QUANTITATIVE VARIATION OF INCENTIVE AND PERFORMANCE IN THE WHITE RAT

By Leo P. Crespi, Princeton University

This investigation is concerned with three specific inquiries under the broad problem of the relationship between magnitude of incentive and performance; (1) What is the relationship between magnitude of incentive and the level of performance? (2) What is the relationship between magnitude of incentive and the distribution of effort (gradients) within performance? and (3) What are the effects of variation of magnitude of incentive upon level of performance?

As the problem of this study is of fundamental systematic importance, it is proper to orient it within the broad context of motivation of behavior. Analysis of the concept of motivation suggests that it involves definitively two complementary aspects. There are, on the one hand, internal physiological states which release energy and lead to activity; and, on the other hand, there are the external stimuli which serve to arouse, direct, and bring trains of activity to a conclusion. The internal motivating conditions of activity, both appetites and aversions, may be termed needs or drives; the external motivating complements may be termed incentives—positive incentives, perhaps, when they complement an appetite, negative incentives when they complement an aversion.

Though all animal responses involve a momentarily dominant physiological state directed generally toward some stimulus-pattern of the external environment in an 'interlocking' relationship, as Stone puts it, the question of the contributions of variation in either the drive or the incentive to variations in the performance of the animal is experimentally meaning-

* Accepted for publication August 1, 1942.

467
ful. If one is concerned with the influence of drive on an animal's responses, obviously the external stimuli or incentives are held constant. Conversely, if one is interested in how the factor of incentive conditions performance, the physiological state or drive is held as constant as possible.

In narrowing its consideration to the latter problem this paper parts company here with the large literature dealing with behavioral results of variation of drive.

The field of variation of incentives is itself wide. How does it break down? First, there is qualitative variation of incentives which further subdivides into what might be termed intermodal and intramodal variation. The former denotes qualitative changes of incentives which involve change in the drive if the new incentive is to be desirable, as for example, a change from food to water. Simmons' early study is a classic in this realm. On the other hand, intramodal variation of incentive denotes qualitative changes which do not disturb the congruency of the drive. This kind of incentive variation is exemplified by change from one kind of food to another, unless, of course, specific hungers are envisaged. Elliot's experiment in which he shifted from sunflower seeds to bran-mash illustrates this type of problem.

Secondly, there is quantitative variation of incentives. This variation embraces both negative and positive incentives. A typical example of the former is electric shock; and considerable study dating from an early period has been devoted to the problem of the effects of the quantitative variation of shock upon the behavior of animals. The researches of Yerkes on mice in 1907, Yerkes and Dodson on mice in 1908, Cole on chicks in 1911, Dodson on kittens in 1915, Dodson on rats in 1917, Dunlap, Gentry, and Zeigler on rats in 1931 all bear on this problem.

The specific field of this inquiry now becomes clear, i.e. the quantitative variation of positive incentives upon the behavior of animals.

---

1. R. Simmons, the relative effectiveness of certain incentives in animal learning, Comp. Psychol. Monog., 2, 1924, No. 7, 1-79.
Until very recently the only study directly concerned with the problem of the influence of quantitative variation of positive incentives on the behavior of animals was one in 1929 by Grindley on chicks.\(^9\) The chicks were required to go down a passage 4 ft. long and 8 in. wide for food placed in a small wooden tray at the end opposite the starting-box. Separate groups of the chicks were each given 0, 1, 2, 4, and 6 grains of boiled rice. On the basis of one trial per day for each chick, learning curves were drawn for the various groups in terms of scores equal to 100 times the reciprocal of the time taken to reach the food (or the tray in case of 0 incentive). According to Grindley, the curves indicated that there is a marked increase in the rate of learning with increase in the amount of reward. He admits later, however, "there were wide individual differences in behavior."\(^{10}\) In a second experiment a simple maze form was employed instead of a runway. Only one and six grains of rice were used as incentives. It was found here that a sixfold increase in amount of reward produced an increase of about 25% in the rate of learning. But, Grindley states "... it must be admitted that the individual differences between the chicks were so great that the experiment (although forty chicks took part in it) was not sufficiently accurate to establish this result with certainty."\(^{11}\)

This study really did no more than pose the problem, for besides being inconclusive by Grindley's own admission, the present paper will later show that it had many serious weaknesses both statistical and methodological. It is difficult to understand why so little further work was done on a problem of such fundamental systematic importance in motivation. It is unlikely that this neglect was due to ignorance as Diserens and Vaughn\(^{12}\) as early as 1931 mentioned in their review of experimental psychology of motivation that Grindley seemed to be the only person who had studied the relationship of food-amount to incentive-value. The inconclusiveness of Grindley's results they clearly intimated with their statement that "individual differences between chicks were so great, however, that it was impossible to establish the results accurately in quantitative fashion."\(^{13}\) Since this review in 1931, various more or less systematic works in comparative psychology such as Tolman in 1931,\(^{14}\) Munn in 1933,\(^{15}\) Stone in 1934,\(^{16}\) Maier and Schneirla in 1935,\(^{17}\) Washburn in 1936,\(^{18}\) Warden, Jenkins, and Warner in 1936,\(^{19}\) Young in 1936,\(^{20}\) and Katz in 1937,\(^{21}\) have each raised, with varying ex-

---

\(^{10}\) Op. cit., 175.
\(^{11}\) Ibid., 180.
\(^{14}\) E. C. Tolman, Purposive Behavior in Animals and Men, 1932, 39-74.
\(^{15}\) N. L. Munn, An Introduction to Animal Psychology. The behavior of the rat, 1933, 321-325.
\(^{20}\) P. T. Young, Motivation of Behavior, 1936, 295-296.
\(^{21}\) Daniel Katz, Animals and Men, 1937, 152-159.
licitness, the problem of quantitative variation of incentives as an integral part of the psychology of motivation. And in the absence of any other studies Grindley's investigation, with all its weaknesses, is more often than not summarized as the definitive inquiry.

It might have been expected that examination of any one of these texts would have indicated the necessity for more adequate exploration of the effects of quantitative variation of incentives (e.g. food) upon performance. In the period from Grindley's study in 1929 up to the time of this inquiry, however, only four papers seem to have touched upon the problem. Two were by Nissen and Elder in 193522 and Cowles and Nissen in 1937.23 These papers were concerned with the influence of amount of incentive on the delayed response performance of chimpanzees. The incentive was food, in the first experiment pieces of banana; in the second, pieces of orange and pieces of banana. Two incentive values were employed, a large and a small.24 The procedure was to place either the large or the small incentive in one of two boxes in view of the chimp. The boxes were then screened for the delay-interval, after which the animal was allowed to open whichever one he chose. It was found in the first experiment that "an increase in amount of incentive used consistently increased the delay limits."25 Decrease in amount of incentive always resulted in lower limits. Delay limits meant the longest delays at which the animals would perform with an accuracy of 80% or better under the conditions of the experiment.

The purpose of the second experiment was primarily to see if the larger incentive was superior merely because it was a bigger and better cue. If this were the case, the results could be explained without reference to any expectation which the animal might have regarding the amount of the incentive to be received. In other words the problem would then be one of perception, not motivation. It was found that the potency of the incentive was a function not only of its size as such, but also of previously experienced consequents of seeing small and large incentives placed in the container, namely, actual reception and ingestion.

Though these experiments were relatively coarse, with their use of only two values of incentive variation, they suggested a relationship between amount of incentive and strength of motivation as inferred from delayed response performances,—what Nissen and Elder term the "mental work method" of measuring motivation. The next reference to quantitative variation of incentive is apparently Fitt's brief abstract in the 1940 program number of the A.P.A.26 This study was concerned with the effect of a large and a small reward as indicated by the resistance-

24 In the first experiment this was; small = 5 grams of banana, large = 2 or 4 5-gram pieces. In the second experiment the small and large incentives varied considerably in amounts on different occasions. They were 60 vs. 12 grams, 35 vs. 7 grams, and 30 vs. 5 grams of orange; and 35 vs. 7 grams of banana. There was also a qualitative variability, some of the oranges and bananas were with skins, others without.
25 Nissen and Elder, op. cit., 71.
26 P. M. Fitts, The effect of a large and a small reward as indicated by the resistance-to-extinction curve for the rat, Psychol. Bull., 37, 1940, 429-430.
to-extinction curve for the rat. The activity was bar-pressing in a modified Skinner problem-box. In the training series the animals were rewarded once a day by a 0.2-gram or a 10-gram food-incentive for pressing a bar which could be moved downward for 10 cm. Points on the resistance-to-extinction curves were established for 1, 5, 10, 20, and 30 reinforcements at each condition of incentive. To quote Fitts: "The curve for the larger reward shows greater resistance to extinction at all points than does the curve for the smaller reward."27 Of course, from an abstract alone no critical appraisal can be made, but if these results were statistically significant, and such factors as drive were properly controlled, then again the suggestion is made that amount of incentive is an influential determinant of motivation.

The most recent paper which is concerned with quantitative variation of incentive is by Wolfe and Kaplon.28 This study considers the effect of amount of reward and consummative activity on learning in chickens. It was in large part a duplication of Grindley's experiment with the aim of distinguishing the effects of consummative activity (number of pecks in this case) from the amount of reward alone. Three situations were employed: a runway, a simple detour setup, and a single-unit T-maze. The incentive was popcorn given as one whole grain, cut into four pieces or as 1/4 of a grain. The scores were in each situation simply the time taken to reach the food. The groups of chickens were given five successive runs per session, two sessions a day. The results from the three experiments together indicate: (1) one large piece of popcorn does not significantly lower times over an incentive 1/4 its size; (2) four quarter-pieces of popcorn result in reliably lower times than one quarter-piece; (3) less reliably, it is suggested that four quarter-pieces produce lower times than an equal incentive-amount in the form of a single large piece.

These results have been derived from the tables of Wolfe and Kaplon's paper; they do not agree entirely with the experimenters' own interpretations because some of these, as the present writer feels, are not justified by the statistics. Results (1) and (2) above indicate that for small incentive-amounts among chickens, a fourfold increase of incentive does not increase speed of performance unless the larger incentive is subdivided. Whether this subdivision gives an impetus to motivation because of additional consummative activity (pecking) occasioned in chickens by subdivision (as Wolfe and Kaplon believe) or because subdivision simply results in an incentive being more readily perceived as larger, apart from eating activities, is not here determinable.

Confusion in the use of the terms 'learning' and 'performance' made it evident that an inquiry into the effects of quantitative variation of incentive could not advantageously proceed without a searching examination of the relationship between these two concepts. Grindley stated that his experiment was concerned with the influence of the amount of reward on 'learning' in young chickens. Diserens and Vaughn,29 Washburn,30 and Katz31 did not, in their considerations of the study, question Grindley's envisagement. And Wolfe and Kaplon,32 in their recent attempt

29 Diserens and Vaughn, op. cit.
30 Washburn, op. cit.
31 Katz, op. cit.
32 Wolfe and Kaplon, op. cit.
to separate out the effect of amount-of-consummative-activity from amount-of-reward, conceived the dependent variable to be learning. Stone, however, in referring to Grindley's experiment looked upon the dependent variable as 'level of performance.' Maier and Schneirla spoke of the dependent variable, in relation to one of Grindley's results, as 'maze performance.' And significantly in this latter case the study was included not in a chapter on maze learning, but in a chapter on performance in learning situations. Young spoke of the effects of incentive-amount-variation not on learning but in 'motivating' learning.

Following somewhat in the steps of Lashley, Blodgett, and Bruce, the writer wishes to make in this paper a distinction between learning and performance. The latter concept is to be taken as the more general. Performance is the behavior of the animal in a specific situation and includes learning as one of its determinants. It includes, however, other determinants, principally those embraced by the term 'motivation.' Learning is an animal's capacity to perform, motivation is, in a sense, his desire to perform.

The conception that the effect of variations-of-incentive-amount is upon learning rather than upon performance probably arises from an uncritical use of the terms, and is in any case gratuitous without rigorous elimination of motivation as a source of variation. The effect of incentive-amount-variation is much more likely to depend rather upon the motivational aspect of performance than upon its learning aspect. Especially is this true if different levels of performance can be demonstrated to be associated in various groups of animals with different incentive-amounts when learning as a source of difference is largely eliminated, e.g. when a relatively simple task has been substantially mastered by all.

For convenience in this inquiry a further distinction will be made, namely, between the learning period and the performance period. In the learning-period the activity is reflecting both learning and motivational factors. The performance-period ensues in the progress of adaptation to a situation when the contributions of learning to the performance come substantially to an end, leaving motivation as the primary determinant.

Apparatus. Several considerations guided the design of our apparatus. First, time rather than errors was to be the dependent variable. Simmons, in her early study, in comparing the relative effectiveness of such diverse incentives as bread-and-milk, escape from apparatus, and a female in heat, found that time-differences among the

33 Stone, op. cit.
34 Maier and Schneirla, op. cit.
35 Young, op. cit.
various incentive groups were much greater than error-differences. She suggested by way of interpretation that the more effective incentives re-enforced impulses which dominate and tend to suppress time-consuming extraneous impulses operating under less effective stimulating conditions. It is possible, she states, that "one of the chief points of difference among the various incentives is their relative effectiveness in suppressing irrelevant activities, and for this reason time may, in this case, be even a better criterion than errors." 39

The superiority of time-differences to error-differences in Simmons' study suggested that time is a more direct index of motivation as reflected in performance; whereas errors may be a superior index of learning. As this investigation primarily concerns itself with the former, the use of time-measurement is indicated. Further, one of the problems deals with gradients of performance. This necessitates a situation having no error pathways for, as Hull remarks, 40 in the maze the situation is so complicated by the various interactions of excitatory and inhibitory factors involved in the elimination of the blind alleys (to employ his terminology) that a possible speed-of-locomotion gradient might be greatly distorted if it were manifested at all.

It must be recognized, however, and Drew has recently remarked upon the matter in a similar context,41 that time suffers from certain disadvantages. Time is much more sensitive than errors to distractions in the experimental situation. Uncontrolled physical conditions of the organism, e.g. parasitic irritation, may make the animal pause to scratch, thus affecting a time-score, whereas an error-score under the same conditions would have little likelihood of being affected. More importantly, uncontrolled extraneous stimuli—such as an unusual sound, a bump on the nose in entering or leaving the starting box, or some fear-producing happening in the home cage—may markedly affect time-measures. Most of these extraneous and uncontrolled factors, it is apparent, act unduly to prolong the time-scores. To counteract this lengthening error, the writer believes that the observer must be watchful of such occurrences in the experimental runs, and where prolonged times are reasonably correlated with observed extraneous factors, the results should be discarded.

A second consideration in the design of the apparatus was simplicity. In endeavoring to investigate motivation as reflected in performance it is advantageous to minimize the learning-factor as far as possible. A simple situation gives both a learning-period and a performance-period within which to view the operation of differential incentive-amounts. A complex situation would put the performance-sector, that is to say, the region where substantially no more improvement is likely, far beyond a convenient experimental time. Further, a complex situation might increase the variability of groups by introducing a factor of varying mastery of the situation. This would, of course, make the presence of a significant influence of amount of incentive more difficult to detect, thus decreasing the efficiency of the experiment.

39 Simmons, op. cit., 76-77.
40 C. L. Hull, The rat's speed of locomotion gradient in the approach to food, J. Comp. Psychol., 17, 1934, 393-422.
A third feature of apparatus-design to be considered was observability. The linear runways employed by Hull\textsuperscript{42} and Drew\textsuperscript{43} in their studies of speed-of-locomotion in the rat satisfied the criteria of simplicity and temporal measurement, but were deficient in this characteristic. The present investigator wished to observe the rat as well as the records. In Hull's apparatus with its cover of heavy-wire screen, observation would, of course, be impossible. Drew also had a wire-covering screen, as well as non-retracing doors at the end of each section of his runway wired to lights on the experimenter's table. This arrangement "allowed the experimenter to check the animal's progress through the apparatus without being visible to the animals," and, one might add, without the animals being visible to the experimenter.

Wire-covering screens and doors for either timing or the prevention of retracing or both, were avoided in the apparatus of this experiment in the interests of observability. This was a fortunate precaution as some of the most interesting results were derived from qualitative observation of the rat, rather than quantitative calculations from the records. Another reason for avoiding the doors was that they might be the source of artifacts. To interpose successive bristol-board barriers to a rat dashing full tilt in order to measure his speed would hardly seem to conform to the basic precept of measurement, \textit{i.e.} that the measurements must not alter the variable measured.

Of course with the elimination of doors in a runway one must be prepared for some retracing of path by the rat. The experimenter feels that in a study endeavoring to get at motivational forces, far from being a difficulty, retracing should certainly be allowed to make its contribution to the index of performance. All other factors being equal, a group of rats which retrace more than another group can be assumed to be less motivated. Simply including the time spent in retracing as part of the time of transit is as expedient a way as any of including this factor.

Again unlike Hull and Drew, and partly in the interest of observability, the food-box instead of being simply a prolongation of an alley 4 or 5 in. in width was expanded to a large square inclosure. This larger setting allowed a freer expression of any behavioral effects of changed conditions of incentive.

Concrete Aspects. A. Runway: The experimental situation was a linear runway 20 ft. in length with 3/8-inch plywood sides extending 1 ft. above the 2 x 4 base. The starting-box was simply a 1-ft. continuation of the runway at its initial end; the foodbox was a 14 x 14 in. square enclosure at the other end. Sliding doors marked off the entrance to the starting-box and the exit into the alley proper. The exit door could be operated by the experimenter by means of a line-and-pulley system from his position overlooking the foodbox. From the same place $E$ could operate a door, swinging in from a lateral position, to close the foodbox from the runway after the entrance of the rat. The entire runway was open to the view of $E$ to enable him to observe the qualitative features of the rat's behavior from the time of his exit from the starting-box until he was removed from the foodbox. To eliminate the possibility of disturbance of the rat by direct observation when in the foodbox, a mirror was adjusted that brought the inclosure clearly into

\textsuperscript{42} Hull, \textit{op. cit.}

\textsuperscript{43} Drew, \textit{op. cit.}
view. Observation was thus direct when $E$ was beyond the rat's distance-perception, and indirect when within that range.

An adequate illumination for observation was provided by four 60-watt bulbs set in 10-in. reflectors which were placed one each above the foodbox and starting box, and the remaining two equally spaced above the alley. The lights were hung 30 in. above the runway in its position on a shelf set 3 ft. above the floor. The entire apparatus was painted a dull non-reflecting black.

B. Timing System: In Exper. I the timing setup was designed to measure the time taken to traverse successive quarter-sections of the 20-ft. runway. It consisted essentially of, first, a simple and efficient kymograph designed by Dr. Bray. This was operated by a constant-speed motor employing a roller which moved the paper $\frac{1}{2}$ cm. a sec. By merely employing the metric scale of an ordinary ruler the records could be easily read to $\frac{1}{10}$ sec. As only two electromagnetic recording siphons were necessary, adding-machine paper of the narrowest width was adequate. One of the ink-recording siphons was connected in circuit with an electric clock which prompted a deflection every half-sec. This time-record checked the paper-speed against any slippage that might occur. The other siphon, powered by a 4-volt bank of Edison cells, was connected in circuit with, successively, a switch on the exit-door of the starting-box, a photo-cell relay, and three treadle-switches. The relay and treadle-switches served the same purpose—to record the passage of the rat. Both were included that $E$ might observe their relative efficiency. Each treadle was simply constructed of an $8\frac{1}{2}$ by $3\frac{3}{8}$ in. rectangle of galvanized iron set upon transverse length of brass $\frac{1}{8}$ in. in diameter. The treadle was attached to the brass roller slightly off center so that at rest the end facing the starting-end of the runway, being slightly heavier, would lie flush with the floor. The other half remained about $\frac{1}{8}$ in. above a copper contact-strip fastened to the floor beneath the treadle. When a rat stepped upon the further side of the treadle his weight would tilt it the $\frac{1}{8}$-in. necessary to complete the circuit and occasion a deflection in the recording siphon. When the rat would step off, the treadle in returning to its former position would break the circuit.

The photo-cell relay was a model 201-R made by the Warner Products Corporation. A 60-watt light-source with a gathering lens was trained on it through the walls of the apparatus via 2-in. circular apertures covered on the light-source side by a deep red Corning filter, and on the other side by a plain glass slide. As the eye of the rat is quite insensitive to deep red light, this beam occasioned no disturbance to the rat while effectively activating the photo-tube. The rat in breaking the light-beam would momentarily close the electric circuit and cause a recorded deflection of the electro-magnetic siphon in the same fashion as the treadle-switches.\(^4\)

The contact made and broken by the opening starting-door and those of the photo-relay and treadle-switches set respectively at the beginning and successive

\(^4\)It may be remarked here that the red-light-source photo-relay arrangement proved an excellent means of recording the passage of the rat past a point. The treadles, while adequate for our purpose, were inferior in two obvious respects. Unlike the red beam to which the rat seemed oblivious, the treadles required for some rats a brief adaptation; secondly, their precision was less. The treadles will record a rat at a point though he vary an inch or so on each side of it, depending on where the foot is placed in depressing the treadle. There is none of this indeterminancy in tripping the photo-relay.
five-foot sections of the runway, were wired to switches placed at E's hand at his
station overlooking the alley. The wiring was such that any contact would be
thrown out of the circuit at will by opening the proper switches. Consequently,
if any rat exhibited an inclination to retrace, the contacts which he had already
traversed would be switched out of the circuit. With this procedure the one
siphon could record for all the measurement-points; for E thus saw to it that
the deflections represented successive contact, not repeated activation of the same
contact. In the quarter-sector time-measurements of Exper. 1, retracing time (as
can be seen from this procedure) was added to the time of the sector from which
the retracing occurred.

In Expers. II and III of this inquiry, where gradients were no longer part
of the aim of the investigation, merely total runway-times were taken. Initially
this was done with the kymograph-system described. It was shortly found, however,
that for these total times a stopwatch had all the accuracy that was necessary. Hence
the treadles were removed and the use of the kymograph discontinued. The photo-
relay was disconnected, but the light-source was retained, as the rat's nose in the
light proved an efficient way of signaling the beginning of his run.

C. Pellet Apparatus: Considerations raised in the discussion of the chosen
experimental values of the independent variable—food quantity—to be found below
under general procedure, suggested that a small unit of food, relatively constant in
size and available in quantities, would be desirable. Skinner's pellets were ruled
out because they were somewhat too large, 1/15-gram on the average. The following
procedure was finally devised. A small grease-gun was employed with its usual
head replaced by a machined nozzle 1¼ in. in length with a 5/32-in. bore. A food
paste consisting of equal portions by weight of powdered Purina dog-biscuit and
water was placed in the gun and squeezed out in cylindrical ribbons. Quantities of
these were then allowed to dry for approximately ½-hr.; they were then gathered
in groups of ten or so placed side by side, and cut into ¼-in. sections. With 10 or
12 units per cut, several hundred could be produced in a relatively short period.
Some care in judgment of length in cutting was all that was necessary to produce
reasonably constant units approximating 1/50-gram in weight. These pellets when
placed in the food cup during the experiments were thoroughly dampened with
water by means of an atomizer.

**General Procedure**

*Equalization of drive.* An obvious necessity of general procedure in any
investigation of incentive as a factor in the motivation of performance
is the control of the complementary internal aspect of motivation—the
drive. Variation of performance cannot be ascribed to variations of
incentive-amounts if drive is varying simultaneously. Only with drive held
constant can this interpretation be made. And, of course, always with the
assumption that some uncontrolled systematic factor did not covertly
intervene to produce the results. There must certainly be a drive; for
incentives can only be defined as objects appropriate to the satisfaction
of a drive; but the amount of drive, however experimentally specified,
should not be a variable factor.
With the various incentive-groups obtaining different amounts of food in their experimental trial, a first feature of drive equalization was to feed each rat individually up to a constant amount (varying with body-weight), including the amount of food obtained in the experimental run. This procedure would leave the rat each day, just before his experimental run, with approximately a twenty-two hour deprivation from a weight-maintenance ration. Also taken into consideration was a factor stressed by Bousfield and Elliot in their discussion of the experimental control of the hunger-drive. These writers hold that the hunger-drive is probably more dependent upon occurrence of hunger-contractions than simply upon the period of deprivation. So the former must not be left to chance but must be established in a period which gives hunger-contractions at the projected time of experimentation, not at varying times related to the food habits of the rats before experimentation. Varying rhythmic effects were eliminated in the present experiments by a ten-day pre-experimental feeding-period to condition stomach-contractions to the anticipated time of experimentation. Bousfield and Elliot in another paper suggested that after the rhythm is established, in maintaining constant drive, food-consumption need not be controlled by weighing the food or the animals. Instead, the animals were to be allowed free access to the food for a definite period; 90 min. was suggested. The present experimenter, however, decided to control both rhythm and food-intake, as he does not have entire confidence in the view and, further, a 90-min. period was not conveniently available in which to feed the animals.

The next feature of drive-control was to introduce a variable and appreciable delay between the experimental run and feeding period, with the aim of maintaining the vitality of the experimental situation, especially for the small-incentive groups. If the animals were fed with their maintenance-ration immediately after their experimental trial, they might in a very few trials be reacting in the runway not to the test-incentive alone, but to the complex test-incentive-plus-subsequent-feeding-ration. This complex being a constant, relative to the weight of the animal, would then tend to incite constant performance, completely obscuring any tendencies the rats might have to manifest differential performance to varying amounts of incentive in the test-situation.

The introduction of a period of delay between the test and the feeding situation diminished the danger of associations between the two among the rats. The delay averaged 3/4 hr. This delay is appreciable, though a longer one would have been desirable had the schedule allowed it. The delay was also variable, ranging from 1/2 hr. to 1 hr. Presumably this would also function to diminish undesirable associations by making more difficult the establishment of any anticipations or temporal conditionings.

Because of considerations of drive-control the present experiments were inescapably limited to one trial per day. With various groups of animals receiving different amounts of incentive in their initial trial, upon any subsequent trial the assumption can no longer be made that drive is constant. On the contrary, the drive will now have varied inversely as the amount of incentive that each rat obtained. Performance after the initial trial, therefore, would be a function not

---


only of incentive-variation, but of drive-variation. The only simple way to circum-
vent this intra-series variation of drive is by limiting the trials to one-per-day for
each rat.

*Values of incentive variation.* A. Discrimination of Quantity: Values of the
independent variable were employed which were sufficiently dissimilar to be
readily discriminable, that is to say, readily discriminable in the sense that if both
were offered in a choice-situation under suitable motivation, other factors being
equal, one, presumably the larger, would be chosen most of the time. It is con-
ceivable, though unlikely, that difference in amounts of incentive not discriminable
in this sense would result in significantly distinct levels of performance. The present
investigation, however, is not concerned with this problem. Easily discriminable
incentive-amount values avoid the complexity of varying discrimination-capacity
among the rats. They have also the advantage that, in investigation of the effect of
shift of incentives for the rat from one amount to another—the third problem of
this paper—, the contrast is perceived.

The literature on the problem of discrimination of quantity in lower animals is
scarce. Katz has discussed the comprehension of quantity and number in his volume
*Animals and Men.* He cites two studies there which bear on the discrimination
of quantity. Experiments by Révész revealed that a hen prefers 3 to 2, 4 to 3,
5 to 4, and 6 to 5 grains of corn. In this case it was of little account how the
individual grains were grouped together. Fischel, in similar experiments with the
goldfinch, found that it was not difficult to train the goldfinch always to choose
the larger of two relatively large quantities. When, for example, the goldfinch
had learned to choose the larger group of a 12-6 pair, it did the same thing when
presented with 16-8, 20-10, and 14-7. But when it had learned to distinguish
two smaller groups in the same two-to-one ratio (e.g. 6 from 3) it needed fresh
practice on the 8-4 pair to choose the larger.

With the hen successfully discriminating six grains of corn from five, on the
one hand, and the goldfinch having difficulty in some cases with a two-to-one ratio,
on the other, extrapolations to the rat could not be other than precarious. The
writer has found no specific published study with rats as subjects, unless one
would consider relevant an inquiry by Kuroda on the discrimination of number
given in the form of an acoustical sequence. This investigation employed a T-shaped
box with a call-bell presenting numbers in the form of successive strokes. The
results showed that though 1 seems to be discriminated with difficulty from 2 or 3,
the white rats were able to differentiate 5 from 1, 2, 3, and some succeeded in the
differentiation of 5 from 4. This result can only be taken as suggestive; the use
of the same term 'number' must not obscure the fact that number of food pellets
may be a far different matter from number of bell-strokes.

More directly pertinent data were obtainable from a previous study of the

---

47 Katz, *op. cit.*
48 G. Révész, Tierpsychologische Untersuchungen, *Zsch. Psychol.*, 88, 1921. [Cited
in Katz, footnote 21.1]
49 W. Fischel, Haben Vogel ein Zahlengedächtnis? *Zsch. f. vergl. Physiol.*, 4,
1926. [Cited in Katz, footnote 21.1]
50 R. Kuroda, Discrimination in number given in the form of acoustic sequence in
present writer, where it was found that in a T-box-choice situation, with training, six segments of meat were discriminable from three. It was a more difficult problem for the rats to discriminate one unit from two. These results and other considerations from the study suggested that a 3-to-1-amount ratio was within the discriminatory capabilities of the white rat. To be well on the safe side in the present experiment, fourfold variations were taken as the experimental steps in the independent variable.

B. Range: The unit-incentive was, as pointed out in the apparatus-section, 1/50th-of-a-gram pellet of Purina dog biscuit. Fourfold progression from this base gives 1-4-16-64-256 unit-incentives. This is as many fourfold increments as can be used without affecting drive. 256 units is somewhat over five grams of dry-weighed food. The average maintenance-ration for the rat-samples, taking all the experiments together, approximated 7½ grams of Purina powder. Hence, to introduce any more experimental steps in the variation of incentive-amounts would give more food in the experimental situation than the rat should have for the day, and would thus dislocate the drive-control procedure.

The reason for employing units in the varying amounts of incentive was, if possible, to re-enforce discrimination. With amount-differences broken into units, the discrimination of amount may be aided by a discrimination of number. At the 64- and 256-levels it is very unlikely that number would offer any aid, so, at these steps, undivided amounts were employed. Furthermore, in the interest of practical considerations of time, the 64- and 256-units were mixed with an equal weight of water to make them more speedily edible. To minimize any qualitative change this expedient might have introduced, in the smaller incentive steps, the food-pellets were thoroughly moistened with an atomizer.

Since the average maintenance-diet for the rats was approximately 7½ grams of powdered Purina biscuit, the experimental values of 1/50 to 256/50 grams of food covers about 70% of the maintenance-range of the appetite.

Experimental design and statistical considerations. Analysis of variance offered an effective statistical instrument to apply in this inquiry, in the form of a simple design in which different treatments, in this instance the varying amounts of reward, would be assigned at random to different groups of rats, all drawn from the same population. The logic of the Null Hypothesis was employed. This hypothesis states that the parameter is zero, which is to hold, in this setting, that the average performance-scores of the different food-amount groups are random samples from the same population. Upon this hypothesis, with any set of r-groups of n-cases each, it follows from the basic proposition of the analysis of variance that one may derive two independent estimates of the population-variance: one which is based upon the variance of incentive-group means, the other on the average variance within the various incentive-amount groups. The test of the Null Hypothesis lies in ascertaining whether the ratio F between these estimates falls above or below

---

81 L. P. Crespi, Gambling behavior: a comparative approach. I. Constant-goal vs. gamble-goal in the albino rat. Paper read at Atlantic City meeting of Eastern Psychol. Assn., 1940.
82 This is E. F. Lindquist's use of the concept (Statistical Analysis in Educational Research, 1940). Fisher (Design of Experiments, 1937) embraces by the term any exact hypothesis one may be interested in disproving, not just the hypothesis that the parameter under consideration is zero.
the value of $F$ in the table\(^a\) which corresponds to the selected level of significance.

This $F$-test of the ratio of treatment-variance to error-variance is a powerful but relatively unspecific tool. In these experiments, it would indicate whether treatments in general (the five different incentive-amounts taken simultaneously) make any difference in performance among the rats; but it would not tell how the individual incentive-values compare among themselves in their determination of performance. However, the purpose of Exper. I was essentially exploratory, to ascertain whether or not there was a difference in performance present in general under conditions of varying amounts of food, such as to warrant more specific investigation. A further aim of Exper. I was to obtain some notion of the magnitude of the possible effects in order to estimate how large the samples had to be reliably to detect with the least expenditure of effort whatever phenomena were present.

In spite of the relatively small number of animals utilized in Exper. I, the results were so gross that the application of an $F$-test to verify performance-differences, taking the five incentive-variations simultaneously, was obviously unnecessary. It was possible to proceed directly to a $t$-test analysis between adjacent incentive-amount steps. The $t$-test is essentially a particularized $F$-test, as the latter was to be employed in this experiment, namely to the treatment- and within-group variances. If there are only two treatments, the $F$-test becomes algebraically equivalent to the $t$-test. The advantage of the $F$-test is that, in applying $t$-tests to all treatment-differences simultaneously, it may reveal significant results where individual $t$-tests would not. Of course with this gain in significance, there is a loss of specificity of conclusion.

The results of Exper. I suggested the variability and order-of-magnitude of results to be such that small-sample theory would be sufficiently precise to answer the specific questions as to significancies of adjacent incentive-steps. Hence in Expers. II and III, though the groups were increased in size, the use of small-sample analysis was continued.

The level of significance chosen in these experiments was 5%. It would have been inefficient perfectionism had we employed a level of significance as conservative as 1% in this inquiry, as the experiments are set up in such a fashion that a number of results bear on the same question. Any generalizations which arise obtain their significance not only from a local result, but in relation to a considerable pattern. In every case, however, the actual probabilities that the result could have arisen by chance will be given and the reader may draw his own conclusions.\(^b\)

In the application of the $F$-test to the ratio of between-groups and within-group variances, a fundamental assumption known as the homogeneity of variance must be made. In other words, it must be assumed that whatever factor is producing significant differences in the means of the various treatment-groups will not produce significant differences in their variability. On the contrary, to apply legitimately the $F$-test in this case, the variability-within-groups must be the same except for

\(^a\) Snedecor's table of $F$s as obtained from Fisher's $Z$-statistic.

\(^b\) The assumption is common, it may be noted here, that it is entirely arbitrary what level of significance is chosen. This assumption is not correct as Lindquist aptly points out. The choice should be guided by consideration of the relative danger involved in accepting (1) the Null Hypothesis as true when it is false, by insisting on a high level of significance, or (2) of accepting the Null Hypothesis as false when it is true, by employing a low level of significance.
chance differences from group to group. The t-test is also affected by the factor of homogeneity in its interpretation. When a value of t exceeds that specified as significant, one may reject the Null Hypothesis that the samples were drawn at random from the same or identical populations. Though this is equivalent to saying that the samples came from different populations, it is not equivalent to saying that the means of these populations necessarily differ. It is possible, though improbable, as Lindquist points out, that the samples came from populations whose means are the same but whose standard deviations differ. In the analyses of results in this investigation, with every application of the t-test there is also a small sample-test for the significance of whatever differences in variability were obtained. This significance-test utilizes the same F-statistic as employed above. The F in both cases is based upon a distribution-function Z as developed by Fisher. In the test of homogeneity of variance within two groups, the ratio of the within-group variances is employed in the test with the larger variance in the numerator. With such a prior analysis of the observed differences in variability, one can rigorously infer with a t-test whether or not significant differences in means exist.

An integral part of the statistical basis of an inquiry is, of course, the sampling and controls employed. In small-sample theory this is important enough to warrant detailed consideration for each experiment.

Experiment I: This has been characterized as the initial exploratory experiment. In it were employed twenty-two albino rats, divided as follows among the five incentive-amount values: six animals which were run for one-unit incentive, and four rats each which were run for 4-, 16-, 64-, and 256-unit incentive-amounts.

There were two limitations on the population from which these samples were drawn, which of course must be taken into account in any generalizations from the samples. The sampling was not of both sexes, but was limited to males alone. The purpose of this exclusion was to avoid any difficulties of analysis which might be introduced by a relationship between oestrous cycle and activity-level. To the extent that sex-differences in the effects of varying incentive-amounts on performance are suspected, then, conclusions obtained upon males could not be extrapolated to females. However, sex-differences in rats are so few in the literature that the risk in such extrapolation would be minor.

A second limitation on the population related to age. The range of the samples was relatively small, from approximately 5½ to 6½ months. Since many experiments have revealed age-differences in the activity of the rat, investigation of some

---

55 However, and here Lindquist is in error, in testing the homogeneity of variance by the variance ratio, the 5% and 1% points listed in Snedecor's table for F become approximately 10% and 2% levels of significance respectively. In the Analysis of Variance the F-table is used to test the Null Hypothesis that the means of various groups are the same by observing whether or not the between-groups variance is significantly greater than the error-variance. The 5% and 1% levels of significance are placed upon the larger end of a sampling distribution beginning at 0. The variance-ratio, on the other hand, with the larger variance placed in the numerator, tests the hypothesis that the ratio = 1. No cases can fall below one, hence the cases from 0 to 1 in the first distribution must be included beyond 1 in the sampling distribution of the variance-ratio. It can be shown that this roughly doubles the number of cases that fall beyond the 1% and 5% points in the Analysis-of-Variance distribution, making them in the variance-ratio distribution approximately 2% and 10% levels of significance respectively.

56 Actually 23 rats began Exper. I and one rat was dropped from the 16-group because of persistent refusal to run associated with inordinate emotionality.
other age-range rather than extrapolation is needed to widen any conclusions. A partial purpose of Exper. II was to do just this.

The a priori possibility was envisaged that runway speed, the dependent variable, might be in part a function of the weight of the rat. Instead of randomizing this variable, an improvement was attempted in the experimental design by equating weights for each incentive-group. The five groups were randomly selected, subject to the restriction that the mean weight of each group be substantially similar. The range of weights for the entire group of rats was from 152 to 220 grams, with a mean of 188 grams. The final variation in means for the various incentive groups was within 5 grams.

The substitution of an equating restriction for randomization in the treatment of weight, had this variable actually proved related to speed, has the advantage that it makes possible the subtraction of the variance due to weight from the within-groups variance and thus improves the efficiency of the experiment by reducing the error variance—the denominator in the F-test.

Frequent comparisons of running speeds with weight in the course of Exper. I, however, revealed no appreciable relationship between the two variables, so a 3-component analysis of variance was not indicated.

The sample employed in Exper. I was from the stock of the laboratory. As these rats had been long inbred, strict attention was not given to complete randomization of litter. Later in the paper some possible effects of litter will be considered.

Experiment II: The primary aim of this experiment was to check and to strengthen the conclusions about the lower levels of incentive-amounts. A total of twenty-one rats were here employed, divided into three groups of seven each. The groups were run with 1-, 4-, and 16-unit incentives.

Again, as in Exper. I, only males were employed. The age-sampling was, however, from a different range in order to enlarge the implication of the results. The range here was from approximately 2½ to 3½ months. Also in the interest of widening the applicability of the results, the stock was not of the Princeton Laboratory, but was obtained elsewhere. Weight was randomized instead of equated, as Exper. I did not reveal any significant relationship between it and speed. However, the variation in the mean weight for the three groups, it may be noted, was only three grams (138-141 grams). The range of weight for the entire group was 121 to 156 grams, with an average at 140. The litter factor was thoroughly randomized.

Unlike the method of Exper. I, the entire group of rats were first run at 16-units reward until their performance approached a level. Then on the basis of the last 6 out of the 21 trials required, the 3 groups were equated for average speed of running. Two groups were then transferred to 1- and 4-unit rewards respectively, and the third remained at 16- for a control comparison. The matching procedure was to assign rats of similar average times for the matching period to the 3 groups, such that the 3 distributions would be substantially similar. The matching means of the 3 groups varied only 1/10 sec.—from 9.37 sec. to 9.47 sec. It was found, however, in Exper. II that the performance of the rats at the various amounts of incen-

---

Exper. II began with 24 rats, 3 were discarded in the matching process described below.

A group was also run with 0-incentive conditions in Exper. II, but as it involved some special considerations it will be described later.
INCENTIVE AND PERFORMANCE IN THE RAT

Incentive was not significantly related to their performance during the matching period. Hence the efficiency of Exper. II was not improved by the matching procedure.

Experiment III: The primary aim of this experiment was a re-examination of the upper incentive levels with larger groups and an attempted refinement of the method of Exper. I. Twenty-four rats were employed in this experiment, ten in a 16-unit reward group, and seven rats each at 64- and 256-units incentive.

Comparable to the conditions of Exper. I, the animals again were all males; the ages were from 6 to 7 months; and the stock was that of the Princeton Laboratory.

The weights in this group varied from 121 to 203 grams with an average of 162. Weight was randomized among the 3 incentive-amount groups. The mean weight of the 3 groups differed, incidentally, by less than 10 grams.

The litter-variable, instead of being randomized, was split up among the 3 groups as equitably as possible. As the variance due to litter was not subtracted from the error-variance, this quasi-split-litter technique had the effect of making the significance-tests in Exper. III more conservative by somewhat over-estimating the error-variance and by reducing the variance-between-groups.

Thirty-one rats actually began Exper. III. On the basis of the first trial (before the experience of differential incentive-amounts) the groups of animals running to 16-, 64-, and 256-unit rewards were equated in speed. By discarding seven rats, the average speeds of the 3 groups were made substantially similar. This was an attempted improvement over Exper. I, where the rats were simply assigned at random. No large hope was held for this attempt because of the probable unreliability of one day's run, and, as the experimenter anticipated, the correlation between the equating performance and final performance was not sizable enough to contribute to the efficiency of the experiment.

A final problem which might advantageously be discussed under statistical considerations is that of metric. Though time-measurements are unequivocal either for the twenty-foot runway, taken as a unit, or for five-foot sections of the runway, speed may have two distinct meanings. For any single measurement, time and speed are mere reciprocals. But if one has a number of time-measurements, say for the running of the twenty-foot path by a group of rats, the average speed of the group will be quite different depending upon whether one divides the distance by the average of the times to obtain the final average speed, or divides the distance by each time, and then averages these speeds to obtain the final average speed. In formal terms, this asserts that the reciprocal of the mean of a set of values is not equal to the mean of the reciprocals of those same values. The former is always smaller and, as the range of the values increases, the discrepancy can become very considerable.

Hull, particularly, in his studies of speed-of-locomotion of the rat in a runway, has understood speed in the former sense. That is to say, the average speed of a group for a particular distance is to be obtained by dividing the distance by the

---

69 The experiment began with 31 rats, 7 were eliminated in the matching process to be described.
60 Such subtraction would have been impracticable as the litters were not split up perfectly, e.g. some 5-rat litters were divided among the 3 groups. Further the litter-distribution was somewhat affected by the discard of several rats in the speed-matching procedure described below.
61 Hull, op. cit., footnote 40.
average time taken by the group to traverse the distance. If distance is made unity, the average speed is simply the reciprocal of the average time. The present experimenter feels that average speed is more appropriately the mean of individual speeds obtained by averaging the reciprocal of each individual time.

Before discussing the considerations bearing on the relative propriety of the two conceptions, a simple example might be considered to sharpen the appreciation of the distinction. Two individuals run the 100-yard dash. One does it in ten seconds and hence runs ten yards per second; the other takes fifty seconds and hence runs two yards per second. What is the average speed of these two individuals for the distance? The present experimenter, averaging the individual speeds, would say six yards per second. Hull’s procedure in dividing the 100-yard distance by the average of the times would yield an average speed of 3 1/3 yards per second.

In Hull’s procedure the assumption is implicit that the difference between a rat running a twenty-foot runway in five seconds and in ten seconds, is the same as the difference between, say, 95 and 100 seconds for the same distance. This assumption is unreasonable. In endeavoring to get at effort as an indication of motivational forces activating the rat, it is much more reasonable to suppose that the difference between ten and five seconds for the running of a twenty-foot distance represents a substantially greater difference in effort than indicated by the difference between 95 and 100 seconds. The difference in effort in the former case is considerable, in the latter case it is probably negligible. The metric average-mean speeds, unlike average speed, embodies this difference in effort; for though the difference in time is the same in the two cases cited, the difference in speed is properly much more considerable in the first case—being a difference of two feet per second—as compared in the second case to a difference of about 1/100 feet per second.

It is fundamental in the application of a metric to a set of data, that, unless systematic differences in variability are quite likely to be present, the metric should yield similar variabilities of scores on all portions of its range. If, on the contrary, the variability of score on the lower end of a scale is widely different from the variability at a higher portion of the scale, the suspicion legitimately arises that the units, instead of being constant, have in a sense stretched or contracted. In the earlier stages of this investigation, time-scores were employed for averaging instead of their reciprocals (mean-speeds). It was found in Exper. II, e.g. that the significance-test for the obtained differences in variances gave an F of 72.4 for the comparison of the 1-unit and 4-unit incentive-amount groups, and an F of 41.2 for the 4-unit and the 16-unit incentive-amount groups. The 2% level of significance with 6 degrees of freedom in the numerator and denominator is 8.47! The possibility is very unlikely, it was reasoned, that the differences in incentive-amounts could occasion such differences in variability. The misgiving arose that time was not an appropriate metric in this context; the units at the larger end of the scale being much smaller, so to say, than the units at the lower end and thus artifactually expanding the variabilities for the larger times. When the mode of analysis was changed to averaging not time but their reciprocals (=speed), the 1-4 comparison-F

42 This term can properly denote the notion of speed which we are employing. In taking the reciprocal of the times required to traverse an interval, one obtains the mean speed over the interval. An average of these yields the average-mean speed.

43 The alternate notion of speed may be termed the average speed. This is proper since it denotes the distance over the average time.
became 1.06, and the 4-16 comparison-F became 1.05. Such Fs, which could readily arise by chance, indicate that in terms of average mean-speed, the variances for the 3 incentive-amount groups were the same except for chance differences. These results are much more believable, and in a sense validate the metric.

The significant variability-differences obtained with the averaging of times instead of speeds would, of course, have made it impossible to employ the t-test for evaluation of the significance of obtained differences in means of performance of the various incentive-groups. It has been pointed out previously that the t-test can only be so employed when there is no significant difference in variance between the groups. There were no significant differences in variances employing the concept of average mean-speed so the t-test was applicable.

RESULTS

PROBLEM I

What is the relationship between magnitude of incentive and the level of performance?

Exper. I. The results of this experiment appear in Table I. The performance-period, namely a block of trials after the point at which sizable learning-changes in the activity are no longer apparent, extended in this case from trials 21 to 25 inclusive. The average mean-speed (in ft. per sec.) of each rat within the various incentive-amount groups for this performance-period are here listed from greatest to least. The means of each incentive-group, it is easily observed, increase from 1-unit incentive-amount to 64-units.

Table II presents the small sample-analysis of the experimentally obtained differences. Each fourfold increase of incentive-amount within the range of Exper. I is paralleled by an increase in level of performance as measured by average mean-speed, which is significant at better than the 5%-level. The difference between the 16- and the 64-unit incentive-groups is especially reliable; the significance being beyond the .1% level.

As has been previously stated, the use of the t-test for the appraisal of the significances of obtained differences in means is only legitimate when the groups compared are homogeneous in variance. In Table II the

<table>
<thead>
<tr>
<th>Incentive-amount</th>
<th>N</th>
<th>Average mean-speed in feet-per-second</th>
<th>M</th>
<th>σ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>.95 .76 .45 .32 .30 .22</td>
<td>.500</td>
<td>.078</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1.33 .95 .80 .73 — —</td>
<td>.953</td>
<td>.070</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>2.06 1.90 1.56 1.08 — —</td>
<td>1.06</td>
<td>.19</td>
</tr>
<tr>
<td>64</td>
<td>4</td>
<td>3.88 3.74 3.51 3.33 — —</td>
<td>3.62</td>
<td>.057</td>
</tr>
</tbody>
</table>
variance-ratio test of homogeneity (F) indicates that the obtained differences in variances within the three incentive-amount comparisons do not differ except by chance.

A 256-unit incentive-group was also run in Exper. I. Its results are not included, however, because of possible bias. An effect of litter upon level of performance was not anticipated at the beginning of our inquiry. Examination of litter-differences later in Exper. I revealed that it might be a factor of some importance. In the assignment of rats to the various incentive-groups care was not taken in every case completely to randomize the litter-variable. The 1-, 4-, 16-, and 64-unit incentive-groups were of varied litter representation, but the 256-unit group, added to extend the incentive-range, was constituted of but one litter. In Exper. III this difficulty is remedied and the level of performance associated with 256 units of incentive is properly indicated.

In Fig. 1 the results of Exper. I are graphically summarized. For the block of trials of the performance-period a positively accelerated curvilinear relationship appears between logarithmic increase in incentive-amount and arithmetic increase in average mean-speed.

Certain significant results of Exper. I were obtained from qualitative observations that were a part of the procedure. In the course of the twenty-five daily trials, behaviors of the following kinds were occasionally noted: peering, scurrying, jerking of head and body, biting of objects, jumping, face-washing, grooming, scratching of the body, prolonged smelling, and retracing. These behaviors in the large majority of instances occurred in the first sector of the runway, particularly about the release door of the starting box. Peering, that is to say, gazing upward toward the open top of the runway while standing against the wall, was found in the early stages in all groups. The other behaviors listed above, however, were manifested almost entirely by members of the 1-unit and 4-unit incentive-groups. The patterns began to appear around the 12th trial and as the daily trials were continued to the 25th became more pronounced in form and more general in the group.

The scurryings back and forth, the quick jerky turns, the hoppings and jumpings, the bitings, the prolonged smellings, the face-washings—all these seem to the writer like nothing so much as frustration. In the discussion, an interpretation of this frustration will be offered in terms of whetted appetite.

### TABLE II

<table>
<thead>
<tr>
<th>Incentive-amount comparison</th>
<th>Diffs in means</th>
<th>d.f.</th>
<th>t</th>
<th>( \sigma_a^2 )</th>
<th>( \sigma_b^2 )</th>
<th>F</th>
<th>d.f.</th>
<th>Level of sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>0.453</td>
<td>8</td>
<td>2.57</td>
<td>5% = 2.306</td>
<td>0.078</td>
<td>1.1</td>
<td>5/3</td>
<td>10% = 9.01</td>
</tr>
<tr>
<td>4-16</td>
<td>0.697</td>
<td>6</td>
<td>2.73</td>
<td>5% = 2.447</td>
<td>0.070</td>
<td>1.90</td>
<td>3/3</td>
<td>10% = 9.28</td>
</tr>
<tr>
<td>16-64</td>
<td>1.97</td>
<td>6</td>
<td>7.95</td>
<td>5% = 5.959</td>
<td>0.190</td>
<td>0.077</td>
<td>3/3</td>
<td>10% = 9.28</td>
</tr>
</tbody>
</table>
Exper. II. With regard to the problem of this section, the relationship between magnitude of incentive and level of performance, the purpose of Exper. II, was further to examine, by the employment of a somewhat different method, the lower incentive-levels. The groups were not assigned to the various incentive-levels before their initial trial (as in Exper. I) but all of them were first run with 16-unit reward until the learning curve for the entire group had substantially reached a point where no more learning-changes were likely. Then, on the basis of the last 6 out of the 21 trials which this learning required, 3 equated groups were formed; a group shifted to 4-unit incentive, a group shifted to 1-unit incentive, and the remaining group allowed to continue at the 16-unit incentive. With these incentive-conditions these three groups were run for 14 more daily trials. Trials 12 to 14 incl. were taken as the performance-period within which the influence of the amount of incentive upon level of performance was to be examined.

It occurred to the writer that a view of the effect of 0 incentive-amount upon level of performance could be obtained simply by switching at this time the control 16-unit incentive group to 0-incentive. The results under 0-incentive conditions would be comparable to the 1-unit and 4-unit incentive-group results because both these groups had also been switched from 16-units, the only difference being that they had been switched earlier. There is no reason for supposing that a somewhat longer period at 16-unit incentive for the 0-incentive group would invalidate the comparison made below between the results of 0-incentive and the results under 1-unit incentive-conditions, especially since the only change in the 14 additional trials at 16-unit incentive for the group switched to 0-incentive was a slight decrease in the average mean-speed of the group—which fact is loaded against the conclusion that was drawn from the results of this change.

The quantitative results of Exper. II for the performance-period are summarized in Table III, while the statistical analysis of the significances of these results appears in Table IV. One may observe from Tables III and IV, and graphically from Fig. 1, that the results for the 1-, 4-, and 16-unit incentive-groups are in strong confirmation of the data in this range from Exper. I. There is an increase in level of performance in terms of average mean-speed with each fourfold increase in incentive-amount. These mean increases are both significant at better than the 5% level, the increase associated with the difference between 4- and 16-units incentive being indeed significant at better than 1%. Also as in Exper. I, it can be seen (Fig. 1) that the curve of relationship is positively accelerated within this range between level of performance measured arithmetically and incentive-amount measured logarithmically.
**FIG. 1. CURVES FROM EXPERIMENTS I, II, AND III**
Relationship between amount of incentive and level of performance (average mean-speed) within performance period.

**FIG. 2. CURVES FROM EXPERIMENT III**
Average mean-speed per trial of 3 groups of rats each receiving the amount of incentive indicated.

**FIG. 3. AVERAGE CURVE FROM EXPERIMENTS I, II, AND III**
Relationship between amount of incentive and level of performance (average mean-speed) as an average of the performance-period results.
It is important here to point out that the curves of Exper. I and II, as indicated in Fig. 1, would probably exhibit an even more striking similarity were the latter curve not depressed at the 1- and 4-unit level by a factor of downward contrast introduced by the procedure of shifting the groups from a prior 16-unit incentive-

**TABLE III**

**EXPERIMENT II. EFFECT OF INCENTIVE-AMOUNT UPON LEVEL OF PERFORMANCE FOR PERFORMANCE-PERIOD (TRIALS 9-11 FOR 0-GROUP, 12-14 FOR OTHERS)**

<table>
<thead>
<tr>
<th>Incentive-amount</th>
<th>N</th>
<th>Average mean-speed in feet per second</th>
<th>M</th>
<th>$\sigma^2$</th>
<th>$1000/M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>1.83</td>
<td>1.52</td>
<td>1.81</td>
<td>.000</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>1.68</td>
<td>.624</td>
<td>.490</td>
<td>.124</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>1.90</td>
<td>1.41</td>
<td>1.37</td>
<td>.644</td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>3.08</td>
<td>2.62</td>
<td>2.58</td>
<td>2.44</td>
</tr>
</tbody>
</table>

* This is the lowest speed possible with the conventions of this experiment. The rat did not reach the food-box within a 420-sec time-limit. It was convenient to call the speed in this case $1/2$ the distance from $20\text{ to }20 = 20 = 0.024$.  
  
A latter portion of this paper will be concerned in detail with this depression-effect. Its presence here is suggested by these considerations: the 16-unit incentive group of Exper. II, it is found, is significantly superior at the 5% level to the 16-unit group of Exper. I. This difference under identical conditions indicates that the rats of Exper. II are faster,—which is not surprising since they are younger.

**TABLE IV**

**EXPERIMENT II. SIGNIFICANCE-TESTS FOR PERFORMANCE-PERIOD OF EFFECT OF VARIATION OF INCENTIVE-AMOUNT UPON LEVEL OF PERFORMANCE**

<table>
<thead>
<tr>
<th>Incentive-amount comparison</th>
<th>Diffs in means</th>
<th>d.f.</th>
<th>t</th>
<th>$\sigma^2_a$</th>
<th>$\sigma^2_b$</th>
<th>F</th>
<th>d.f.</th>
<th>Level of sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1</td>
<td>.64</td>
<td>12</td>
<td>2.33</td>
<td>3% $= 2.179$</td>
<td>.392</td>
<td>.146</td>
<td>6/6</td>
<td>10% $= 4.28$</td>
</tr>
<tr>
<td>1-4</td>
<td>.70</td>
<td>12</td>
<td>2.30</td>
<td>2% $= 2.681$</td>
<td>.327</td>
<td>.309</td>
<td>1.06</td>
<td>6/6</td>
</tr>
<tr>
<td>4-16</td>
<td>1.08</td>
<td>12</td>
<td>3.68</td>
<td>5% $= 2.179$</td>
<td>.327</td>
<td>.309</td>
<td>1.06</td>
<td>6/6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2% $= 2.681$</td>
<td>.327</td>
<td>.309</td>
<td>1.06</td>
<td>6/6</td>
</tr>
</tbody>
</table>

* As circumstances necessitated the halting of the o-incentive group at the end of 11 trials instead of 14, this 0-1 comparison is for trials 9-11 inclusive. However, the 1-group mean for these trials is practically identical with its mean for the trials 12-14; being in the former case .578, in the latter case, as seen in Table III, .560.

When these animals are shifted downward to the 4- and 1-unit incentive-levels, however, this superiority in speed disappears.

The variance-ratio analyses of homogeneity of variance in Table IV, in yielding only results which could readily arise by chance, indicate that the t-statistic is properly applied in Exper. II as a test of the significance of differences in means.

An interesting feature of Exper. II is the comparison of the levels of performance associated with unit-incentive conditions, approximately 1/50th of a gram,
and 0-incentive conditions. The rats run significantly faster to no-incentive at all than to a very small incentive. When the results of the three experiments are correlated, this finding will be interpreted as quantitative evidence for the factor, frustration of whetted appetite, which was inferred from the qualitative observations of Exper. I.

Before considering the qualitative results of Exper. II, it is interesting to note (Table III) the coefficients of variation. This quantity is a measure of the relative variability of groups. From the coefficients of variations of Exper. II the generalization can be made that within the performance-period for the various levels of incentive-amounts (other than 0) the relative variabilities are inversely related to the levels of performance.

About four trials after the shift of rats in Exper. II from 16-units incentive to 1- and 4-, the patterns of behavior began to appear which have been suggested as diagnostic of frustration. In the 1-group, and to a slight extent in the 4-group, from their initial appearance at trial 4 to trial 14, these frustration-indicators became progressively more pronounced in form and more widespread in the group. Though no formal quantitative treatment of these observations was made, there were substantially more frustration-incidents in the unit-incentive group of Exper. II than in the unit-incentive group of Exper. I. This was particularly true for jumping near the starting box and biting of the apparatus. In contrast to the unit-incentive group, among the group shifted to 0-incentive conditions there was very little evidence of frustrational behaviors. Subsequently these facts bearing on frustration will be collated to see what conclusions they suggest.

Exper. III. The principle aim of this experiment was to pursue further the consideration of the relationship between amount of incentive and level of performance at the upper incentive levels. The results of this inquiry for a performance-period cross-section, in this case trials 15 to 20 inclusive, appear in Table V. The analysis of the significance of these results is summarized in Table VI. These two tables indicate that, though the differences in levels of performance associated with fourfold increases in incentive-amount are substantial in magnitude, they do not quite achieve the same levels of significance as in Expers. I and II. This result is mainly occasioned by the fact that the variabilities of the groups in Exper. III were larger than the variabilities in either Exper. I or II. Though the rats of Exper. III may simply have been more variable than

---

64 As here employed this statistic was based on the variability of the group of animals for each incentive-amount on the average trial of the performance period, not on the average of the variabilities for each of the group of trials of the performance period.
the samples of the other two experiments, it is quite probable that the variabilities in Exper. III are not as large as they appear to be. In this experiment, it will be recalled, the rat-litters were for the most part split among the three incentive-amount groups, instead of assigned at random. And with the litter-variable of some importance in the determination of speed, this procedure would act to inflate somewhat the estimate of the within-group variabilities, the error-variances in the significance-tests, and would thus make them appear larger than they actually are, their proper size being, of course, their values under conditions of random-assignment.

The significances indicated in Table VI for the increase in levels of performance with increase in incentive-amount, approximately the 8% level for the 16-64 incentive-increment and a 24% level for the 64-256 increment, are excessively conservative, if one takes into account this probable exaggeration of the variabilities. The actual chances that the differences may have arisen by chance are probably appreciably less. This consideration of inflated variabilities hardly needs to be made for the 16-64 difference in levels of performance, since the establishment of the difference at the 8% level in Exper. III, when coupled with its establishment in Exper. I at a level of significance better than 0.1 of 1%, justifies entire confidence in its reliability. The argument is important, however, in that it increases confidence in the presence of a rise of level of performance with incentive-amount increase from 64- to 256-units.

Examination of the results of Exper. III as graphed in Fig. 1 suggests that within this incentive-amount range, the increments of average mean-speed with a logarithmic increase of incentive-amount are negatively accelerated. This finding is in contrast to the positively accelerated relationship exhibited in the lower incentive-amount range. Wherever it begins, either beyond the 64-unit incentive-level, as suggested by Exper. I, or somewhat before the 64-unit level, as suggested by Exper. III, the negative acceleration probably reflects a greater difficulty of the rats in increasing their speeds when they are already approaching their physiological limit. With smaller differences in speed possible because of this factor of diminishing returns, it is quite understandable why the significance of
the 64-256 difference should fall somewhat below that of the other increases in incentive-amount. With more animals to offset the constricted range of possible differences, quite possibly the reliability of the 64-256 difference would have been equivalent to the others. It is to be noted here that the 16-256 difference in level of performance is significant well beyond the 1% level.

Needless to say, in Exper. III, as in the two prior experiments, homogeneity of variance was demonstrated before the t-test was applied to any group-differences in means with the object of investigating their significance. The variance-ratios necessary for this demonstration in Exper. III are listed in Table VI, along with the degrees of freedom and levels of significance associated with each.

The coefficients of variation recorded in the last column of Table V

<table>
<thead>
<tr>
<th>Incentive-amount comparison</th>
<th>Diffs. in means</th>
<th>d.f.</th>
<th>t</th>
<th>Level of sig.</th>
<th>$\sigma^2_a$</th>
<th>$\sigma^2_b$</th>
<th>F</th>
<th>d.f.</th>
<th>Level of sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-64</td>
<td>.78</td>
<td>15</td>
<td>1.89</td>
<td>10%$=1.753$</td>
<td>.770</td>
<td>.658</td>
<td>1.17</td>
<td>6/9</td>
<td>10%$=3.37$</td>
</tr>
<tr>
<td>64-256</td>
<td>.53</td>
<td>12</td>
<td>1.27</td>
<td>30%$=1.083$</td>
<td>.770</td>
<td>.450</td>
<td>1.71</td>
<td>6/6</td>
<td>10%$=4.28$</td>
</tr>
<tr>
<td>16-256</td>
<td>1.31</td>
<td>15</td>
<td>3.51</td>
<td>1%$=2.947$</td>
<td>.658</td>
<td>.450</td>
<td>1.45</td>
<td>9/6</td>
<td>10%$=4.10$</td>
</tr>
</tbody>
</table>

are consistent with the generalization that was proposed on the basis of the relative variabilities for the lower incentive-amounts in Exper. II, namely that within the performance-period the relative variabilities of the various incentive-amount groups are inversely related to their levels of performance.

To the performance-period cross-sectional view that has been given in Expers. I, II, and III of the relationship between incentive-amount and level of performance, a supplement is added in Fig. 2 (See p. 488) of a longitudinal view based upon the data of Exper. III. In this graph one can observe how the relationship develops from the first trial to the end of the performance-period. It has been explained in the account of the design of this experiment that the three incentive-amount groups, 16-, 64-, and 256-units respectively, were equated in average mean-speed on the basis of their initial run before they experienced differential-incentives. Hence, it can be seen (Fig. 2) that the three curves start from substantially the same point going into trial 2 on the second day. By trial 4 the curves are in the hierarchy that they maintain without exception to the end; the
256-group fastest, the 64-group next, and the 16-group slowest of the three. The primary concern in this paper is with the influence of incentive-variation upon performance in the performance-stage of the process of adaptation to a situation, namely a sample-period after the point where substantial learning-changes in the activity are no longer apparent. Fig. 2 shows clearly, however, that incentive-amount differences seem to occasion considerable differences in level of performance not only in the performance period but substantially throughout the entire range of the learning period.

Correlation of results and discussion. Taking together the data of Expers. I, II, and III, what final generalization can be formulated as an answer to the problem of the relationship between magnitude of incentive and the level of performance? Based as it is upon the most numerous and varied samplings, the most appropriate generalization is probably to be founded on an average of the results of the three experiments. Whatever small differences in procedure obtained in the various experiments are not important in the face of their fundamental similarity. In each experiment a determination was made of the level of performance, in terms of average mean-speed, associated with various incentive-amounts for a performance-period block of trails. This period was defined similarly in each case as a sample-period in the progress of adaptation to a situation taken when changes in performance attributable to learning appeared to be largely at an end.

In Table VII the data of the three experiments are collected and the overall averages defining the levels of performance for each incentive-amount level are indicated. The final values appear in Fig. 3 (See p. 488). This figure becomes, then, a graphical presentation of the most general answer to the problem. On the basis of this graph it may be stated that within the performance-period of an activity, as incentive-amount is increased logarithmically within approximately the first 70% of the appetitive range of the rat the level of performance, in terms of average mean-speed measured arithmetically, describes a flattened sigmoid curve. The curve in the middle range, from 4- to 64-units incentive is substantially linear. The non-linearity appearing beyond 64-units incentive probably reflects the diminishing returns as the animals approach their physiological limit of speed. The non-linearity in the 1-to-4-unit region, and especially the actual change in direction of the curve between 1 and 0 incentive, reflects—as will be developed later in this discussion—the presence of a frustrational factor.

Though this paper concerns itself primarily with the effects of incentive-amount variation upon performance when learning-contributions to the
activity are negligible, longitudinal views of the results like that of Exper. III (see Fig. 2) give reasonable basis for extending the scope of the conclusions. These results make it appear that the amount of incentive is a determinant of the behavior of the rat from the very beginning of his activity in a new situation; that is to say, throughout the period when learning is a part of his performance, as well as when learning is largely over.65

Examination of the coefficients of variations for Expers. II and III discloses a generalization which holds for the range of incentive-amounts from 1 to 256 units. Within the performance-period the relative variabili-

<table>
<thead>
<tr>
<th>TABLE VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFFECT OF INCENTIVE-AMOUNT UPON LEVEL OF PERFORMANCE AS AN AVERAGE OF THE PERFORMANCE-PERIOD RESULTS OF EXPERIMENTS I, II, AND III</td>
</tr>
<tr>
<td>Incentive-amount</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>64</td>
</tr>
<tr>
<td>256</td>
</tr>
</tbody>
</table>

| 65 What would happen to the distinct levels of performance associated with different incentive-amounts after many trials cannot, of course, be said on the basis of these data. Would the curves of Fig. 2, for example, diverge, converge, or remain as they are? In attacking this problem there would be a distinct danger of artifactual convergence. For with many repetitions of trials the rats begin to integrate the test-situation with the post-test feeding-situation. Taking the two together, the amount of food is equal for the members of the different incentive-groups. Equality of performance to the total situation, then, would not be surprising.
In the introduction to this investigation the few studies which had more or less relevance to the problem of quantitative variation of incentive and performance were briefly considered. None of these could be accepted as even approaching a generalized view of the relationship since consideration clearly revealed that the treatment of the independent variable in these inquiries was either very limited or very coarse. Grindley's study on chicks was one of the former. His employment of several incentive values, 0, 1, 2, 4, and 6 grains of rice, gave a fine graduation, but the appetitive range explored was almost negligible. Wolfe and Kaplon's study of chickens with their range from \( \frac{1}{4} \) to 1 grain of popcorn was even narrower in scope. Nissen and Elder's study had a range of 5 to 10, on some occasions, 20 grams of banana, which is probably not very considerable with chimpanzees. Furthermore they employed but two incentive-values, the coarsest graduation possible. The follow-up experiment on chimpanzees by Cowles and Nissen was just as coarse in its incentive-graduations and, though the large incentive was as large as 60 grams of orange on some occasions, generally it was no more than 35 grams of orange or banana. Fitts' study with rats covered a wide range, 0.2 grams to 10 grams, but with a minimum of incentive-precision as only these two values were used.

Besides these criticisms of scope and precision which apply to all the experiments, Grindley's study had serious structural weaknesses. This investigation has been cited so often as definitive, probably in the absence of any other, that it is worthwhile to list briefly some of its defects. (1) Deficiencies of drive-equalization. There was no apparent control of amount of rice eaten. All that Grindley says is that "when they had been tested they were given a large meal of rice. At other feeding times they had no rice." Further, there was no assurance that stomach-contractions were controlled; the statement appears that "the experiments were always made at a time when the chicks were hungry." This is an interpretation, not a statement, of conditions. (2) Sampling weaknesses. Thirteen chicks were thrown out because "they were found to develop a dislike for rice during the experiments." This suggests that the complexity of varying preference existed which would operate to expand differences obtainable by chance alone. There was no account of how the various groups were divided. There is no consideration of possible influences of sex. The age range of the chicks was from 5 to 10 days, a very narrow base from which to draw generalizations. (3) Significance-tests. There was absolutely no statistical analysis of the results, and with the observation that there were "wide individual differences in behavior" even Grindley thought that the experiment was not sufficiently accurate to establish the results with certainty.

In view of these experimental obscurities and inconclusive results it was certainly premature of Katz to adduce a law of relativity based considerably on Grindley's findings. This law states that "the greater the hunger the greater the incentive-value of the same food, and, under constant conditions of hunger, the

---

66 Grindley, *op. cit.*
67 Wolfe and Kaplon, *op. cit.*
68 Nissen and Elder, *op. cit.*
69 Cowles and Nissen, *op. cit.*
70 Fitts, *op. cit.*
71 Grindley, *op. cit.*., 175.
better the food the greater its incentive value.” The term 'better food' means more of the same food. For the generalization with respect to constant hunger supposedly summarizes the quoted and similar studies. Grindley's is the only quoted study concerning varying incentive with constant drive. Even if Grindley's findings could have been accepted as reliable, Katz's law is still an unwarranted extrapolation from such a minute range of incentive-variation.

The deficiencies that have been pointed out in the various experiments from the point of view of answering the question of the relationship between incentive-variation and performance, prompted efforts toward more precise controls and a more equitable balance between precision and range in incentive-variation. For only in this manner could a generalized picture be obtained.

The functional relationship between incentive-amount and level of performance shown in Fig. 3 has an importance not so much critical as systematic. The findings point out basic characteristics of animal motivation and performance which should not be overlooked in theories of animal behavior. The results are not lacking in immediate implications, however. One may consider, e.g. the implications for the methodology of learning experiments.

In many studies of the ability of animals to master various situations, 'learning' curves are drawn which purport to depict progressive adaptation with successive repetitions of a situation. Two kinds of curves are frequently given, one based upon the number of 'errors' occurring in successive trials, the other based upon the expenditure of time in the successive trials. With the validity of the learning curve based upon errors this paper is not specifically concerned. The writer agrees, however, with Maier and Schneirla's contention that an error due to insufficient learning is one thing and an error due to inattention, distraction, or lack of motivation is quite another. Consequently before it can be concluded from error-curves that a difference in learning ability exists between two groups, it must be shown that the errors are not due to other than the factor of learning.

Upon curves of learning founded on time-measurements the results have a definite bearing. In studies employing such curves in the comparison of the learning ability of various animals, it seems to be commonly assumed that a considerable change in motivating conditions would not occasion any substantial change in the results. Often at most, modality of drive employed as hunger or thirst and, perhaps, time of deprivation are specified as gestures toward comparability of motivation. The results of these experiments, particularly the longitudinal observations of Exper. III, indicate that with drive held strictly constant, incentive-variation can occasion substantial changes in performance throughout the entire period of adaptation. The implication is apparent. If animals are to be validly compared as to their ability to adapt to a situation there must be a precise specification of the quantity of incentive present with the groups compared. This is not to say that other factors which may influence motivation are not also to be specified. But these experiments show that without the specification of inventive-amount the results with respect to the problem of learning become indeterminate.

It is perhaps needless to remark upon an obvious application of the curve delineated in Fig. 3; namely, with regard to the contribution of incentive-amount to strength of motivation, the most efficient incentive in terms of high speed of

---

32 Katz, op. cit., 157.
performance and low relative variability is the abscissa-value corresponding to the end of the linear rise in the curve. For Purina biscuit, a widely used incentive, this amount is approximately 1.3 grams dry weighed. Very probably this amount would hold as a good approximation for many other foods as well. To speak of the 64-unit incentive as the most efficient is not to say that it produces actually the maximal speed and minimal relative variability of performance. It is rather to imply that further improvements in these variables would necessitate incentive-amount increases requiring impractical extensions of eating-time.

There is yet to be considered the interesting findings both qualitative and quantitative associated with very small incentive-amounts, findings which have been taken in this paper to be indicative of frustration. The qualitative behavior-patterns observed involved the following components; peering, scurrying, jerking of the head and body, biting of objects, jumping, face-washing, grooming, scratching of the body, prolonged smelling, and fitful retracing. Their appearance in the various small incentive groups was distributed as follows; (1) The unit-incentive group of Exper. I and to a smaller extent the 4-unit incentive-group, began to manifest the frustration-indicators around the 12th daily trial. From that time to the end of the experiment the behaviors became more pronounced in form and general in the group. (2) In Exper. II, about 4 trials after the shift from 16-units incentive these same indicators began to appear in the unit-group and to a smaller extent in the 4-unit group, growing more pronounced in form and general in the group to the end. The prevalence and amount of these behaviors, particularly of the more violent kinds, as jumping to get out of the apparatus near the food box and biting the apparatus, were appreciably greater in the 1- and 4-unit incentive-groups of Exper. II than the similar groups in Exper. I. (3) In the group shifted in Exper. II from 16 to 0 incentive-amount, there was hardly any appearance of the frustration-indicators, particularly of the more violent forms.

The quantitative evidence bearing on the present discussion is that the comparison in Exper. II of the level of performance of the 1-limit and the 0-unit incentive-groups shows the latter group to be superior at better than the 5% level.

This finding suggests that there is a frustration-effect specifically associated with very small incentives, particularly 1-unit and to a lesser extent 4-units, which is not associated with large incentives nor with complete absence of incentive. The very few occurrences of the frustration-indicators in the 0-incentive group is adequately accounted for on other grounds, in terms of the procedure of Exper. II. This experiment, unlike Number I, involved a shift from 16-units incentive downward to 0-, 1-, and 4-units. This downward contrast produces on its own account (as will be demonstrated in another section of this paper) a 'depression-effect.' That this depression-effect was present in Exper. II, in addition to the direct frustration-effect of small incentives, is indicated by two considerations. First, the number of frustration-behaviors in the 1- and 4-unit incentive-groups of Exper. II was substantially greater than that in Exper. I, where only the direct frustration of small incentives was operating. Secondly, though the rats of Exper. II were natively faster than those of Exper. I, as revealed by a significant difference in speed under the same conditions at the 16-unit incentive-level ($t = 2.26$ with 9 d.f.), their speed-advantage, presumably because of the depression-effect of downward contrast, was lost after their shift from 16-units incentive down to 4 or 1.
Granting that there is a frustration-factor associated with the receipt of a very small incentive which is not associated with the complete absence of an incentive, what can it be? A plausible theory is as follows. Eating the small incentive serves to stimulate and increase desire or tension in the rat without, however, improving the chances for obtaining more food. Hence, the rat after obtaining the small incentive must remain in a state of heightened tension for a (to him) indefinite period (it will be remembered that the rats were fed an appraisable and variable time later after their daily test run was over). This state of heightened tension is unpleasant. Therefore, though food is present and the animal has had no food for some twenty-two hours, the situation is labelled 'frustrating.' The reactions the rats expressed to this frustration, principally in the first section of the runway, seemed to vary, if some interpretation be allowed, from an agony of indecision indicated by halting excursions back and forth, peering, and the like; through distraction indicated by grooming, scratching of the body, and face-washing; to definite avoidance indicated by purposeful attempts to escape from the apparatus by jumping near the starting box.

Now that the thesis has been formulated that a frustration-factor of whetted appetite is associated with very small incentive-amounts, how does it fare when applied to Grindley's and Wolfe and Kaplon's data, the only other experiments in which very small incentives have been employed. In Grindley's study the significant statement is found that "some of the chicks were found to develop a dislike for rice during the experiments." Thirteen chicks or more than 25% of the entire group were discarded because of this dislike. It is extremely likely that these chicks were only expressing the behavioral results of frustration upon receiving such small incentives (1, 2, 4, or 6 grains of rice). And Grindley, instead of studying this phenomenon on its own account, threw these chicks out of the investigation.

No facts from which frustration might be inferred were reported in Wolfe and Kaplon's study. Frustration could hardly be expected, however, for though the rewards were small (¼ to 1 grain of popcorn) the chickens were given five successive trials, twice a day. Further, when one considers the eating habits of chickens, it is suggested that, unlike rats, they would become accustomed to receiving their food in comparatively small amounts.

**PROBLEM II**

*What is the relationship between magnitude of incentive and distribution of effort (gradients) within performance for the (a) learning and (b) performance period?*

One of the purposes of Exper. I was to throw light upon this question. Before presenting the results that were obtained, it is necessary to indicate some procedural considerations that have not yet been discussed because they were more specifically relevant in this context.

73 Grindley, *op. cit.*, 175.
74 Wolfe and Kaplon, *op. cit.*
Procedure. The main orientation in the procedure with respect to this inquiry concerning gradients was to make the conditions comparable to those of Hull.\textsuperscript{75} As Hull reports uniformity of results despite variations of procedure and apparatus, it is indicated that only rough similarity is necessary to achieve comparability.

The data bearing on this present problem were the times taken in Exper. I at successive five-foot sections of the twenty-foot linear runway. Time spent in the starting-box after the door was opened was also recorded, but as these times were very variable and apparently affected by purely chance factors (\textit{e.g.} in which direction the rat was pointed when the door was opened) they were discarded. In order to prevent any possible influence of odor in heightening the speed-of-locomotion as the animal neared the food, cups of fresh food were placed on the shelf to the side of the runway, one at the starting-box door, and one each in the middle of the five-foot sections, as in Hull’s procedure. Further, the animals were run in a direction opposite to their food-cages.

The apparatus in this experiment, unlike Hull’s, had no retracing valves for reasons that have been discussed. However, Drew\textsuperscript{76} states that in one of his experiments the results given by a group of rats in a runway with the retracing doors removed were identical with those obtained with the doors present, with the sole exception of being slightly faster. This suggests that the difference may be overlooked.

Hull’s procedure involved five trials per day, whereas the daily regimen in the present experiment, because of considerations of drive-equalization, could only allow one trial. The only effect this difference should have is to speed up the practice-effects on the goal-gradient, since spaced practice is superior to massed practice. To compensate for this difference, trials 4 to 7 were taken in this experiment as a basis for the learning-period gradient, whereas Hull took trials 1 to 10. This is in fact more compensation than it appears, as trials 1 to 10 in Hull are actually trials 5 to 15 if one includes adaptation-trials, as was done in my count.

Certain statistical considerations were incident to the treatment of the data. (1) To avoid dislocation of the schedule a time-limit of 420 sec. was set up for the rats to traverse the runway. When the animals took longer than this they were gently pushed the remainder of the distance into the food-box, and their sector-times were not included in the gradient data. (2) In the pooled trials of both the learning-period (trials 4 to 7) and the performance-period (trials 21 to 25), those observations were excluded from the population of observations defining the speed in a particular sector for a particular incentive-amount whose deviations from the mean of that population were such as to arise by chance less than 1\% of the time. As has been discussed before, time is a blanket index that is sensitive to a host of extraneous factors. Therefore when it is employed as the dependent measurable

\textsuperscript{75} Hull, \textit{op. cit.}
\textsuperscript{76} Drew, \textit{op. cit.}
in any experiment, one must be prepared to exclude the 'wild observations,' as they are termed in astronomy. In his data Hull also remarks that "occasionally there appears an extremely large reading." 77 (3) The question of metric must be brought up again. In the general procedure of this inquiry, it was discussed at length why, in the opinion of the experimenter, average mean-speed was a superior metric to average speed. The latter, however, is employed by Hull. In order to test one of his predictions on his own ground, therefore, the average speed concept must be used. It is also desirable to compare the results with those of Drew and he too employs this metric. 78

The results of this inquiry 79 into the effect of variation of incentive-amount upon the distribution of effort within a performance are sum-

TABLE VIII

EXPERIMENT I. GRADIENTS OF AVERAGE TIME AND SPEED FOR INITIAL LEARNING-PERIOD
(Trials 4-7)

<table>
<thead>
<tr>
<th>Section of runway</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incentive-amounts</td>
<td>N</td>
<td>Mean t (t)</td>
<td>Mean s (s)</td>
<td>Mean t</td>
</tr>
<tr>
<td></td>
<td></td>
<td>in seconds</td>
<td>in feet-per-sec.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>49.8</td>
<td>.10</td>
<td>39.4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>34.5</td>
<td>.145</td>
<td>25.2</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>29.1</td>
<td>.172</td>
<td>7.80</td>
</tr>
<tr>
<td>64</td>
<td>4</td>
<td>6.56</td>
<td>.762</td>
<td>3.69</td>
</tr>
</tbody>
</table>

These results may not appear to be based upon as large a number of cases as might be desirable. It should be noted, however, that the implications drawn from them are not derived from separate curves, but from the entire family of curves taken as a group. The hierarchy in these families is in every case precisely in accord with the statistically significant hierarchy found in Exper. I, where speeds were taken for the runway as a whole.
graphed in Fig. 5. Both these expressions for the same data are included because the first form bears directly upon a prediction of Hull’s; the second form bears directly upon a generalization of Drew’s.

The observation may be made from the smoothed curves in Figs. 4 or 5 that the shapes of either the speed-of-locomotion gradients or the locomotion-time graphs vary consistently as a function of amount of incentive. In the case of the speed-of-locomotion gradients this generalization can be drawn: in the early learning-period as the amount of incentive is increased in successive groups the shape of the gradient changes from negatively accelerated, through linear, to positively accelerated.80

The results for the distribution of effort within the runway for a per-

---

80 It may be remarked here that the results of the 256-unit incentive-group which were excluded from Exper. I because of faulty randomization of litter were in accord with this generalization, having, like the 64-unit group, a definite negative acceleration.
formance-period are listed in Table IX. Again both average times and average speeds in each sector are given for each incentive group. Only the locomotion-time graphs are shown graphically, however, as this is the pertinent form for the comparison of these particular findings with those of Drew.\(^8\) It is seen that these curves, which are very regular without any smoothing, express a consistent relationship between their configurations and incentive-amount. To generalize: in a performance-period of an activity, as the amount of incentive is increased in successive groups, the locomotion-time graphs vary progressively all the way from marked negative acceleration to complete flatness.\(^8\)

**Discussion:** In one of his papers, Hull is led to a consideration of the effects of increasing the amount of incentive upon the excitatory-gradient. The shape of this gradient he takes from various considerations to be a positively accelerated curve tentatively expressible by the equation \(E = a - b \log D\), where \(E\) is the strength of the excitatory tendency, \(D\) is the distance from the goal, and \(a\) and \(b\) are positive constants. The effect of increasing the excitatory potentialities of the goal by increasing the amount of reward is as follows, to quote Hull. "Other things being equal, this gradient of excitation grows higher and steeper with the increase in amount of reward."\(^8\) He gives as an example of doubling the strength of a lure by increasing its amount the equations \(E_w = 10^{-4} \log D\), and \(E_s = 20^{-8} \log D\).

If these are Hull's theoretical expectations as to the relation of the excitatory gradient to incentive-amount, how can the prediction be tested? Unfortunately, Hull, despite his avowed stress upon the necessity of empirically verifiable consequences in scientific hypothesis, apparently neglects to state precisely what the prediction will be in an actual situation, e.g. the linear runway. The excitatory

---

\(^8\) Drew, *op. cit.*

\(^8\) Again for whatever significance it may have the discarded 256-group was in complete accord with this generalization. The shape of its performance locomotion-time graph was practically identical with that of the 64-unit incentive-group.

\(^8\) Hull, *The goal-gradient hypothesis applied to some 'field-force' problems in the behavior of young children, Psychol. Rev.,* 45, 1938, 271-299.
gradient is not a measurable, it is an inferred, entity. Hence Hull’s generalization as quoted above is not testable until it is formulated in terms of measurables.

In lieu of a definite statement, Hull’s prediction can possibly be dissected out in this manner. The positively accelerated curve of excitatory tendency was inferred, in Hull’s paper upon the speed-of-locomotion of the rat, from a “substantially rectilinear speed gradient” on the assumption, supported from several sources, that a much greater increase in amount of excitation is required to increase the vigor of muscular action in the higher ranges of activity than in the lower.\textsuperscript{84} In brief, a positively accelerated curve of excitatory tendency was derived from a linear speed of locomotion-gradient. If this derivation is reversed, it is found that a higher and steeper positively accelerated curve of excitatory tendency, consequent upon increase in incentive amount, reflects—in the realm of the measurable—a higher and steeper linear speed-of-locomotion gradient. This deduction is not rigorous, but the original reasoning was not intended to be mathematically precise, so reversing the procedure cannot be expected to come out with any more determinate relation.

The effect of increasing the amount of incentive in a linear runway, then, is—as Hull’s theory implies—to displace a linear speed-of-locomotion gradient upward and increase its slope. Either this is implied in Hull’s discussion or, it would seem, the discussion fails to make any verifiable statement.

Fig. 4 indicates that, when incentive-variation is examined over a substantial range, Hull’s expectation does not obtain. With increases in incentive-amount there are definite and progressive changes, not only in the ordinate placement of the curves and their slopes, but in their total shapes. It is possible that increases of incentive-amount within the medium range (around 16-units incentive) would result in a family of linear speed-of-locomotion gradients successively displaced upward and successively steeper in slope. This can in no sense, however, be taken as typical of the change in gradients associated with variation of incentive-amount.

Hull must be left now, for the moment, to bring the findings of the experiment to bear on certain major conclusions of Drew’s experiment.\textsuperscript{85} In this study Drew gives a result which is “completely incompatible with the goal-gradient hypothesis.” This is “the broad fact that the Hull type gradient [which is for Drew a negatively accelerated locomotion time-graph] is only found, assuming the rats are strongly motivated and receive an appropriate reward, which is the only situation in which the goal-gradient hypothesis can be held to apply, when a massed-trial technique is used, a flat graph resulting from spaced trials. . . .”

Fig. 5 unequivocally demonstrates that, with suitable variation of incentive-amount under spaced-trial conditions (24-hours apart), in which Drew asserts a Hull type negatively accelerated locomotion time-graph does not appear but only a flat-speed graph, there can be obtained: (a) a positively accelerated curve, (b) a Hull type curve, and (c) Drew’s ‘flat’ speed-graph. The writer would not call this latter curve flat; Fig. 6 reveals that such curves can become much flatter. And if the former curve is labeled flat, what can one term the latter curves?\textsuperscript{86}

In generalizing his results Drew has erred on two counts, then, both in stating

---

\textsuperscript{84} Hull, op. cit., see footnote 40, p. 418.

\textsuperscript{85} Drew, op. cit., 366.

\textsuperscript{86} Drew employed trials 1-5 for his learning-period sample. This is comparable to trials 4-7 employed in the present experiment, for Drew does not include in his trial-count the preliminary training, consisting in some cases of 5 forced trials.
that a Hull type gradient is only found under massed-trial conditions and in stating that it cannot be found under spaced-trial conditions. With large incentives Drew would be correct in saying that Hull's gradient cannot be found under spaced-trial conditions; with medium incentives Hull would be correct in his expectation of a Hull type gradient; with small incentives neither would be correct. What gradient really appears under spaced-trial conditions? Is there any better example of the necessity of specifying the covert variable of so many experiments, i.e. incentive-amount. Had Drew done so, he would not have said of his flat-speed graph in describing performance in a linear runway, "because it is typical not only of spaced practice, which is the most efficient method of learning in this type of apparatus, but also of the first trials of any series under massed-practice condition, it

![Graph showing locomotion-time graphs for performance-period (Trials 21-25)](image)

**Fig. 6. Locomotion-Time Graphs for Performance-Period (Trials 21-25)**

Groups of rats running to the amounts of incentive indicated.

would seem that the basic gradient in speed for this type of apparatus is a flat one, or in other words, the speed of running throughout an apparatus of the kind used here is approximately constant, when 'goal attraction' is the only factor operating."[^87]

[^87]: Drew, *op. cit.*, 370.


Drew verified Hull's finding that under massed-trial conditions there is a progressive flattening of the locomotion-time graphs with practice. His data under spaced-trial conditions, however, lead him to state that "the improvement in the total score is a generalized function under these conditions as opposed to the marked improvements localized at the beginning under massed-trial conditions."[^88]

The locomotion-time graphs of Figs. 5 and 6 are based upon, respectively, a
group of trials prior to any substantial practice (the learning period) and a group of trials after considerable practice (the performance period). Does comparison of these curves bear out Drew's generalization about the effect of practice under spaced-trial conditions? It is clear that his characterization of improvement as a generalized function does not hold in any case, except perhaps as a rough approximation to practice-changes under large incentive-amount conditions. In the results of this experiment under spaced-trial conditions, most of the improvement with small incentives is localized in the middle sectors. With the large incentives, most of the improvement is localized in the first sector. At no incentive-amount is practice-change best described as a generalized function.

Hull's description of improvement as a progressive flattening with practice would approximately apply to incentive-amounts in the medium range (16-units).

The intention in this preceding discussion was to point out the pervasive influences of incentive-amount variation— Influences so potent that the variable cannot be left unspecified without making the results of an experiment indeterminate. It was not the intention to be concerned here with the theoretical soundness of the goal-gradient hypothesis, or with such problems as to whether the speed-of-locomotion gradient is appropriately characterized as a goal-gradient, as Hull would have, or an entrance-gradient as Dennis would suggest. When the term gradient was employed, it was not to imply goal-gradient; its use was divorced from any theoretical preconceptions. A more neutral term is desirable, one strictly descriptive, since the empirical gradients have not so far been unequivocally accounted for theoretically—although suggested by the goal-gradient hypothesis.

**Problem III**

What are the effects of variation of magnitude of incentive upon level of performance, namely, (a) an upward shift of incentive-amount, and (b) a downward shift of incentive-amount?

It was part of the purpose of Expers. II and III to reveal the effects upon performance of conditions (a) and (b) respectively. The results obtained under each of these conditions will be presented in order, prefacing them with relevant specific procedural considerations not included previously under the general procedure of these experiments.

*Upward shift in amount of incentive.* Procedure: In Exper. II, it will be remembered, twenty-one rats were run to 16-units incentive until the learning curve had substantially flattened. The group was then divided into three equated subgroups of seven rats each; one was allowed to continue at 16-units incentive, the other two were switched to 1- and 4-unit

---

incentives respectively. The last three trials out of fourteen under these conditions were chosen as a performance-period in which to ascertain the influence of these different incentive-amounts upon level of performance (Problem I of this inquiry). At this juncture it occurred to E that the circumstances were opportune for examination of the effects of upward shift of incentive-amount. So after a three-day intermission, the 1- and 4-unit incentive-groups were shifted upward to 16-units of incentive. It was not deemed necessary to continue the 16-unit control-group any longer, for they had already gone 14 trials beyond the point where their learning curve had flattened out and, further, their performance had become very consistent. For example, for trials 12 to 14, the three trials of the performance-period, the average mean-speed was 2.39, 2.34, and 2.26 feet-per-second. It was possible, therefore, to use an extrapolation of the average of this extremely flat performance-period in the shift-comparison in order to release the 16-unit control-group for the 16 to 0 incentive-change which had significant implications for the phenomenon of frustration of whetted appetite discussed under Problem I.

It was anticipated that any difference that might be occasioned in level of performance between the control-group and the shifted-groups by the changes in incentive-amount could easily be of a rise-and-fall nature, such that averaging the difference for a long series of trials might lose an effect entirely. Since the question is whether or not an effect would occur somewhere in a brief series of post-shift trials and not how long the changes might persist, the indicated statistical procedure was to apply a significance-analysis to the region where any obtained effect appears most strongly marked. It was decided in advance, therefore, to apply the t-test to the average of some sample three-trial period at such a region.

Results. The results of shifting groups of rats adapted to 1- and 4-units incentives up to 16-units of incentive appear in Fig. 7. Both of the shifted-groups quickly surpass in average mean-speed the performance-period control-group average, and appear to be becoming progressively faster to the very end of the 9 post-shift trials where the experiment was discontinued.

The block of post-shift trials from 5 through 7 was taken for significance-analysis. The data for this block and the small sample-analysis are summarized in Tables X and XI. Both the fourfold and the sixteen-fold upward-shifts result in highly significant 'elation' effects. The 4- to 16-unit incentive-shift is significant at better than the 1% level; the 1- to 16-unit incentive-shift, at better than a

---

90 This is not strictly the region of maximal effect but if the differences here are significant, larger differences would only be more significant.
level of 0.1 of 1%. The significance-analysis of homogeneity of variance shows that the t-test is here legitimately applied to the obtained differences in means.

The last column of Table X reveals that in the case of a sixteen-fold upward shift in amount of incentive the relative variability falls in accord with the inverse relationship found to hold in Problem I between the levels of performance and the relative variabilities of groups. Though the fourfold upward-shift in incentive (from 4- to 16-units) raises the level of performance above the control-group, nevertheless it does not drop the relative variability of performance below the control.

The three-day period intervening between the shift of the 1- and 4-unit incentive-groups to 16-units of incentive indicated a result which may have some significance. When the 1- and 4-groups were run for the first time at 16-units of incentive they did not of course know that the incentive-amount had been changed. Hence their

<table>
<thead>
<tr>
<th>TABLE X</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXPERIMENT II. ELATION EFFECT. COMPARISON OF POST-SHIFT RESULTS (TRIALS 5-7) WITH PERFORMANCE-PERIOD CONTROL-RESULTS</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incentive-amount</th>
<th>Average mean-speed in feet-per-second during performance-period for control 16-unit group; during test-period for shifted-groups M</th>
<th>$\sigma^2$</th>
<th>100$\sigma^2$/M</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>7</td>
<td>3.08</td>
<td>2.62</td>
</tr>
<tr>
<td>4 to 16</td>
<td>7</td>
<td>4.70</td>
<td>3.93</td>
</tr>
<tr>
<td>1 to 16</td>
<td>6</td>
<td>3.92</td>
<td>3.79</td>
</tr>
</tbody>
</table>

* One rat at the latter part of the unit-reward period refused to run at all. This qualitative difference from the other rats in the group suggested that it might be advisable to include him in the post-shift inquiry. At the last post-shift trial, however, he was running at a speed greater than the mean of his group.

<table>
<thead>
<tr>
<th>TABLE XI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXPERIMENT II. SIGNIFICANCE-TESTS OF ELATION-EFFECTS (TRIALS 5-7)</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incentive-amount</th>
<th>Diff. in means</th>
<th>d.f.</th>
<th>t</th>
<th>Level of sig. $\sigma^2$</th>
<th>$\sigma_1^2$</th>
<th>F</th>
<th>d.f.</th>
<th>Level of sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-(4-16)</td>
<td>1.16</td>
<td>12</td>
<td>3.20</td>
<td>$1%=3.055$</td>
<td>.632</td>
<td>.293</td>
<td>2.16</td>
<td>6/6</td>
</tr>
<tr>
<td>16-(1-16)</td>
<td>1.12</td>
<td>11</td>
<td>4.61</td>
<td>$1%=4.437$</td>
<td>.293</td>
<td>.152</td>
<td>1.93</td>
<td>6/5</td>
</tr>
</tbody>
</table>

performance upon that trial—labeled (—) in Fig. 7—revealed the effect of a three-day intermission upon level of performance. On the first trial after this intermission both the 4-group and the 1-group ran slightly faster than their performance-period average; an improvement of 1.26 to 1.76 feet-per-second for the 4-group, and of 0.56 to 0.68 feet-per-second for the 1-group. Remembering that these were being subjected to the frustration-effects of small incentives at the 1- and 4-unit incentive-levels, the suggestion arises that this increase in speed represents some sort of decay of inhibition, or in the jargon of conditioning, spontaneous recovery.

Under the conditions of this investigation, no consistent differences were observable in the quality or magnitude of the elation-effects consequent upon fourfold as compared to sixteen-fold upward-shift. The curves as seen in Fig. 7 are very much the same.
The magnitude of the elation-effect may be accounted for in part by the fact that the shift was not simply from smaller to larger incentives, but from incentives that were frustratingly small-to-larger.

There was little distinctive qualitative behavior correlated with the quantitative crossing of levels of performance that we termed the elation-effect. In general, the qualitative results were that the 1- and 4-unit incentive-amount groups ceased the frustration-behaviors that characterized their performance at these small incentives—jumping, biting, face-washing and the like—and by the 5th trial after the shift to 16-units reward were running smartly with little if any extraneous activity.

**Fig 7. Curves from Shifted Groups**

Effect of upward shifts of amount of incentive in comparison with performance-period control-results.

**Fig 8. Curves from Shifted Groups**

Effect of downward shifts of amount of incentive in comparison with non-shifted control-group.

*Downward shift in amount of incentive.* Procedure: In Exper. III, it has been previously indicated, 10 rats were run to 16-units incentive and 7 rats were assigned to 64- and 256-units incentive respectively. After the establishment, for the purposes of Problem I, of the relative levels of performance within a block of trials chosen as a performance-period, the 64- and 256-units incentive-groups were shifted downward to 16-units incentive-conditions. There was no delay between the performance-period and the post-shift trials other than the 24-hours that was part of the 1-trial-per-day procedure. The 16-unit incentive-group was held at the same
INCENTIVE AND PERFORMANCE IN THE RAT

value in order to serve as a 'comparisons' control. Continuing the control-group, instead of extrapolating the performance-period average, was necessary in this experiment because the rats did not stabilize sufficiently well to make possible a secure extrapolation of their performance.

As in the elation-inquiry, it was decided in advance that significance-analysis would be applied to some sample three-day period at the maximum of any obtained effect. It was anticipated that if any 'depression' effect appeared, i.e., a dropping of the performance of the shifted-groups below the level of the 16-unit controls, the effect might very well be transitory. In this case a significance-test applied to an average of too long a range of trials would be inappropriate.

Results. The results of shifting groups adapted to 256-units and 64-units of incentive down to a 16-unit incentive-amount appear in Fig. 8. The level of performance of both shifted-groups falls below the controls on the first trial after they have experienced the lower incentive. With one exception in the 64-unit group, they remain below for the duration of the experiment. The two shifted-groups manifest no consistent or substantial difference between one another in terms of average mean-speed in their reaction to change of incentive-amounts.

The block of trials 4-6 was chosen for significance-analysis. The data for this analysis appear in Table XII; the results of the significance-tests in Table XIII. As was the case with upward-shifts in incentive-amount, fourfold and sixteen-fold downward-shifts both occasioned very significant effects—under these conditions a 'depression.' The fourfold-shift is significant at almost the 2% level, the sixteen-fold downward-shift effect is significant well beyond the 1% level of confidence. As before, the t-test was applied to the obtained difference in means only when homogeneity of variance had been established.

Observation of Fig. 8 might suggest that the depression in level of performance resulting from downward-shift of incentive-amount would disappear in time as the
animals may adapt to the new incentive-amount. The experiment was discontinued at trial-8, however, short of the investigation of this possibility because the animals were beginning to integrate the test-situation with the subsequent feeding-situation. This integration was inferred from the fact that the rats, who would jump about in aroused fashion during the course of the experimentation proper, would not, as was previously their wont, become relatively quiescent in the pre-feeding delay-interval, but would remain excited and active up through the feeding-situation. Any integration of the experimental run with the post-experimental feeding would of course tend to converge the groups, as the amount of incentive taken together is the same for all groups (relative to the weight of each animal).

Examination of the last column of Table XIII suggests that downward-shift of incentive affects the relative variability of performance as well as the level. Both shifted-incentive groups showed a change in their coefficients of variation in the direction expected by the inverse relationship found previously to hold between the levels of performance and the relative variabilities of groups.

The qualitative observations may properly be considered here to see what light they shed upon the interpretation of the significant depression-effects that have been observed. In the 16-unit incentive control-group most of the rats went directly down to the food-box. Occasionally there was pausing here and there in the runway, but there were very few long steps or retraces. Peering had by this time become rare.

The first experience of change in the seven rats shifted down to 16- from 64-units incentive brought general excited peering in the food-box, general delayed-eating, one refusal to eat, and one persistent attempt to jump out of the food-box. In subsequent trials considerable peering and retracing around the starting-box quickly arose, as did also retracing and jerky, hesitant locomotion. Among some rats prolonged smelling of the pathway and biting of the starting-door were in evidence.

The first experience of change among the seven rats shifted downward to 16- from 256-units of incentive occasioned general frantic peering, general delayed-eating, repeated jumping-attempts to escape the food-box in three cases, and refusal to eat all or part of the incentive in three cases. Upon subsequent trials there was prolonged peering at the starting-door, jerky hesitant locomotion, retracing, prolonged smelling, and door-biting. On the first experience of the incentive shift, at least, the 256-group exhibited substantially more disruption of behavior than did the 64-unit group.

These qualitative observations would tend to indicate that the depression-effect is a quantitative manifestation of some degree of frustration. The frustration-indicators were not as general and considerably less intense than those characteristic of the receipt of very small incentives, as described under Problem I, but enough of them were present unequivocally to label the depression as grounded upon frustration.

Discussion. The discovery of an 'elation' and a 'depression' effect consequent upon upward-shifting and downward-shifting of incentive-amount respectively, forcibly demonstrates that with drive-held-constant the level
of performance associated with an incentive-amount is not determined by the quantity alone, but is also determined by the preceding experiences with quantities. The elation and depression effects are an adequate experimental basis for defining a variable within the rat which may be termed on the basis of human analogy an 'expectation.' This interpretation is influenced in part by Tolman's concept of cognitive expectation of goal object.91

These results, with rats as subjects, agree with the findings of Nissen and Elder92 and Cowles and Nissen93 on chimpanzees. In addition to the fact that delay-limits consistently varied with the amount-of-incentive, it was found in the first inquiry that amount-of-incentive not only affected response in the trial following the observation of the size of the incentive and preceding its ingestion, but perseverated to affect subsequent responses. The explanation was suggested that an 'expectancy' or 'set' was produced by the large incentive which perseverated until, gradually, it was replaced by one which conformed to the new (smaller) incentive.

In the second experiment, size of the stimulus was ruled out as the determining factor in the effect of increased incentives. Incentive expectancy, according to the authors, was again suggested by the data. In this study the large and small incentives visibly placed in the containers had little effect upon performance when only constant-sized incentives were actually obtained. When large and small incentives were obtained as well as seen, the differential performance appeared.94

Before theorizing on the concept of expectation it would be well to widen its experimental basis by a consideration of whatever data may exist which bear upon the parallel problem of qualitative variation of incentives.

Elliot in an experiment involving qualitative change of incentive found quantitative evidence, in both time- and error-curves, of the depression effect. In this study, two groups of rats were given one trial-per-day in a 14-unit T-maze with bran mash and sunflower seeds for incentives respectively. On the 10th day, the bran-mash group was changed to sunflower seeds, a reward less desirable in terms of the performance of the two groups during the ten-day period. After this qualitative shift in incentive, the performance of the original bran-mash group became with respect to both time and errors significantly worse than the straight-through sunflower-seed group. These reliable differences may be interpreted as depression, which defines in this instance a dimension of expectation, namely an expectation of food of a particular quality. Whether the sunflower-seed group would have shown quali-

91 Tolman, op. cit., 71.
92 Nissen and Elder, op. cit.
93 Cowles and Nissen, op. cit.
94 Certain experiments, e.g. those of Tolman and Honzik (Introduction and removal of reward and maze performance of rats, Univ. Calif. Publ. Psychol., 4, 1930, 241-256), Sharp (Disintegrative effects of continuous running and removal of the food incentive upon a maze habit of albino rats, J. Comp. Psychol., 9, 1929, 403-423), and Bruce (footnote 38) bear tangentially upon the problem of this section. These are the studies involving reward and non-reward groups, and the introduction and removal of rewards. But as the amounts of reward are unspecified, the situations complex, and needed controls often lacking, there is probably little to be gained for the purpose of this discussion from an analysis of these studies.
tative elation had they been shifted to bran mash cannot be said, as Elliot's study does not supply these data. In accounting for his results, Elliot stated that "rats running the maze under the drive of hunger were learning to expect a specific reward rather than mere satisfaction of hunger."95 [Italics mine.]

Supplementing Elliot's quantitative data, a number of studies have yielded pertinent qualitative observations upon the effect of contrast-variation of incentive-quality. One of Tinklepaugh's procedures in his analysis of representative factors in monkeys was to substitute in a simple delayed-reaction situation an incentive different from the one the monkey had seen placed there before the two cups were screened for the delay-period.96 When a less desirable food (lettuce) was substituted for a more desirable (banana), the animals, as soon as they picked up the cup and found lettuce there, would typically look around on the floor; look in, under, and around the cup; glance at the other cup; look under and around themselves; look and shriek at any observers present; and finally walk away leaving the lettuce untouched. On the other hand, when the monkey found lettuce which he had seen placed under the cup he would seize it and rush away with it in his mouth, paying no attention to the other cup or the setting. It was the opinion of W. R. Miles who served as an observer in some of these experiments, that "there was clear evidence in the animal's behavior that the preparatory state of having seen the banana had predisposed him to a certain expectancy which was frustrated on uncovering the lettuce" (p. 228). [Italics mine.] Tinklepaugh spoke of the same behavior as indicative of 'disappointment.'

When the more desired banana was substituted for the less desired lettuce, the food was seized without noticeable signs of hesitation or any particular emotion. There must have been 'surprise' here, as Tinklepaugh believes; but this was not discernable in the subjects' behavior. In a similar shift in our elation inquiry (it may be remembered) there were no distinctive behavioral results apart from the rise in level of performance.

Tinklepaugh also tried a simple quantitative substitution. When two pieces of a given food were shown to be placed under one of the two cups, and one of the pieces was secretly removed, the monkeys would often search in and about the proper container just as they had in the qualitative substitutions. This Tinklepaugh believed indicated 'representations' in the monkey for quantitative aspects of the reward.

Maier and Schneirla report an experiment by Fischel97 on goats which obtained results similar to Tinklepaugh's. Two boxes were placed before the animal with lids that could easily be raised by the goat's nose. One box contained bread, the other grain. Bread was preferred in the sense that it was always eaten first when there was a choice between them. It was found that whenever the box containing the grain was opened first, the goats did not trouble to eat it, rather they merely turned their heads to the side and opened the box containing the bread. As ordinarily when goats find food they eat it as they find it, Maier and Schneirla were led to comment "for the goats actually to ignore food of a particular kind they

95 Elliot, op. cit., 29.
97 W. Fischel Weitere Untersuchung der Ziele der tierischen Handlung, Zsch. vergl. Physiol., 11, 1930, 523-548. [Cited in Maier and Schneirla, footnote 17].
must have had a memory of the other food, or we might conveniently say, they must have expected bread in the other box.\footnote{Maier and Schneirla, op. cit., 412.} [Italics mine.]

A last pertinent reference bearing upon the empirical basis of the notion of qualitative dimensions of expectation is a brief note by Lorge and Sells.\footnote{Irving Lorge and S. B. Sells, Representative factors in the rat under "changed-incentive technique," \textit{J. Genet. Psychol.}, 49, 1936, 479-480.} In a group of three problem-boxes, rats were rewarded by escape and whole-wheat-bread-and-milk for particular responses, 'standup,' 'face-washing,' and 'begging,' respectively. In a late portion of the study the reward was changed to sunflower seeds. It was found that all the animals on reaching the food cup after making the response appropriate to the particular problem-box would stop, sniff around, seize a sunflower seed, throw it aside, run back to the food-cup and repeat until all the seeds (usually 5 or 6) were thrown away. These authors believe that the rat \textit{"anticipated} the specific food of bread and milk." [Italics mine.]

The demonstration of an elation and a depression effect with contrast-shifts in incentive-quantities and the foregoing account of characteristic disturbances of behavior with contrast shifts of incentive-qualities, with shift from preferred to less preferred, argues, in the opinion of the writer, for a two-factor theory of incentive-value. Incentive-value is profitably viewed as proportional to the distance between level of expectation (both of quantity and quality) and level of attainment. The attainment of incentive-amounts and -qualities below the level of expectation, this thesis holds, is frustrating in proportion to the degree of negative deviation; the attainment of qualities and amounts above the level of expectation is elating in proportion to the degree of positive deviation.

This concept of level of expectation is similar to the principle of expectancy as formulated by Hilgard and Marquis.\footnote{E. R. Hilgard and D. G. Marquis, \textit{Conditioning and Learning}, 1940, 75-103.} Each \"may be thought of as a secondary or learned motive, differing from a drive in that it is characteristic of an individual and dependent on his prior experience, rather than being unlearned and universal for a species.\"\footnote{Op. cit., 87.} According to the principle of expectancy, however, \"re-inforcement must be such as to confirm an expectancy.\" The concept of level-of-expectation does not entirely agree with this notion. Re-inforcement may occur without specific confirmation of an expectation, in fact, if an expectation is pitched at a low level by virtue of preceding experiences, re-inforcement will be greater if it does not confirm the expectation. This is shown with respect to food-amount expectation by the elation-effect. In the parallel case, if the incentive is again not what was expected, this time below the level of expectation, re-inforcement will still occur,—but to a lesser degree than would have been the case had the incentive been congruent with the expectation.

In the course of two experiments, Bruce observed the effects of shifting rats from their customary diet to a variable diet of different food-qualities. He reported a \"breakdown of habitual behavior, the first time a drive or reward was shifted, becoming less and less evident with continued shifting.\"\footnote{Bruce, \textit{A further study of the effect of variation of reward and drive upon the maze performance of rats, J. Comp. Psychol.}, 20, 1935, 157-182.} He believed that this phenomenon is closely analogous to, or identical with, habit interference, and interpreted it in that light. The rats had built up specificity of drive, or an 'ex-
pectation,' in our terminology, because they had been habitually fed with one kind of incentive. This previous habit conflicts with the new habit demanded in the variable incentive-situation. The notion of level of expectation is not in accord with this interpretation. To employ the concept of habit-interference would necessarily imply behavior deterioration with any change from the previous incentive-habit. In contrast, the level of expectation-principle holds that if there is a shift from less desirable to more desirable incentive-quality, there will be no disruption of behavior no matter how habituated the less demanded incentive may be. This fact was shown in Tinklepaugh's experiment when more desirable banana was substituted for less desirable lettuce. There was none of the behavior-disruption that was characteristic of a change of incentive-quality in the opposite direction. This observation on quality-change cannot be verified in Bruce's experiments because he unfortunately included no consideration of relative desirability of the various kinds of food which he employed in his incentive-quality variation. As far as habitual incentive-amounts are concerned, the present paper demonstrates that, far from producing deterioration of performance, some changes from the habitual incentive (namely an increase) will raise the level of performance.

Hull includes in his conceptualization of animal behavior an hypothetical basis for the 'realization of an anticipation.' This is the goal stimulus (sG), or the proprioceptive backflow from the fractional anticipatory goal-response (rG). This construct enables us to understand, for example, why withholding the usual reward at the end of an accustomed maze-run will cause a disintegration of that particular habit-sequence while leaving the organism free to pursue alternative sequences based on the same drive. It offers an explanation of why, during a maze-learning process, the substitution of one reward for another presumably of about the same attractiveness should produce a transitory slump in the learning-scores. It throws light on why an animal evidently motivated by the anticipation of one kind of food will leave untouched a different but otherwise acceptable type of food which has been surreptitiously substituted (504-505).

The goal-stimulus cannot be accepted as a theoretical basis for the notion of level-of-expectation advanced in this paper and must be rejected for precisely the same reason as Bruce's 'habit interference.' The goal-stimulus is the proprioceptive backflow of the portion of the goal-response to a particular incentive which in a fractional form has been drawn to the beginning of a behavior-sequence and is maintained throughout it by the action of the drive-stimulus. These fractional anticipatory goal-responses are different, according to Hull, for different incentives, and hence their consequent goal-stimuli (the basis of expectations) are different. Any change of incentive will prevent a particular rG from occurring, and, in thus eliminating a particular sG, must by theory produce a disintegration of a particular habit-sequence. But any shift of incentive quality does not produce the disintegration of a particular habit-sequence or a transitory slump in learning-scores. Only shifts from desirable to less desirable incentive-qualities have been shown to do so; shift in the contrary direction in Tinklepaugh's experiment certainly produced no such effects. In such cases, the animals did not 'leave untouched a different but other-

---

103 Tinklepaugh, op. cit.
105 Hull, op. cit.
wise acceptable type of food which has been surreptitiously substituted," as Hull says, but, to quote Tinklepaugh, "in these cases [substitution of more desired banana for less desired lettuce] the subjects made their choices and seized the food without noticeable signs of any particular emotion, and without hesitation."

The present inquiry demonstrates that expectations are present for incentive-amounts as well as for incentive-qualities. If the concept of goal-stimulus as a basis for expectation is applied to these new data there is more difficulty. First, one is forced to assume that there are different goal-responses to different quantities of incentive as well as to different qualities. This ad hoc extension greatly strains the concept. Secondly, shifts in quantity-of-incentive even more unequivocally than quality do not necessarily produce a deterioration in the particular habit-sequence present before the shift, or produce a transitory slump in learning-scores. The elation-effect disproves such an assumption completely.

White, with his completion hypothesis, takes the position more explicitly and perhaps more firmly than Hull that reward or re-inforcement only arises as a completion of a fractional anticipatory-reaction—as a fulfillment of an expectation. This seems to be an expression in Hull’s system of the principle of expectancy as expressed by Hilgard and Marquis. The principle, it will be remembered, held that “re-inforcement must be such as to confirm an expectancy.” The denial of this thesis (that was presented in our discussion of the expectancy-principle) holds also for White’s completion-hypothesis. Re-inforcement may arise without the fulfillment of an expectation, and if fractional anticipatory-reactions are the basis of expectations, re-inforcement may arise without the completion of this fractional response.

**SUMMARY**

Problem I. What is the relationship between magnitude of incentive and level of performance?

1. Within the performance-period of an activity (namely a sample-period in the progress of adaptation to a situation taken when changes in performance attributable to learning appear to be largely at an end), as incentive-amount is increased logarithmically within approximately the first 70% of the appetitive range of the rat, the level of performance, in terms of average mean-speed measured arithmetically, describes a flattened sigmoid curve (Fig. 3).

2. A longitudinal view of the activity of the rat indicates that differences in amount of incentive occasion considerable differences in level of performance not only in the performance-period of an activity, but substantially throughout the entire range of the learning-period (Fig. 2).

3. Within the performance-period of an activity, the relative variabilities of performance of the various incentive-amount groups are related inversely to their levels of performance. The larger the amount of incentive with the consequent higher average mean-speed, the less variable is the group relative to the size of its mean.

---

106 Tinklepaugh, op. cit., 230.
107 R. K. White, The completion hypothesis and reinforcement, *Psychol. Rev.*, 43, 1936, 396-403. What is denoted in this paper by the term expectation White would probably term wants. He employs the term 'expectation' also, but it is not entirely clear just what is meant by it.
108 Hilgard and Marquis, op. cit., 87.
4. Within the performance-period of an activity rats run significantly faster to no-incentive-at-all than to a very small incentive.

5. Qualitative observations of behavior-patterns indicative of frustration characteristically associated with receipt of very small incentive-amounts, when contrasted with the relative absence of such indicators upon receipt of larger amounts of incentive or no-incentive-at-all, suggest an interpretation of the above quantitative finding (4) in terms of the factor, frustration-of-whetted-appetite.

6. The relationship found to hold between amount of incentive and level of performance (see 1, above), with drive held constant, indicates that experiments upon learning as inferred from performance are indeterminate without precise specification of amount of incentive employed.

7. The most efficient amount of incentive indicated in these experiments under one-trial-per-day conditions, in terms of high level of performance, low relative variability, and length of time necessary for eating, is approximately 1.3 grams (dry-weight) of Purina discuit.

Problem II. What is the relationship between magnitude of incentive and the distribution of effort (gradients) within performance?

1. In the early learning-period of an activity, as the amount of incentive is increased in successive groups, the shape of the speed-of-locomotion gradient changes from positively accelerated, through linear, to negatively accelerated.

2. In the early learning-period of an activity, as the amount of incentive is increased in successive groups the shape of the locomotion-time graphs changes from a slight positive acceleration through strong negative acceleration to near-flattness.

3. Within the performance-period of an activity, as the amount of incentive is increased in successive groups, the locomotion-time graphs vary progressively from marked negative acceleration to complete flatness.

4. Comparison of the locomotion-time graphs of the early learning-period with those of the performance-period indicates the effect of practice to be as follows; (a) For small incentives most of the improvement is localized in the middle sectors, (b) For medium and large incentives most of the improvement is localized in the first sector.

5. The first conclusion above (1) contraverts an expectation implied by Hull that increasing the amount of incentive in a linear runway would displace a linear speed-of-locomotion gradient upward and increase its slope.

6. The learning-period locomotion-time graphs obtained with varying amounts of incentive under spaced-trial conditions (see 2, above) refutes Drew's contentions that (a) a Hull-type gradient (negatively accelerated locomotion-time graph) is only found under massed-trial conditions and (b) that it cannot be found under spaced-trial conditions.

7. The practice-improvements in locomotion-time graphs obtained with varying amounts of incentive under spaced-trial conditions (see 2, above) refutes Drew's contentions that improvement in total score is a generalized function under spaced-trial conditions.

Problem III. What are the effects of variation of magnitude of incentive upon level of performance, namely (a) an upward shift of incentive-amount, and (b) a downward shift of incentive-amount?
1. Both fourfold and sixteen-fold upward shifts in amount of incentive occasion significant 'elation' effects. That is to say, the levels of performance of these groups shifted from a smaller amount of incentive to a larger become significantly superior to the level of performance manifested by rats receiving the larger amount of incentive who have not had the prior adaptation to a smaller amount.

2. No consistent differences were observable in the quality or magnitude of the elation-effect consequent upon fourfold, as compared to sixteen-fold, upward shift.

3. With a sixteen-fold upward shift in amount of incentive, the relative variability of performance decreased in accord with the inverse relationship found to hold in Problem 1 (see conclusion 3, above) among the various incentive-amount groups between the levels of performance and the relative variabilities of performance.

4. Both sixteen-fold and fourfold downward shifts in amount of incentive occasion significant 'depression' effects. That is to say, the levels of performance of these groups shifted from a larger amount of incentive to a smaller become significantly inferior to the level of performance manifested by rats receiving the smaller amount of incentive who have not had the prior adaptation to a larger amount.

5. No consistent differences were observed in the quality or magnitude of the depression-effects consequent upon fourfold as compared to sixteen-fold downward shift.

6. Both sixteen-fold and fourfold downward shifts of amount of incentive prompted increases in relative variabilities consistent with the inverse relationship found previously to hold among the various incentive-amount groups between the levels of performance and the relative variabilities of performance.

7. Qualitative observations of the reactions of the rat to downward shift in amount of incentive suggest that the depression-effect is a quantitative manifestation of some degree of frustration.

8. The elation- and depression-effects are taken as an experimental basis for defining a variable within the rat which may be termed on the basis of human analogy an 'expectation.'

9. The elation- and depression-effects associated with quantitative expectations coupled with the phenomena associated with qualitative expectation argue for a two-factor theory of incentive-value. Incentive value is profitably viewed as proportional to the distance between level of expectation (both of quantity and quality) and level of attainment. The attainment of amounts of incentive and qualities below the level of expectation is frustrating in proportion to the degree of negative deviation; the attainment of amounts and qualities above the level of expectation is elating in proportion to the degree of positive deviation.

10. This expectation-level-of-attainment hypothesis has important disagreements with the following theories bearing upon the phenomena of changes of incentives and re-inforcement: (a) The principle of expectancy as formulated by Hilgard and Marquis,109 (b) Bruce's110 notion of 'habit-interference' in accounting for the effect of shifts in incentive qualities, (c) Hull's111 concept of the goal-stimulus (s_g) as the hypothetical basis for the 'realization of an anticipation,' and (d) White's112 completion hypothesis of re-inforcement.

109 Ibid.
110 Bruce, The effect of lessening the drive upon the performance by white rats in a maze, J. Comp. Psychol., 25, 1938, 225-248; see also Footnote 102.
111 Hull, op. cit., footnote 83.
112 White, op. cit.