An abbreviated version of this paper was presented to the National Meeting of the Institute of Management Sciences, San Francisco, February 5, 1965. This study was initially supported by a grant from the National Aeronautics and Space Administration, and since November 1963, by a grant from the National Science Foundation. The author gratefully acknowledges the assistance of Maurice P. Andrien, Jr., Richard J. Bjelland and William D. Putt, and of the many companies, project managers and engineers who must remain anonymous but without whose help the study could not have been conducted.
ABSTRACT

Three pairs of parallel R & D projects are examined. The data analyzed were gathered by means of Solution Development Records—a form which provides a weekly estimate of the probability of adoption of the approaches under consideration as possible solutions to a technical problem. It is found that the longer an approach is indicated by these forms to be in a favored position, the more difficult it is to reject. Furthermore, the number of alternative technical approaches considered bears a relation to judged solution quality. Groups producing higher-rated solutions generated fewer approaches during the course of the project, and they more closely approach an ideal strategy of approaches off on a two-at-a-time basis than do their poorer performing rivals.
The research and development process provides a challenge to the behavioral scientist who wants to study the process of human problem solving in vivo. Engineers and scientists in R & D laboratories devote their careers to the solving of rather sophisticated problems, and the study of their behavior should tell us much about the higher mental processes.

Such a study, however, presents the researcher with two difficulties which he must resolve. The first of these is access. Access must be gained, not only to the laboratory under study, but to the subtle mental processes of individual scientists. This is a standard problem in psychology, but the real-life environment serves to compound it. Observable indices that in some way provide valid, reliable estimates of the actual cognitive data must be established. In addition to validity and reliability the indices must further be highly efficient in terms of the time required on the part of subjects. The latter requirement is imposed by the field situation, and is absolutely essential to the recruitment of cooperative subjects.

The second difficulty to be faced is that of problem comparability and solution evaluation. In the psychological laboratory, this is easily managed by assigning the same problem to large numbers of college sophomores. The substance of the problem is controlled so that a valid comparison can be made of the subjects' approaches. In addition, the problems are generally of a type which has but one correct solution. The sample of problems can then be split into those having correct and incorrect solutions and approaches may be related to success rate.
METHOD

While we, as researchers, have not reached the level of affluence where we can hire a number of engineers and assign them the same problem, there are organizations such as N.A.S.A. and the Department of Defense who can. For this reason, instances in which two or more R & D laboratories have been awarded a contract to perform the same preliminary design or research investigation were sought out for study.

At the present time, about a dozen sets of two or three parallel projects have been located and studied. This paper will present the results of an analysis of three of the earliest completed sets.

Once a parallel project has been located, its work statement is obtained and analyzed and factored into a reasonable number of subproblem areas (generally subsystems). The breakdown is then checked with the technical person who prepared the work statement, and data collection forms based upon it are designed. After all data have been collected from the contractors, the technical monitor is revisited and asked to provide a confidential evaluation of each lab's performance on each subproblem. Data are gathered by four means: (1) time allocation forms, indicating the amount of time each engineer spends on the job in several activity categories; (2) before and after interviews with individual engineers; (3) periodic tape recorded progress reports by the project manager; and (4) solution development records.

The Solution Development Record, which is the principal source of
the data presented in this paper, is a research tool which provides a record over time of the progress of an individual engineer or group of engineers (or scientists) toward the solution of a technical problem. The lead engineer responsible for each subproblem is asked to provide a weekly estimate, for each alternative approach under consideration, of the probability that it will be finally chosen as the solution to that subproblem.

Referring to Figure 1, if at some point in the design the respondent were considering two technical approaches to rendezvous at Uranus, and he were completely uncommitted between the two, he would circle 0.5 for each, as shown. Eventually as the solution progresses, one alternative will attain a 1.0 probability and the others will become zero. By plotting the probabilities over time, we obtain a graphic record of the solution history. Alternative approaches are identified from the contract work statement, when so specified, or from the responsible engineer when he is interviewed prior to beginning the task. Blank spaces are always provided so that new approaches may be reported as they arise. In cases where a respondent believes there is some probability of choosing an approach which he cannot clearly specify at the time, he is instructed to assign a probability to an approach which he may call "other."

A copy of the form is mailed every week to each respondent. The forms are sufficiently flexible, so that new alternatives may be incorporated, old ones dropped, and in instances in which an early solution is reached and "frozen", subproblems at the next level may be substituted.
The Solution Development Record, by economizing on the respondent's time, provides a quite efficient record of a problem history. When the project is completed, each respondent is presented with a time-plot of his probability estimates, and is interviewed at some length to determine causes and effects of design changes reflected in this record. The plot thus provides a stimulus to the man's memory and assists the investigator in gathering a detailed record of each project.

The three projects under consideration, involved in the following general problems:

1. The design of the reflector portion of a rather large and highly complex antenna system for tracking and communication with space vehicles at very great distances.
2. The design of a vehicle and associated instrumentation to roam the lunar surface and gather descriptive scientific data.
3. An investigation of passive methods for transfer of modulation between two coherent light beams.

Only the first of these will be discussed in any detail. All three are considered in the aggregate statistics of our data.

RESULTS

The plot of Solution Development Record points over time (Figure 2)
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Solution Development Record

Manned Uranus Landing in an Early Time Period Study
General United Aerospace Corporation

Name ___________________________ Date ___________________________

Estimate of Probability that Alternative will be Employed

<table>
<thead>
<tr>
<th>Subproblem #1: Method of rendezvous at Uranus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative approaches:</td>
</tr>
<tr>
<td>orbital rendezvous mission with excursion vehicle</td>
</tr>
<tr>
<td>orbital rendezvous mission without excursion vehicle</td>
</tr>
<tr>
<td>direct mission</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subproblem #2: Design of the electrical power supply subsystem for the space vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative approaches:</td>
</tr>
<tr>
<td>hydrogen-oxygen fuel cell</td>
</tr>
<tr>
<td>KOH fuel cell</td>
</tr>
<tr>
<td>Rankine cycle fast reactor</td>
</tr>
<tr>
<td>Rankine cycle thermal reactor</td>
</tr>
<tr>
<td>Brayton cycle reactor</td>
</tr>
</tbody>
</table>

FIGURE 1
provides a rather interesting perspective on the history of a project and illustrates the intimate relation between technical information inputs and problem solving process.

We can see here the approaches followed by two engineering groups (labelled Lab A and Lab B) in the design of the reflector surface for an antenna. While both teams ultimately decided upon the same general approach, they arrived there by quite different routes. A brief summary of the history of the two solution processes for this subproblem will explain much of this difference, and will provide some background substance for much of the data to be presented later.

Rather early in the study, as indicated by flag notes A-1 and B-2, the customer agency supplied both contractors with the results of an experiment to determine the wind loadings which would be experienced by the antenna. As a result, approach \( \beta \) rose in favor at both labs. Prior to this time, Lab A showed considerable vacillation among the three alternatives; Lab B, during this early period, conducted an intensive literature search but failed to uncover any evidence of empirical or analytical work having been done with flat aerodynamic shapes at the low air speeds under consideration.

At A-2 Lab A's aerodynamic staff reported wind load moments for approach \( \gamma \) to be about twice that for \( \alpha \) or \( \beta \). At the same time, the electromagnetic staff reported that approaches \( \alpha \) and \( \beta \) satisfactorily met electrical requirements up to 3 gigacycles.

An electromagnetic analysis by Lab B shows that approach \( \alpha \) failed to satisfy the electrical performance specification. This is indicated by
FIGURE 2

DESIGN OF ANTENNA RADIATION SUBSYSTEM
flagnote B-3. During the same period (B-4), this laboratory attempted to extrapolate from previously acquired data (an earlier antenna) to estimate wind loads.

About the twelfth week (A-3) Lab A conducted a wind tunnel study, which showed that approach $\beta$ resulted in a wind torque considerably larger than that predicted by the customer's data (A-1). Since approach $\delta$ did not perform as well as $\beta$, electrically, one of three things now had to be done. Either the aerodynamic specification had to be relaxed, a penalty in electrical performance accommodated, or a new alternative meeting both the aerodynamic specifications and the electrical performance level of approach $\delta$ had to be generated. The latter possibility did not present itself, so negotiations with the customer over specifications were pursued.

The customer subsequently allowed a relaxation of the electrical specification (A-4) and approach $\delta$ rose in favor. The change in specification was provided to Lab B as well, but there was no consequent change in the probability level of approach $\delta$. This undoubtedly resulted from the fact that B did not have as complete information as A regarding approach $\beta$. The brief drop in $\beta$'s position at Lab B (B-5) was a result of some doubts which this lab had concerning the feasibility of the approach, but as far as can be determined this was not based upon hard data. Following this brief period of skepticism, $\beta$ rose rapidly to a 1.0 level (B-6) and was further established there when, as indicated by B-7, information concerning special fabricating machinery became available.
At Lab A, meanwhile, approach A encountered some difficulty with the cost analysts. And we see a resulting dropoff at note A-5. Lab A remained indifferent between A and B for quite some time while trade-off studies were pursued (A-6). Numerous contacts were made with vendors to determine the costs associated with various elaborations of the two approaches.

Finally, at A-7, information was obtained from the Weather Bureau which allowed a 20% reduction in the wind loading specification. This information was instrumental in Lab A's decision to adopt approach B.

The work statement for the subsystem under consideration suggested three technical approaches, and these were the only ones considered by the two teams. That this is not always the case can be demonstrated by a look at another of the antenna subsystems (Figure 3).

Here we see design histories for the position feedback subsystem. Space does not permit a detailed discussion of the protocol for this subproblem, so we will merely point out the introduction of new approaches during the course of the project. The work statement for both labs suggested approaches A, B and C. Both rejected these, however, and generated two new approaches each (D, E, F, and G). In both labs one of the new approaches resulted from difficulties incurred by the currently preferred approach; the other resulted from receipt of new information, and was independent of the state of approaches currently under consideration.
FIGURE 3

DESIGN OF SUBSYSTEM TO DETERMINE ANTENNA POSITION
Decay Time for Rejected Approaches

There are a number of ways in which the data contained in these plots can be analyzed. Looking at the time required for rejection or introduction of approaches, we see that old ideas are not dispensed with very abruptly, nor do new ideas rise very rapidly. Once an R & D group has committed very much time to the design of a particular approach, it is not easily convinced that this approach should be dropped. Analysis of 31 instances shows that a once dominant idea takes about six or seven weeks on the average to be completely rejected. Dominance is defined as being in the preferred position in the probability plot. Decay time is measured in two ways:

1. if the old approach is replaced by a new idea, decay time is the period between introduction of the new idea and the point at which the old approach reaches a probability of zero and stays there.

2. if the old approach is replaced by one of the approaches initially considered, decay time is the period between the point at which the two approaches trade dominance or reach a tied position, and the point at which the rejected idea reaches zero and stays there.

There is a clear relation between decay time and time in dominance. \( r = 0.54, p < 0.001 \) The longer an idea is preferred, the more inertia is

---

Pearson Product Moment Correlation Coefficient \( (r) \) with a probability of occurrence \( (p) \) of a coefficient this high, under the null hypothesis of no relation, less than 0.001. The correlation coefficient varies in magnitude from zero to one and is an index of goodness of fit to a least-squares straight line.
built up, and the less easily is it rejected. This is independent of
the absolute level of preference (r=0.03) and of the number of other
alternatives being considered (r=-0.04).

As engineers invest time and effort in the formulation and deve-
lopment of a technical approach, they become more and more committed
to that approach, and hence more resistant to disconfirming informa-
tion. An analogy may be drawn here to Bruner's (1957) paradigm of
the decision sequence in perception. During the initial consideration
of a new idea, the engineer is relatively "open" to stimulation from
his information environment. As the amount of confirming information
increases, he raises his probability estimate for the alternative and
modifies his search behavior. The "openness" to new information de-
creases sharply now. The longer he retains the approach in a dominant
position with respect to its alternatives, the more he attenuates
his search behavior. At this point he enters a stage quite analogous to
Bruner's final stage in perceptual identification, where:

"...openness to additional cues is drastically reduced,
and either normalized or 'gated out.' Experiments...
suggest that once an object has been categorized in a
high-probability, good-fit category, the threshold for
recognizing cues contrary to this categorization increases
by almost an order of magnitude." (p. 131)

What is being suggested here is that an engineer tends to develop
a similar threshold as he becomes committed to a technical approach, and
that this threshold severely inhibits the effect of information which
should tell him that the approach is defective in some way. In addition,
it appears to gate out information related to new alternative approaches.
Examination of protocol data reveals that the vast majority of new alternatives introduced during the course of a problem result from an active search resulting from some defect in the currently preferred approach. Seldom does a new idea enter the system solely on its own merit.

From studies of perception, and the receipt by an engineer of new technical information is certainly a complex form of perception, there is a great deal of support for this threshold hypothesis. McGinnies (1949) has demonstrated a physiological defense mechanism in the human organism, and a tendency toward higher recognition thresholds for taboo as compared to neutral words. Bruner and Postman (1949) have demonstrated a tendency to fixate upon an early perceptual hypothesis and to subsequently resist the recognition of disconfirming information. Postman, Bruner and Walk (1951) have shown the operation of such a threshold to gate out those signals which would tend to disconfirm a perceptual hypothesis.

Relatively few data are available in the purely cognitive area. One may, however, hypothesize that the operation of threshold development in this context is a result of the progressive buildup of cognitive organization, i.e., of the multiplexity and interconnectedness of the engineer's cognition of his design.

This can be illustrated by considering an engineer designing, for example, a power supply. At the start of the problem he may know little concerning specific alternatives, except that they are different and have certain very general characteristics which may or may not be ex-
pected to fulfill the needs of the problem. As he gains more information concerning an alternative, he shifts, gradually, from thinking of it as simply a "radioisotope power supply" and comes to think of it as a "radioisotope thermionic supply": employing Cm 242; having a certain amount and kind of lead shielding; weighing four pounds; generating twelve watts of effective power; and having an operating temperature of 1500°K. In addition to this increase in multidimensionality or multiplexity of the concept, it becomes progressively more interconnected or integrated into the design of the system. Several other subsystem designs will usually be dependent upon the solution to a particular subproblem such that there is a great deal of interdependence among the subsystems. As the design progresses these interdependencies become stronger and more pronounced. Various decisions in, for example, the instrumentation subsystem of a space vehicle are dependent upon the power supply design. Changes in that design imply changes in instrumentation which in turn imply changes elsewhere and so on. Such changes become more difficult to make as time progresses.

All of this is reflected to some extent in the cognitive structure of the engineer designing the power supply. As the organization of his cognitive map progressively increases, it becomes more and more difficult for him to envisage fundamental changes. It seems perfectly reasonable then that his cognitive system should become more resistant to change and that a threshold against information implying fundamental change should build up as part of this process. As a solution alternative becomes more embedded in a growing cognitive organization, it develops a greater resistance to change.
This phenomenon can be seen quite often in the protocol data. An engineer decides rather early to pursue a particular alternative and formulates his design to a considerable degree of sophistication. Then he is confronted with information that something is wrong with one of the dimensions: perhaps it begins to appear to be too costly, or the design has become too complex and has sacrificed reliability. This information may be ignored at first, or an inordinate amount of effort may be expended in attacking this aspect of the problem (which might readily appear intractible to an equally experienced outsider). At length he comes around and accepts the necessity for basic modification in the design, but the process usually requires a considerable amount of time and energy for its accomplishment.

In all of this, it is not our intent to deny the affective component. Despite the norm of dispassionateness, engineers do become committed to their own designs. This commitment undoubtedly develops with time. But such commitment should also be a function of the uniqueness of the design, and in the instances under examination few of the rejected approaches could be considered to be really very novel. Furthermore, the engineers were in a position to claim full credit for the development of the replacement idea and had little to lose and actually something to gain through rejection of the old ideas.

For a fascinating discussion of this problem, the reader is referred to an article by T. C. Chamberlin originally published in 1890 and reprinted in *Science*, 148, 7 May 1965, pp. 754-759.
Number of Alternative Approaches Considered

Engineers considered from one to eight approaches to a subproblem with a median of three. (Table I). Two of the three approaches were considered from the beginning of the project, with a median of one new approach arising during the project's course.

TABLE I

<table>
<thead>
<tr>
<th></th>
<th>median</th>
<th>range</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>3</td>
<td>1-8</td>
</tr>
<tr>
<td>initial</td>
<td>2</td>
<td>1-5</td>
</tr>
<tr>
<td>additional</td>
<td>1</td>
<td>0-5</td>
</tr>
</tbody>
</table>

The ideas which come up later in the project have a success rate of about 15%, compared with a success rate of 39.5% for initially considered approaches. Eighty per cent of the approaches finally adopted were among those initially considered. Only 20 per cent of solutions originated during the course of the study.

Considering once again the total number of alternative approaches considered by the engineers, it was hypothesized that this number should
bear a direct relation to solution quality. More specifically, the hypothesis stated that the greater the number of alternative approaches considered to a subproblem, the better the quality of the solution to that subproblem.

On eight of the 20 subproblem pairs in the sample, relative evaluations of the solutions were obtained from responsible technical monitors in the government agency. In the remaining twelve, scores were either tied or no evaluation was available. This relative evaluation permits the use of simple tests between the number of alternatives considered and judged quality of the solution.

Not only is the hypothesis concerning total number of alternatives, not supported by the data, but as can be seen from Table II, the relation is in the opposite direction.

Engineers submitting higher-rated solutions actually consider fewer alternatives than their lower-rated competitors. This not only contradicts the hypothesis but appears to be in direct conflict with previous findings. Allen and Marquis (1964) in a study of two R & D proposal competitions found that teams considering a larger number of approaches to a proposal subproblem were better able to overcome negative transfer from related prior experience. Furthermore, going back to the data from the earlier study, it is found that proposal teams submitting successful solutions had considered 2.13 approaches, on the average, compared with a mean of 1.71 for teams with unsuccessful solutions. "Success" was determined by technical evaluators in the customer agency.
Since proposal competitions represent a relatively early phase in the R & D process, the possibility remains that the relation between number of alternatives and quality is a function of problem phase, and that this relation shifts direction as the project progresses beyond its initial phase. To test this possibility the number of approaches considered initially are separated out and compared. As can be seen in Table II, this difference is still in a direction which favors the lower-rated subproblems, but is reduced in magnitude and far from statistical significance. The last line in the table shows that the real distinction between the two sets lies in the number of alternatives generated when a team encounters trouble with one of their initial approaches, and this difficulty is reflected in a greater number of alternatives generated during the later stages of the problem and in the quality of the solution.

TABLE II

Mean Number of Alternative Approaches
Considered per Subproblem
(Eight Subproblem Pairs)

<table>
<thead>
<tr>
<th></th>
<th>subproblems with higher-rated solutions</th>
<th>subproblems with lower-rated solutions</th>
<th>p (as determined by t-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>2.88</td>
<td>4.37</td>
<td>0.06</td>
</tr>
<tr>
<td>initial</td>
<td>2.38</td>
<td>2.63</td>
<td>0.10</td>
</tr>
<tr>
<td>additional</td>
<td>0.50</td>
<td>1.75</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Number of Alternative Approaches Considered Simultaneously

In Figures 2 and 3 it may be noted that there is a distinct tendency for engineers to consider but two or three alternatives at any given time. During the course of a problem, they may consider as many as eight alternatives, but they seldom report considering all of these at once. The mean number of alternatives simultaneously considered ranges from one to five with a mean of 2.20. These statistics are based upon those weeks prior to the point at which one of the approaches attains a probability of 1.0 and stays there. Inclusion of reports received beyond that point would obviously lower the mean value.

Figure 4 provides a relative frequency distribution of simultaneously considered alternatives and is based upon weekly reports. The strong central tendency around two alternatives, and the sharp cutoff above three can be seen quite readily from this figure.

Since evaluations of solution quality are available for eight of the 20 subproblem pairs, a comparison was made between the means for groups submitting lower and higher-rated solutions. On the average those producing higher rated solutions considered 2.08 alternatives at a time; lower rated teams considered 2.42 alternatives at a time. The difference between means is statistically significant at the 0.001 level.

Frequency distributions of alternatives (figure 5) for the relatively higher and lower rated performances show that those teams turning in better performances come much closer to the ideal of two at a time. In fact, 81% of the time they report considering only two alternatives
(58% for lower-rated) and they never report considering more than three alternatives. A Kolmogorov-Smirnov two-sample test (Siegel 1956) rejects the null hypothesis that these two samples are drawn from the same universe (p<0.001).

It thus appears that more successful engineers differ from their brethren who are poorer problem-solvers in the manner in which they trade off and consider alternative solution possibilities. Further, the principal difference appears to be the tendency toward limiting their consideration to two alternatives at a time.

Why should this be so? A number of investigators, (Miller, 1965; Hayes, 1962; Santa Barbara and Paré, 1965) have demonstrated a finite limitation on human information processing capacity. Neimack and Wagner have shown further that amount of information gathering is a linear function of log \(2^n\) (where \(n\) is the number of possible solutions). While their measure of information gathering is the number of discrete steps taken in a rather limited type of problem, it does seem clear that the amount of information which must be processed should increase with the number of alternatives being considered. Hayes (1962) has shown quality of solution in a decision-making context to be independent of the amount of information (number of characteristics) provided on each alternative, but to be significantly decreased as the number of alternatives is increased from four to eight. Increasing the number of characteristics increases the dimensionality of the judgement. Miller (1956), in his summary of a large number of studies has shown that in-
creasing the number of dimensions operates through the intervening variable of codability to improve information transfer capacity at a somewhat less than linear rate. Increasing the number of alternatives, on the other hand, does not ease codability and merely adds directly to the amount of information to be processed. In this same line, De Groot (1964), in a study of chess strategies, finds that grandmasters rather severely limit the number of alternative moves which they consider in formulating strategy. Yntema and Torgerson (1961) in a discussion of future research on evaluative decisions predict the key concept to be that of problem simplification. They mention a number of possible strategies for reducing the number of facts which must be considered at one time, and thereby lessening "cognitive strain".

From the data reported here, it appears that engineers operate to reduce "cognitive strain" through the heuristic of reducing the number of alternative solutions under consideration at any time. Furthermore, this appears to be an effective heuristic since the data show a clear relation to solution quality.

The better-performing engineers in our sample, when confronted with this difficulty managed to cope with it through the heuristic of comparing two alternatives at a time, temporarily rejecting one while bringing up another for comparison, thus proceeding through as many as eight alternatives on a two-at-a-time basis.
SUMMARY AND CONCLUSIONS

In summary, the paper has demonstrated the feasibility of a new technique for the study of the R & D process. This technique employs the vehicle of parallel projects to provide a control over the substance of the problem and to enable a relative evaluation of solutions. Data are gathered by means of Solution Development Records, which report the probability of adoption associated with each solution alternative on a weekly basis, and by before and after interviews with the engineers.

The principal conclusions of the study are:

1. When a technical approach to an engineering problem becomes preferred over any other, it is not easily rejected; and the longer it is in a dominant position, the more difficult it becomes to reject.

2. Groups producing higher-rated solutions generated fewer new approaches during the course of the project. There is some indication that these arise when the favored approach encounters difficulty, and are probably symptomatic of poor early choices.

3. R & D groups trade alternatives off on a two-at-a-time basis. And the better-performing groups come closer to this ideal than do their rivals. This appears to be the result of limitations on the ability of humans to process information. The consideration of more than two multidimensional alternatives of the sort studied may approach channel capacity.
REFERENCES


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


FIGURE 4

Proportion of Time that N Alternatives are Considered
(40 sub-problem pairs; three parallel projects)

Number of Alternatives Considered Each Week (N)
FIGURE 5

PROPORTION OF TIME THAT N ALTERNATIVES ARE CONSIDERED
(8 SUB-PROBLEM PAIRS WITH RELATIVE EVALUATIONS)

teams submitting higher rated sub-problem solutions

teams submitting lower rated sub-problem solutions

NUMBER OF ALTERNATIVES CONSIDERED AT ONE TIME (N)

FRACTION OF TIME