Estes Memorial Speech

APS Convention

Gordon H. Bower

Thank you. I'm honored to speak in remembrance of Bill Estes. I was his friend and great admirer and we stayed in touch until the end of his life. He had a very profound influence on me, providing me with inspiration and counsel. (SLIDE of Title) I'll talk about some early work we did together in relating his learning theory to that of Clark Hull's.

Let me tell you a story about how I came to the topic of today's lecture. As an undergraduate at Western Reserve University, I had become interested in the learning theory of Clark Hull. That interest was further developed when I entered the graduate psychology program at Yale University in 1955. (SLIDE of Miller & Logan) I had received a research assistantship with Neal Miller who became my major scientific mentor. I also worked with Frank Logan, a devoted Hullian theorist, who had studied at Iowa with Kenneth Spence, who was the premier Hullian devotee of that era.

At that time (SLIDE Hull) Clark Hull was the dominant theorist in learning and conditioning. Yale was the citadel where all we students were immersed in his version of S-R theory. This portrait of Hull adorned the front lobby of the

Psych department, confronting us every morning as we walked into our offices. We felt as though he was commanding us to keep the faith and carry forward the banner of Hullian theory.

On the other hand, (SLIDE of Estes) I had become very impressed by the writings of Bill Estes. I met Bill in person when I attended a workshop on mathematical learning theory the summer of 1957. (SLIDE of Bush) Bill and Bob Bush were leading the workshop that met for 8 weeks at Stanford University. Bush's teaching assistant was a young Harvard kid named Saul Sternberg. Another workshop being held next door that summer was on psycholinguistics (SLIDE of George); its leader was the young George Miller who brought along a young TA by the name of Noam Chomsky.

The members of the math models workshop were some of the young hotshots of that field. Here's (SLIDE Restle) the only photo I have in my files from those days. We were on a coffee break. This photo shows, from the left, Frank Restle, Bill, Pat Suppes, Duncan Luce, and Saul Sternberg who's pointing in the front. Here're (SLIDE) some other participants   -- Norman Anderson, Dave LaBerge, Dick Atkinson, Gene Galanter, and Danny Berlyne; others whose pictures I couldn't find were Clete Burke and Murray Glanzer. They comprised a high-powered group for a 2nd year graduate student like me. But I

didn't know about any of them and was very naïve and fearless. So I just did what the rest of them were doing, giving talks about the math models we were each working on. We even published a book (SLIDE of book) of some of our seminar papers. Bill in his quiet way helped me shape my two theoretical papers that were included in this book. (SLIDE OFF---BLACK)

Many years later, while reminiscing, Bill told me I had been invited to that workshop as a graduate student, on the expectation that I would sit back, listen quietly and be enthralled by the great works of the senior scientists. As some of you know, that quiet retiring style has never been a prominent feature of my personality.

After the workshop, I returned to Yale for my 3rd year of graduate study. Bill and I were interested in each other's ideas, so we carried on a lively correspondence between Yale and Indiana. Recently I came across a few of those handwritten notes and thought that they might form the basis for a talk like this.

The notes showed that although I was deeply embedded in Hull's learning theory at Yale, I was increasingly attracted to Estes' stimulus sampling theory. The dissonance created by my juggling those two approaches led me to try to re-cast Hull's

theory into Estes' framework. It is those efforts, undertaken in 1958 and '59 partly in consultation with Bill, that I will speak about today.

To set the context, this was in the days of the Great Learning Theory Wars (SLIDE—Spence). A highly publicized dispute was raging between the Yale & Iowa supporters of Hull's S-R theory (like Kenneth Spence) versus the UC Berkeley supporters of Edward Tolman's cognitive theory (this is Tolman). A rash of contentious experiments were being published almost weekly, arguing for Hull's or Tolman's approaches. For most of this time Estes' theory was pretty much ignored, was hardly on the radar screen. So in my schoolboy enthusiasm, I figured that if I could just translate the basic principles of Hull's theory into Estes' framework, it would give a big boost to the standing of Estes' theory among the Big Boys. If I could show that Estes' theory could explain the findings supporting Hull's theory, I thought that should bring more recognition and honor to Estes and his theory. Such was the naiveté of the young graduate student!

I'll begin by telling you a bit about Clark Hull and his theorizing. Hull was a leader in the behaviorist tradition inspired by (SLIDE #9 of Pavlov) Ivan Pavlov, Edward Thorndike, and later, Fred Skinner. Following the path forged

by Charles Darwin and John Watson, these three men aimed to make psychology an objective, biological science. So their initial plan was to uncover basic principles of behavior, mainly by studying how lower animals learn simple behaviors.

Their favorite paradigms (SLIDE dog) were classical conditioning, such as Pavlov's familiar studies, and (SLIDE MAZES) instrumental conditioning where animals, usually lab rats, received discrete trials in learning to run to a goalbox for a reward. Their learning would be indexed by their increasing speed of getting to the goal box.

Let's consider some details of Hull's theory (SLIDE OF HULL) Hull's began theorizing in the mid-1930s and continued until his death in the early 1950s. He wanted to pull together hundreds of studies being done on simple animal conditioning. Experimenters were generating a huge catalog of lawful relationships between training conditions and indicators of learning, but the field was growing haphazardly without organization.

Hull believed that learning followed just a few fundamental principles, and once he'd identified those principles he could derive from them explanations for many other behavioral phenomena. But how did he identify the basic principles?

We can imagine that one day Hull sat down and began to list (SLIDE variable lists) the kinds of training variations being studied (shown here on the left), such as the amount of training, the intensity or distinctiveness of the conditioned stimulus, the animal's degree of deprivation, the magnitude, quality, the delay of reward for the animal's response, and so on. Hull could also list the major indicators of learning that experimenters were using, such as the relative frequency or probability of the critical response, its amplitude, its preference in a competitive choice, its latency and speed, its resistance to extinction once rewards were stopped, and so on.

Hull said, "OK --- Let's simplify it". So he wrote down everyone's intuition, (SLIDE degree of conditioning), namely, that we should be able to summarize the effect of these conditions of learning into a single entity, call it the degree of conditioning. Moreover, this degree of conditioning would be related in turn to the several indicators of learning, such as response probability and response speed. An implication is that if a training variable like Drive increases one indicator like response speed, it should also produce a comparable increase in any other indicator like resistance to extinction.

Hull called the degree of conditioning excitatory potential, symbolized as sEr. He assumed that each of these input

variables (SLIDE of names) would give rise to named internal factors that contribute their part to the degree of conditioning. Hull summarized the effects of these input variables in his main equation shown here (equation slide).

$$sEr = sHr \times VxD \times K \times J.$$

That is, on a given trial, strength of the animal's response in the situation, sEr, was assumed to vary with its habit strength due to past training, sHr, the distinctiveness or intensity of the CS, V, the animal's relevant drive level, D, the magnitude and quality of reward, K, and the delay of reward following the response, J. Hull assumed these several factors would multiply together to convert habit into performance..

In Hull's theory (SLIDE of growth curve) the primary factor changing with training was habit, sHr. Using some students' data from hungry rats running to a food reward, Hull calculated that habit strength would increase with training trials according to the popular exponential learning curve as shown here. For later comparisons, I've written out the equation describing how sHr varies with its training trials. Here n is the training trial number and theta is the learning rate, and habit strength is multiplied with those other factors to get reaction potential, sEr.

Hull always viewed his set of principles not as an end in themselves but as the basis for deriving explanations for many other behavioral observations. Included was phenomena like response chaining and the goal gradient, preference for the shortest or least effortful path to a goal, response to a combination of conditioned stimuli, and so on. In this regard, he and his followers were fairly successful.

Now, let's consider Estes' theory. (SLIDE of Estes intervening variable) In preparing this talk I came to realize that the stimulus sampling approach was similar to Hull's in that I end up relating input and output variables together through a single internal factor that I called response probability. Moreover, using stimulus-sampling theory I'll be getting to an almost identical equation as Hull's for performance of a conditioned response. That in fact is the main goal at which my translation was aimed.

To begin, Estes noted that the typical learning apparatus used with animals, such as Skinner boxes and straight runways provide stimuli to which the animal learns to associate a successful response that brings him reinforcements. That is fundamental to the S-R approach. In talking about stimuli, Estes' theory becomes more abstract as he represents a given training situation like a runway as giving rise to a large

population of dozens of little stimulus aspects that he called stimulus elements. Even a simple tone or light used in a conditioning experiment was assumed to give rise to a couple dozen of little stimulus elements --- --- which is decidedly not the way laymen would describe it. But bear with me awhile because Estes' theory begins to do wondrous things with these abstract stimulus elements.

On any given trial, only a sample of these elements will be noticed and effective (SLIDE #15 marbles). This slide uses the analogy that views the stimulus population like a bag of marbles from which a random sample or handful is drawn on each trial. We let theta θ be the probability that any particular element becomes active in the sample on a given trial.

This is an S-R theory, so for a given experimental setup, Estes views each stimulus element as being connected --or not -- to some reference response or critical response that the experimenter is measuring, such as the animal's lever pressing or running to a food box. In this slide elements labeled 1 denote those connected to the critical response, and elements labeled 2 are not connected to it --- perhaps they are associated to some random, incompatible behavior. On the trial depicted here, 4 elements are sampled and become active; two are connected to response 1, so 2 out of 4, or 50% are

connected to the critical response. Estes considered that 50%, that is, the percent of conditioned elements in the sample, to also be the probability that the critical response will occur on this trial.

In this theory, learning occurs by the stimulus elements being sampled on a trial and becoming associated to the response reinforced on that trial. I will assume that only a certain fraction, K, of the sampled elements on each trial become correctly associated. We'll see later how this K-factor is used for other matters.

(Learning slide). We can describe the trial-by-trial change in conditioning by the equation in this slide. Reinforcement increases response probability from one trial, n to the next trial, n+1, moving it a fraction of the way towards an asymptote of K%. K will usually be a high number, like 70 to 90%.

(SLIDE OF CURVES). If the animal is consistently rewarded over trials, that trial-by-trial equation generates an overall learning equation like that at the top of the slide, which describes curves like those pictured at the bottom of the slide. This is the analog in Estes' theory to the growth of habit strength in Hull's theory. The graph shows three curves of response probability, starting at an initial value of 15%, rising with training to an asymptote of 85%, moving up at the rate of theta

equal to .10, or .15, or .25. This equation describes how any of the stimulus sources are becoming conditioned as training proceeds.

Let's now consider the several sources of stimulation on a trial in order to describe overall performance. I will identify 3 different sources of stimuli, as shown here [SLIDE of 3 boxes].

First are stimuli arising from the experimental apparatus, such as the runway, or from conditioned stimuli presented by the experimenter, such as the bell in Pavlovian conditioning. These stimulus elements are depicted as little c's.

Second are stimuli arising from the subject's internal milieu, such as his feelings of fatigue, or his state of satiation or deprivation. Examples are feeling full or feeling hungry or thirsty ---- such as a parched, dry mouth from thirst, or stomach pangs from hunger. These elements are labeled with little d's. Note that I let $P_{c,n}$ and $P_{d,n}$ denote the proportion of these stimulus sources that are conditioned by trial n.

Third, Estes assumed there would be a set of extraneous stimuli that recur rarely throughout the experiment, such as random outside noises or flickering lights in the lab, stray odors, ticks, flea bites, itches, or whatever might distract the animal. Estes assumed that extraneous cues never recur, which implies that their conditioning remains at zero. That's why $P_{E,n}$ is set to zero. Extraneous cues act as a sort-of drag or an

impediment that must be overcome by the conditioning of the CS and drive cues.

This next SLIDE (Combining…#18) shows how total performance is determined by these several stimuli, each contributing many or few elements to the sample. I let C represent the proportion of the sample that comes from the CS and apparatus cues, and let D be the proportion of elements that come from the Drive cues. So this equation says that response probability is determined by the weight and conditioning of the apparatus cues, and the drive stimuli.

Because the apparatus and the drive cues are presented every trial during training, I will simplify by assuming that they are equally conditioned. If Pd is roughly equal to Pc, and we substitute the earlier result for the proportion of elements conditioned by trial n, we arrive at the bottom equation here. This is how total response probability should vary with training trials. This is the fundamental equation for Estes theory, and is the analog of Hull's reaction-potential equation. In a few minutes I'll show that it's practically a simple translation of Hull's equation.

Okay. Let's begin applying some of these ideas to behavioral predictions.

First, (SLIDE of 3 boxes) it was well known that in Pavlovian conditioning a more intense conditioned stimulus such as a louder bell leads to better conditioning. In Estes' theory a more intense CS would increase the prominence and weight of the CS in the stimulus complex relative to the zero-conditioning of extraneous, distracting cues. (SLIDE on CS Intensity) That increase in C will be reflected in a larger amplitude of responding to the CS. So the simple theory explains stimulus intensity effects on conditioning.

Second, in a 1958 paper, Estes proposed a theory of how variations in an animal's drive level should affect its performance. For example, in a representative experiment by John Davenport in 1956 (Davenport SLIDE #19), rats were trained to run to a food reward under different drive levels --- high, medium or low hours of food deprivation. They gave running speed curves that looked like this

In Estes' theory, (3 sources SLIDE) food deprivation increases the number of hunger drive-stimuli and their sampling probability, thus increasing their weight in the overall sample relative to the extraneous cues. Therefore, Estes' theory predicts performance curves like Davenport's, as shown in this SLIDE (Drive CRs). Also, as the theory predicts, if, after training, the animal is shifted from one drive level to

another, its performance shifts rather quickly to the level appropriate to the new drive.

Estes theory works the same for describing the behavior of animals that are motivated to escape from painful external stimuli, such as happens when they are running to escape electric foot-shock or to escape loud noises. As the aversive stimulus increases, so does its salience and weight in motivating the animal to escape more quickly.

Furthermore, (new SLIDE of Hunger vs Thirst) different drive states like hunger and thirst have different internal stimuli, represented here by small h's or small t's. In Estes' theory, these distinctive drive stimuli provide the basis for animals to learn to respond appropriately to different drives. On this basis, a rat could learn to turn left in a T-maze for a drink of water when thirsty, and to turn right for food when it was hungry.

Although Estes' theory of drive yields many other confirmed predictions, it's time that I move on to other matters.

So far I've dealt mainly with response probability. But the most common measures of learning as we've seen were response amplitude, latency, and speed. So how does Estes' probability theory make contact with those measures?

Let's start with salivary Pavlovian conditioning. As this SLIDE (Time Bins) illustrates, I will divide the overall CS-UCS interval of length T into small units of time of length h, such as a half second, during which the recording apparatus can record at most one drop of saliva, or not. This slide shows that on this trial the dog produced 6 drops over the 12 time bins. We let our learning variable, P_n, represent the probability that the animal produces a drop of saliva in each of the little bins of time. If there are T/h time bins within a trial, this model expects an average of P-times-T/h drops to be emitted during the CS-US interval on a single trial. (SLIDE of CR Amplitude) Therefore, as P_n increases over training trials, so will the amplitude of the conditioned salivation, as shown here.

A similar story can be developed for amplitude of a conditioned motor reflex such as leg flexion or an eye blink to an air puff, but I haven't time to describe those ideas here.

Let's next consider response time (RT latency SLIDE), say, of a hungry rat learning to run down a runway to a food reward. In a 1950s paper, Estes showed how to convert response probability into response time. He imagined that at the beginning of a trial, we put the rat in the start box, open the gate, start a clock and divide time elapsed into small time-units

of length h milliseconds. We let Pn denote the likelihood that the animal advances out of the start box and into the next runway section in any given time unit. Then 1-P is the likelihood that he makes some competing response that delays his advance out of that section, such as stopping to sniff the floor. The average number of failures that occur before the rat moves out of the start box will be 1/p, and so his time will be h/p. So that's describes moving him just out of the startbox.

So, what about getting the animal down a longer runway? The next slide (SLIDE 5 segments) illustrates an animal that has to traverse 5 runway sections to get to the goal box. In each section he may carry out 0, or 1, or more competing responses that will interrupt and delay his progress. Open dots here denote such delaying responses and a filled dot denotes his forward movement into the next runway segment..

This slide depicts a trial when the rat wastes time on 8 such competing actions before completing his fifth goal-directed movement, so his total time is 8 + 5, or 13 times h milliseconds. Assuming that forward movements occur in each of the sections with probability P, then his total time to get to the goal box is given by the negative binomial distribution that has an average time of h times the number of sections divided by p.

Since speed is the reciprocal of time, the derivation means that the average speed will be proportional to Pn. Thus, the theoretical equations we've derived earlier for Pn may be considered as linear predictions of response speed, and that was the performance measure most frequently used by Hull and others back in those days.

So we've now seen how Estes' statistical theory relates response probability to speed and amplitude measures of performance. So, moving on, how does the theory explain the effect on performance of varying the magnitude and delay of reinforcement?

The means for doing so are in this equation that you saw earlier, where consistent rewarded trials (Magnitude equation slide) causes P_n to increases to an asymptote of K percent. So I will now simply assume that K will reflect the magnitude and quality of the reward. An implication is (SLIDE - Magnitude Speed Curves) that animals trained with large, medium, or small amounts of reward will produce performance curves like these; and such performance curves were frequently observed in experiments of that era. As you might expect, a random mixture of some large and small reward trials throughout training will produce intermediate performance. Also, changes in the amount of reward magnitude, say, from a small to a large

reward, will cause animals to shift gradually to the appropriate level of performance.

(Slide Utility curve) We can think of K or incentive as a kind of subjective utility, and suppose that it is a concave downward function of the reward magnitude. That is a well-known utility function, often used to describe people's satisfaction stemming from variations in monetary rewards.

Let's move on. What happens if we delay the reward after the critical response? Consider a rat learning to press a lever in a Skinner box, and we impose a delay between his lever press and reward. Typical performance curves for rats receiving short, medium, and long delays of reward often look like this (SLIDE for Delays).

A naïve rat put in the Skinner box doesn't know which of his activities produce reward. He probably makes many movements before he accidentally hits the lever and starts the delay of reward. But it's difficult for him to connect the lever pressing with the reward that arrives later. He may think the reward comes about because of superstitious responses he makes during the delay interval.

We can formulate this process as follows: (SLIDE of Competing J). Let s be the likelihood that in each small unit of delay time the animal performs a competing response. The

likelihood that he makes no competing responses over a delay of T/h time bins is (1-s) raised to the power T/h. So that is the likelihood, J, that the correct response is nevertheless reinforced by the delayed reward. That function is graphed at the bottom of this slide. On the other hand, 1-J is the likelihood that competing superstitions will occur closer to the reward, and so they rather than the correct response will be strengthened when the reward finally arrives on this trial.

This competing-response explanation was bolstered by several animal studies in the 1950s. Those studies increased the opportunity for animals to perform an incompatible response during the delay.. Such studies found indeed that a given delay of reward had a more devastating effect on performance when the apparatus encouraged more competing responses during the delay.

Okay, let's now combine our result on delayed reward with the earlier one on reward magnitude. (SLIDE Equation) This gives us the top equation for the expected trial-by-trial change in the average probability, viz.,

$$P_{n+1} = P_n + \theta(KJ - P_n)$$

You can see that an effective reinforcement of the desired response requires two things --- a sufficient reward magnitude,

K, and a delay of reward interval during which no incompatible responses occur before the reward, the J.

If the reward conditions remain constant throughout training, then the learning curve is described by the second equation in the slide, viz.,

$$P_n = KJ \, (1 - (1-\theta)^{n-1}).$$

That is, conditioning rises to an asymptote equal to K times J--- that is, the size of the reward and its delay.

Recall how we earlier showed that the CS and Drive cues combine in the sample to yield total performance. That led us to the third equation here --- total performance is (C + D) times the conditioning of each of those stimulus sources.

We then substitute the learning equation from the third line for Pn to arrive at the equation at the bottom of the slide.

$$P_n = (C + D)KJ \, [1-(1-\theta)^{n-1}].$$

(SLIDE of rearranged terms) We can rearrange terms to arrive at the top equation here. This is the equation we have been aiming for. It says that the conditioning of the desired response in stimulus-sampling theory varies with the CS intensity, C, with drive level, D, with the magnitude of reinforcement, K, and with the delay of reward, J. In nearly all respects, this equation is almost the same as Hull's

fundamental equation, which I've repeated here on the last line:

$$sEr = \ sHr \times V \times D \times K \times J.$$

The difference is that my formulation is written in terms of response probability that is close to observable performance, whereas Hull's sEr is a theoretical abstraction that requires some elaborate measurements. This correspondence means that Hull's principles have effectively been translated into stimulus sampling theory.

The equation above describes the probability or strength for a single response, indicated, say, by the animal's response speed in a runway. But how does this approach deal with animals being given explicit choices such as in T or Y-mazes ? (SLIDE of Y maze) We may think of a Y-maze as just two runways arranged side by side, with the animal choosing the momentarily stronger of the two response alternatives.

A model of choice behavior I proposed at that 1957 workshop (RATIO SLIDE) implied that such choices depend on the ratio of the two strengths, as shown in the top equation. For making predictions, it's convenient to form the Odds ratio in the middle row, dividing the probability that Option 1 beats Option 2 by the probability that Option 2 beats 1. When preparing this talk, I realized that I could plug in our equations

for the two response strengths, as shown in the bottom of this slide. That approach makes several predictions about asymptotic choice percentages. I know of one that has been confirmed in experiments, namely, that choice percentages are independent of the animal's drive level. Other predictions are that if the two reward magnitudes are equal, or the two reward delays are equal, they cancel out in the Odds ratio; that means that choice percentages should be the same regardless of how much reward is being delivered on both sides or how much delay of reward is imposed on both sides. I don't know of any data on those predictions. It was only later a couple years later that I also realized that my choice model implied Luce's theory of choice. That was lucky since Luce's theory has always been a powerful and fruitful way to think about choice, far better than Hull's. (SLIDE OFF)

OK, that completes my translation of Hull's theory into stimulus sampling theory. For me its significance was that Estes' statistical theory should be able to explain many of the findings that supported Hull's theory. That should put stimulus-sampling theory on an equal footing with the theories of Hull and Tolman that were so dominant in those waning days of the Grand Learning Theories. So, in the theoretical competition of those days, Bill and I could sing a private duet

using the lyrics of Irving Berlin's song, "Anything you can do, I can better."

OK, I'll wrap up. This has been a history lesson, discussing some very early work that Bill Estes and I did, relating Hull to stimulus sampling theory. Although such speculations might have been noteworthy at the time, I never published any of it and Bill published only a little bit on the Drive theory. In retrospect the speculations are very much confined to their historical time capsule. Moreover, beginning in 1960, Bill and I both joined the cognitive revolution and essentially moved on to studies of human memory and category learning.

Nevertheless, these exchanges with Bill were very important for me for three reasons. First, they fostered the burning passion I was developing for mathematical theorizing. Some such burning passion is needed to motivate young research students. Second, Bill became my mentor and encouraged me in my later research on human memory, even though all my Yale training had been in animal conditioning. He modeled how to move on into new fields of research and theorizing.

Third, Bill and I continued our discussions for many years about animal conditioning, a field from which both of us had come. Long after we had moved into studies of human memory

and category learning, we nonetheless enjoyed these sideline discussions of new findings from animal learning. Those conversations are some of the fondest recollections I have from my interactions with Bill.

So I present this little history as a way of rounding out the early days of Bill Estes as we interacted some 59 years ago. Not only was he a superbly creative scientist, he had an enormous impact on our field. I feel lucky to have known him so well for so many years. (SLIDE of Bill & Kay).

Thank you.

As a side comment, this response-time model claims that the reason animals get to the goal box sooner with more training or more favorable reinforcement is not that they are actually running faster; rather, they decrease their time largely by reducing the number of competing interruptions as they traipse down the runway. John Cotton, another of Estes' students at Indiana, reported strong evidence for this view in a 1952 JEP article. (Cotton SLIDE) He tested two groups of well-trained rats under different drive levels and found, as expected, that high drive levels did indeed cause animals to reduce their total time in getting to the goal box --- that's the top curves here. Importantly, Cotton also closely observed how the animals got themselves down to the goal box. He observed that they cut down their time almost entirely by eliminating behaviors that interfered or competed with running. On trials when the rats made no observable competing responses (that's the lower flat lines), they took about the same time to get to the goal box regardless of how hungry they were. So that flat line supports Estes' competing-response theory about total response time.

*We can derive a time-discounting function from more elementary considerations. Specifically, consider a rat traversing a runway to a goal box where he encounters a delay of T seconds before he gets his food reward. I will suppose that during the delay the animal might perform one or more behaviors incompatible with running forward, so that those incompatible responses would be reinforced when the reward finally arrives. Thus, we need to calculate the likelihood that NO incompatible response will occur during the delay. (SLIDE delay) If we let **s** denote the probability of an incompatible response occurring in each small time unit of length h, then over T/h small intervals there is probability 1-s raised to the T/h power, or J =(1-s)$^{T/h}$ that the reference response will receive an effective reinforcement on this trial. In contrast, with probability one minus J, one or more competing responses will occur during the delay, so the competing response would be reinforced instead. The curve at the bottom of the slide shows how reinforcement probability and performance will decline with longer delays.*

This incompatible-response explanation was bolstered by several animal learning studies in the 1950s (SLIDE on K vs Delay). Those studies increased the opportunity for animals to

perform an incompatible response during the delay, and found indeed that any given delay of reward now caused a much greater drop off in performance.

There is one final topic I'd like to address, namely, how this formulation deals with explicit choice situations such as rats choosing in a T-maze. (SLIDE Y-maze) We may think of a T-maze as just two runways laid out side-by-side and the animal's choice is determined by the strength of the two response alternatives.

This allows me to introduce the other math model I was working on at that 1957 workshop, the VTE model for choice behavior. A diagram of the model is shown here (SLIDE) and it's easy to read. The animal starts in state 0 at the choice point of the T-maze. He orients towards alternative 1 with probability o_1 or option 2 with probability o_2. Once he's looking at Stimulus 1, he considers its attractiveness and with probability p1 goes for it and terminates the trial. Or he may hesitate and go back to the beginning state and orient to the other alternative and may decide to choose that one with probability p2. Or he may hesitate and go look at the other one again and maybe choose it.

This back-and-forth oscillation continues until the animal eventually chooses one of the options. If the orienting probabilities are equal, then the likelihood that the trial ends with choice of option 1 just depends on the ratio of the two

strengths, as shown in this SLIDE. If we plug in the PT equations from before, we see that the odds ratio depends on the ratio of the two habits, their incentive magnitudes, and their delays of reward.

An interesting initial fact is that the drive levels cancel out, so choice is independent of drive --- and that is mainly true across animal experiments. This slide shows two experiments: one by Peter Carlsson showing equal T-maze learning for high- and low- drive rats; the second by me in 1965 where hungry rats under high or low drive received equal exposure to 4 vs 3 food pellets in the two-arms of a T-maze. After training they settled into choosing the 4-pellet side at 84% regardless of their drive level.

Another observation (SLIDE) is that if we force rats to have equal exposure to the two sides of a T-maze, their free choices will reflect the ratio of the reinforcement probabilities and incentive values of the two options. For example, varying the likelihood of reinforcement on the two sides will cause the animals' choice percentages to eventually match the reinforcement rates on the two sides.

The VTE model (SLIDE) is about the same as Duncan Luce's choice axiom. A first implication is this: given pairwise choices among any three options, the likelihood of choosing option one

over option two is predicted by knowing the pairwise likelihoods of choosing options 1 and 2 over any third option, X. A second implication is that the odds of choosing option 1 over option 2 is independent of how many other options are competing against them. I mention these features because Luce's theory of choice was very strongly supported by a variety of data from human and animal subjects. (SLIDE OFF)

A Markov choice model I proposed at that 1957 workshop (RATIO SLIDE) implies that such choices depend on the ratio of the two strengths. This is most clearly seen when we examine the odds ratio of choosing option 1 over option 2, as shown here. We can then plug in our equations for the two response strengths and proceed to predict choice probabilities. That approach makes many predictions about choice that have been confirmed in experiments, such as that choice percentages are independent of drive level. Luckily, my model also implied Luce's Choice Axiom. And Luce's Choice Axiom has always been a powerful and fruitful way to think about choice, far better than Hull's approach to choice behavior. (SLIDES OFF – BLACK)