A MECHANICAL MODEL FOR HUMAN ATTENTION AND IMMEDIATE MEMORY

D. E. BROADBENT

Applied Psychology Research Unit, Cambridge, England

Many people have a natural distaste for model building. A mechanical model is essentially a theory expressed in material parts rather than in abstract symbols such as words or mathematical expressions. Its logical standing is the same: that is, it stands or falls by the degree to which it fits the results of experiments on human or other animals. Yet many models in the past have been somewhat undistinguished in the closeness with which experiment has been considered in their design. It is difficult to avoid feeling that this is because a model is unduly laborious to build as compared with a verbal theory, so that the builder tends to become obsessed with the properties of his model rather than those of the organism. Consequently there is much to be said for building theories verbally, and especially for using the qualitative terms of information theory in the hypothetico-deductive fashion ably set out by Mackay (29). Such an approach has certain advantages over the other popular alternative of quantitative S-R terms; these advantages have been considered elsewhere (11).

For example, the writer holds that the human perceptual system has a limited capacity, that in consequence a selective operation is performed upon all inputs to the system, and that this operation takes the form of selecting all inputs having some characteristic in common.

Such an operation extracts little information from the signal and thus should be economical of nervous mechanism. Characteristics on which the selection can operate may be named "sensory channels." The particular selection made at any one time will depend partly on characteristics of the input itself (physical intensity, earliness in time, absence of recent inputs on that channel, position of the channel in the hierarchy of all channels) and partly on information in a more permanent store. The change from one selection to another will take a determinate time.

Incoming information may be held in a more temporary store at a stage previous to that of the selective operation. Such information will pass through the perceptual system on the next subsequent selection of the sensory channel of its arrival, if it is still in store; but the probability of the latter condition being fulfilled will decline with time spent in store. After passage through the perceptual system, information may be returned to the same temporary store, the selection of information for such return being determined by information in a more permanent store.

But although such a purely verbal theory may fit experimental results, it is difficult to communicate to others without putting them to the trouble of learning the necessary vocabulary. And if the theory is rephrased, still abstractly, it is open to misinterpretation; thus Deese (19), in a paper which makes a number of valuable contributions to the theory of prolonged work, has described the writer as postulating an inhibitory construct in human per-
formance. This misunderstanding probably arises from the fact that the theory outlined in the last two paragraphs is intended to apply both to conditioning and to human watch keeping; in neither case is it thought necessary to find an inhibitory construct of orthodox learning theory type. To say that conditioning and perception are related is not to apply a particular interpretation of the former to the latter.

Clearly, then, some sort of expository device is needed for an abstract theory using unfamiliar terms. And it is even difficult for the theorist to remember in abstract form the results of the many different experiments which a good theory should consider. A simple mechanical model has the virtue of avoiding these difficulties. It has other vices: it may have accidental properties which mislead research. Perhaps the best compromise is to state a theory in abstract terms, and also to give a model which can be described by the same verbal theory. Information concepts are applicable to any system, whatever its physical nature, and so may equally fit a model or a man. This is the approach of Deutsch (20, 21), and it has real advantages, independent of the value of his particular theory. The present paper is therefore intended to describe an extremely simple model of the human perceptual system. It may serve both as an easy introduction to the formal theory in information flow terms and also as a convenient mnemonic for the results of a number of experiments.

The Basic Model

The necessary requirements are a Y-shaped tube (Fig. 1) mounted vertically, and a set of small balls. Each ball bears a number so that all are individually recognizable. The Y tube has a narrow stem which will just take

only one ball, though the branches are wider. At the junction of stem and branches is a hinged flap which normally hangs straight downward, but which can be pivoted about its upper edge so as to close off either of the branches of the Y. This pivoting can be done by a handle from outside the tube, controlled by stored information; purists may wish to control the handle by a punched-tape system, but a human being is an adequate substitute. When the handle is left alone the flap moves freely so that a ball dropped into one arm of the Y will knock the flap aside and fall into the stem of the Y.

In this model the balls represent the information from various stimuli. The branching arms represent different sensory channels; thus one might be one ear and the other the other ear. Alternatively, one might be the ear and one the eye. (Sensory channel is not, however, quite equivalent to sense-organ, since we would treat sounds localized in different places as being on different channels.) The bottom of the Y represents a response output, so that the process of dropping a ball into the arms and observing its emergence at the bottom is analogous to that of delivering a stimu-
lus and observing a response. The behavior of the model resembles that of man in the following ways:

(a) If two balls are dropped simultaneously, one into each of the branches, they will strike the flap on both sides; it will not move and therefore they will jam in the junction. Numerous experiments show this “distraction” effect, but there are certain advantages in citing an auditory experiment (8, 33). These same experiments show that if the handle is used to shut off one branch before the balls are inserted, then the ball entering the other branch will emerge successfully, which is analogous to previous instructions.

(b) If the two balls are not strictly simultaneous, the first to arrive will obtain an advantage by knocking the flap over and shutting out the other. This had been shown to be analogous with competing auditory stimuli by Spieht, Curtis, and Webster (36).

(c) If the Y is not perfectly vertical, the ball in the more vertical branch will have an advantage over a simultaneous ball in the other because the door will hang to one side. Equally, one sensory channel may have an advantage over another, as has been shown for high-pitched noise as compared with low (10).

(d) If one ball is flung violently down its branch, it may succeed in forcing over the door against the unassisted weight of a ball on the opposite branch. Equally, an intense stimulus may have an advantage (2, 10).

(e) After a single ball has been passed through the system, the door will swing back from the position into which it has been pushed. Naturally it will overswing, and temporarily close the branch which has just been used. A stimulus has similarly an extra advantage for response if it comes on a previously quiet channel as opposed to a previously busy one. This has been shown by Poulton (34) for auditory signals, and a related finding is that of Hyman (25) for visual reaction times. In the latter case, stimuli of different frequencies of occurrence were delivered, and it was found that average reaction time to a set of stimuli was proportional to the information conveyed by them; but the infrequent signals, while giving long reaction times, did not give times as long as those to be expected from information theory calculations. Note, however, that the time taken for a swing is important in the model; this will probably also be true in man, to judge from data which are best considered below. It may also be connected with Hyman’s finding that, for numbers of alternatives greater than two, the second of two identical signals received an unduly fast reaction.

(f) If a given number of balls are to be put through the tube, it is best to deliver them asymmetrically, the majority to one branch. There is then less risk of jamming than if equal numbers are admitted by both branches. The analogous finding for auditory messages has been reported by Webster and Thompson (37). This point is related to another, that if balls are being inserted into one branch at random intervals the effect of increasing the rate of delivery of balls through that same branch (the “speed” of work) is not the same as that of adding the same number of extra balls to the other branch (the “load” of work). There is more risk of jamming in the latter case. This point was first clearly emphasized by Conrad, using visual signals (16). Obviously the effect of using two branches rather than one will be more serious if the rate of delivery of balls is high, since this increases the probability of a jam. Conrad showed such an interaction of speed and load. Mackworth and Mackworth (28) have shown a similar
effect in a different type of task, and have demonstrated that the fluctuating difficulty of this complex task at any instant can be represented by the amount by which each signal is overlapped by other simultaneous ones.

At this point we may pause to consider the cynical reader who is wicked enough to be doubtful of the existence of a Y-shaped tube somewhere in the region of the thalamus. Such a reader will probably have noticed already a situation which will follow when two balls are dropped into the model and jamming is avoided by the door being to one side for one of the reasons listed. The favored ball will descend the stem of the Y, but the impeded ball will not therefore disappear. It will emerge later, when the door next swings back to the opposite branch. Surprisingly enough, this also happens with man. Simultaneous stimuli either jam or produce successive responses. It has been shown for three different combinations of sensory channels (7, 8), and is probably an effect identical with the “prior entry” of classical psychology. In the same way Conrad has not only shown complete failures to respond (17) but also shifts in the time at which responses appear (18).

Having thus demolished the cynic, we must add that a slight complication should be added to the model in order to cover perfectly the results of experiments on successive response to simultaneous stimuli. But this will be left until the section on immediate memory; let us first deal with the effects of speed and of prolonged performance on the simple model.

(6) As the stem of the Y is so narrow, there will be a certain amount of delay between the insertion and emergence of each ball. At slow rates of insertion each ball will, however, emerge before the next is inserted. At faster rates there will develop a lag such that one ball may be inserted before the previous one emerges. With still faster rates balls will begin to accumulate in the branch; response will get further and further behind stimulus as the task proceeds. But this is a desperate expedient and will lead to breakdown when the branch is full unless the rate slows down again. These stages have been shown by Vince (35) for visual stimuli; a related effect has also been shown in less detail in hearing (4). A result of the piling up of balls in the branch when the rate is too high is that a ball inserted at a very short interval after another will remain in the tube longer than is normal: the “psychological refractory period” (38). It must be remembered that balls are analogous to information and not to stimulation; highly probable stimuli will not give an unduly long reaction under these conditions (22).

(6) Now suppose that we wish to operate the handle in such a way that one branch always has priority: prolonged performance of a task involving only one sensory channel. If there is a fairly rapid flow of balls down the selected branch, the handle will hardly need holding after the initial setting, since the beginning of any swing back will probably be checked by another ball. But if the selected branch is not very busy, the weight of the door must be held by a positive force on the handle. In this case the fingers holding the handle will fatigue; for purists, the

---

2 At the risk of complicating the issue, we must say that “stimuli” in this sentence means “stimuli not of low information content.” A familiar predictable sequence of stimuli may quite well produce responses simultaneous with other responses (1). And the same S-R unit interferes less with another task when drawn from a smaller ensemble of possibilities (12). This point is of course implied in the statement that the perceptual mechanism has a limited capacity in the sense of information theory.
punched-tape machine will only hold one branch shut for a limited period. When the handle is released the door will swing back to close the opposite branch, and then return pendulum-wise to the desired position when it will again be held. The result will be that prolonged tasks in which unexpected stimuli appear for less than a certain period (the "swings of the door") will show marked decrements in performance. But similar tasks in which the stimuli are present for longer times will show much less decrement.

The earlier evidence for this view of prolonged performance has been summarized elsewhere (6). But earlier presentations have misled Deese (19) into supposing that this view implied a decrement with continued performance of any vigilance task. He employed tasks in which a signal was either presented repeatedly until seen, or else painted on a tube face by a sweep line which did not return to obliterate the signal for nearly three seconds. It was in fact predicted by the writer (6, p. 300) that such tasks would show little or no decrement, and this was found by Deese to be the case. The latter type of task gave more signs of decrement, probably being a borderline case, since the trace left on the phosphor screen decayed until scarcely visible at the end of the three seconds. It should be noted that an expectancy theory of the type favored by Deese, while undoubtedly applicable to some aspects of vigilance, is not able to account for the effects of varying the length of signal presentation. In our Y-tube model, the role of expectancy is incorporated by describing the balls as representing information rather than stimulation; the more probable signals receive more efficient response. But the swinging of the flap is also necessary.

To summarize the writer's present views on vigilance, the efficiency of a man asked to detect infrequent signals should be described by both a mean and a variance. The mean level is determined by such factors as the rate at which signals arrive but not necessarily by the length of time since the session began. The variance, on the other hand, increases as the session progresses, short intervals of extremely low efficiency being interspersed with fairly long periods of normal or supernormal efficiency. The score from any given task may depend on one or on the other of these quantities. Thus, for example, the earlier British work has been mostly concerned with the instants of very low efficiency; the type of results reported by Deese and by later British workers (6) mostly with the mean over appreciable portions of time.

To return to our model, two further points should be made. The first is that the door need not swing the whole way over when the handle is released; it may reach a central position just as two balls arrive. There will then be a jam rather than passage of the wrong ball. Equally, failure of reaction to a task need not imply overt reaction to some irrelevant stimulus.

Secondly, as the balls represent information rather than stimulation, tasks in which the sequence of stimuli is predictable will not show fatigue decrements of this type. To show decrements, the signals must be unpredictable either in content, as were those of Bills (3), whose "blocks" are instances of this effect, or else in time of occurrence, as were those of Mackworth (27); this, however, is only one of the necessary conditions.

**Immediate Memory**

Our Y tube does seem at this stage to have related a number of facts about perception and put them in a way which most people can understand. It is admittedly ludicrous as a description of
what really happens in the brain, but this is a positive advantage. Psychologists are not likely to mistake this model for speculative neurology, and so they should concentrate their experiments on the essentials of the theory rather than the irrelevant properties of the model. As a device for communicating the outline of the theory, however, the model seems sufficiently adequate to justify an extra complication in order to express a theory of immediate memory. This theory is in a slightly different position from the views on perception which have been given so far; the latter are entailed by the experiments, but the theory of immediate memory is not the only conceivable explanation of the observed facts. Yet it has a fairly high probability and is worth discussing.

The complication is twofold. First, a device must be supposed fitted to each branch of the Y tube, such that if any individual ball remains in the branch for more than a certain time continuously it is removed from the system completely. This could be done by filing the tube with acid, but the writer does not wish to encourage the development of a race of fingerless psychologists. Mechanical devices are quite possible, though complicated to describe, and the details will therefore be left unspecified and available from the author. The second complication is that from the foot of the stem two return tubes lead back to the branches (Fig. 2). Admission to these return tubes is controlled by a lower door which again is operated by a handle. The latter is dependent on an outside operator, or the familiar punched-tape machine—in brief, some form of stored information. Finally, as the return tubes are operating against gravity they must contain some form of conveyor, but this has no particular psychological significance.

In immediate memory experiments, as usually performed, a stream of stimuli is delivered completely before any overt response is required. Thus all the information in the stream is within the organism at one time; all the balls are somewhere in the tubes. Yet with small numbers of balls the time between insertion and emergence can be varied without any apparent effect on efficiency (13). Let us consider some of the methods which may be adopted.

(a) In the first place, let us suppose a short series of balls to be waiting in one of the branches. If they stay there indefinitely, the critical time will be exceeded and the ball which has been there longest will leave the system. Thus the branch will not serve as a store during an unlimited period of delay before response is allowed. What is perfectly possible, however, is for the series of balls to pass down the stem and, by the lower door, back up a return tube to the other branch. As soon as the full series has completed this round the process can be repeated, so that the set of balls can be kept in the system for any desired time. But note that insertion of any other ball during that time will either mean that the extra ball will never emerge, or else that the circulation of the series will meet inter-
ference. Note also that the system will only operate with a limited number of balls; above that number each ball will have to wait in the branch for more than the critical time, and there will therefore be a severe breakdown of storage. These are well-known characteristics of immediate memory. Less obvious is the suggestion that the interfering effect of an extra ball during the delay will depend on the size of the series circulating, being most serious near the limit. This has been shown by Brown (14). He also demonstrated that the interval before or after the interpolated stimuli had little effect, as one would expect, and that if the extra stimuli came before the memory span stimulus there was little effect. It is Brown’s explanation of his results which has prompted the present account.

(b) If both branches contain balls, movements of the door will put one group into circulation before the other; to move the door over while a group is actually passing would be to risk a jam. So all balls on one branch will pass down the stem before any on the other. As has been said, the memory analogy is in fact true, if a digit memory span is obtained with half the digits on one ear and half on the other, or half on the eye and half the ear (7, 8). In addition, the first branch to be dealt with undergoes less risk of any ball reaching the critical time; equally, performance is better on the first set of digits to be recalled (9, 13).

(c) So far we have considered only the case in which the branches start with balls in them. But in practice these balls will be inserted into the branch one after another, and the extra time intervals thus produced will have their effect. Thus, for example, if balls are inserted rapidly into both branches simultaneously, the difficulty of moving the door will apply, as mentioned in the last paragraph; and so all the balls in one branch will be dealt with before any in the other. But if there is a sizeable interval between successive balls in each branch, the door may be swung back and forth so as to deal with each branch alternately. A slow presentation of stimuli may equally allow them to be dealt with in the actual order of arrival rather than channel by channel (7, 8). It should be noted that the time allotted to “swings of the door,” as measured by these experiments, agrees with that determined by the experiments on prolonged performance mentioned previously (6).

(d) Again, suppose a lengthy series of balls is inserted through one branch, the door being held open for them by the handle. If a couple of balls are meanwhile delivered to the other branch and wait there till the stem is clear for them to pass, there is naturally some risk that this extra pair of balls will exceed the critical time. The risk will be greater if they are inserted with the earlier balls of the long series than if inserted with the later balls. This also is true of immediate memory for spoken digits arriving at the two ears (9). A point of interest concerns the effect of “irrelevant” balls which arrive on the second branch during a long series on the first. If all the balls in the second branch are irrelevant, the door may be kept closed against them and they will eventually exceed the critical time and be removed (15). But if some relevant balls are also on the second branch, the door will have to be opened and the mixture of relevant and irrelevant balls passed down the stem. The former can be recirculated in the usual immediate memory fashion, and the latter removed by the lower door. The presence of irrelevant balls either before or after relevant ones will therefore produce a greater risk of jamming or of the critical time being exceeded. This also is true of immediate memory (9).
(c) Now consider cases in which the order of the balls emerging at the bottom of the stem is different from that of their insertion. The extreme case is that of "backward memory span." With our simple model, the easiest situation to consider is that in which the first half of a series of balls is to emerge after the second half. This can be arranged on the first circulation by passing the first balls down the stem and back up the return tube to the branch opposite to that in which they were inserted. The later balls are returned to the same branch which they entered, and on the second circulation are the first to be admitted to the stem. After they have passed the flap the other balls can follow them. This will mean that the first balls have stayed in the branch longer than we normally be necessary for simple recirculation, and so are more likely to exceed the critical time. Therefore such a rearrangement of order will reduce the "memory span." Furthermore it will alter the order of difficulty, since the first members of the series will suffer more than the later ones. These effects have been shown by Kay and Poulton (26), and are supported by Brown (13). The reduced memory span for a rearranged list is of course familiar from intelligence testing, but the change in the serial-order effect is a more important deduction.

(f) Serial-order effects will clearly depend very considerably on the rates of presentation of stimuli and of required response, as compared to the rate of recirculation. They will also depend on the way in which stored information is used to operate the two handles—that is, on the strategy of the subject. They will certainly not be completely determined by the primitive learning process, as is supposed by some existing theories. In fact it has been shown by Kay and Poulton (26) that the serial-order effect is altered by the subject's absence of knowledge about the order in which recall will be required, even though the actual order remains the same as in a control experiment. (Knowledge of the amount to be recalled will also affect the efficiency of recall as well as the order, as has been shown by Brown [13]. This is because the uninformed subject must recirculate material which is not in fact to be recalled later; retention is an active process.)

Some general points about serial-order effect may be made, however. First, if the rate of response is below that of recirculation, but not so slow as to allow complete recirculation between responses, the first items in response will have stayed in store less long than the later items. So, as in (d) above, the earlier part of a list recalled in the order of presentation will be better recalled than the later part. Second, if the time taken to respond is eliminated—for instance, by requiring recall of only one item in the series—then the time taken to present the stimuli will be the chief factor influencing serial-order effects. The last stimuli to arrive will then be those which have been stored for the shortest time, and will be best recalled. This is the result of Gibson and Raffel (23).

Third, a special case of some importance arises when fresh material is being presented while earlier material is being recalled. Considering the model, suppose a short series of balls are inserted and recirculated to the branch opposite that by which they arrived. If they are now to be passed down the stem during the arrival of fresh balls on the original branch, the earliest of the second series of balls will suffer the longest delay. Consequently the end of the first series and the beginning of the second series will be the points of greatest difficulty. If the two series are considered as one long one, the familiar U-shaped serial-order effect will appear.
This has been demonstrated by Poulton (32). In conventional learning experiments subjects are not instructed to rehearse (recirculate) the earlier items during presentation of the later ones. But it seems plausible that, during the presentation of a long series at a medium rate, the presence of the earlier balls in the second branch should normally encourage a tendency to recirculate them while the later balls are still arriving. This would be comparable to the higher priority of a previously quiet channel, which was mentioned in the section on perception. It would result in the U-shaped curve of difficulty being the usual one for serially presented material; but the curve will be subject to great modification by instructions. Furthermore a very slow rate of presentation, resulting from allowing recirculation during gaps between presentations, will minimize the effect. This is the case (24).

LIMITATIONS OF THE MODEL

Certain properties of the model are likely to be misleading. Of these the most important has been stated several times above, but it is worth repeating. The balls represent information, not stimulation. The reader must not contemplate the Y tube and decide that two stimuli cannot be dealt with simultaneously. They can if they convey sufficiently little information. Clearly, many reflexes are compatible with one another, and it is likely that simple “voluntary” reactions are equally capable of being carried on simultaneously. It is only with unpracticed reactions involving a choice between several alternatives that we find an interference between two stimuli; but this is a very normal case outside the laboratory. Theories such as that put forward by Welford (38) for the “refractory period” undoubtedly need the qualification that highly probable stimulus sequences may not show these effects (22). But this does not disprove their general value.

A related point is more serious. The length of the immediate memory span is roughly constant whatever the size of the ensemble from which the items are chosen (31): one cannot remember enough binary digits to make the information in immediate memory equal to that stored in memorizing ordinary decimal digits (30). Yet the model makes the limit on memory span dependent on the time taken to pass the items through a limited capacity system, which will in turn depend on the information per item. Perhaps this difficulty may be resolved by suggesting that each possible item represents an extra branch on the stem, and an incoming ball is always recirculated to its appropriate branch. The time taken to withdraw all the balls again would then depend on the time taken by the flap to operate, which is not dependent on the information per item. But this is clearly leading us into complications; for the present we may merely note that the point is an important one about immediate memory but not impossible to handle with a model of this type. A minor caution which should be added is that the model is deterministic while all the experimental results quoted are statistical.

Finally, we have put forward this model as one which may be described by an exactly worded theory which applied also to man. The present paper is directed largely at those who find such a theory unintelligible in its original form, but it should be borne in mind that the theory under test is the abstract one given in the first section. Otherwise the error of identifying the model with the organism may be made, if only to discredit the theory by its obvious absurdity. The formulation given ear-
lier is not complete, but it indicates the way in which the model and the man may be described by the same abstract theory. There should therefore be no excuse for treating the Y tube as anything more than an expository device and a mnemonic for recalling the results of numerous experiments. The writer's freedom from such an error is demonstrated by the fact that he has never built his model in any physical sense.

SUMMARY

A mechanical model is described, to act as an easy introduction to a formal theory of attention and immediate memory in information theory terms. A number of deductions from the theory which agree with experimental results on human beings are given as descriptions of the behavior of the model.

REFERENCES


30. Miller, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.


(Received November 30, 1956)