Copy 1 DEWEY

THE EFFECTS OF GOAL DIFFICULTY ON PERFORMANCE: A FIELD EXPERIMENT

by

Andrew C. Stedry* and Emanuel Kay**

November, 1964 #106-64

* Graduate School of Industrial Administration
Carnegie Institute of Technology

**Behavioral Research Service, General Electric Company, and Lecturer,
Sloan School of Management, Massachusetts Institute of Technology

The research for this paper was undertaken while Stedry was on the faculty of the Sloan School of Management, and was fully supported by the Behavioral Research Service of the General Electric Company. Support for the preparation of thia report was provided by the Office of Naval Research at the Graduate School of Industrial Administration, Carnegie Institute of Technology, Management Sciences Research Group (Contract NONR 760(24), NR 047-048). Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government. The authors are indebted to Professor A. Charnes of Northwestern University for aid in planning the experiment and in the interpretation of the data.

Goal Setting - Research and Practice

It might be expected that the relationship between performance goals and subsequent performance would be an interesting subject for psychological research. Within organizations, at least, individuals are surrounded by performance goals -- quotas, standards, budgets, "bogeys," deadlines -- which, whether set by themselves or others, are expected to influence their performance. Performance goals have, indeed, received considerable attention in both the psychological literature and the standard business budgeting literature. A common misconception, perhaps emanating from the sheer mass of the literature is that the problem addressed in this paper, i.e., the effect of goal difficulty on performance -- has been thoroughly investigated. In presenting what is, admittedly, an exploratory study, we feel it is necessary to attempt to demonstrate that it has not.

Let us define a goal as a level of performance whose attainment is associated with "success" and non-attainment with "failure." Thus defined, it will include the psychological aspiration level[1/] provided one interprets success and failure as the individual performer's own subjective feelings associated with attainment. Consistency with the budgeting literature is more difficult to achieve because of the general confusion among forecasts, plans, standards,

---

1/ This quantity has been variously defined but the concept of subjective feelings of success and failure depending upon whether the exact level aspired to is reached appears to be ubiquitous, at least since the publication of the definitive article of Lewin, Dembo, Festinger, and Sears (1944).

budgets for control, etc., to be found there. [1] It is probably adequate, however, in our organizational context to include as goals performance levels whose achievement is recognized by some individuals in the organization -- perhaps including the individual or group whose performance is being measured -- as constituting success. Relying on constructs presented by Simon (1957), a goal may be viewed as the point on the performance continuum which separates "satisfactory" from "unsatisfactory" performance -- i.e., a performance point associated with a discontinuity in a satisfaction (or utility) function.

Budgeting literature and practice have implicitly assumed that goal-setting has a beneficial effect on performance. Writers differ on the conditions required for these beneficial effects,- e.g.,"goals should (or should not) be associated with specific rewards" -- so that many variations may be found of (frequently elaborate) goal-setting mechanisms which are assumed to benefit performance. It is also universally assumed that "properly set" goals -- i.e., according to the procedure advocated -- will be accepted by their recipients. This is equivalent to the assumption that the goal-setting mechanism produces goals which are internalized as the individual aspiration levels of the recipient -- i.e., that he will experience subjective feelings of "success" or "failure" on the basis of attainment or non-attainment. It has become common recently to assume that participation of the recipient in the goal-setting process is both a necessary and/or sufficient condition for goal acceptance and subsequent high performance. The evidence for these assumptions is generally confined to having used the advocated set of procedures with some success in one or another organization. Since the installation of goal-setting procedures is necessarily accompanied by alterations in the organization's

[1] For a more lengthy discussion of this problem, see Stedry (1960), especially Chapters 1 and 6.

information and reporting systems and not infrequently results from a change in management, it is impossible to determine from such evidence which of the assumptions upon which the system is based are valid.

The variable of particular interest here is the difficulty level of the goals arrived at. Where the budgeting literature treats of difficulty at all, some variant of the statement "goals should be attainable but not too loose" will be found. Lacking an operational definition of "attainable" -- e.g., a frequency with which a goal should be attained -- goal-setting in practice seems to have evolved as a process of extrapolation of past data to obtain a "goal" which is a prediction of performance. Performance ceteris paribus, might be expected to fall short of the prediction about as often as it exceeds it. However, these goals are attained with a sufficiently high frequency that non-attainment is often referred to as an "exception" — an indication that management attention is required in the area in which the exception occurred.[1/] Even if one assumes that the goals set are accepted without question, it seems reasonable to expect that a system operating in the presence of a significant amount of random noise can attain goals with such regularity only if considerable slack is present — i.e., if the goals are set quite low relative to possible levels of attaintment. Consideration of this possibility seems to be absent from the standard literature.

[1/] This principle of "management by exception" would seem to contain an assumption, frequently made explicit, that the exceptions should be relatively few in number, lest the (budgetary) system fail in its function as an allocator of scarce higher management resources.

In spite of the theory and evidence indicating that organizations formulate goals for their activities [1] and the almost ubiquitous presence of budgets and quotas in organizations to influence behavior, systematic investigation of the effects of goals on behavior is rare. Psychological research on aspiration level has focused on how performance affects the aspiration level determination or how the latter is affected by personality variables or group influences. References will be found in Rotter (1954) and Lewin et al (1944) to studies which relate performance in other tasks to the aspiration level formation process used on the experimental task. Neither those works nor the more recent summary of Starbuck (1963) appear to indicate specific investigation of the effect of the aspiration level on performance in the same task.

Exceptions to this general lack of interest will, of course be found. Bayton (1943) appeared to have found some effect but it is unclear as to whether the statistical techniques he used were appropriate. Siegel and Fouraker(1959) found that presenting subjects in a bargaining experiment with different aspirations for the outcome of the negotations effectively determined the outcome. French, Kay, and Meyer (1962) found that performance improvement was observed only in areas where specific goals were set as opposed to admonitions that improvement was required. Haberstroh (1961) found safety performance in a steel mill changing in response to changes in goals set as part of a safety improvement program. None of these studies,

---

[1] See Cyert and March (1963)

however, provide results specific to the problem of appropriate goal difficulty level. Having (we hope) established both a need for research in the area and a lack of firm empirical evidence upon which to base experimental hypotheses, we will attempt to provide a modicum of theoretical justification for the hypotheses to be proposed for this study.

Relevant Theory

Simon (1957) and March and Simon (1958) have postulated that a comparison between current performance and a goal by an individual will elicit search behavior [1] if the comparison indicated performance to be below a satisfactory level. Thus, one might conclude that the higher a goal is set, the more likely search behavior will be elicited.

Stedry (1960), however, postulates a maximum tolerance level for a discrepancy between a goal for future performance and current performance. If this gap is too great, discouragement will result, accompanied either by no search behavior or, perhaps, self-defeating frenzy rather than successful improvement-directed search. Experimenting with various goal difficulty levels in a laboratory problem-solving situation, he found "high" and "medium" goals to produce significantly better performance, on the average, than "low" goals resembling the "attainable-but-not-too-loose" goals recommended in the budgeting literature.[2] High goals (which were almost

---

[1] I.e., a search for alternative modes of behavior which will result in an increase in performance. "Effort," throughout this paper may be interpreted as the effort required for this search.

[2] The average performances observed with "medium" goals were better than those observed with "high" goals but pairwise comparison here did not yield significant differences.

impossible to attain) appeared to produce extremely good performance where the goal was presented prior to a questionnaire for determining the subject's aspiration level on a particular trial;where the questionnaire was presented first, subjects committed themselves to much lower goals than those which the experimenter subsequently presented (and which subjects probably rejected) resulting in extremely poor performance. Similar extremes in behavior were not observed for medium or low goals suggesting sensitivity of high goal to other conditions which might cause rejection of a goal -- e.g., perception of the goal as "impossible" of attainment based upon a comparison with other estimates of the situation.

Reasoning that in any real situation not one, but several goals will exist for comparing performance in various dimensions with desired attainments, Charnes and Stedry (1964) postulate a "quasi-rational" goal recipient who will allocate his effort so as to maximize the expected number of goals attained or a similar function with relative weights attached to each goal.[1/] Unless weights are specified, quasi-rational behavior dictates allocating effort only to those areas where a reasonable probability of attainment exists although, within this set, more difficult goals will receive more effort. Thus, reasoning from what might be at least reasonable behavior in the multi-goal situation, increasing goal difficulty in an area already receiving effort should increase the effort allocated to the area as well as the expected

---

[1/] Thus, "quasi-rational" behavior would be perfect rationality if the reward function an individual faced was precisely of this form. However, in the absence of such information, the behavior cannot be described as rational in the economist's sense.

performance in that area. Beyond a certain point, however, further increase in difficulty would drive the area out of the set of those receiving effort with resultant decrement in performance in the area. Clearly, an increase of decrease in effort in one area would result, respectively, in a decrease or increase in effort available for allocation to others assuming total effort remains constant. A model of the effect of goal difficulty on total effort where more than one goal is presented does not now exist, a problem which will be expanded upon later in the paper. The postulated effects on the performance areas taken individually are generally consistent with the experimental work of Stedry described above — i.e., difficult goals produce better performance unless they are raised to the point where, in the proper circumstances, they are rejected as being too high relative to performance expectations.

Thus, confining the above arguments to the case where only two performance areas and goals are assumed to be measurable and goals in non-measurable areas and total effort are assumed to be unchanged we would obtain the following:

Hypothesis 1: In each measurable performance area, a goal which represents an increase in performance over previously attained levels if it is perceived as having a "sufficiently high" probability of attainment will improve performance relative to a goal which represents little or no change from previous levels. If, however, the perception of the probability of attainment of the increase in required performance is not "sufficiently high" performance will diminish relative to a goal which represents little change from previous levels.

Hypothesis 2: Performance in one of the measurable areas will be higher where the effort allocation to the other measurable area is less. I.e., where the goal in the other area is perceived as being so high that no effort is justified, all of the effort is available for allocation to performance improvement in the first; less effort will be available for the first where the other goal represents little change from past performance; still less will be available where the other is a goal requiring an increase

in performance which is perceived as having a sufficiently high probability of attainment.

Hypothesis 3: Where increases in performance are required by goals in both of the measurable areas, the perceived probabilities of the attainment of one will be likely to be adversely affected by the perceived effort required to attain the other. Thus, a greater proportion of insufficiently high probabilities of attainment would be expected.

Performance Measures

In the experimental field situation, data on only two performance areas were sufficiently consistent in their measurement to be used. These were productivity, the percentage of actual cost represented by the standard cost of labor and material usage and rework, the combined labor and material (actual) cost of repairing, reassembling, or replacing products returned by inspection or other departments as unsatisfactory for use. These quantities were computed weekly. 1/ The experimental subjects, nineteen foremen in a department of a manufacturing plant which manufactures precision components for complex machinery, had been consistently measured on both of these quantities. However, prior to the experiment, the foremen's supervisors, unit managers, received reports on individual foremen and set goals only in productivity; the rework reports, although collected, had not been distributed to foremen or unit managers. 2/

1/ Thus, at 100% productivity, actual and standard are equal; where actual exceeds standard, productivity will be less than 100% and conversely. Productivity records varied generally between 60% and 95%. Rework cost varied greatly both by week and by individual; weekly averages varied from a few to several hundred dollars.

2/ The authors obtained reports through the regular accounting mechanism, not desiring to provide a Hawthorne effect during the period of analysis of historical data. Only when discussing the goals for manipulation with unit managers did we discover the lack of distribution. The unit managers promptly demanded (and received) these reports, thus introducing a Hawthorne effect of another kind. Our previous experience would, however, indicate that it is impossible for an experimenter to avoid increasing the data flow in an organization.

Although we had originally intended to study forty foremen, preliminary discussions revealed that managers of some units did not feel that these measures, albeit collected periodically, were meaningful for their subordinates. Whatever the validity of this opinion (since the unit managers both set goals for and evaluated performance of their subordinates) its effect would be to render any results obtained from these units spurious. The data were considered meaningful for twenty-three foremen but one unit manager refused to cooperate in the study reducing the sample to nineteen.

An attempt to extend the manipulation to eight other foremen in two units where managers suggested substitute measures they considered meaningful was thwarted by transfer of one unit manager early in the experimental period and, in the other unit, by an insufficient number of readings taken on the substitute measure which was not normally collected through the department accounting system. [1/]

Experimental Design

An arbitrary assumption was made that average performance for a period

---

[1/] Whatever the failings of data collected in financial terms and the assumptions made in the classification of data according to accounting principles, it is generally impossible to collect any consistently reliable performance data except through the regular accounting system. The constraints of double-entry, audit, and check may appear oppressive but the presence of a set of figures appearing week after week without missing observations (e.g., we were too busy to worry about it that week) and without weekly variations in what the figures measure seems to be available in no other way.

of six months prior to the experiment represented "normal" performance [1] and that a goal set to maintain that average performance over the experimental period of three months would be perceived by the recipients as "easy" to attain. A goal to achieve significantly higher-than-average performance, specifically at the first quartile of the twenty-six weekly performance figures for the six-month period,[2] was considered to be "difficult" to attain and, it was assumed, would be perceived as "challenging" or "impossible" to attain. For simplicity, we refer to normal (N) and difficult (D) goals and easy (E), challenging (C) and impossible (I) goal perceptions.

The foremen were divided into four conditions: (1) N productivity, N rework; (2) N productivity, D rework; (3) D productivity, N rework; and (4) D productivity, D rework. Thus, a foreman in the first group had goals based on an extrapolation of past performance. One in the second or third group had a goal in one area which, provided it was not perceived as impossible, would allocate more effort to that area than he had previously. In the fourth group, increased effort allocation to both areas is called for which must emanate from other "immeasurable" performance areas if both are to be accepted as challenges.

The allocation of foremen to groups was not done randomly but by a process which satisfied two criteria: (1) at least one of each of the four unit managers' subordinates must be in each group; and (2) the average age of foremen in each of the four groups must be approximately the same. [3] The experimental

---

[1, 2] In a few cases it was necessary to make adjustments to these figures where marked changes in performance occurred during the three-month "base period" which elapsed while the data were analyzed and the unit managers prepared for their role in the experiment.

[3] A formal minimization was not carried out but the averages differed by less than a year with a range over the sample of 30 years.

design, eliminating two foremen who may have misunderstood their goals,[1] is shown in Table 1.

Productivity Goal \ Rework Goal	Normal	Difficult	Combined
Normal	4	4	8
Difficult	3	6	9
Combined	7	10	17

Table 1

Experimental Design

The foremen were not informed that they were participating in an experiment. The goals we had determined were presented to the foremen by their unit managers. The latter were requested to present them as they would in their normal fashion of presenting goals. We emphasized our desire not to have the unit managers change their leadership style because of the experiment.

Each foreman was interviewed within two weeks after the manipulation.[2]

---

[1] This was concluded on the basis of a post-manipulation interview as will be described more fully below.

[2] Simultaneously, a series of interviews was being conducted with foremen in another part of the plant in connection with a management program designed to replace performance appraisal sessions with constructive goal planning. This program served as a "cover" for the interviews conducted as part of the current study.

The purposes of these interviews were: (1) to determine whether the manipulation had been successful; and (2) to determine whether the foreman perceived his goals as easy, challenging, or impossible. The interviews were largely open-ended. However, a specific attempt was made to ascertain, by direct question, the number of "chances in 10" the foreman perceived to be his likelihood of attainment of the goal. This answer was accepted as a *prima facie* categorization and, unless we were thoroughly convinced from other features of the interview that an individual was obviously misplaced, a perception of 8 chances in 10 or better was considered "easy," 3 to 7 "challenging" and 2 or fewer "impossible."[1/] Exceptions to the categorization were made only where individuals who said something like "oh 50-50" after an impressively long list of rationalizations as to why improvement in the area couldn't be made were classified as having an impossible perception or where individuals who emphatically stated "100 percent" along with a long list of things that would need to be done or happen for the goal to be made were classified as have a challenging perception.

Considering the arbitrariness of the classification, the success of the manipulation was remarkable. All twenty difficult goals were classified as

---

[1/] It should be noted that these probabilities are significantly higher than would be predicted on the basis of sampling theory. The N goals, averages of past performance, would be expected to be attained 50 percent of the time, assuming independence of performance and goal. Similarly, a D goal, determined as the first quartile of performance for 26 weeks and set for *average* performance in a 13-week period, would, with appropriate assumptions of normality, and independence, be approximately equivalent to a 3-sigma limit. As indicated earlier, either because goals do serve to effect their own attainment and/or because of a substantial difference between potential performance and past performance the levels of attainment predicted by the usual statistical means appear to be irrelevant except, perhaps, as indicators of relative attainability.

having either challenging or impossible perceptions. Fifteen of the eighteen normal goals were classified as having easy perceptions and two of the remaining three represented one individual. Thus 17 of the 19 foremen had perception classifications consistent with the manipulation. It was decided at this stage to eliminate the 2 inconsistent respondees on the assumption that the erroneous classification might have disclosed misunderstanding of the goals.

Results

Average performances in both productivity and rework were observed for a thirteen-week base period prior to the manipulation and an experimental period of thirteen weeks immediately following. Denoting by P_o and P_1 the productivity averages during the base and experimental periods, respectively, and by R_o and R_1 the corresponding rework averages we define improvements I_p and I_R as:

$$I_p = 100 \times \frac{P_1 - P_o}{P_o}$$

and

$$I_R = 100 \times \frac{R_o - R_1}{R_o}$$

Productivity increase and rework decrease represent improvement; hence, negative values of I_p and I_R indicate deterioration in performance. Since the distribution of I_R appears to be highly skewed and any assumptions relating to normality would be difficult to justify in any event, non-parametric methods have been used for all statistical analyses. For completeness, however, the actual percentage improvements have been shown in Table 2 with

Rework			Productivity		
Rework Rank	Goal Perception	% Improvement	Rework Rank	Goal Perception	% Improvement
1	C	50.76	14	E	6.70
2	C	49.67	13	E	6.68
3	C	44.90	3	C	6.02
4	E	42.01	4	C	4.71
5	E	36.49	10	E	4.28
6	C	17.70	9	C	4.10
7	E	16.94	8	C	3.92
8	E	14.05	11	E	3.87
9	C	13.40	6	E	2.34
10	C	9.11	12	E	1.50
11	E	4.05	7	E	.22
12	C	1.78	15	E	-1.09
13	E	.96	1	C	-2.80
14	E	-2.93	17	C	-4.43
15	I	-5.07	5	C	-4.89
16	I	-6.87	2	I	-6.35
17	I	-153.56	16	I	-9.45
Averages	C	26.76		C	.95
	E	15.94		E	3.06
	I	-55.17		I	-7.90
All Foremen		7.85	All Foremen		2.75

Table 2

Goal Perceptions and Performance Improvement by Foreman

Note: "Foremen numbers" were arbitrarily selected as equivalent to rank of performance improvement in rework.

the mean performance improvement shown by goal perception.

In terms of these data, hypothesis 1 relating to the superiority of challenging over easy (C>E) and easy over impossible (E>I) goals in effecting performance improvement may be tested using performance ranks. Of 14 challenging and easy rework goal perceptions, the former ranked 1, 2, 3, 6, 9, 10, 12; under the null hypothesis (C = E) the probability of this ranking or one more extreme occurring is .130. [1] Comparing either C or E with I, the probability that I rework perceptions will rank 8, 9, and 10 of 10 (the most extreme case) is only .008 under the null hypothesis. Examining productivity improvements, it is clear that C > E cannot be confirmed; if anything, the data indicate E > C. [2] Comparing C or E with I, however, the probability that the I productivity goal perceptions will rank lowest is .028 (C > I) and .022 (E > I). Thus, confirmation [3] is provided for the inferiority of goals perceived as impossible. Only the weakest form of confirmation, and only in one of the areas tested, could be found for the superiority of goals perceived as challenging over those perceived as easy:

---

[1] The Mann-Whitney U-test is used for all rank comparisons between pairs of groups. The overall group differences in ranked improvements for both productivity and rework are significant at better than the .05 level by Kruskal Wallis one-way analysis of variance. A description of these tests will be found in Siegel (1956), pp. 116-127 and 184-193.

[2] Were this tested, the probability would be .164. This result has no significance other than to serve as a caution against drawing any firm conclusions from C > E for rework with p = .130.

[3] Subject, of course, to the obvious limitations occasioned by the small sample size.

For the purpose of testing hypothesis 2, the ranked data are grouped as shown in Table 3. Here again,

Rework Goal Perception	Challenging	Easy	Impossible
Individual Productivity Improvement Ranks	3 5 6 9 10 13 16	1 2 4 7 8 11 15	12 14 17
Sum of Ranks	62	48	43
Average Rank	8.88	6.86	14.33

(a)

Productivity Goal Perception	Challenging	Easy	Impossible
Individual Rework Improvement Ranks	1 3 4 5 8 9 17	6 7 10 11 12 13 14 15	2 16
Sum of Ranks	47	88	18
Average Rank	6.71	11	9

(b)

Table 3

IMPROVEMENT RANKS COMPARED BY COMPETING GOAL PERCEPTIONS

The data do not conform to the hypothesis. While the relationship between productivity improvement where challenging and easy rework goals are perceived is in the predicted direction the relationship is reversed for rework improvement. However, it may be presumed from confirmation of the portion of hypothesis 1 that performances observed in the area with a goal perceived as impossible required little effort. Whether because effort was expended uselessly or total effort expenditure went down in response to an impossible goal perception or for some other reason, the performances recorded in the competing area did not reflect a greater (successful) usage of effort.

Support for, but not confirmation of, hypothesis 3 is provided by the data. Four of the five cases in which impossible perceptions were recorded occurred in the group which had two difficult goals. Two of these, however, were recorded for the same individual, lending further support to the notion that, in place of hypothesis 2, a composite hypothesis incorporating the effect of an impossible goal perception on total effort as well as effort allocation must be proposed.

Ex-Post Analysis

Failing to confirm any of the ex ante hypotheses but observing a pattern to their failure, we analyzed the data further in an attempt to formulate alternatives which might be useful for planning additional

research.[1/]

Re-examination of Table 1 reveals that whatever the perception assumed for the difficult goals, they tend to be concentrated at the ends of the scale. If we were to state, rather than hypothesis 1 (C > E > I), a more general form, viz., difficult goals produce extremes in performance. This is tantamount to assuming that goals as difficult as those chosen will produce either very good or very poor performance but it is impossible to predict, a priori, by interview or other means at which extreme performance associated with a particular goal (or individual) will fall.

Depicting the data as shown in Figure 1,

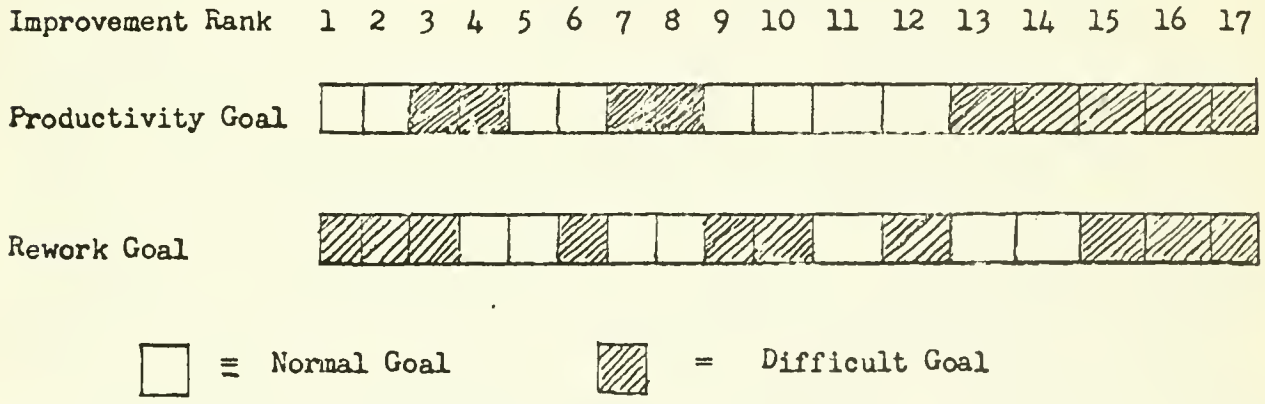


Figure 1

Goals for Productivity and Rework and Associated Improvement Ranks

In view of the small size of the sample it is unlikely that "confirmation" of ex ante hypotheses would have yielded much more. While it is true, of course, that if the researcher formulates and tests a sufficient number of hypotheses he will necessarily find some that attain statistical significance whether or not he formulated these hypotheses before examining the data. The formal testing of ex post hypotheses immediately suggested by the form of the data does not produce any great discrepancy between the "real" significance levels and the normal levels provided by the statistical analysis.

it will be observed that the "span" of improvement ranks from the highest to the lowest is greater in each area for difficult than normal goals - e.g., 1 to 17 as opposed to 4 to 14 in rework. Using a test for extreme reactions developed by Moses, [1] the probability that extremes as marked as those occurring by chance under the null hypothesis in the distribution of productivity scores is .0880; for rework this probability is .0814.

Assuming, for the moment, that difficult goals do in fact produce extremes, the next logical step is to test the proposition that performance extremes observed in response to difficult goals in an area will be marked where a difficult goal has been presented in the other area - i.e., that the extreme-producing effects will be intensified by a difficult goal elsewhere.

For the purpose of this analysis performances in each area were examined separately for groups which had normal and difficult goals in the other. For this purpose the data were grouped as depicted in Figure 1. It will be observed immediately that the performance extremes observed to be associated with difficult goals are more marked in each area where a difficult goal was imposed in the other. The probability of observations as extreme as those observed for difficult goals in productivity in the group that also had difficult goals in rework is .119 under the null hypothesis; for

---

[1] Cf. Siegel, (1956), pp. 145-152. The Moses test allows for discarding in advance a number, h, of the highest and lowest observations in the group assumed to be non-extreme adjusting the occurrence probabilities accordingly. All tests used here, however, were made with $h = 0$.

rework the corresponding probability is .083. By contrast, improvements with difficult goals do not show this tendency where normal goals are presented in the other area. 1/

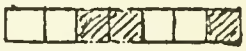
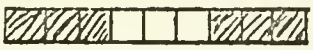
Goal in Competing Area ⟶	Difficult	Normal
Productivity Improvement Rank	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7
Productivity Goal		
Rework Improvement Rank	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8
Rework Goal		

Figure 2

Responses to Difficult Goals in an Area Within Groups Having the Same Goal Difficulty in the Competing Area

▨ ≡ Difficult Goal □ ≡ Normal Goal

Turning to another facet of the data, the failure to confirm hypothesis 2 suggested that the assumption of constant effort across the two areas was untenable and that the combined effort may be affected by the amount of challenge perceived in one or the other area. Examination of the data in pairs as depicted in Figure 3, reveals apparent regions within which the two

---

1/ Productivity improvements with normal goals are in fact more dispersed in the sample. Rework improvements, although the span is greater with difficult goals, could have this or a more extreme distribution with probability .5.

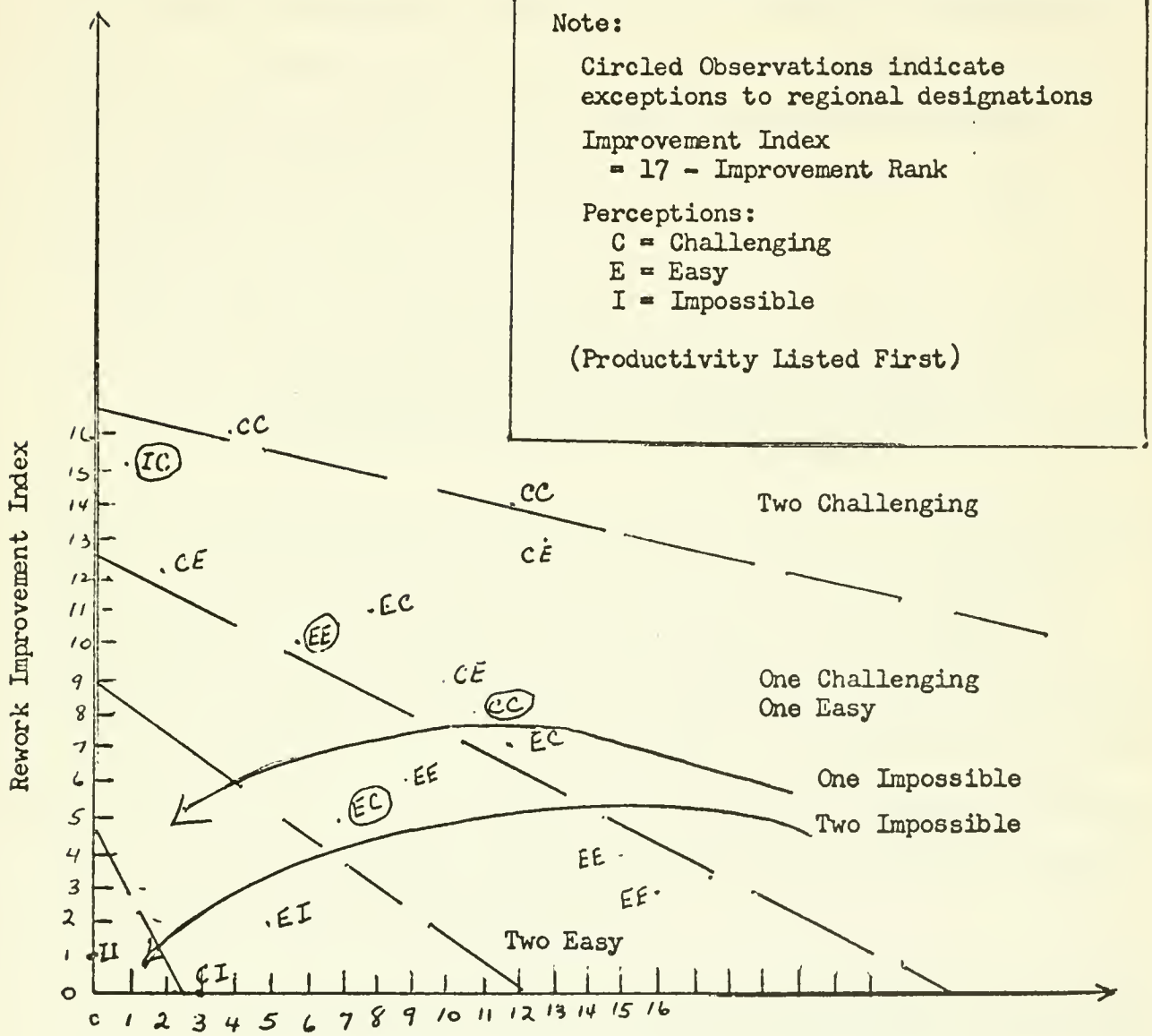


Figure 3

Goal Perceptions Arranged by Productivity and Rework Improvements

performance areas compete for effort. [1] The fact that the lines delineating the regions cannot be drawn as normals to a 45-degree line drawn from the origin suggests that greater effort is required for improvement in one area -- rework -- than the other relative to the improvments made by others. The non-parallelism of the lines that can be drawn suggests a greater impact of challenge in one area - rework - on total challenge perceived in the situation. [2]

Let us define a simple trichotomy on the goal perceptions such that:

$$u^j = \left\{ \begin{array}{l} 1 \text{ when goal in jth area is perceived as challenging} \\ 0 \text{ when goal in jth area is perceived as easy} \\ -1 \text{ when goal in jth area is perceived as impossible} \end{array} \right\} j = p,r$$

where the superscripts p and r will refer, respectively to productivity and rework. We may loosely consider $u = u^p + u^r$ as the "total challenge" represented by the pair of goals. If so, it will be observed that the regions in the graph can be relabeled so that they correspond, starting from the top, to u's of 2, 1, 0, -1, and -2 with the exception that, with this designation, the point (3,0) is misplaced. Using this designation, however, none of the exceptions noted on the graph is in a total challenge category which differs from its own by more than 1.

---

[1] The improvement "index" used is simply a transformation of the improvement rank such that higher numbers are associated with greater performance improvement.

[2] I.e., it is impossible to draw parallel lines of 45-degree slope which divide the set of points as well.

All of the foregoing lead to possibility that total performance and total challenge are related. Let us define:

$y^p \equiv$ productivity improvement index = 17 - (productivity improvement rank)

$y^r \equiv$ rework improvement index = 17 - (rework improvement rank)

and, ignoring possible differences in performance difficulty in the two areas:

$y = y^p + y^r \equiv$ total improvement index.

The relationships suggested by the data are:

$$y = B_1 + B_2 \; u$$

or

$$y = \gamma_1 + \gamma_2 \, u^p + \gamma_3 \; u^r$$

where, in the second instance, we are assuming differential effects of challenge in the two areas. Using the usual least-squares regression techniques, the relationships are found to be:

$$\hat{y} = \underset{(\pm 1.29)}{13.30} + \underset{(\pm 1.11)}{5.10}\, u$$

and

$$\hat{y}' = \underset{(\pm 1.28)}{13.45} + \underset{(\pm 1.72)}{3.55}\, u^p + \underset{(\pm 1.56)}{6.40}\, u^r$$

In either relationship the regression coefficients (adjusted for the mean - i.e., B_2, γ_2 and γ_3) are highly significant ($p \leq .001$). The first, using only "total challenge" accounts for 58.5 percent of the variance and the

second 61.2 percent - an insignificant difference. Thus, little advantage appeared to be gained by discriminating between the areas in which challenge occurred. It is probable that an index which gave more weight to the rework performance would improve the relationship but, considering the small amount of variance remaining to be explained, the improvement could only be slight. The actual and predicted values of performance are shown in Table 3.

Summary

The hypotheses formulated ex ante could not be confirmed. Perceptions of a goal in an area as challenging, easy, and impossible did not produce performance improvement in that order in the area; nor did performance in the other area follow the reverse order. Sample size was insufficient to test the greater effect of two difficult goals on production of impossible perceptions. Support was provided for the proposition that difficult goals perceived as impossible produced poorer performance improvement than goals perceived either as easy or challenging.

Further analysis of the data revealed that, had the performance hypotheses been stated in the absence of perception — i.e., difficult goals will produce either very good or very poor performance relative to normal goals — confirmation would have been obtained. Furthermore, difficult goals in an area appeared to be associated with extremes in performance when the other area had a difficult goal, but this effect was not observed with a normal goal in the second area. Finally, although perceptions were not successful in predicting performance in one area, the combined performance in the two areas could be represented quite well as a linear function of the goal perceptions in the two areas.

Index Total Perf (y)	Prod. Challenge (u^P)	Rework Challenge (u^r)	Total Challenge (u)	Regression Estimates	
				$\hat{y}_i = 13.30+5.10\ u_i$	$\hat{y}_i = 13.45+3.55\ u_i^?+6.40u_i^r$
28	1	1	2	23.50	23.40
20	1	1	2	23.50	23.40
19	1	1	2	23.50	23.40
19	0	1	1	18.40	19.85
19	0	1	1	18.40	19.85
12	0	1	1	18.40	19.85
26	1	0	1	18.40	17.00
19	1	0	1	18.40	17.00
14	1	0	1	18.40	17.00
19	0	0	0	13.30	13.45
19	0	0	0	13.30	13.45
16	0	0	0	13.30	13.45
15	0	0	0	13.30	13.45
16	-1	1	0	13.30	16.30
3	1	-1	0	13.30	10.60
7		-1	-1	8.20	7.05
1	-1	-1	-2	2.10	3.50

Table 3

Performance as a Function of Degree of Challenge - Actual and Estimated Values

Discussion

The failure to confirm the ex ante hypotheses could be attributed to the inability of the interview technique to satisfactorily distinguish between challenging and impossible perceptions. A thorough re-examination of the interview records undertaken after the initial tests revealed that this was unlikely. The interviews of those judged as impossible differ markedly from those judged as challenging. The difficult distinction to draw is that between easy and challenging. It is also interesting that were the low-producing individuals who were classified as challenging to be reclassified as impossible the relationships depicted in Figure 3 would be adversely affected to a remarkable degree. If anything, recategorization of the low-scoring individuals with difficult goals from challenging to easy seems more consistent with the interviews and would improve both the relationships in Figure 3 and the regressions. (This recategorization would also provide data which confirm $C > E$ in each area.)

Such a recategorization would not only vitiate any scientific import in the study. It would detract from the basic generalization to which the data all seem to point. Supervisors using some kind of goal-setting procedures are probably no more capable than experimenters of interpreting the statements of their subordinates as to challenge perceived associated with a goal. While it is reasonable to expect better communication between foremen and their supervisors than with an interviewer, the former is subject to bias of another kind — e.g., if the supervisor is perceived as rating the response as well as performance. It is quite possible that an individual does not know in advance how he will respond to a particular goal in which case mechanisms which incorporate this a priori response - e.g., participative budgeting — would seem

to offer little more promise than other schemes.

Given that difficult goals perceived as challenging did appear to enhance overall performance, and difficult goals perceived as impossible appeared easy enough to detect, the use of difficult goals with a mechanism for revision if an impossible perception is detected seems worthy of further investigation. The use of difficult goals in a particular area with the expectation of improvement in that area seems problematical.

What is strongly suggested for future theoretical work is the need for models which incorporate "total challenge" as some function of all of the goal perceptions involved. Within each of these "total challenge" determinations it is not unreasonable to assume competition between performance areas for total effort. [1] Further research is, of course, also called for to determine the long-run effects of different kinds of stable goal-setting procedures. The current study underlines the inadequacy of assuming that a particular kind of policy will produce certain effects without a thoroughgoing investigation of the real effects as reflected in performance.

---

[1] It will be observed from Figure 3 that within each region the two performance improvements are, in general, negatively correlated.

REFERENCES

1. A. Charnes and A. C. Stedry, "Exploratory Models in the Theory of Budgetary Control," in W. W. Cooper, H. J. Leavitt, and M. W. Shelly (eds.), New Perspectives in Organization Research. New York: John Wiley and Sons, Inc., 1964.

2. J. A. Bayton, "Interrelations Between Levels of Aspiration, Performance, and Estimates of Past Performance," Journal of Experimental Psychology, Vol. 33, No. 1, 1943, pp. 1-21.

3. R. M. Cyert and J. G. March, Behavioral Theory of the Firm. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1963.

4. J. R. P. French, Jr., E. Kay, and H. H. Meyer, "A Study of Threat and Participation in an Industrial Performance Appraisal Situation." New York: General Electric Company, 1962.

5. C. J. Haberstroh, "Administration of Safety in the Steel Industry," Management Science, Vol. 7, No. 4, July 1961, pp. 436-444.

6. K. Lewin, Temara Dembo, L. Festinger, and Pauline S. Sears, "Level of Aspiration," in J. McV. Hunt (ed.) Personality and the Behavior Disorders. New York: Ronald Press Company, 2 Vols., 1954.

7. J. G. March and H. A. Simon, Organizations. New York: John Wiley and Sons, Inc., 1958.

8. J. B. Rotter, Social Learning and Clinical Psychology. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1954.

9. S. Siegel, Non-Parametric Statistics. New York: McGraw-Hill Book Company, 1956.

10. S. Siegel and L. E. Fouraker, Bargaining and Group Decision Making: Experiments in Bilateral Monopoly. New York: McGraw-Hill Book Company, 1960.

11. H. A. Simon, Models of Man. New York: John Wiley and Sons, Inc., 1957.

12. W. H. Starbuck, "Levels of Aspiration," Psychological Review, Vol. 70, No. 1, January 1963, pp. 51-60.

13. A. C. Stedry, Budget Control and Cost Behavior. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1960.