applicable to higher as well as lower levels on the hierarchy, we have in this study concentrated on a lower level and have analysed the internal structure of a single 'action pattern', drinking in chicks. This is because the lower levels have been relatively neglected by statistical ethologists.

The second omission from the statistical analyses of behaviour which have been done so far is that there is insufficient emphasis on the degree of pattern which exists, as a phenomenon in its own right. Techniques such as calculating transition probabilities give an idea of the type of relation which exists between one behaviour and another, but tend not to reveal that certain parts of the behaviour sequence may have more pattern or predictability than other parts (but see ALTMANN, 1965; FENTRESS, 1972). The degree of pattern is an important attribute of behaviour since it is one more feature which neurophysiology will have to explain. Moreover, there would seem to be possible functional reasons why the predictability structure of behaviour might be important. If one were trying to design an efficient and economical animal one might well programme a limited repertoire of discrete stereotyped behaviour subroutines, rather than the capacity for infinite variation which might raise more formidable problems of nervous control.

The need in neurophysiology for a measure of predictability or stereotypy of behaviour is illustrated by BULLOCK (1961). He argues that each behaviour pattern must have a single 'decision' unit somewhere in the nervous system (possibly even a single neurone) which lies at the head of a hierarchy of motor command, and on which a hierarchy of afferent inputs converge. This is an important piece of logic for neurophysiology, but as BULLOCK rightly stresses, it applies only to behaviour which occurs in an all or none fashion, "a large class of behavior, including the 'take-off' of a fly, the onset of instinctive acts .......". But we do not
really know how large this class of behaviour actually is, nor do we have any agreed methods for classifying behaviours for this purpose. It would seem that here is an area in which ethology could contribute something useful to neurophysiology.

Our aims in this paper then are two-fold. Firstly we are trying to show how the statistical methods which are normally used to detect temporal patterning between already recognised behavioural acts, can also be used on a finer time scale to detect patterns within those acts. Moreover this is how those acts are recognised and defined anyway. Secondly we are arguing that among the attributes of behaviour which are worth measuring is one which may be called 'decisioniness' or 'information-content' or 'uncertainty'. We are presenting some data at this rather preliminary stage for purposes of illustration only.

PROCEDURE

Six 23-day old chicks (Cobbs) were placed individually in a long narrow filming cage (90 x 10 x 30 cms high). The front wall of the cage was made of perspex with a 1 cm grid so that the height of the bird's head could be measured. The cage was sufficiently narrow that the chick had to stand broadside on to the camera and could not turn round. This did not apparently interfere with its drinking behaviour as thirsty birds tend to remain in one place while they drink. In the centre of the cage was a glass dish containing water, and the camera was focused on this. The behaviour of the chick was recorded from the time it approached this dish until it left the picture.

Water but not food had been removed from the chicks' living cage 16 hours before these observations and all the chicks had previously been made familiar with the filming box.

Behaviour was recorded on a Sony CV 2100 ACE videotape recorder using two Sony T.V. cameras. One camera was directed at the bird, the other at an Advance TC 9A Timer-Counter functioning as a digital clock. By means of a Sony CMW 110 CE Wiper, the two pictures were 'mixed' onto the same tape. This meant that when viewed, each frame of the videotape had a unique time label at the bottom of the split screen. This was very important for getting information off the tape. Any frame could be quickly reidentified for checking or if the analysis had been broken off.

One measure of the behaviour was selected from the videotape record for fine analysis. This was the height of the bird's eye as measured from the grid. The videotapes were analysed frame by frame by manually moving the tape through the recorder and reading off the heights and
the digital clock time when the height changed. We recorded the times to the nearest 20th of a second. In the rest of this paper we shall use the word 'frame' to refer to these 20th second time slices, rather than to the physical frames on the videotape itself. Our analysis is based on some 50,000 of these 'frames'.

The data were punched on computer tape, and analysed on a PDP 8 L small computer. The computer was used not only as a time-saving convenience, but also in the interests of objectivity. Since one of our main purposes was to make statistically explicit the detection of pattern in the data, we wanted to reduce the role of the human eye as much as possible. Once we had used the human eye to read off the heights of the bird's head in relation to the water, and the times, no more features or patterns were recognised which could not be specified in an algorithm for the computer. Of course the human eye comes in again in the evaluation of the graphs which the computer generated, but this evaluation is public and available to anyone who reads the paper. The reader does not have to trust the authors' powers of observation.

**ANALYSIS**

As seen in a graph of eye height against time (Fig. 1), a typical drink has a rather rapid downstroke, punctuated by one or more short pauses, then a period when the bill is in the water, followed by a smoothly decelerating upstroke. There is then a longer pause before the start of the next drink. We shall begin by treating each whole drink as a unit and looking for long term patterns as shown by the intervals between drinks. We will then dissect the drink into its constituent parts, ending with single frames, the smallest unit available to us.

![Fig. 1. Superimposed graphs of the eye height during three successive drinks, lined up on the moment when the bill strikes the water (time 0).](image-url)
For the first gross analysis we take one clearly identifiable moment, the moment when the bill strikes the water, to stand for each drink. Fig. 2 shows that there is a tendency for drinks to occur rhythmically and that this rhythmic process is fairly stationary over time. It is a plot of the frequency of intervals not merely between successive drinks, but between every drink and every other one. This type of plot has been variously called an autocorrelation, renewal density, or intensity function (Moore et al., 1966; Burns, 1968; Delius, 1969). Roughly speaking it is a plot of the probability (strictly the raw frequency), given that a drink has occurred at Time 0, that another drink will occur after various time lags. With a little imagination it can be seen as a diagram of the reverberating after-effects of a drink. This graph shows more than a simple frequency distribution of successive interdrink intervals would. It shows that if you know the time of occurrence of any one drink, you can predict not only the time of occurrence of the next drink, but also that of the next but 12 and all intermediates. In decision terms this means that to some small extent the decision when to drink may have already been taken some 12 drinks or 60 seconds earlier. However it is important to stress that these 'predictions' are only statistical. Any individual prediction of this nature would very probably be wrong. Much of the appearance of smoothness in the graph is achieved by the dubious technique of sliding averages. All we can say is that our uncertainty about exactly when the next 12 drinks will occur is fractionally reduced by knowledge of the time of the present drink. There is still a large amount of uncertainty. When we start to dissect the drink in more detail we shall want to know whether there are some parts of the whole sequence which are relatively more uncertain than others.

![Graph of 'Renewal density' plot](image)

Fig. 2. 'Renewal density' plot showing the number of drinks occurring at various time intervals after drink i where i is in turn all drinks. The results from all 6 chicks are summed. The intervals were classed in 1/4 second units and the figures graphed are sliding averages for successive groups of 5 units.
Figure 1 shows the form of 3 successive drinks as expressed in the one measure of eye height plotted against time. Using the moment when the bill strikes the water as a synchronisation point, the three graphs have been superimposed to give a visual impression of the relative variability of the different parts of the drink. The impression gained is that the downstroke is more variable than either the upstroke or the phase at the bottom when the bill is in the water. Later we shall show that there is some quantitative basis for this belief.

For analysis purposes, we defined a set of four classes of 'landmark points' by means of which the record could be divided into phases objectively. The landmark points were:

D. Start of Downstroke. A sequence of 4 frames in which the eye height in the 4th was at least 3 cm lower than in the first, and where the height in the first was at least 10 cm above the ground.

W. Bill strikes Water. A frame in which the bill tip was below the water level, given that it was above the water level in the preceding frame.

O. Bill comes Out of water. A frame in which the bill tip was above the water level, given that in the previous frame it was below water level.

U. End of Upstroke. A sequence of 4 frames in which the eye-heights were all the same and at least 10 cm high, given that the previous frame was lower, and that the following frame was not higher.

In addition, all the definitions contain the rule that after one type of landmark has been recognised, that same one cannot be recognised again until one of the other 3 has intervened, i.e. there cannot be runs of the same landmark: obviously any one Downstroke can only start once, and so on. When the definition refers to more than one frame, the time of the landmark was taken to be the time of the last frame referred to.

If we ignore time, and look only at sequences of these ‘landmarks’, the behaviour is very determined. Table 1 shows the frequency with which each landmark is followed immediately by each other one. The data are summed for all chicks.

<table>
<thead>
<tr>
<th>Preceding</th>
<th>Following</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D W O U</td>
</tr>
<tr>
<td>D</td>
<td>× 250 o</td>
</tr>
<tr>
<td>W</td>
<td>o × 259 o</td>
</tr>
<tr>
<td>O</td>
<td>o o × 258</td>
</tr>
<tr>
<td>U</td>
<td>292 9 o</td>
</tr>
</tbody>
</table>

TABLE 1
Table 2 shows 5th order transition frequencies, *i.e.* the frequency of occurrence of sextuplets, again summed for all chicks. Only those sextuplets beginning with D (start of Downstroke) are shown. Of the possible sextuplets in this category, only 6 different types occur, and the vast majority (209 out of 290) are the sextuplet DWoudw, the normal drinking sequence.

<table>
<thead>
<tr>
<th>Sextuplet</th>
<th>Frequency</th>
<th>Expected Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWoudw</td>
<td>209</td>
<td>292.1</td>
</tr>
<tr>
<td>DWoudu</td>
<td>35</td>
<td>36.4</td>
</tr>
<tr>
<td>DUDWOU</td>
<td>30</td>
<td>36.4</td>
</tr>
<tr>
<td>DUWoud</td>
<td>9</td>
<td>1.3</td>
</tr>
<tr>
<td>DUDUDW</td>
<td>5</td>
<td>5.4</td>
</tr>
<tr>
<td>DUDUDU</td>
<td>2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

The expected frequencies on the right are calculated on the basis of the doublet frequencies in Table 1. The observed distribution as a whole does not differ significantly from expected (p > 0.05, Kolmogorov-Smirnov test, SIEGEL, 1956). However the sextuplet DUWoud appears to occur more frequently than would be expected on the basis of the doublet transitions which it contains. Examination of triplet frequencies shows that it is the triplet DUW which accounts for this high result. The doublet UW is a rare occurrence in which an upstroke is followed by the bill entering the water without a detected downstroke. Obviously there must be a downstroke of sorts, but it is too gradual to be recognised by the computer. Of the 9 occasions when this ‘slow downstroke’ happened, all 9 were immediately preceded by a DU doublet. This is also rather rare, representing an unfinished drink, in which the downstroke is broken off prematurely before the bill strikes the water. It appears that these two rather anomalous movements, the ‘aborted downstroke’ and the ‘slow downstroke’ are disproportionately likely to occur in sequence together.

These ‘aborted downstrokes’ are interesting from the decision point of view. It appeared that there might be certain favoured points in the downstroke where decisions are taken whether to continue with the drink, or break it off, rather analogous to the specific points in the count-down before launching a space rocket when the decision whether to continue or abort the mission can be revised. We looked to see whether there was any tendency for reversals in downstrokes to occur at particular favoured eye-heights. We defined as a ‘Minimum’ any frame or sequence of frames of constant eye height, in which the height was lower than that of both
the preceding and following frames. We then counted the frequency with which various eye heights occurred in minimum frames. The results are shown for all 6 chicks in Figure 3. A normal drink would have its minimum at 3 or 4 cm, the height of the eye when the bill was in the water. Not surprisingly the bulk of the distributions in Figure 3 fall at this level. A temporary reversal of direction, or a full scale 'aborted mission' would show up as a minimum at a higher level. If there were certain favoured decision-points at some specified level on the way down, this should show up as a bulge at that height in Figure 3. There is indeed a suggestion of such an effect, at least for those chicks with the most data. Chick 1 for example is rather likely to complete each Downstroke with a real drink,

**TABLE 3**

*Variability of different phases of the drink, expressed as the Coefficients of Variation (sd/mean)*

<table>
<thead>
<tr>
<th>Chick</th>
<th>No. of Drinks</th>
<th>Interdrink</th>
<th>Total Drink</th>
<th>Downstroke</th>
<th>In-water</th>
<th>Upstroke</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53</td>
<td>0.66</td>
<td>0.34</td>
<td>0.86</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td>2</td>
<td>43</td>
<td>0.44</td>
<td>0.16</td>
<td>0.35</td>
<td>0.33</td>
<td>0.27</td>
</tr>
<tr>
<td>3</td>
<td>51</td>
<td>0.94</td>
<td>0.60</td>
<td>0.95</td>
<td>0.58</td>
<td>0.85</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>0.82</td>
<td>0.25</td>
<td>0.49</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>5</td>
<td>43</td>
<td>0.83</td>
<td>0.29</td>
<td>0.40</td>
<td>0.51</td>
<td>0.40</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>0.88</td>
<td>0.23</td>
<td>0.45</td>
<td>0.34</td>
<td>0.42</td>
</tr>
</tbody>
</table>

The differences between Interdrink and Total Drink are significant for each individual chick (p < 0.01, using the value \( c = (CV_1 - CV_2)/\sqrt{Scv_1^2 + Scv_2^2} \). The probability associated with this value is obtained from Table III of Fisher and Yates (1963)). Differences between downstroke and upstroke are significant only for Chick 1 (p < 0.01).
but there is a distinct possibility of a reversal or hesitation at around 10 cm. For Chick 1 the 10 cm area might be called a decision zone. If it gets past this to down around 5 or 6 cm it would appear to have left the decision zone, and be fully committed to a complete drink.

Sequence analyses of the sort described, of however high order, are rather crude as they do not consider time. One way to bring time in is to measure intervals between landmarks, or in other words, durations of the various phases into which the behaviour can be divided.

Five phases were named as the periods between pairs of landmarks as follows:

D to W: Downstroke
W to O: In-water
O to U: Upstroke
U to D: Interdrink  (W to W — the period used for Figure 2)
D to U: Total drink

Figure 4 shows for one chick the frequency distribution of periods of various durations, for all the 5 classes of period. Interdrink duration is far more variable than the durations of total drinks themselves, and this is not merely because they are absolutely larger, as the figures for Coefficient of Variation in Table 3 show. That is to say it is easier to predict when a drink will end given that you know when it started, than it is to predict when a drink will start, given that you know when the previous one ended. But there is still some uncertainty within drinks, and our impression from Figure 1 and from the minimum analysis was

![Fig. 4. Frequency distributions of durations of phases for Chick 6.](image-url)
that the greater part of this is contributed by the Downstroke. Figure 4 and Table 3 seem to confirm this. Furthermore for all 6 chicks total drink duration is more strongly correlated (Spearman rank) with downstroke duration than with the duration of either the in-water phase or the upstroke. This is not due to the greater absolute duration of the downstroke, because if the other 2 phases are added together to form an ‘end phase’ (which is longer on average than the corresponding downstroke), the correlation of end-phase duration with total drink duration is less than that of downstroke with total drink duration.

The total drink then is a unit of decision in the sense that its beginning is more uncertain than its end—i.e. there is a decision to be made when to start a drink, but the end of the drink is relatively determined by the beginning. But this determination is not absolute: there are some decisions taken during the course of a drink, especially during the downstroke. Some of this uncertainty is about whether the drink will ‘abort’, but this is rather unlikely in our tests, so the uncertainty is not great. Given that this does not happen, there is still some uncertainty about exactly when the downstroke will end. It is these decisions about the timing of the downstroke which largely determine the duration of the whole drink. We need to do some further dissection to find out more exactly when the downstroke decisions are taken, for example whether the greatest uncertainty attaches to how many pauses there are (decision to begin a pause) or to how long pauses last (decision to end a pause).

Rather than continue the dissection by defining yet more rather arbitrary ‘landmarks’, such as ‘start of pause’ and ‘end of pause’, we have switched to an attempt to measure uncertainty continuously, or at least single frame by frame, and dispense with defined landmarks altogether. Ideally, following the arguments in the introduction, natural landmarks should emerge from this as the delimiters of highly determined sequences of frames. This has not convincingly happened in our preliminary analysis, and there are grounds for some dissatisfaction. Thus there are many possible methods of assessing the uncertainty, and an element of arbitrariness comparable to that involved in the defining of landmarks, enters into the selection of a precise method. However the results are presented for what they are worth.

The general principle of the method is to associate with each frame a set of prior information, a collection of facts about that frame and preceding ones. One has to set an arbitrary limit to the kind of prior information used, and to how far back in time one goes for it. Then, on a second pass through the data, the computer ‘looks’ at each individual
frame and 'attempts to predict' the probable outcome of the frame on the basis of its 'experience' of the outcomes of all those frames with the same sort of prior information. An uncertainty value for each prediction is assessed in Bits of Information, and this is the index which is finally plotted against time. The nature of the prediction attempted, and how far ahead it looks in time, are also arbitrarily chosen.

In detail the prior information associated with each frame was:

1. The height of the eye in the frame. Each frame was placed in a 2 cm eye-height class.
2. How long the eye had been at that height up to the present frame. The time since the eye arrived at that height was expressed logarithmically, i.e. the results were placed into time classes which were equal on a log scale, rounded off to the nearest integral number of frames.
3. Whether it arrived at the present height from a previously lower or higher position.

Every frame was thus placed on these criteria into a class with (usually) many other frames. For example in one frame from Chick 3, the eye height was 4 cm, this was the 9th frame in a row at that height, and the previous height before this row was higher. The class into which this frame fell was that of all frames with eye heights of 3 or 4 cm, where the time since the eye first arrived at that height was either 8 or 9 frames, and with the previous eye position higher. There were 116 frames in this class in the data from Chick 3.

On the second pass through the data the specific 'prediction' made at each frame was about the mean height of the next 6 frames. The possibilities were reduced to 3: that this future mean height would be higher, lower or the same as the present height, again working in 2 cm height classes. The probability of each of the 3 outcomes, \( p_1 \), \( p_2 \) and \( p_3 \), given the prior information associated with the present frame, was calculated (see Figure 5). The uncertainty value of the frame was then calculated from Shannon's formula \( \sum p_j \log_2 (\frac{1}{p_j}) \) where \( j \) is successively 1, 2 and 3. For example, of the 116 frames in the particular class mentioned, 5 were followed by an average downward trend in the next 6 frames, 65 by an upward trend, and in the remaining 46 cases the mean height of the next 6 frames stayed the same. The 3 respective probabilities therefore were 0.0431, 0.5603 and 0.3966. The sum of \( p \log_2 (\frac{1}{p}) \) for all these 3 p's is 1.1930. This is our measure in Bits of the average uncertainty of outcome of frames in the class of the frame under consideration. We call it the uncertainty value of the frame. The minimum
possible uncertainty for a frame is 0 Bits: its outcome is certain. The maximum possible uncertainty in our framework of 3 possible outcomes is \( \log_2 3 = 1.585 \) Bits; here all 3 outcomes are equally probable, and prediction is at its most difficult.

A highly stereotyped behaviour pattern should appear as a sequence of frames all with low uncertainty. Following the arguments in the introduction, a Decision should appear as a precipitous drop in the graph of uncertainty against time, followed by a sequence of low uncertainty frames. Long sequences of medium uncertainty frames will indicate that as far as our methods have been able to show, the animal does not organise its behaviour by means of clear cut major decisions; rather it governs its behaviour by a continuous series of small decisions.

![Diagram](image)

**Fig. 5.** Diagrammatic representation of changes in the computer's predictions over time during one particular drink. The direction in which the arrows point represents the direction in which it is predicted the eye will move; the length of the arrows represents the probability associated with each of the three outcomes. For example, for most of the frames of the upstroke, the probability is 1.0 that the average eye height in the next 6 frames will be higher.

One way to present the data would have been to tabulate systematically for all possible combinations of antecedent conditions, the associated uncertainty. However, as explained, we have instead made a second pass through the frames, 'looking up' in such a table the uncertainty of the individual frames as they actually occur. This enables us to plot an actual graph of measured uncertainty against time. However it means we can only plot a very small portion of the results, the total number of frames being so large. The graphs therefore refer to only one or two individual drinks. However the antecedent information which was used to plot them is summed for all drinks and all chicks, and in this sense the figures represent a large amount of data, unlike the example above where the antecedent information was taken from Chick 3 only.
Figure 6 is a plot of uncertainty against time for 2 consecutive drinks from Chick 2, on the same time scale as the corresponding eye height graph. It might be interpreted roughly as follows. The downstrokes are relatively uncertain. At the end of the downstrokes the uncertainty drops briefly, as it is relatively certain the bill will stay in the water for at least the next 6 frames. However, uncertainty about exactly when the upstroke will start increases steadily. There is another drop in uncertainty as soon as the upstroke actually does start, and the whole course of the upstroke is relatively certain. There is some uncertainty however about exactly where the upstroke will end. About 1 second into the interdrink period the uncertainty has dropped to zero—it is certain that nothing new will happen in the near future. This period of certainty only lasts about ½ second however, and uncertainty increases concerning when the next downstroke will start.

![Fig. 6. Part of a continuous record showing variations in eye height (line) and uncertainty (histogram) over time for two successive drinks from Chick 2. The uncertainty shown here is calculated on the basis of antecedent information summed from all 6 chicks.](image)

Figure 7 presents the same kind of data in a different way. Eye height is plotted against time for 4 successive drinks. The thick black lines represent frames or sequences of frames whose uncertainty was lower than an arbitrary threshold value of 0.4 Bits. Ideally this should be a way of picking out natural 'features' of behaviour, frequently recurring sequences of posture, each initiated by a decision. These are in fact only moderately convincing, and it seems that, at least according to the sort of measures we have used, drinking in chicks does not have a very rigid decision structure. However there is some justification for calling the upstroke a natural unit of behaviour.
Fig. 7. Graphs of eye height against time for 4 successive drinks from Chick 2, showing (thickened line) where uncertainty falls below an arbitrary threshold of 0.4 Bits. Antecedent information summed from all 6 chicks.

DISCUSSION

It is obvious that the amount of uncertainty which is measured is a function of how efficient the methods of prediction are. A period of high uncertainty might completely disappear if more antecedent information were taken into account in calculating uncertainty. For example, it might prove that just before starting the downstroke of a drink, chicks always close their eyes. If we included whether the eye was open or closed as part of the antecedent information, the graph of uncertainty against time might have looked very different. It is therefore essential to make further measurements of different aspects of the birds' behaviour besides the one variable, the height of the eye, which we have described so far. But however fine the detail with which we describe the behaviour, it must be emphasised that the resulting measure of uncertainty is an operational measure. It is the uncertainty which an observer in possession of a particular set of facts has about what an animal will do next. Our hope is that if we make an honest, thorough attempt to predict behaviour from moment to moment, there will be some variation in the uncertainty over time which is not wholly attributable to artifact: in other words, some periods of behaviour really are more unpredictable than others and this reflects real uncertainty in the animal, not only uncertainty in the observer.

What this 'real' uncertainty in the animal might be due to we do not know. One must always be prepared to accept the most economical
hypothesis that it is merely due to factors too complex for the observer to disentangle. However the possibility that nervous systems may be subject to more fundamental physical uncertainties (Pringle, 1963; but see Bullock, 1970) is an exciting and philosophically reassuring one. We have not so far investigated the role of external stimuli in the control of drinking and for this reason have refrained from applying the term Fixed Action Pattern either to a whole drink or to any part of the drink (Hinde, 1970). However, although finding that certain sequences of behaviour are highly predictable does not in itself tell us the nature of their control (Moltz, 1965), stereotypy of parts of the drinking is a phenomenon of interest in its own right. What it means is that despite the summation of ‘antecedent information’ from 6 different chicks, and despite variations in both internal (e.g. amount of water in the crop) and external conditions (e.g. extraneous noise), certain parts of the drink were always much the same. They seemed to be fairly resistant to variations in the prevailing stimulus conditions. An interesting next step would be systematically to manipulate environmental disturbance and internal state to see whether the effects of these are concentrated on the relatively more variable parts of the sequence which we have called the decision times. For instance, a distracting loud noise might have an immediate effect if applied during a downstroke, but might be ignored during an upstroke. Fixity in the face of a wide range of experimental disruption would seem to need a rather different sort of explanation from that needed for parts which are easily disrupted (c.f. Russell et al., 1954).

The main question which we have asked is ‘When are decisions taken?’ However the decision concept as used here could be the basis of other kinds of question. ‘Who takes the decisions?’ might be a meaningful question in social contexts. In an analysis of sequences of behaviour involving a courting pair, if the next action of the female is highly dependent on the preceding action of the male, but the male’s behaviour is difficult to predict from that of the female, we would say that the male is taking most of the decisions.

‘Where is the decision taken?’ is a neurophysiological question which in some cases might be made meaningful in terms of measured uncertainties. Marler (1961) has discussed the location of stimulus ‘filters’. He suggests that peripheral filtering is involved in the responsiveness of male silkworm moths to female sex-pheromone, because electrodes placed

---

in the antennal nerve record activity only when the correct substance is present in the air. We would say that the decision is taken in the antenna if it is possible to predict whether the male will behave sexually given that we have information only from the electrode in the antenna. It could be that whether the male actually responds with sexual behaviour to the pheromone depends partly on other variables which might be unknown, or which might be measurable. In this case the quantitative 'decisioniness' which can be attributed to the antenna is measured by the reduction in our prior uncertainty about whether the male will show sexual behaviour, which is obtained if one takes into account information about activity in the antennal nerve. If it were possible to record from several points along the chain from receptor to effector one could measure how much uncertainty about overt behaviour is reduced at each point compared with the preceding one.

We claim no more for this study than that it is a preliminary attempt to make sense of the large amounts of data which result from any analysis of the fine structure of behaviour with film or videotape. Accurate and complete description must precede attempts to generate models of underlying mechanism since neurophysiologists must know what has to be explained. We hope to have suggested that this 'ethological spadework' may be most economically and fruitfully done by including a measure of the decision structure of the behaviour.

SUMMARY

Behaviour can be described in terms of its changing uncertainty or decision structure over time. Such a description is economical, maximally informative and may well be of importance neurophysiologically. We try to show that the methods which are normally used to detect temporal patterning between already recognized behavioural acts can also be used on a finer time scale to detect moment to moment patterns of posture within those acts. From such analyses, it is possible to calculate the 'decisioniness' or 'uncertainty' of different parts of the behavioural sequence. We illustrate this by an attempt to describe the decision structure of the drinking behaviour of domestic chicks, using frame by frame videotape analysis. For example, it appears that the first downstroke phase of each drink is more uncertain as to outcome than the other phases, suggesting that 'decisions' are taken during the downstroke. We end with an attempt to plot a continuous graph of behaviour uncertainty against time sampled at 50 msec intervals.

REFERENCES


WIEPKEMA, P. R. (1961). An ethological analysis of the reproductive behaviour of

ZUSAMMENFASSUNG
Verhaltensabläufe können als Zeitfunktionen ihrer Ungewißheit oder Entschei-
dungsstruktur betrachtet werden. Diese Art der Beschreibung ist ökonomisch,
maximal informativ und wahrscheinlich auch neurophysiologisch von Bedeutung.
Wir wollen zeigen, daß statistische Methoden, die normalerweise angewandt werden
um die zeitliche Organisation von Verhaltensabläufen zu beschreiben, mit einer
feineren Zeitskala auch dazu benutzt werden können, den Ablauf von Einzel-
handlungen zu erkennen. Von solchen Analysen ausgehend kann man ein Entschei-
dungs- oder Ungewißheitsmaß für die verschiedenen Phasen einer Verhaltenssequenz
berechnen. Als Beispiel versuchen wir, die Entscheidungsstruktur des Trinkver-
verhaltens von Hühnerküken zu beschreiben unter Benutzung einer Einzelbildana-
lyse von Video-Magnetbandaufnahmen. Die erste Abwärtsbewegung eines jeden
Trinkaktes scheint, bezüglich der ihr folgenden Bewegungen, ein größeres
Ungewißheitsmaß zu besitzen als alle anderen Phasen des Trinkaktes. Dieses deutet
an, daß Entscheidungen hauptsächlich während der Abwärtsphase getroffen
werden. Zuletzt zeichneten wir versuchsweise eine kontinuierliche, zeitabhängige
Verhaltensungewißheitsfunktion dieses Trinkverhaltens mit einer 50 msec. Auf-
lösung auf.