The sequence of alternative guessing under the repetitive two-choice situation has been much studied after Humphreys' verbal conditioning experiment (6). The experiment is usually run under the following design: Some number of Ss are noticed that either one of the two alternative symbols will be successively presented, and that they are expected to guess individually as correctly as possible which the next symbol is before each presentation of the symbol. When the experiment is over, the ratio of the number of Ss who guessed a particular symbol is calculated for each trial, resulting in a sequence of these ratios called the guessing sequence. The order of the presentation of symbols is ordinarily random or quasi-random while the relative frequency of the symbols is the main experimental variable, not necessarily 50-50.

A rather remarkable fact thus far observed in guessing sequences is that the ratio almost unexceptionally approaches a level closely near the relative frequency of the symbols. This fact is remarkable not only because Ss are consistently irrational. The criterion of rationality conceived by most authors in this area seems to be whether S's responses accord with the strategy of maximizing the number of correct predictions. This strategy, as now commonly understood, is given by keeping to guess the same symbol of greater probability of occurrence as soon as such a symbol is found out. Though a problem remains as to when and by what amount of evidence Ss collected they regard the probability of occurrence of a symbol as greater than the other, it is true that there can be no asymptote of the response ratio other than 100% or 0% as far as Ss follow this kind of strategy (4).

One could draw two alternative conclusions from this fact: The guessing behavior is irrational. Or, the criterion of rationality above stated is too restrictive and inappropriate to apply to the guessing behavior produced by the above stated type of guessing experiment. Jarvik (7), in effect, stood on the side of the first alternative in speculating that his Ss might have misconceived it a good strategy to mix their responses with the same frequencies as those of symbols. Hake and Hyman (5)
stood on the same side but, by the reason entirely contrary to that of Jarvik. They found by analysing their experimental data that each S's predictions systematically and in a quite simple way correlated to the symbol presented on the just preceding trial and to the success or failure of his preceding prediction. The mystery appeared to be solved then. The response frequency approaches the symbol frequency just because Ss respond to the symbol sequence in such a simple way that the statistical structure of the symbol sequence automatically reflects in the response sequence. Now that they have evidence, it might seem quite natural for them to conclude that man is none such rational mechanism as the statistical automata.

Nevertheless, an objection can still be raised. Hake and Hyman's conclusion is right, provided that Ss should maximize the number of correct predictions. But why should they? The only condition as far as I can cover that bounds Ss is the instruction given to them, which is in most experiments believed to be the replication of or essentially the same as Humphreys' instruction of "Do your best to answer correctly." *(6)* Evidently, to predict as correctly as possible does not tautologically mean to maximize the number of hits. It is quite probable that Ss' evaluation of correctness differs between the symbols frequently presented and those infrequently presented. Then they might hunt for infrequent symbols, and this kind of behavior does not run against the given instruction if strictly comprehended. In a sense, they were quite faithful to the instruction if they behaved so, since they were correct even in situations in which it was very difficult to be correct.

Even when S is instructed to maximize the number of hits or the given instruction is comprehended in this sense, it is still possible that S's passion of going his way exceeds the restrictive power of instruction; they might be often frustrated, giving rise to a frustrated behavior; they might be bored of making the same response. Every one of these factors would contribute to introducing a random element into the guessing frequency, causing a centrality bias within the guessing ratio.

To speak more precisely, let me describe the situation in terms of the theory of decision making (1). Though my line of thinking is a little different (10, 11) from the conventional model, the difference would not be essential in the following discussion.

The theory, based upon the subjective probability-utility model, establishes first of all a proposition, which can be stated as follows: Person chooses among alternative acts the one that promises him the maximum value of the subjectively expected utility; where utility means, roughly speaking, the desirability or the valence of the state to which

* This was pointed out by Goodnow (2).
he would possibly be led by that act. According to this proposition, the statement that an instruction has a restrictive power means that being faithful to the instruction promises him a greater expected utility than being unfaithful. It is then expected that, if an instruction is given to force Ss to maximize the number of hits, an utility field would be generated in such a manner that the larger the number of hits, the greater the utility payoff. If this is the only utility payoff and if most Ss are so rational as being capable of finding out how to maximize the number of hits, the guessing sequence could have no asymptote other than 100% or 0%, no matter what the probability of the symbol sequence may be, except for 0.5. However, the first condition would never be strictly fulfilled. There could be many uncontrollable miscellaneous utilities such as the positive utility* of predicting rare event, the negative utility of regret, the negative, decreasing utility of making the same kind of responses, the positive utility of saving the hard work of prediction by taking some easy-going way and so on, as stated above. The proposition of decision making tells that Ss maximize the expected value of the whole set of utilities, regardless of those imposed by E, say formal utility, or those of Ss' own, say informal utility. Now assume that the formal utility is roughly proportional to the number of hits. Suppose the symbol frequency is 90-10. If a S follows the above-stated rational strategy, he will get 90% hits and the corresponding formal utility, but with a severe loss in informal utility. If he pursues entirely the informal utility, he will get every obtainable informal utility at the cost of the formal utility corresponding to in average about 40% misses. Then the strategy of the optimal compromise would locate the guessing ratio somewhere between 100% and 50%. The more the symbol frequency approaches 50%, the more will this ratio of compromise approach 50%, since, for example, when the symbol frequency is 70-30, S will lose the formal utility corresponding to only 20% misses in average, even if he pursues entirely the informal utility. If this hypothesis really be the case, the congruence between the symbol frequency and the asymptotic value of guessing ratio is only a coincidence, and despite of all the apparent irrationality, man could still be regarded as rational.

All of these considerations can be centered upon the hypothesis that the asymptote of the guessing sequence is the result of the compromise between the gain of formal, controllable utility and the loss of informal, uncontrollable utility. This hypothesis is experimentally testable as far as the one of the

* The origin of utility scale is arbitrary in nature, since utility can be measured only by interval scales. Here I assume, only temporarily, the state of gaining nothing and losing nothing has zero utility.
determinants is controllable. If the formal utility payoff is increased, the more the increase, the closer to 100% or 0% would be set the asymptote. This is the prediction that the present experiment is aimed to test.

**METHOD**

*Apparatus*—The symbol which was presented to Ss was lighting on of either right or left of two 100-watt electric bulbs set 1.30 m. apart. At the middle of the two bulbs was a tiny warning lamp. An assistant attended in each room holding five Ss, and he turned on the switch connected with the corresponding sign lamp in the control room, as soon as all the five Ss in the room finished their predictions. After every sign lamp in the control room was lighted on, the control room operator switched on the warning lamp and, 3 sec. later, either one of the stimulus lamps according to the predetermined schedule. Thus the intertrial interval was not held strictly constant, in average 30 sec., but, since the lighting of lamps in all the rooms were entirely synchronized, the experimental condition was common in this respect for every S in one experimental session.

*Subjects*—30 Ss were used for the first experiment, of which randomly selected 15 were used for the present purpose. They were randomly split into three experimental groups, Group 1, Group 2 and Group 3, of five Ss each. For the second experiment, five Ss were newly adopted, of whom Group 4 was constituted. Every S was an undergraduate student of Hokkaido University invited for two hours work as S in a psychological experiment with reward of 150 yen, which amounts to about 42 cents, a little better wages for their work than the usual standard.

*Procedures*—All the 20 Ss were given the same sequence of 101 symbols, of which the first 100 symbols were constituted of 75 rights and 25 lefts. The order of presentation of the two symbols were randomised according to the table of random numbers, modified a little. The modification was done so that the symbol frequency was kept 75–25 within every block of twenty trials; attention was kept, however, not to keep it so within shorter blocks, since such over-homogeneity of symbol sequence adds information to symbols more than that provided by the probability, reinforcing the so-called negative recency effect, so making the strategy of sticking to the same prediction no more the best.

In the first experiment, members of the three groups used were introduced to the three separate rooms, seated side by side before a desk, on which were placed before each S a little rubber block and a sheet of cardboard on which were marked two squares right and left, and the right and left sides of the sheet were turned up to serve as screens. S's prediction was expressed by placing the rubber block on either one of the right and the left squares according as his prediction was right or left. Immediately after the presentation of each symbol, Ss were required to draw back the rubber block and then to replace it according to their prediction of the next symbol. As soon as all the Ss in a group had replaced the blocks, the assistant in the room switched on the sign lamp in the control room, and then recorded Ss' responses on the recording sheet.

Only difference in the experimental conditions among the first three groups lay in the conditions of utility-payoff. The Ss of Group 1 were paid 150 yen in advance of the experiment and instructed simply as "Predict!" The Ss of Group 2 were paid also 150 yen but after the experiment, and told that they were the players of a guessing game, to compete among them for the number of correct predictions, and the more the correct predictions a player made, the greater the winning he obtained. On each trial, the total number of hits of each S up to the trial was shown on a blackboard. The Ss of Group 3 were instructed that they would be paid three yen per each hit; therefore, the total earnings of each S was the total number of hits times three yen. Like Group 2 the total scores of Ss up to each trial were shown on a blackboard.
Information could be drawn from the scores on the blackboard on one another's response pattern in making predictions by the Ss of Groups 2 and 3. This might have a sort of group effect on the Ss of Groups 2 and 3. The Ss of Group 1 could, however, know one another's response pattern if they wanted to, since the screens were quite tiny, no more than symbolic screens. Therefore, the Ss of Group 1 could be expected to be under the similar, though not entirely the same, group effect.

The experimental conditions imposed upon the Ss of Group 4 differed from the other three in two ways: First, the group effect was removed; each S was separated from others by huge wooden screens, which substantially exterminated the interaction in any sense between Ss. Secondly, the utility-payoff condition were further strengthened. Behind each S, stood an assistant, recording the predictions, paid the S a ten-yen coin at each of his correct predictions and drew back a coin from the S's possessions at each of his wrong predictions. Each S was lent 100 yens in coin in advance of the experiment and paid back that amount from his possessions after the experiment.

RESULT*

The data of this experiment are $4 \times 5 \times 101$ predictions, from which guessing ratios of predicting "right" are calculated for each group in each 20-trial block, excluding the beginning trial. The predictions made at the beginning trial, specified as the zeroth trial, have some specificity compared to others; since at the beginning Ss had no information about the nature of the symbol probability, no strategic considerations due to the differing conditions of formal utility payoff could reflect upon their predictions, which, then, could serve only to detect the group specificity, if any, with respect to the response preference. The guessing ratios of the four groups on the zeroth trial are .6, .2, .2 and .6 in this order, none of these values being statistically significant beyond chance, and fortunately no correlation being found between the orders of them and those of the values in the final trial block. For all the Ss the guessing ratio on the zeroth trial is .4, a value being a little in favor of the response "left", but also insignificant.

* This result was reported orally at the 18th annual meeting of JPA (13).

The four guessing sequences calculated are plotted in Fig. 1. All the four ascending curves end on the final trial block keeping the order expected from the magnitudes of the formal utility payoff. This impression has to be tested statistically, and for this purpose the $X^2$-test is applied to each pair of guessing ratio in the final trial block. As tabulated in Table 1, the result is
generally in favor of the theoretical prediction; though no differences between the values of adjacent groups are significant, all the others are significant.

**Table. 1.**

THE RESULT OF $\chi^2$-TEST ON THE TOTAL NUMBER OF PREDICTIONS IN THE TRIAL BLOCK 81—100 BETWEEN EACH PAIR OF GROUPS

<table>
<thead>
<tr>
<th>Group</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of Prediction</td>
<td>R</td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td>Group 4</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Group 3</td>
<td>*</td>
<td>†</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>†</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† Not significant.
* Significant at 5% level.
** Significant at 1% level.

admit that the guessing curve is roughly monotonic growth curve, however at least the curves of the first three groups seem to have reached their asymptotes in the final trial block. If the same assumption is to be applied, we have to expect that the asymptotic value for the fourth group is greater than the value in the final trial block, since the curve seems still ascending thereon. To confirm these speculations, Kendall's serial correlation coefficients (8) are calculated and tabulated in Table 2 for the sequences of guessing ratios on single trials within trial blocks 61—80 and 81—100.

The extreme smallness of the serial correlation coefficients in the table for the first three groups in the final trial block indicates that the guessing patterns of these groups were kept stationary during this period. On the other hand, the figure for the group 4 is fairly big, though not significant at 5% level of confidence, suggesting that the curve still has an increasing tendency.

Naturally, this evidence will lack a sufficient persuasive power, unless it is shown that the order in the final trial block is not incidental, since the theoretical prediction concerns only with the eventual asymptotes. In view of the result, 101 is certainly too small a number of trials. As far as we block indicates that the guessing patterns of these groups were kept stationary during this period. On the other hand, the figure for the group 4 is fairly big, though not significant at 5% level of confidence, suggesting that the curve still has an increasing tendency.
Together with the evidences supplied by Tables 1 and 2, we may conclude that the question raised about the tentative hypothesis has been answered favorably to the extent that the present experiment could tell.

Another important information is contained in Table 2. As a glance at Fig. 1 will show the differing formal utility payoff does not seem to affect the slope of the curves, but does seem to affect the period when the curves reach asymptotes. The figures, particularly probabilities, in Table 2 clearly illustrate this; they suggest that the more the formal utility payoff, the later the curve reaches its “ceiling”. Naturally, the later the “ceiling” comes, the asymptote will generally be higher, as far as no systematic change appears in slopes.

So far, we analysed our data primarily as sequences of guessing ratios. It might as well be possible and worthwhile to examine the effect of differing utility payoff upon the guessing patterns of individual Ss. As previously mentioned, Hake and Hyman (5) found that there was a definite sequential dependencies among the responses of their Ss, who are presumably comparative to our Ss in group 1. Following their notation, we shall write as P_{RC}(R), P_{LI}(R), and so on, the relative frequency within a trial block that a correct prediction of “right” was followed by  

<table>
<thead>
<tr>
<th>Probability</th>
<th>P_{RC}(R)</th>
<th>P_{RI}(R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>1--2021--4041--6061--80 81--100</td>
<td>1--2021--4041--6061--80 81--100</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.53</td>
<td>.55</td>
</tr>
<tr>
<td>2</td>
<td>.37</td>
<td>.50</td>
</tr>
<tr>
<td>3</td>
<td>.27</td>
<td>.53</td>
</tr>
<tr>
<td>4</td>
<td>.52</td>
<td>.80</td>
</tr>
</tbody>
</table>

* No entries, because of too few samples.
the "right" prediction, that an incorrect prediction of "left" was followed by the "right" prediction and so on, respectively, which are tabulated in Table 3. $P_{RC}(R)$ and $P_{LI}(R)$ are also plotted in Fig. 2 and 3. The curves of $P_{RC}(R)$ and $P_{LI}(R)$ are more or less ordinary type of growth curves, whereas $P_{RI}(R)$ and $P_{LC}(R)$ in Table 3 show no significant systematic change from the beginning to the end. It should be remembered that these values are all relative frequencies, calculated from different size of samples—different for every entry in Table 3; the sample size is particularly small for every $P_{LC}(R)$, for $P_{RI}(R)$ in the beginning trial blocks and for $P_{LI}(R)$ in later trial blocks.

Hence, if we want to say anything definite on this kind of sequential dependency, we must use the exact numbers of responses that fall into the $2 \times 4$ categories, produced by the two response alternatives, and the four alternative combinations between the preceding response and its hit and miss. The result of the $\chi^2$-test is presented in Table 4, which will be summarized as follows: For the first three groups, no significant structurization of responses against S's states on the just preceding trial is observed during the first two trial blocks, except for the group 1 on the first trial block. This exceptional structurization is, however, appeared in

![Successive 20 Trial Blocks Fig. 2.](image1)

![Successive 20 Trial Blocks Fig. 3.](image2)
the opposite direction—in the direction that $P_{RC}(R)$ is less than $P(R)$. In the succeeding trial blocks all the three groups have normal structurization except for group 3 in the last trial block, where the structure again vanishes. The structurization of group 4 in the first trial block is also in the opposite direction. Then it is switched to the normal direction in the second trial block, again vanishing in the last trial block.

**DISCUSSION**

The most important result of this experiment is the finding that a bigger payoff of utility brings about a more rational response pattern of Ss. Independently of the author, however, Goodnow obtained by a similar gambling experiment the same result, which was published recently (2). Though the present experiment had been completed before her publication, the lack of communication between her country and mine quite incidentally made my experiment to serve as a post-test for the confirmation of her result. One thing which is new in mine is that stepwise increased utility payoff results in stepwise increasing asymptote. This fact adds to Goodnow’s findings something which would throw a light upon the hidden mechanism of guessing behavior.

It is true that the light is not quite bright. The present result also confirmed Hake and Hyman’s finding, but their conclusion drawn from the same finding brings about a little confusing contradiction if checked up with our main result. It is doubtlessly true that, as they say, Ss do not first perceive the rules by which the symbol series were constructed; their responses are largely dependent upon their previous responses and whether those responses were correct or not (5). It is nevertheless doubtful that therefore the guessing sequence is simply by-products of this sequential dependency, if “by-product” is to mean that the rule of the symbol sequence and the nature and the amount of sequential dependency are the only parameters determining the guessing sequence. Because, first of all, this idea cannot account for Goodnow’s and the author’s results; it must at least be admitted that the sequential dependency is a function of utility payoff. Further, though it is true that the sequential dependency develops and is kept as developed for Ss in groups 1 and 2, who are comparable to Hake and Hyman’s Ss, this is not true for the other groups who were in gambling situations. For them the dependency seems to vanish in the last trial block or at least becomes less conspicuous than before.

This is one of the facts to be accounted for, and others are found in Table 3 and Fig. 2 and 3. The possibility that the sequential dependency is innate to Ss is first ruled out, since, noting $P_{ij}(R)$’s are relative frequen-
cies, it is at all possible that $P_{ij}(R)$'s are constant from the beginning even if Ss have no response preference at the start. On the contrary, the dependency in the first trial block is clearly the converse; $P_{RC}(R)$ and $P_{LI}(R)$ are little and $P_{RI}(S)$ is great. From this original level, Ss start to learn, so to speak, the sequential dependency—separately for each of the four as was shown in Table 3.

Now we shall see what will come out of the theory of decision making briefly introduced in the preliminary remarks, if it is to account for those findings. First, we could assume that the probabilities in the first trial block roughly represent the response pattern which maximizes the informal utility only, since thereon no sufficient information is provided by the symbol sequence to pursue the formal utility. This assumption would also be supported by an intuitive reasoning; when a S has missed a prediction, he would probably be in a state of slight frustration, and however slight it might be, it would probably affect his next response in such a way that hitting the same symbol which has caused frustration in him has greater informal utility than hitting the different symbol. And naturally, as trial proceeds the utilities of hitting the common symbol and the rare symbol, and of missing the common symbol and the rare symbol will become different from each other. As far as this speculation is taken for granted, it is quite natural that the four sets of response probabilities develop corresponding to the four different utility status. As soon as the observed number of symbols becomes fairly large, Ss can pursue the formal utility, and the strategy of maximizing the informal utility only is no more good. Ss have to modify their strategy according to the accumulation of information, in order to maximize the whole set of utilities. For this purpose, it is evident that to increase $P_{RC}(R)$ is most efficient, as far as the symbol probability of R is greater than that of L. The high level of $P_{RC}(R)$ is not sufficient, however, to maintain a fairly high rate of hit; it is also necessary that $P_{LI}(R)$ is great too. The values of $P_{RI}(R)$ and $P_{LC}(R)$ are only of secondary importance. Hence, the original structurization of response pattern, if any, is first broken down, and then the structurization in the opposite direction is established and kept developing. If the development is carried to the extreme, the structurization ceases to be statistically significant as in case of groups 3 and 4, at least at the sample size we dealt with, since most samples fall into the single category of RC. Naturally not all the groups are to be carried to this extreme. If the formal utility payoff is not very great, the structurization will stop on its way of development, where the whole set of utility payoff is maximized. The curves for group 1 may be typical;
and the difference between the asymptotic values of $P_{RC}(R)$ and of $P_{LI}(R)$ for this group would in some way reflect the difference in the informal utility payoff.

These discussions are no more than a mere conjecture. There might be many other informal utilities than that caused by frustration; experimental evidences are also not enough to either affirm or negate the conjecture. Nevertheless, such speculation would show a possibility of explaining the complexity of guessing behavior hidden behind its simple outlook.

Before closing this discussion let me add an observational record on the group 4 Ss. Near the ending of the experiment when they are consistently predicting R, most Ss showed quite frequently some irrelevant responses such as cratching their head, shaking their head and so on at each time they did not change the next prediction in spite of their present miss. This seems to tell us something relevant to the above discussion.

**SUMMARY**

Four groups of Ss were required to predict on each of 101 trials which one of the two alternative symbols would appear on that trial. The symbol series used was the same for all groups, a random series of probabilities .75 and .25. The four groups were different in the degree of strength of the experimental condition—the utility payoff—that forced Ss to adopt the "maximizing hits" strategy. Ss of group 1 were paid 150 yens in advance and were instructed only to predict. Ss of group 2 were made to compete among them for the numbers of hits and were paid 150 yens after the experiment. Ss of group 3 were paid 3 yens on every hit. Ss of group 4 were paid 10 yens on every hit and fined 10 yens on every miss.

The results were as follows:

1. The order of guessing ratios of the four groups was the same as the order of the utility payoff in the final twenty-trial block, where the guessing ratios of the first three groups seemed to reach their asymptotes and that of the group 4 seemed still increasing. This result was essentially in accord with Goodnow's result.

2. The differing utility payoff had influence not upon the slope of guessing curves but on their period of reaching the asymptote; the greater the payoff, the later the period.

3. The analysis of sequential dependency of Ss' predictions produced the comparable result to Hake and Hyman's. The sequential dependency was significant in the later trial blocks, but in the final block it became insignificant for the groups 3 and 4, whereas still significant for the groups 1 and 2.

4. Concerning this dependency, not only $P_{RC}(R)$, the conditional probability of predicting the symbol "right" when the preceding prediction was correctly "right," that was dealt with by Hake
and Hyman, but also every other conditional probability was calculated. Among them $P_{RC}(R)$ and $P_{LI}(R)$ showed typical learning curves, whereas the other two showed no systematic change during the experiment. For those hidden complexity of human Ss' guessing behavior, a possible explanation from the standpoint of the theory of decision making was pointed out.

REFERENCES


* Japanese text with English abstract.

** Japanese abstract.