data which encourage the belief that we can distinguish between the "transmitter" and "translator" categories by means of such attitude profiles, and also by another set of profiles between those who originate new codes, and those who do not. Unfortunately neither the attitude profiles nor the experimental analysis has been finished to the point where these may be presented as definite conclusions; they are discussed here because of the importance we feel is attached to this type of measurement.

Summary.

A guiding principle of our work has been to attempt a mathematical analysis, of simple group situations, and, following this, to attempt the construction of mathematical models that will reproduce the performance of the group from a small number of parameters governing the behavior of individual group members. The effort has been at least partially successful in some simple cases of information handling. This paper has presented some results of preliminary studies of situations in which the handling of information is more complex—the reduction of errors due to noise and ambiguity and the problem of filtering.

In addition, we have presented at the close of this paper the outlines of the first preliminary attempts to deal with the problem of individual differences. Although the results to date are few and tentative, they present the beginning of a method that may make it possible to gain sufficient knowledge of the properties of individuals to generate the behavioral probabilities needed for any predictive mathematical model. In addition, they open the way toward schemes of selection of individuals for service in a particular group that will make possible given types of performance. Some work has already been done in this area, and the results are promising (21).

The fact that there are relevant individual differences contributes to the variability of the results of experimental work with groups and this has necessitated that a large number of groups be run in each experiment in order to obtain results by treating this variability as sampling variation. It would be desirable to provide for the individual differences in the model of the group process rather than in the statistical treatment of the data. In real-life situations, one of

the major factors in the behavior of any specific group is the character of the individuals composing it. In the long run, any realistic theory of groups must take this into account.

IV. SOME ASPECTS OF TIME AND DECISIONS.*

Introduction.

During the past fifteen years certain theoretical and mathematical investigations have been directed toward a better understanding of some aspects of human behavior; much of this attention has been focused upon decision making and the closely related phenomena of information processing. The most interesting theories to have been developed, or extended, are utility theory, information theory, game theory, statistical decision theory, and the programming models. It is of some interest, and we believe importance, that the central concern of these studies has been with what may loosely be called the structural aspects of decision processes and that there has been a marked avoidance of the temporal aspects. For example, in utility theory there is no concept of time at all, while in the other theories mentioned there is rarely introduced more than a notion of temporal ordering.

The problem we raise—time—must certainly have been one of the major stumbling blocks prior to creation of these theories, and one which had to be put aside to develop them. One must emphasize the genius behind these theories that rejected the obvious importance of time and of the timing of decisions in ordinary experience and went on to develop essentially timeless theories that do illuminate major aspects of communication and decision making. The brilliance of the strategy concept in game theory and of the timeless notion of noise in information theory should not be dulled by familiarity. Since their existence forcefully demonstrates that complicated and fruitful theories can be created without a stronger notion of time than temporal ordering, there can be little sympathy for the person who now argues that time must be reintroduced. The burden of proof rests upon his shoulders.

* By R. Duncan Luce.

The need is most sharply demonstrated by reminding the reader of one or two of the well known problems that seem to elude solution in our present theories. First, information theory demands that the output of a source be of a stationary statistical character, and so, whenever the source is a human being passing through a learning phase, we must quietly turn our attention to other issues until the learning is completed. To fuse information theory and some aspects of learning theory, a more time-dependent information theory appears to be needed. As a second example, many would agree that the central failure of game theory to date is its apparent inability to formulate and to solve the dynamics of the bargaining problem and the closely related phenomenon of coalition formation in the n-person theory. To my mind the difficulty centers about the timing of decisions, and game theory as currently formulated, while not completely excluding timing problems, seems to be best suited to other problems.

With these examples as an indication of a need to introduce some timing considerations, let us re-examine how it has been possible to leave out time at all. The desirability of omitting time is fairly apparent; the real point is how has it been done. There is little object in examining the special details of each case; rather we shall simply deal loosely with "decisions." For a single decision, or for a set of decisions in prescribed order, one may think of the problem as consisting of a fixed set of alternative choices, a fixed set of possible states of affairs or nature in some appropriate sense, a payoff matrix whose rows and columns are the two sets just mentioned, and information arriving sequentially that alters the *a priori* probability distribution over the states of nature. While in any specific problem the arrival of information has a particular time structure, it is generally assumed that only the order of time of arrival is pertinent to the decision made, and in many problems, though by no means all, it is assumed to arrive in one "package." Many psychologists would claim that it is still necessary to consider the order in which the decision maker considers the contents of the package, for it appears that, on the basis of slim and incomplete evidence, human beings generate hypotheses that are not readily dislodged even when they prove to be in direct contradiction to information received or examined after their formulation. Such a remark is only important for descriptive theories and not for normative ones.

Such then is the structure of a single decision, or of a prescribed arrangement of decisions, and there does not appear to be any need to introduce time except if one would care to know how long the decision process takes. But what about the structure of a problem that rests upon a complex of intimately interrelated decisions? In any group problems in which a decision depends on information available and that, in turn, depends both on the decisions of other people and when they were made, timing quite obviously influences the over-all behavior of the group. The approach to this problem has generally been to "divide and conquer," with most of the emphasis on dividing and relatively little on conquering. The division is executed by introducing certain plausible, though not generally completely valid, assumptions, which are not always made explicit but which can be detected by examining the procedures of analysis. The set employed in the analysis of the experimental material we shall discuss is one which is often used and which leads to a fairly conventional breakdown into disciplines. These assumptions are:

(1) An individual decision depends statistically on the information available at the time it was made and on a fixed payoff function, which very often is culturally determined and is treated as a characteristic of the human being.

(2) The information present at any point in the complex at a given time depends on the time and the nature of certain decisions.

(3) The time taken to make an individual decision is governed by a distribution, stimulated by one or more prior decisions, whose form and parameters depend both on the individual and in some fashion on the information available to him at the time the decision is reached.

In conventional terminology the study of phenomena covered by the first assumption is that of decision processes, by the last is that of reaction times, and by the second is that of the construction of analytic models of information flow and the empirical study of gross group behavior. It is, of course, the second that may be called the conquering of the over-all problem. There is little question that the first is the most elegantly formulated as a model, the third is the most completely and easily studied experimentally, and the second is a very poor third on both counts.

What justification is there for these assumptions? Admittedly, there is little beyond intuitive considerations based on experience with complicated decision-making and information-processing organizations and the empirical fact that certain complicated problems appear to make some, though by no means complete, sense when analyzed in terms of these assumptions. They were the assumptions behind the analyses carried out on the group experiments run at MIT by Christie, Luce, and Macy (7), and by these three in conjunction with Hay (16). It is part of the latter experiment that we shall discuss later.

It is also clear that these assumptions do not always hold. For example, when one can distinguish between trivial and weighty decisions, very often the time for a decision will depend upon its importance; *i. e.*, upon implications of the decision that are not a part of the formal structure of the problem as it is now conceived. In such a case, the third assumption does not hold. In our experiments the decisions were not weighty and the subjects wished to conclude the experiment rapidly, and so we may hope that the third assumption will hold. One can prepare similar justifications of the other two assumptions.

Lest the reader feel that, by raising the whole question of time and of the timing of decisions and then turning to a group communication problem, we are suggesting that this is the most important or most fruitful way to try to generalize decision theory, let us now make our disclaimer. Group communication is at best a messy problem full of complications and pitfalls and one not chosen in order to achieve an elegant theory quickly; rather, it seemed sufficiently interesting and important to deserve study in its own right, using whatever tools and techniques were available.

We propose in this part of the present chapter to examine some of the time data from the final completed study in our program at MIT and to indicate some of the lines of research that it has suggested in the areas of the second and third assumptions: analytic models of information flow and the study of reaction times. We do not propose to present these data in great detail, for that has already been done (16); rather, we shall sketch the outlines of our analysis and give the central ideas behind some of our arguments. To some degree the discussion will rest on arguments and results given elsewhere in this chapter, and this knowledge, as well as the general structure of the experiment, will be assumed to be known.

Latency Distributions.

Before we can present any time data, we must consider briefly the problem of presenting such data; this is not the trivial matter it may seem, for we doubt that we could have pushed our analysis as far as we did had we employed the conventional presentation. Latency distributions arise from situations in which an unambiguous stimulus triggers a decision process, the reaction time being the clock time from the stimulus to the decision. When the same situation is replicated many times, either with the same subject (if no interaction can be assumed between occurrences) or with a population of subjects, then it is possible to form a histogram of the decision latencies that may be considered to be an approximation to the probability density governing reaction times. In other words, if $f(t)$ is postulated to be the density of decisions at time t when the stimulation occurred at time 0, and if the time axis is divided into intervals with end points a_i, the data grouped by intervals are considered to be estimates of

$$F(a_{i+1}) - F(a_i) = \int_{a_i}^{a_{i+1}} f(\tau)\, d\tau$$

where

$$F(t) = \int_0^t f(\tau)\, d\tau.$$

The plots resulting are generally skewed and have a single maximum.

The question we raise is whether this convention is best suited to our purpose, or whether some function of the distribution might be more revelatory. Whatever we may do, one thing is certain: We must deal with approximations to F or some simple function of F, and not directly with f, for the latter would require the numerical differentiation of data, a process that is notoriously subject to error.

We might simply assert that rather than use F we shall employ.

$$\Lambda(t) = -\ln [1 - F(t)], \qquad (1)$$

defending this choice on the grounds that from experience we know that this function is very nearly a straight line. This has the very practical merit that we may discuss our data in terms of deviations from straightness and in terms of approximate slopes and how these vary from case to case. However, it is somewhat more persuasive if a simple model is introduced. It must be admitted that, although the model has not been adequately justified, it was by means of it that we were led to Equation 1.

Let us suppose that a stimulus occurs at time 0 and that by time t no decision has been reached. It is reasonable to postulate a probability that the decision will be reached in the interval from t to $t+\Delta t$, this probability depending both on t and on Δt. If we suppose that for a small value of Δt this dependence can be written as $\lambda(t)\Delta t$, which, of course, is not the most general case and so it is an assumption, then it is easy to show (1) that

$$\begin{aligned} f(t) &= \lambda(t)e^{-\int_0^t \lambda(\tau)d\tau} \quad \text{for } t \geqq 0 \\ &= 0 \quad\quad\quad\quad\quad\quad\quad \text{for } t < 0. \end{aligned}$$

If we let $\Lambda(t) = \int_0^t \lambda(\tau)d\tau$, then Equation (1) follows by a few more manipulations (4). Some attempts have been made to justify the basic assumption of the model as well as the empirical observation that $\lambda(t)$ is more often than not nearly a constant, but we doubt that any argument is known at present that really accounts for the observed phenomena of nearly exponential decay. But in any case, we shall deal with $\Lambda(t)$ as given by Equation 1 rather than $F(t)$ directly.

All the above discussion presupposes that we know what the stimulus is and that it occurs at time 0. It is a trivial shift of the time axis to move any unambiguous stimulus to 0, and so the crucial question in the analysis of our experiments is whether we can find unambiguous stimuli. Let us suppose for the moment that subject A receives at time 0 a message from another member, say B, of a group under consideration. If this message contains information new to A that he believes is not known by C, then we might suppose that this message serves as a stimulus for A to send a message to C. Now if between the receipt of this message and his sending the information on to C nothing of significance happens to A, then the above formulation of the latency problem appears to apply. If, however, there are intervening stimuli, such as other messages that do not cause the process to be begun again, but which do have a perturbing influence, then it is by no means clear how to handle the data. For example, A may receive a message from C which he reads, and which, if it does not contain the information he has just received from B, will reinforce his decision to send to C. But the time required to read this message may alter the time between the stimulus and the sending from what it would otherwise have been. The difficulty in dealing with the wide variety of cases that can arise is apparent; so, we are either forced to abandon this method of analysis or may, with some reserve, attempt an analysis in the hope that the effect described will be small and possibly countered by other effects. We took the latter course and our results appear to justify our choice to some extent.

Even if we elect to disregard as second order effects any perturbing influences, the question remains as to what the primary stimulus is for a message. One might attempt to argue, as we did initially in our work, that the message received immediately prior to the message under consideration should be taken to be the stimulus. However, if our discussion above is not completely empty, this cannot always be the case. If not, and if it is an incoming message that is the relevant stimulus, then how can we determine which one it is? This we do not know, but possibly we can argue that we need not know. Our argument about the perturbing but not decisive effect of many messages suggests the answer. If the density of load is sufficiently high, then we may suppose that in most cases the subject is already engaged in the preparation of a message when a message arrives that introduces new information or demands an answer. If this causes only a perturbation, then the *effective* cue for beginning a new time cycle is the completion of current activities. This, then, suggests that in a reasonably high load situation the individual time data will best be understood by taking as the stimulus for a given message the message *sent* by the same subject immediately prior to it, or, in the case of first messages, the beginning of the trial. We shall in fact not only discuss the data for this case, which will be noted by S for "send," but also the data obtained when the immediately preceding received message is treated as the stimulus. The symbol R will be used in this case. Our intuitive argument that the S data are the more basic will be strengthened by our partially successful attempts to account for the characteristics of the R data in terms of the S data.

In addition to this dichotomy, at least three other classifications of the data are needed if it is to assume any coherence. First, the test phase of the experiment consisted of fifteen trials and so the possibility of learning must be admitted, which led us to group the data into five blocks of three trials each. Next, when each message is prepared the subject has before him from one to five pieces of input information—it will be recalled that there were different inputs to each of the five subjects. It is plausible to suppose that reaction time depends upon the complexity of the decision, and so a separation should be effected. Finally, in the communication net studied, there are two classes of positions: the center position, C, and the peripheral positions, P. The net used, diablo, is shown in Figure 2, Network c.

With these distinctions in mind, the data may be characterized by a symbol of the form (x, y, z) where $x = P$ or C, $y = 1, 2, 3, 4, 5$, and $z = R$ or S. Thus $(C, 4, R)$ means the data for the *center* subjects who at the time of sending a message had *four* out of the five pieces of information, and we are measuring reaction time from the immediately preceding message *received* by the subject.

For each such symbol, the data consist of five sets of points, one set for each trial block. Each point falls at some multiple of fifteen seconds since, for simplicity of analysis, the subjects were constrained to send messages only at fifteen-second intervals. Each such set of points, which we shall join together by a line and call a curve, is based on several hundred cases.

Some Qualitative Aspects of the Time Data.

There is no point is presenting all of the time data here, for that has been done (16). As typical examples, the cases $(C, 4, R)$ and $(P, 5, S)$ are presented in Figures 36 and 37. From an examination of all the different curves we have prepared nine statements that characterize qualitative aspects of these data that struck us as interesting. These statements are of necessity somewhat idealized and another person examining the data may feel that we have oversimplified the picture or that we have omitted from consideration other important features. We can recommend only that the reader consult our report (16) and draw his own conclusions. After noting these phenomena, we felt certain that they could not all be independent

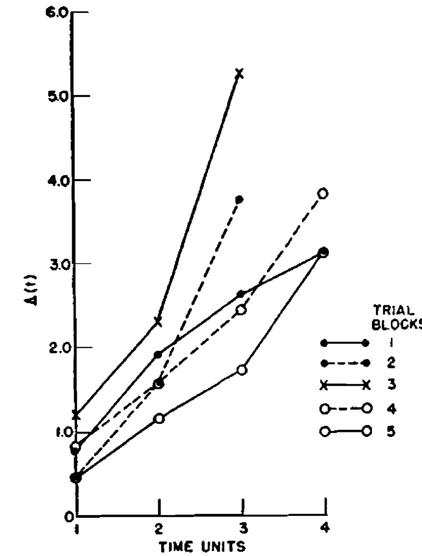

Figure 36. Plot of $\Delta(t)$ *vs* Time Units for $(C, 4, R)$.

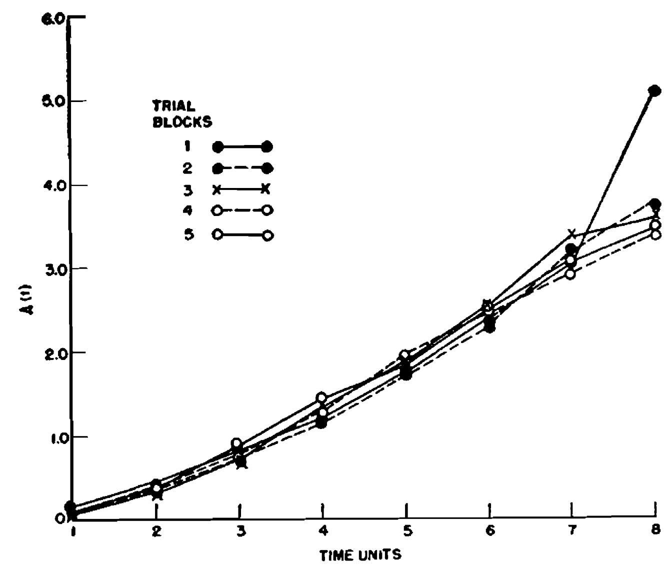

Figure 37. Plot of $\Delta(t)$ *vs* Time Units for $(P, 5, S)$.

of one another, and so we felt obliged to establish some as basic phenomena and to indicate how the others depend on them. The following three sections will be devoted to an attempt to explain, or really to rationalize, the last five statements in terms of the first four.

1. The $(\ , 1, S)$ data generally consist of a single point; *i. e.*, the first message was generally sent at the first opportunity to send (fifteen seconds) after the beginning of a trial.

2. The $(\ , 2, S)$, $(\ , 3, S)$, and $(\ , 4, S)$ curves are approximately straight lines, though a case can be made that $(\ , 4, S)$ is concave down, especially in the final two or three trial blocks. This case is materially strengthened by Statement 9 and its attempted explanation in the section on "Rationalization of Statements 8 and 9," below.

3. The $(P, 2, S)$ and $(P, 3, S)$ curves increase slightly in slope with trials, whereas $(P, 4, S)$ and $(P, 5, S)$ remain roughly constant.

4. For the S data, the C curves have a slightly larger slope than the corresponding P data.

5. For the C data, the slope of each R curve is larger than the slope of the corresponding S curve; this is most marked in $(C, 3, \ )$. For the P data, the slopes of the R and S data are about the same except for $(P, 5, \ )$ where R is the larger.

6. For the R curves, there is a decrease in slope with increasing time.

7. The $(P, 5, R)$ curves increase in slope with trials.

8. For the S data, the slopes of the corresponding curves decrease with increased information present; *e. g.*, $(P, 3, S)$ has a larger slope than $(P, 5, S)$ for a given trial block.

9. The $(\ , 5, S)$ curves are definitely concave down.

From these statements it is easy to exhibit two arguments that suggest that the R data are not basic. First, if they were, Statements 3 and 7 imply that we must account for a constancy over trials in terms of a phenomenon changing over trials, for with five pieces of information present the S data exhibit no learning while the R data do. Of course, this can occur, but it is generally much easier to combine variation with constancy to yield variation than it is to combine variation with variation to yield constancy. Second, Statement 6 implies that for the R data the probability of reaching a decision, assuming it has not already been reached, decreases with time. While we can all recall personal examples of difficult or embarrassing decisions that we have delayed, and then delayed them even more because of our initial hesitancy, it is unlikely that a simple and comparatively emotionless experimental situation could produce the same reaction. Rather, the desire of the subjects to conclude the experiment would, if anything, cause an increase, not a decrease, in slope.

Assuming that we shall be able to rationalize the last five statements in terms of the first four, what interpretation can be given to the first four? The first is certainly the easiest: There is no real difficulty in preparing and deciding where to send the first message, and so it is done with dispatch. The second is simply an observation, once again, that the latency of simple decisions is approximately exponential. Of course, because of the discrete time scale, we cannot argue that the rising limb of the distribution is really as sharp as an exponential—which is certainly doubtful from other data—but rather that the decay portion of the curve is approximately like an exponential. We note as a point to be discussed that, as the information to be processed increases, there appears to be a deviation from the exponential; there is no question of this in the case of five pieces of information (Statement 9). The third statement indicates that for the peripheral subjects some learning occurs over trials for those situations in which small amounts of information are available, but not when more information is present. This indication, coupled with observations presented elsewhere in this chapter, suggests that the subjects learn to routinize the early moves into a pattern that is efficient for the given network, much in the same way that chess players learn to pattern their opening moves. The later situations with four and five pieces of information must continually introduce special problems that require some thought, and so little learning can occur. The fourth statement indicates that the center subjects reach their decisions a little more rapidly on the average than peripheral subjects in corresponding situations; this probably reflects the greater work load and stress upon the center subjects. The latter point is supported by the results of an analysis of questionnaire results (16).

Rationalization of Statements 5 and 6.

Let us consider the pattern of message destinations at the beginning of a trial. We know from Statement 1 that there is a high probability that each subject will send a message at the first opportunity, and from the results on message destinations (16) we know that the probability a peripheral subject will send to the center is roughly 0.5. Thus, the most likely cases to arise are the four shown in Figure 38. In this figure we have not distinguished the particular

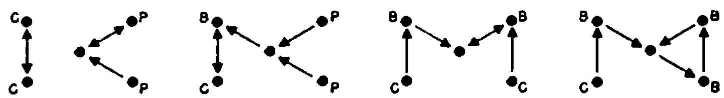

Figure 38. Probable Sending Patterns.

subjects in the four peripheral positions, but rather we have shown only the four topologically different cases. If we consider each peripheral position in turn we may determine whether, as a result of this first exchange of messages, he has new information to send to the center, to the other peripheral position to which he can send, or to both. We have accordingly labeled the nodes C, P, or B. Now, from our work on message destinations we know that, when a subject knows he has new information for exactly one other subject to whom he is connected, the probability is 0.97 or thereabouts that his next message will go to him. In the B case the probability is about 0.65 that he will send to the position to which he did not send his previous message. If one approximates 0.97 by 1.00, then a simple calculation shows there is a probability of 0.75 that either one or two of the peripheral subjects will send their next message to the center position.

Keeping this in mind, let us now turn to the center position. Let δ denote the number of time units between his first and second messages; according to Statement 2, δ is distributed approximately exponentially. Let us suppose C has selected δ according to this distribution. If δ is fairly large there is a reasonably high probability that one of the peripheral subjects will have sent him a message in the time interval from 1 to $1+\delta$. Thus, the time between the last message received and the given message will often be less than δ, when δ is large. If, on the other hand, δ is small, then the time from the

last message received cannot be greater than δ, which is already small by assumption. Thus, the R data should include more small values and fewer large values of time than the S data. If we let λ_R denote the initial slope of $(C, 3, R)$ and λ_S the slope of $(C, 3, S)$, then a detailed computation shows that, when one peripheral subject sends to C before C sends, $\lambda_R/\lambda_S = 1.54$, and, when two do, $\lambda_R/\lambda_S = 2.08$. The observed ratio is 1.63. This accounts for Statement 5.

If the S curves are straight lines (Statement 2) and if the R curves are formed by making some of the large values of time small ones, then it is evident that the slope of the R curves should be a decreasing function of time, which is Statement 6.

For the peripheral subjects the picture is far more complex, but it appears that the same general argument should apply, except that the density of incoming messages is less and therefore the effect should be less pronounced, and quite possibly not detectable at all.

Rationalization of Statement 7.

Statement 7, by itself, is innocent enough, for one can remark that it simply reflects learning over trials. But, as we pointed out, if Statement 3 is taken as basic, then for the S data no learning is exhibited when five pieces of information are present, and it appears that there may be a difficulty. In fact, if we were able to convince ourselves that the learning suggested by the $(P, 5, R)$ data had actually directly influenced the decision when five pieces of information are present, then I believe that we would have to drop our tentative hypothesis that the S case is the more basic of the two. The alternative is to show that the apparent learning in $(P, 5, R)$ is an artifact stemming from other learning in the process.

Subjects in the center position exhibit comparatively little learning with trials in the S data, whereas the peripheral subjects do exhibit learning for the cases of two or three pieces of information present, though not for more. Now, from our work on message choice we know that there is little change shown over trials, and so a pattern of message sending is just as likely in the later trials as in the early ones. Consider any such pattern, say the one shown in Figure 39, and let δ denote the number of time units between the message from C that gives P the final piece of information and the next message sent by P. In later trials the same pattern of sending will arise just about as often as in early trials, but it will be changed in time structure in accordance with the learning that has occurred, which, of course, will affect only the time between the first few messages. This, however, has the effect of shifting the last message to the left as shown in Figure 40, which in turn reduces the average number of time units

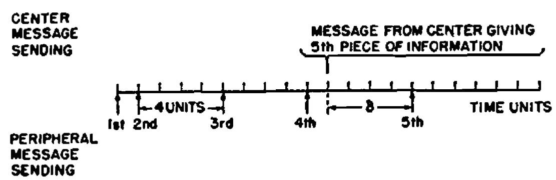

Figure 39. Pattern of Message Sending.

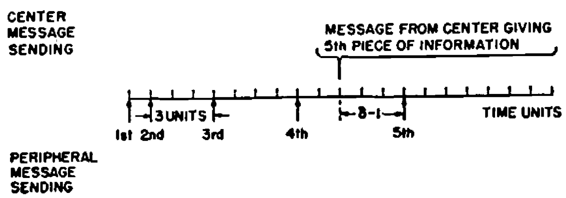

Figure 40. Effect of Shifting Last Message to the Left.

from the message from C to the last message sent by P to something less than δ. In other words, the R data with four and five pieces of information present can be expected to reflect the learning exhibited in $(P, 2, S)$ and $(P, 3, S)$, and so Statement 7 is rationalized.

Rationalization of Statements 8 and 9.

The qualitative content of Statement 8 seems reasonable, for we are all used to the fact, that, generally, the more information there is to be processed the longer it takes. The exact numerical changes are another matter, and at present we know of no comprehensive model to account for changes in time constants with changes in information. The change in shape of the distribution with information (Statement 9) is, in a way, more disturbing and demanding of an explanation. We do not truly have such an explanation; however, we can present a simple calculation which, if it is not fortuitous, is intriguing.

Suppose that the time distribution when there are two pieces of information present is given by $f = \lambda e^{-\lambda t}$, where from the data $(P, 2, S)$ in the first trial block we find $\lambda = 0.67$. Suppose that in the more complicated situation of five pieces of information the problem is divided by the subject into two *independent* parts, each of which can be handled in a manner similar to the case of two pieces of information, and that the two parts are dealt with successively by the subject. If this were the case, then it is easy to see that the distribution would be

$$f_2 = \int_0^t f(t-\tau) f(\tau)\, d\tau.$$

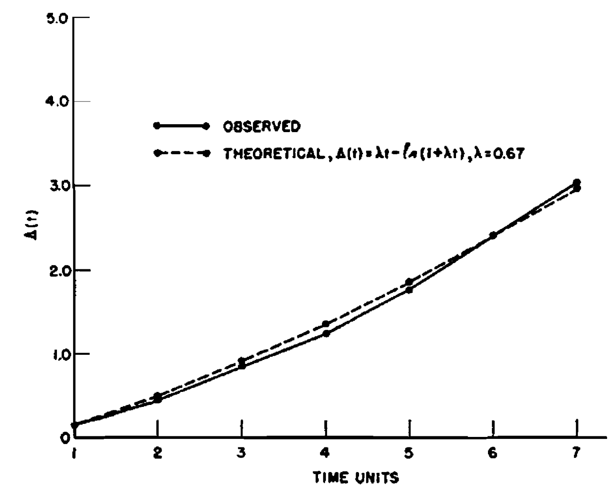

Figure 41. Observed and Theoretical Curves for $(P, 5, S)$ in Trial Block 1.

Substituting f and solving we find

$$\Lambda(t) = \lambda t - \ln[1 + \lambda t].$$

Using $\lambda = 0.67$ and plotting the result along with the $(P, 5, S)$ data for the first trial block we obtain the curves shown in Figure 41. Their similarity is remarkable, especially considering that there are no adjustable parameters. The difficulty is to know whether this rough analysis has some basis in fact or whether it is entirely accidental; we shall return to this problem later.

Analytic Models of Information Flow.

So far, we have devoted our attention to dissecting the group processing of information into "elementary events" with comparatively simple properties. The real test of such an analysis is whether, from it, one can once again reconstruct the group behavior, for if this is not demonstrated one cannot be sure whether certain crucial phenomena have not been omitted. As suggested in the introduction, this phase of the work is by far the least complete, mainly because it creates so many mathematical difficulties, but partly because certain phenomena are still not understood.

Let us try briefly to summarize some of the principal things that we know, and some that we know we don't know. As to the choice of destination for the first message of a trial, it appears to be a chance event, heavily biased toward doing what was done on the preceding trial. For later messages in the trial there appears to be little direct dependence on the preceding trial, but rather a fair appreciation of the logical structure of information flow. If one considers for subject i the set of information A^i_j he is able to deduce what subject j must know (j may in fact know more) then the sets $A^i_i - A^i_j$ (where minus denotes the logical difference of two sets) appear to be extremely relevant to his choice of message destination. We know, for example, that if among the j's to whom i can send a message exactly one of these sets is non-empty, then with probability nearly 1.0 the message will be sent to that j. However, what happens when two or more are non-empty is not nearly so certain. Undoubtedly there is a large role for chance, but in addition behavior appears to depend in a complicated way upon several other factors. We had the feeling that the factors we examined were not, however, basic to the phenomena and that possibly we were obtaining artifactual effects. Thus a major open problem in this work, and we judge in many other areas, is the question of how a subject selects among several alternatives when each maximizes his payoff function; or put another way, what other factors come into play that are normally subordinate to a known payoff function.

Within the realm of time we appear to have shown that, if the work load is ample, then the stimulus for the preparation of a message may be taken to be the last message sent. The latencies appear to be

approximately exponential when the decisions are simple and to deviate from it in both mean and shape when the decisions become more complicated. There seems to be some reason to suppose that the more complex cases are serially composed from the distributions of the simpler ones, but this is by no means firmly demonstrated. Thus a second major problem is to acquire more detailed and precise information as to how decision latencies depend upon the information available at the time the decision is taken. With respect to this problem we shall have a few more comments in the next section.

When these two problems are resolved, then there is no reason why we could not assume these statements as the axioms of a model for information flow. There is every reason to suppose that such a model cannot at present be solved analytically, but with some labor one could code it for a digital computer and carry out the necessary Monte Carlo runs to obtain stable estimates of any of the group parameters one cares to compare with experimental data.

One reason for supposing the model cannot at present be dealt with analytically is the result of an attempt (15) made to set up the equations for a simpler case; there it was patently obvious that closed solutions are most unlikely to be forthcoming except for very simple special cases. The case studied met the following conditions:

(1) The communication network has all channels open.
(2) Time distributions for the injection of input information into the group are given.
(3) Decision latencies are stimulated by message sendings; *i. e.*, preceding decisions.
(4) The latency distribution for decisions are exponential with parameters that are constant for the problem.
(5) Message destination probabilities depend only on the set of information available to the subject at the time the message is sent.

We see that Assumptions 1, 4, and 5 are much simpler than those arrived at for our experimental situation.

As a final problem complicating a model builder's task, we should mention our unease at the assumption that an incoming message has no other effect than to change the information state of the subject. It appears obvious that in some cases certain incoming messages may change significantly the process with which a person is currently engaged, but so far we have developed no suitable experiments or techniques to get information on this problem.

Reaction Time Experiments.

Let us, in this final section, return to the time problem, now in its purest form: the reaction time experiment. In such an experiment, of course, one carefully eliminates many of the difficulties that have plagued us in group experiments; the stimulus is in fact unambiguous, care is taken to prevent any intervening external stimuli prior to the decision, the conditions of the experiment (such as the information available to the subject) are carefully controlled, the experiment is replicated many times, and time is measured in milliseconds. And still there have been difficulties, one of the principal ones being to separate the time required to make the decisions from the time required to carry out certain motor activities, such as signaling the decision to the experimenter. Until this separation is effected, it is difficult, if not impossible, to determine exactly how the decision part of the latency depends on the decision itself. Because of this, the study of reaction times has fallen into disfavor, though in the past few years there has been a tentative resumption of interest because of some new ideas arising from information theory.

Whether or not the study of latencies is looked upon with favor, it is clear from our remarks above that much of complex decision making and information processing will not be understood until the dependency of decision latencies on information present at the time of decision is better understood. Our realization of this, plus the hint given by the apparent serial composition of latencies, led Christie and Luce (6) to put forth a suggestion concerning the analysis of reaction time data. Briefly, the idea, which depends heavily on the fact that the Laplace transform of a convolution integral (*i.e.*, an integral of the form $\int_0^t F(\tau)G(t-\tau)\,d\tau$) is simply the product of the two Laplace transforms of the given functions.

If F is a (reasonably well behaved) function that is 0 for $t<0$, then the Laplace transform of F, $L(F)$, is defined by

$$L(F)=\int_0^\infty e^{-st}F(t)\,dt,$$

which we see is a real-valued function of the real variable s. For example, if $F=\lambda e^{-\lambda t}$, then a simple computation shows that $L(F)=1/(s/\lambda+1)$. Now suppose t is a reaction time measurement distributed according to $f(t)$, and that t is composed of a base time t_b and a decision or choice time t_c; *i.e.*, $t=t_b+t_c$. If these two component times are *independent*, that is, if one can change the motor reactions, and thus t_b, leaving t_c unaffected, or change the choice without affecting the motor reactions, and if t_b is distributed according to f_b and t_c according to f_c, then it is easy to show

$$f(t)=\int_0^t f_b(\tau)f_c(t-\tau)\,d\tau.$$

By one of the central properties of the Laplace transform it follows that

$$L(f)=L(f_b)L(f_c).$$

Now if one supposes that he has a set of experimental conditions that holds f_b constant and changes f_c by changing the complexity of the decision to be made, then for two such cases f and f' we see that

$$\frac{L(f)}{L(f')}=\frac{L(f_c)}{L(f'_c)}$$

and all mention of base times is dropped.

In particular, let us suppose that for a certain class of experiments the decision latency results from a serial composition of "elementary" decision latencies whose form is exponential with constant λ. If one experiment, with over-all latency f_n, requires n elementary decisions, and another, with latency f_1, only one decision, then it is easy to show that

$$\frac{L(f_n)}{L(f_1)}=\frac{L(\int_0^t f_n(\tau)\,d\tau)}{L(\int_0^t f_1(\tau)\,d\tau)}=\frac{1}{(s/\lambda+1)^{n-1}}.$$

At the other extreme, we might suppose a decision process requiring the execution of n elementary decisions in parallel; *i.e.*, all n-component decisions are initiated simultaneously and the process is concluded when all n decisions have been made. In that case

$$\frac{L(\int_0^t f_n(\tau)\,d\tau)}{L(\int_0^t f_1(\tau)\,d\tau)}=\frac{n!\Gamma(s/\lambda+1)}{\Gamma(s/\lambda+n+1)}$$

where Γ is the gamma function; *i.e.*, $\Gamma(x) = (x-1)!$, where x is an integer, and for all x, $\Gamma(x+1) = x\Gamma(x)$. For any other particular arrangement of decisions similar calculations are possible (6).

One might hope that by obtaining latency data it would be possible simultaneously to select parameters and to choose among several hypotheses as to the arrangement of elementary decisions, assuming such exist. Unfortunately, however, if one calculates the two functions given above and plots them against s/λ, it will be seen that the curves, in the region where they overlap, are very similar indeed. Of course, any pair of similar curves has different values of n, the parallel case being the larger of the two. This means that extremely detailed and precise data will be required to decide between cases. Nonetheless, the need for adequate knowledge about the structuring of decisions in terms of the logic of the decision to be made is so great that it may be worth while spending some effort to see if this procedure can be made to work.

V. TASK TYPES AND REQUIREMENTS FOR ORGANIZATION.*

Introduction.

The preceding parts of this chapter have discussed human information processing in the light of data from a series of experiments with little concern for the relation of one experiment to another. Various tasks have been required of the subject groups in these experiments, each task designed for the purposes of a particular investigation. The contribution to knowledge that the series makes is no more than the sum of the contributions made by each experiment in isolation unless we can give an integrated account of the whole. To accomplish such an integration we need a theory of tasks; to quote Luce, *et al.* (16), ". . . this problem must be given serious consideration, or the study of information processing groups will result in an elaborate but incoherent bibliography."

In this chapter a start will be made on a theory of tasks. The

* By Lee S. Christie.