FACTORS RELATED TO

SCIENTIFIC RESEARCH PERFORMANCE

Robert C. Davis

Reprinted from:

Interpersonal Factors in Research: Studies in Selected Aspects of Performance, Communication and Attitudes, Part I.

by Robert C. Davis
Glen Mellinger
Donald C. Pelz
Howard Baumgartel

Conducted at the National Institutes of Health by the Institute for Social Research, University of Michigan, Ann Arbor, Michigan

October 1954
PREFACE

The Human Relations Study of NIH was undertaken to serve two related objectives: an "applied" and a "basic" research goal.

The "applied" goal is to assist NIH in diagnosing and solving some of its human relations problems, so as to maintain high work motivation, reduce frustrations which may hinder performance, and improve relations between individual and groups. The "basic" research goal is to augment scientific knowledge of how large organizations function, and how they can function more effectively.

To meet both of these objectives, several types of data were obtained. In the fall of 1952 all employees were asked to fill out a written questionnaire, concerning attitudes toward the job itself, supervision, directors and administrative officers of the NIH and the institutes, promotional and salary systems, working conditions, auxiliary services, and other aspects of the working environment. A copy of the questionnaire for intramural professionals of the laboratory research is appended. Other categories of employees received slightly different forms.

In January of 1953 a second set of data was obtained. Within each of the major laboratories, and within each broad field of work or discipline, several qualified investigators were asked to evaluate the scientific performance of individuals within that laboratory or field. An overall judgment of relative standing in the group was requested, rather than a series of ratings on specific qualities. Details are given in Appendix A.

Supplementary interviews were also held with a limited number of individuals and small groups.

All of these data were obtained with the understanding that no individuals would be identified. Opinions were to be reported only for groups. The evaluations of scientific performance would not be reported for individuals, and would be used for analysis only.

The first year of the study constituted a diagnostic phase. An extensive analysis was carried out to answer such questions as:

What are strong or weak points in the working environment at NIH? What are major sources of frustration?
What groups within the organization are affected most strongly? For example, in what institutes, among what employee categories, is the greatest dissatisfaction or tension expressed?

This analysis resulted in a General Report of the first years findings. It was aimed mainly at the applied objective.

The second year of the project has stressed both practical and theoretical goals. On the practical side detailed reports of opinions in each laboratory or branch were prepared and discussed with directors and laboratory chiefs in each institute, over a period of several months. Excerpts from the personal interviews and from written comments were also circulated.

On the more theoretical side, several analyses have been carried out in order to understand how the organization operates and how its effectiveness might be improved. These results are the subject of the present report. Analysis along similar lines is being continued, and a supplementary report (to be issued as Part II) is planned for the spring of 1955.

The paper by Davis considers several conditions in the individual or in the social environment which may lead to a higher or lower level of research performance. The paper relates the assessors' evaluation of each individual to several questionnaire items obtained from the same person. That motivations in the scientist, and what relations between the scientist and his supervisor, are more effective in terms of the subordinate's research accomplishment? Davis finds important connections between performance and the scientist's personal emphasis on scientific contribution, the degree of influence he has on his chief's activities, and the degree of liking or confidence which the scientist feels toward his chief. Significant is the fact that these connections are not simple; rather, a complex set of conditions is required for high performance.

Mellinger's paper considers another aspect of organization functioning: that of communication. In an effective organization there must be some general agreement on its goals and the policies for reaching them. To what extent does discussion among members result in closer agreement, or in greater understanding of the other person's viewpoint? Mellinger finds that communication as such may not have these results; interpersonal factors such as liking

and confidence or trust have an important bearing on the outcome.

The chapter by Pelz takes up the specific topic of auxiliary services. Where should services be located, and what channels used to obtain them, for maximum effectiveness in the eyes of scientists? What part do the institute administrative offices play in obtaining services, and how does their role affect the adequacy of results? Pelz finds that interpersonal factors such as strains and frictions between scientists and administrative officers are more significant than the particular channels used.

Baumgartel's chapter presents some results in which the laboratory is used as the unit of analysis, rather than the individual scientist. Laboratory scores on some two dozen questionnaire measures of motivation and satisfaction were obtained, to answer questions such as: Does a laboratory's emphasis on health problems reinforce or inhibit its emphasis on basic science? What factors in the work itself lead to satisfaction with scientific leadership or with promotional opportunities? Two distinct clusters of data are found, one an orientation toward basic science, the other an orientation toward both health and prestige. Some of these same items were found useful in Chapter I for understanding individual scientific performance. Baumgartel's results are part of a larger analysis to determine how the laboratory leadership can influence the motivations and satisfactions of the group.

While the aim of these analyses was to establish general principles, they also have practical implications. The latter are particularly stressed in Chapter III on auxiliary services. The first two chapters have implications for scientific supervision and for methods of communication. The fourth chapter will be useful in understanding the meaning of the specific laboratory data which were discussed with each group this past year.
A major concern of the NIH study is to determine what kinds of interpersonal relations are associated with effective research.

More specifically, we want to answer such questions as these: does it make a difference, in terms of performance, toward what general goals the scientist is motivated? How do patterns of mutual influence and personal feeling between the scientist and his chief relate to performance? Are the same decision-making procedures related to performance in the same way for all levels of research personnel?

These and related questions formed the basis for the selection of the social and psychological variables in this analysis of research performance.

Methods of assessing research performance

The term "research performance" as used in this study denotes an overall concept which includes both the ideas of "productivity" and "creativity." That the assessment of current research performance is a reasonable aim is based on the fact that, in the conduct of everyday work, scientists informally assess the relative value of the research of their colleagues. In fact, on the basis of these judgments research is planned, funds allocated and prestige accorded.

In developing methods for measuring performance, consultation with advisory groups of scientists led to two conclusions: first that the judgments of fellow scientists would be the best way to measure the performance of individual researchers, and second, that a single overall assessment should be made. It was felt that an attempt to rate specific dimensions separately — such as "creativity", "soundness", etc. — would create a sense of artificiality for the assessors.
To make the rating process concrete the assessor was asked to think of each scientist on his list as a potential candidate for a research grant, and to select those who were equipped to turn out the best possible research within the next few years. In making this evaluation, assessors were asked to keep in mind a variety of elements that go into effective research work. Some persons may be valued more for creativity, others more for persistence and efficiency. The specific guidelines for assessors are reproduced in Appendix A. Because age, grade, etc., could be corrected for later, raters were asked not to make allowances for such factors.

Each researcher was rated by at least two assessors from each of two groups: one group familiar with his laboratory, the other with his field of work or discipline. The advisory committees of scientists suggested the initial group and the latter in turn suggested additional assessors. Qualifications sought in assessors were: competence in a given area of research, familiarity with the individuals to be rated, and the ability to make relatively objective judgments of others. Care was taken to draw assessors from both supervisory and non-supervisory ranks.

Lists of scientists by laboratories and fields were given to the respective assessors, who crossed off unfamiliar names as well as their own. The assessors were asked to divide the list into two or more categories on the basis of the criteria already mentioned. Some chose to divide their group into "higher" and "lower" performers; some preferred to distinguish several categories of performance.

Two methodological steps remained: combining the multiple ratings into a single laboratory and discipline score respectively for each individual, and consolidating these two ratings into a final combined score. */ The details of these procedures will be found in Appendix A; a brief summary will suffice here.

A preliminary inspection of ratings eliminated a few assessors whose ratings were markedly divergent from the others. Within each assessment group the ratings of the various assessors were converted into comparable scores. At this stage an examination was made for widely discrepant patterns. The pattern of ratings for each person assessed was examined, and those few individuals having widely discrepant pat-

*/ An exploratory scaling, using only a high-low division, is reported in Human Relations in a Research Organization (2 vols., Institute for Social Research; Ann Arbor, 1953), Chapter 8 and Appendix C. This preliminary system has been superseded by the present nine-point scale, which provides greater sensitivity and flexibility in analysis.
terns were eliminated. The scores of the several assessors for each individual were then averaged, thus giving a laboratory and a discipline score.

These two scores were correlated to see if both the laboratory and discipline assessors were in reasonable agreement. When it was determined that they were, the scores were merged into a final nine-point scale. */

Criteria for choosing scientists for the sample

It was felt that in order to reduce the range of differences in performance which might be ascribed to education or experience, we should limit the present inquiry to those scientists holding doctoral degrees, and to those with at least a grade of GS-9, or its Corps equivalent. Further, we included only those intramural scientists who are primarily engaged in active research or who directly supervise it. Laboratory chiefs, scientific directors, and institute directors were omitted because in general they are not supervised by scientists engaged in research.

Out of 33½ persons who were given Form 1 (intramural professional), 23½ scientists met these criteria. We obtained a final performance score on 20½, or 6% of the selected group. On 26 of the cases there was either no assessment made (due to unfamiliarity, too short a time at NIH, etc.), or assessment by only one assessor. Four cases were eliminated because of widely discrepant ratings.

The analysis presented here is based primarily on the 20½ cases, reduced in some instances by non-response on given items which are being related to performance.

*/ As to the validity of the scale, there is one independent piece of evidence which should be mentioned. On fifty individuals we have information on the number of publications in the last three years. We find that those who are rated high on the performance scale may have either a large or small number of publications, but those who are rated low have, almost without exception, a low number of publications. Because of this type of relationship it seems safe to infer that the ratings were based, in accordance with the instructions, on criteria which include "productivity" but go beyond it.
Adjustment of the performance measure

In order to study the relations of interpersonal factors to research performance, it is necessary to control, as far as possible, performance differences associated with age, experience and achievement. These variables are reflected in grade to a large extent, and grade, as we would expect, is correlated with performance. It was therefore decided to "control" grade by means of a statistical adjustment of the performance scores.

The procedure of adjustment was as follows: the entire distribution of scores in a grade category was simply moved up or down until the average performance fell at the middle of the nine-point scale. Thus, the average performance in each grade category has the adjusted score of five.

The performance data presented in this report are in terms of the adjusted measure. The adjustment made it possible to study those variations due to environmental factors, and to make generalizations which include the entire population of scientists.

It should be kept in mind that the variables in this analysis are derived from two independent sources: the performance measure from the groups of assessors, and the attitude variables from the questionnaires filled out by the individual scientists. The independence of these two kinds of measures adds weight to the findings in which they are significantly related.
B. Analysis and Findings

In examining the factors which facilitate or inhibit the translation of research impetus into research performance, we decided to begin with the motivations of the scientist in his role as researcher and work "outward" into the immediate environment in which he must work. In doing so we are quite conscious of the fact that we are examining a process which is in reality circular. Motivation is not a constant quantity, but it is affected by social factors external to the individual. However, we are examining the individuals and situations at a single point in time. For this reason it is convenient to start with motivation as it exists at that point.

Measuring motivation
in terms of value orientation

Motivation, as the term is most generally understood, signifies the purposive, goal-striving aspect of human behavior. It involves a selective orientation to a goal, and the degree of intensity of the desire to reach the goal.

In the present analysis we are interested in what the scientist wants to get out of his role as researcher, and how important it is to him to get what he wants. Accordingly, we set up measures of orientation to two sets of goals or "rewards" which are potentially available in the role of researcher in a large organization.

The first may be called a science orientation. It involves the degree to which doing research has an intrinsic value for the scientist. The second is an institutional orientation. It concerns the degree to which rewards derived from the organization, such as advancement and prestige, are of importance.

We asked these questions about factors contributing to satisfaction with the job to get at the science orientation (item 50 in the questionnaire):

"How important is each one to you?

---Contributing to basic scientific knowledge
---Freedom to carry out my own ideas; chance for originality and initiative
---Chance to use my present abilities or knowledge"

*/ These concepts are modifications of those used by D. Marwick: Career Perspectives in a Bureaucratic Setting (University of Michigan Press: Ann Arbor, 1954).
and these to get at the institutional orientation:

"How important is each one to you?

—Having an important job in the organization
—Association with high-level persons having important responsibilities
—Sense of belonging to an organization which has prestige in the lay community."

There is a sizeable intercorrelation between the three items in each set of questions. Therefore we were able to make indices of high to low science and institutional motivation respectively, from the two sets of questions.

Most of the scientists show a high degree of science motivation, but a moderate or low degree of institutional motivation. It was decided to cut each index at a point which divided the group roughly in half. Each scientist therefore received a score of high or low on each value, depending on which half of the total group he fell in. The important thing to keep in mind is that the terms "high" and "low" are relative. One should be especially careful in interpreting labels. The only differentiations which can legitimately be made are those which denote the relative emphasis in orientation toward the two values.

Validity of the value measures

The value measures have internal consistency and seem to represent logical clusters of items. However, if the measures are to be considered valid, they should show predictable and meaningful relations to other variables.

We would expect the indices to differentiate between scientists who prefer a relatively pure research setting with comparatively little emphasis on institutional rewards, and those who prefer the opposite kind of setting. To test this hypothesis, the scientists were asked where, if they were to leave NIH, they would prefer to work (Question 1). The alternatives were universities, industry, private practice, other PHS, other government, etc. We would predict that the university setting would be the most congenial to the person highly motivated to science, whereas the person highly motivated toward institutional values would choose other settings. It is evident from research at NIH and elsewhere that the university is perceived by sci-
entists largely in terms of the values we have called the science orientation.

The results are shown in Table 1-1. As we expected, significantly more of the high science people prefer universities than do low science people, and significantly more high institutional people choose "other" rather than universities, compared to those holding low institutional values. The differences are even greater when both value orientations are studied jointly. Among those holding high science, but low institutional values, 85% prefer the university setting. However, for those having the low science and high institutional motivations, only 42% choose universities. (See Table 1-1.)

These data lend support for the validity of the value measures, by confirming the hypothesis concerning preference of organizational settings. We may predict further that those having high rather than low institutional motivations will be more willing to accept promotions contingent on assuming supervisory or administrative functions.

The following question (item 62) was asked:

"Would you be interested in a higher-level job at NIH if it meant doing less of your present work and more of something else? For example, would you be interested in a higher-level job which required spending a large part of your time on the activities below? . . .

—Professional leadership: stimulating or advising subordinate professionals about their work
—Administrative planning or coordination: allocation of funds, recruitment of personnel, expediting of services, etc."

*/ In the Steelman Report, for instance, 64% of the scientists who said "freedom in research" was important to them preferred the university setting. Other such data are found in the President's Scientific Research Board, Science and Public Policy (5 vols., GPO: Washington, 1947), Vol. 3, Appendix III, pp. 205-252.

**/ The joint use of two or more variables in relation to another, as in the present case, illustrates how relationships may be "sharpened" by stating the conditions under which the relationship occurs. The procedure is used throughout the analysis. It serves two functions: it is a means of statistically "controlling" variables, and it serves as a constant reminder of the complexity of social data.
### TABLE 1-1

Motivations Related to Type of Setting Preferred if Leaving NIH

<table>
<thead>
<tr>
<th>Preference for:</th>
<th>Science Motivation</th>
<th>Institutional Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>University setting</td>
<td>78%</td>
<td>50%</td>
</tr>
<tr>
<td>Other settings</td>
<td>22</td>
<td>50</td>
</tr>
<tr>
<td>N =</td>
<td>125</td>
<td>82</td>
</tr>
<tr>
<td>Chi-square:</td>
<td>16.98</td>
<td></td>
</tr>
<tr>
<td>Sig. at:</td>
<td>.001</td>
<td>*/</td>
</tr>
</tbody>
</table>

#### Combinations of the Two Motivations

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>University setting</td>
<td>85%</td>
<td>72%</td>
<td>62%</td>
<td>42%</td>
</tr>
<tr>
<td>Other settings</td>
<td>15</td>
<td>28</td>
<td>38</td>
<td>58</td>
</tr>
<tr>
<td>N =</td>
<td>54</td>
<td>71</td>
<td>50</td>
<td>32</td>
</tr>
<tr>
<td>Chi-square:</td>
<td>20.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. at:</td>
<td>.001</td>
<td>*/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*/* The significance level indicates the probability of the differences occurring due to chance alone. Thus, .01 ("significant at the .01 level of confidence") means that there is only one chance in 100 of a difference this large being attributable to chance. A confidence level of .05 (1 chance in 20) is generally the lower limit for "significant" differences.
Table 1-2 shows that significantly more scientists in the high, compared to the low institutional group, are willing to take on science supervision activities. However, there is no appreciable difference between those of high and low science motivation.

**TABLE 1-2**

Motivations and Interest in a Higher-Level Job Involving a Large Amount of Time in Supervising Scientists

<table>
<thead>
<tr>
<th>Interest</th>
<th>Science Motivation</th>
<th>Institutional Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Yes, or possibly</td>
<td>81%</td>
<td>80%</td>
</tr>
<tr>
<td>No</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>N =</td>
<td>120</td>
<td>81</td>
</tr>
</tbody>
</table>

Chi-square: .009
Not significant
Chi-square: 10.39
Sig. at: .01

One step beyond direct supervision of research is taking on additional administrative duties not of a direct research character. Again we would predict that high institutionalists would be the most willing to accept these duties, while the researchers with high science motivation would be the least receptive. As Table 1-3 shows, the hypothesis concerning the institutional measure is confirmed. No significant difference is revealed by the science measure.

In view of these relationships, it seems safe to conclude that our two measures are in fact measuring two qualitatively different sets of values held by scientists.
TABLE 1-3

Motivations and Interest in a Higher-Level Job Involving a Large Amount of Time in Administrative Work

<table>
<thead>
<tr>
<th>Interest</th>
<th>Science Motivation</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>Diff.</td>
<td>High</td>
<td>Low</td>
<td>Diff.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes, or possibly</td>
<td>23%</td>
<td>27%</td>
<td>-4%</td>
<td>30%</td>
<td>17%</td>
<td>13%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>77</td>
<td>73</td>
<td></td>
<td>70</td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N =</td>
<td>113</td>
<td>77</td>
<td></td>
<td>110</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-square:</td>
<td>0.45</td>
<td></td>
<td></td>
<td>3.89</td>
<td>Sig. at:</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Motivation and research performance

Other things being equal, we would expect performance to increase as motivation increases. In the present case, this means that the greater the science motivation, the higher the research performance is likely to be. This is in fact the case, as Figure 1-1 shows. The figure shows a significant increase in performance with increase in science motivation. However, examining the same group on institutional motivation, no significant difference is found.

At first glance it would seem to follow from this finding that institutional motivation, in itself having no relation to performance, need not be considered further. However this is not the case. We know that there are some researchers with high science motivation who have institutional motivation as well. If we control on institutional motivation, we may find that science motivation is related to performance in a different manner.

Figure 1-2 shows the results. For researchers with low institutional motivation, the more science motivation the higher the performance. But for researchers with high institutional motivation, performance does not vary significantly with the degree of science motivation.
FIGURE 1-1
Motivation, as Related to Research Performance

Research performance

Science Motivation
- t of difference: 3.80
- Level of significance: .001

Institutional Motivation
- t of difference: 0.11
- Level of significance: Not sig.
FIGURE 1-2
Science and Institutional Motivation, as Related to Research Performance

<table>
<thead>
<tr>
<th>Institutional Motivation</th>
<th>t of difference</th>
<th>Level of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.61</td>
<td>Not sig.</td>
</tr>
<tr>
<td>Low</td>
<td>2.82</td>
<td>.01</td>
</tr>
</tbody>
</table>
The most important fact presented in Figure 1-2 is that high science motivation combined with high institutional motivation is associated with lower research performance than is a single-minded science motivation. 

The interpretation which seems most meaningful is that this situation represents a conflict between two strong goals in which part of the science motivation is channeled into the pursuit of institutional ends. In concrete terms, the researcher may be using some of his time and energy in activities related to his institutional interests.

This is a relatively simple explanation of what is, doubtless, a complex process. And furthermore, no attempt has been made to trace the functions that institutional motivations may have in the organization. Nevertheless, pending further evidence, the conflict interpretation stands as the most meaningful in conjunction with the other findings.

Influence as a factor in performance

We have related certain motivational measures of the individual to his performance without regard for the context in which he operates. One of the major conditions of the social environment is the mutual influence of chief and subordinate with respect to research matters. Scientists were asked two related questions (items 10 and 11):

"To what extent do [the chief's] activities or decisions affect your work, directly or indirectly?"

"... to what extent do you feel you could influence him in regard to these activities (if you wanted to)"

* It might be expected that the Commissioned Corps and the Civil Service personnel would vary on the two value orientations, but there is no significant difference between them. Furthermore, satisfaction with the promotion systems of the Corps and Civil Service is not significantly related to the values held by their respective members.

** We do know, however, that those scientists who express an interest in higher-level supervisory jobs do not necessarily have the same motivation as scientists who actually hold these supervisory positions. Neither science nor institutional motivation varies significantly among non-chiefs, unit, section, and laboratory chiefs.
The distributions of responses were such that they could be dichoto-
mized between "a great deal" and "moderately", thus forming rela-
tively "high" and "low" influence groups. It is important to note
that the measures represent the subordinate scientist's perception
of the chief's influence and of his own influence. */ Each scientist
is considered as a subordinate, even though he may supervise
other scientists.

Neither chief's nor subordinate's influence alone is significantly
related to performance. The trend is clear, however: the higher
the subordinate's influence and the lower the chief's influence,
the higher the subordinate's performance; but the differences are
not significant.

The implications of this finding are developed further in Figure 1-3,
where the relative balance of chief and subordinate influence, not
merely the level of each one separately, is related to performance.
Equal influence, whether equally high or low, forms the midpoint of
a three-point scale of the balance of influence. When the balance
is favorable to the chief, the performance is significantly lower
than when it is in favor of the subordinate.

There are two alternative interpretations, each based on different
assumptions, to explain the relationship of influence to performance.
One is that subordinate influence is a sort of autonomy which con-
stitutes a reward for past research performance. The better a sci-
entist's research, the more influence he is given. The other inter-
pretation is that influence may be a facilitator of performance;
that is, the more elbow room the scientist has — the more control
over his own work and over decisions affecting him — the better
job he will do. In all probability a combination of both series of
events would better reflect the reality of the situation. But be-
because we have data at one point in time we can only infer the pro-
cess — which means, in effect, weighing the evidence for the two
major interpretations.

*/ The two measures of influence are not significantly related to
each other; that is, high chief influence is about equally likely
to go with high or low subordinate influence. Nor do the two
influence measures vary significantly with grade, or between non-
chiefs, unit chiefs, and section heads.
Balance of Influence Between Chief and Subordinate, as Related to Research Performance

F ratio: 5.11
Level of significance: .01
Motivation, influence and performance

We know that science motivation is related to research performance. Our problem now is to specify the conditions under which the relation holds, and those under which it does not. Subordinate's influence — the variable we have just examined briefly — is important in this connection. In essence it is a general measure of the scientist's control over those things important to the research process.

Figure 1-4 relates science motivation to performance under conditions of high and low influence. For scientists with high influence, the degree of science motivation is significantly related to performance. However, for low influence scientists there is no significant difference between those with high or low motivation. In other words, high motivation toward science, when coupled with low influence, is not associated with performance appreciably above that of a low science, low influence researcher.

The gist of the figure is this: either high science motivation or high influence separately is not sufficient; only when they are joined are they associated with outstanding performance. This effect suggests the facilitative function of influence. Influence, in this instance, appears to operate as one of the necessary conditions favorable to effective research performance. It is difficult to interpret these results in terms of the alternative hypothesis that influence is mainly a reward following high performance. If this hypothesis were correct, one would not expect low as well as high performers to be accorded high influence.

Confidence and liking in chief-subordinate relations

The emotional tone in interpersonal relations is a difficult thing to measure. It is nonetheless important, especially when the interplay of influence is involved in making decisions.

We asked two questions which were designed to get at a general "feel" of the chief-subordinate relations (items 13 and 12, respectively):

"To what extent do you have confidence in the chief's intentions and motives? Do you feel he is always sincere in his dealings with others? Does he really mean what he says?"

"How strongly do you enjoy your contacts with [the chief] — whether you like him personally, gain professional stimulation from him, or enjoy contacts with him for any other reason?"
FIGURE 1-4

Subordinate's Influence and Science Motivation as Related to Research Performance

<table>
<thead>
<tr>
<th>Subordinate's Influence</th>
<th>t of difference</th>
<th>Level of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>3.26</td>
<td>.01</td>
</tr>
<tr>
<td>Low</td>
<td>1.50</td>
<td>Not sig.</td>
</tr>
</tbody>
</table>
The answers tend to pile up at the "positive" end. In order to get a high-low division, it was necessary to consider an answer which expressed less than complete confidence or very strong enjoyment as "low". At first glance this may seem to be an unrealistic breaking point, but it became evident that any deviation from the statement of absolute confidence or enjoyment represents more negative interpersonal feeling than the verbal categories imply. */

We find, from data not shown here, that neither confidence nor enjoyment as such is significantly related to performance. In conjunction with other measures, however, these variables show clear relationships. Figure 1-5 shows subordinate's influence as related to performance under high and low confidence. The performance of the low confidence group varies strikingly with the degree of influence. But no such variation characterizes the higher confidence group. It is clear that ability to influence the chief is more important, in terms of performance, to those subordinates who have low confidence in the chief. The same conclusion follows from Figure 1-6 for the low enjoyment group. Only under low enjoyment is influence significantly related to performance.

The meaning of these findings seems to be that high influence can serve as a defense against potential interference by a chief who is not trusted. The chief may be only mildly annoying, or he may seem actually threatening, but the subordinate in the low confidence group is likely to consider the chief's decisions and actions as unwelcome intrusions. And, if our interpretation is correct, a sense of control over the situation mitigates this apprehension.

We may call this the defensive function of influence. Like the facilitative aspect, it is clear that the defensive function can be operating only if influence is viewed as a condition of performance, rather than a reward which follows performance. Again the latter interpretation is difficult to draw from these data. Why should high performing scientists be given less influence by a chief whom they trust or enjoy a great deal? If anything we might expect the reverse on the basis of the reward hypothesis.

Decision-making about new research projects

Having considered the relation of subordinate influence to research performance in a general manner, let us turn to a specific decision-making situation. A question was asked (item 22):

*/  This conclusion is substantiated in Chapter II, when the same variables are related to communication.
FIGURE 1-5

Influence of Subordinate and Confidence in Chief's Motives, as Related to Research Performance

<table>
<thead>
<tr>
<th>Degree of Confidence</th>
<th>t of Difference</th>
<th>Level of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.06</td>
<td>Not sig.</td>
</tr>
<tr>
<td>Low</td>
<td>2.67</td>
<td>.01</td>
</tr>
</tbody>
</table>

Research performance
Subordinate's Influence, as Perceived by Subordinate
FIGURE 1-6
Influence of Subordinate and Enjoyment of Contact with Chief, as Related to Research Performance

<table>
<thead>
<tr>
<th>Degree of Enjoyment</th>
<th>t of difference</th>
<th>Level of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.5</td>
<td>Not sig.</td>
</tr>
<tr>
<td>Low</td>
<td>2.68</td>
<td>.01</td>
</tr>
</tbody>
</table>

Research performance

Subordinate's Influence, as Perceived by Subordinate
"What are the actual . . . relations between your chief and you, in determining what concrete work problems or assignments or follow-up steps you (or your staff) will work on next? . . ."

The four alternatives, in brief, were:

A. Subordinate makes such decisions himself, with the routine approval (if any) of the chief.
B. Subordinate and chief jointly decide.
C. Chief consults thoroughly with subordinate and then decides.
D. Chief decides himself as he feels best.

Respondents rank-ordered the alternatives in terms of frequency of occurrence. By an analysis of the combinations of rank-orders, it was determined that the alternatives can be considered as lying along a single dimension; that is, the results permit the items to be arranged in the order shown above, from "non-directive" behavior of chief to "directive." If an individual's first choice is A, his next choice will nearly always be B, rather than C or D. Similarly, if an individual chooses D, his next choice will generally be C, rather than B or A.

By a method developed for this study, the rank-order data were converted into a seven-point scale. For convenience, the dimension revealed by analysis of the rank-orders is called "directiveness" of chief.

Taking all researchers, irrespective of grade, the directiveness of the chief, as perceived by the subordinate, is related to performance. The more directive the chief the lower the research performance of the subordinate. This finding raises the question of cause and effect, but the analysis by grade which follows will tend to clarify this problem.

It should be noted that 77% of the scientists report their chiefs to be in the 1-4 range of the directiveness scale. This fact points up a major characteristic of the supervisory structure of the organization: the bulk of the decisions in this work area are made by the

*/ On the questionnaire these alternatives appeared in the order C, D, B, A.

/** The numerical scale has this general equivalence to the four alternatives: 1 2 3 4 5 6 7.

A B C D

The midpoint, 4, contains some tied scores.
subordinate alone, or jointly with the chief. */

Directiveness by grade

Figure 1-7 shows the directiveness data, by grade, as related to performance. The directiveness scale had to be collapsed because of insufficient numbers toward the directive end. It was collapsed in such a way as to preserve the qualitative distinctions: subordinate decides, joint decision, and chief consults or decides. In effect, a three-point scale was made from the seven-point scale. The groups shown on the graph include two groups of non-chiefs: those at GS-9 and 11, and those at GS-12 and above (all of whom, it will be recalled, have doctoral degrees). Unit chiefs were made a separate category, but section chiefs were omitted. The decision not to show section chiefs was based on the observation that decisions about their work do not center primarily on their own personal research assignments.

Figure 1-7 indicates an interesting difference in performance by grade. The non-chiefs at GS-9 and 11 have their peak performance under conditions of joint decision-making, do somewhat less well under complete autonomy and below average under the more directive conditions. These variations are statistically significant. In contrast, non-chiefs at GS-12 and above show no significant difference in performance under the three conditions. This does not necessarily imply that all three conditions are equally enjoyable; it is rather that they are not markedly related to performance. The unit chief, finally, shows a performance curve inversely related to directiveness to a significant degree. /**

Again the complexity of the supervisory process is emphasized by these data. Here the necessity for flexibility of behavior seems to be the central theme. There appears to be a need to progress from joint decision-making at the beginning levels to autonomy at the higher grades, especially in the latter case when supervisory responsibilities are involved. It becomes abundantly clear that problems of supervision are not necessarily all solved by adopting a laissez-faire approach. The prime consideration appears to be that the degree of "directiveness" must be suited to the needs of the particular scientists involved.

*/ See Human Relations in a Research Organization, pp. 118-152 and 275-279 for additional data.

/** Ibid., page 279, Figure 8-2 shows roughly parallel results from the earlier high-low performance scaling.
FIGURE 1-7

Directiveness of Chief in Work Decisions, as Related to Research Performance, by Grade

Research performance *

Directiveness of Chief, as Reported by Subordinate

<table>
<thead>
<tr>
<th>Grade</th>
<th>F ratio</th>
<th>Level of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS 9-11, non-chief</td>
<td>5.12</td>
<td>.05</td>
</tr>
<tr>
<td>GS 12+, non-chief</td>
<td>0.10</td>
<td>Not sig.</td>
</tr>
<tr>
<td>Unit chief</td>
<td>6.35</td>
<td>.01</td>
</tr>
</tbody>
</table>

* Because the performance measure is controlled for grade the lines show performance high or low relative to a specific grade; no comparisons of the levels of the three lines to each other is legitimate.
Subordinate influence by grade

An illustration of the point that the "same" supervisory behavior has different consequences on different groups of scientists is afforded by Figure 1-8. Here we go back to the subordinate's perception of his influence with his chief. We see that only for the scientists at GS-9 to 11 is degree of influence significantly related to performance; the more influence, the higher the performance. That is, the performance of these scientists is closely connected to the degree of influence accorded them by the chief. At the same time, to judge from previous data, influence does not mean complete autonomy, but rather a share in joint decisions. It appears that the effects we have been analyzing are particularly relevant in the supervision of the scientists at the bottom of the grade hierarchy.

Sense of belonging and research performance

There are other interpersonal variables which bear on research performance. Thus far, however, none of them appears to be as important as the supervisory factors we have just outlined.

One factor, sense of belonging to the organization, has played a part in previous studies in other settings. Part of the analysis of research performance was therefore planned along this line.

Identification with the organization, or part of it, is generally considered one aspect of "morale." Sense of belonging to the organization or group is the general meaning in which identification is used here.

The analysis of the sense of belonging to section, laboratory, and institute shows interesting trends, but none of the relationships goes beyond marginal statistical significance. However weak statistically, they do shed some light, as a few trends will show.

There is a consistent but slight tendency for performance and sense of belonging to be positively related at the section level, unrelated at the laboratory level, and inversely related at the institute level. In general then, high section belonging goes with slightly higher performance, and high institute belonging with slightly lower performance.

This would seem to imply that identification with immediate work group is a factor related to performance, and, perhaps that identification at more than one level is not desirable. The first hypothesis is supported in that performance is positively (but not significantly) related to degree of identification with the smallest, the most immediate work group (whether unit, section, or small
Subordinate Influence by Grade, as Related to Research Performance

<table>
<thead>
<tr>
<th>Grade</th>
<th>t of difference</th>
<th>Level of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS 9-11</td>
<td>3.49</td>
<td>.001</td>
</tr>
<tr>
<td>GS 12</td>
<td>0.40</td>
<td>Not sig.</td>
</tr>
<tr>
<td>GS 13+</td>
<td>0.02</td>
<td>Not sig.</td>
</tr>
</tbody>
</table>
laboratory). As Figure 1-9 indicates, high belonging to section tends to be positively related to performance only under conditions of low belonging to laboratory.

The implication of this finding appears to be that it is more important to feel part of the immediate group which forms the interpersonal setting for research, than of the larger or broader levels of the organization. Identification with the work group may reflect the involvement with a specific scientific problem to be solved in this group. In general, group attitudes and characteristics are found to be only moderately related to research performance. Data which are not presented here contain leads which will be followed up in later analyses.

* In addition, data bearing on the characteristics of group attitudes at the laboratory level are presented in Chapter IV.
FIGURE 1-9
Identification with Section and Laboratory, as Related to Research Performance

Research Performance

Degree of identification: 1.0

with Section

with Laboratory

Low Low High Low High

Sig. at: .10

Not sig.

Sig. at: .10

Not sig.
C. Summary

The analysis presented in this chapter was designed to examine research performance in relation to the social setting in which the scientist operates. More specifically, we were interested in determining what interpersonal factors facilitate or inhibit the scientist in translating his motivation toward research into actual research performance.

To begin with, motivation as such was examined. Two motivational dimensions were set up, on the basis of measures of orientation toward two sets of goals potentially available to the scientist in his role of researcher in a large organization. These were the science and the institutional orientations. High performance is associated with high science motivation, but this is not the whole story. High science motivation may be checked to some degree by conflicting motivation toward institutional goals.

Next the analysis proceeded to examine the situations in which these motivations are translated into performance. In studying the interpersonal setting, one area stood out as meriting major emphasis — that of chief-subordinate relations. Within this area, the factor of subordinate's influence — his degree of control over the chief's decisions affecting his work — was found to be highly important, and to have either facilitative or defensive functions. High science motivation apparently is facilitated by a high degree of influence. On the other hand, in chief-subordinate relations marked by some negative feeling (such as lack of confidence in the chief), a high degree of influence apparently has a defensive function as well. In view of these relationships it seems safe to interpret the role of influence as that of a necessary condition to effective performance, rather than a reward for high performance.

In a more specific influence situation — decision-making about new research projects — it was clear that there are distinct patterns of performance linked to the degree of directiveness exerted by the chief at different grade levels. In the lower grades somewhat less than complete autonomy is associated with the peak of performance, but at higher grades freedom to make work decisions is related to high performance.

The findings as a whole indicate the complexity of the supervisory process, and point to a need for flexibility of behavior within it, particularly in terms of the motivations and research experience of the scientist. The laissez-faire approach to supervision does not solve all the problems involved.
Preliminary examination of evidence concerning sense of belonging as it affects performance indicates that identification with the small work group has more relationship to performance than has identification with broader levels of the organization. Further analysis is necessary to explore fully the meaning of belongingness and to integrate it with the other variables.
This appendix contains additional information on the assessment procedures. It includes the instructions to assessors, a list of the groups assessed, details of the method of combining the assessment ratings into a single score for each person, and data on the sample of scientists used in analysis of research performance.

Instructions for assessors

In Appendix C of *Human Relations in a Research Organization* (1953) the entire set of instructions to assessors is reproduced. This includes instructions for Assessment I, current scientific performance, and Assessment II, "The scientist's current scientific performance, as compared with the maximum performance of which he is capable in view of his own background and experience."

Only data collected in Assessment I have been scaled and utilized in the present analysis. Accordingly, only the instructions applicable to it are reproduced below.
INSTRUCTIONS FOR ASSESSORS

The names of the scientists you are to assess within a particular laboratory or a particular discipline are listed alphabetically on the attached sheets. You will be asked to make two separate assessments on each set of names; therefore duplicate lists are provided. We would like your opinion as to the relative standing of these scientists with respect to each other. All may stand high; but within this list we want you to indicate which ones are higher, which lower.

You may wish to indicate:

- two groups: (1) Higher; (2) Lower
- or three groups: (1) Highest; (2) Next highest; (3) Third highest
- or four groups: (1) Highest; (2) Next highest; (3) Third highest; (4) Fourth highest, etc.
- or more.

We suggest you read through the list and mark as "1" the names that stand highest; then mark "2" the names that stand next highest; and so on.

Use as many groups as you need to show the differences that exist; in general you may need three or four. You may assign several people to the same group, providing you feel they are all about equal.

Cross off the names of any whom you feel you do not know well enough to assess; and cross off your own name without assessment.

Following are the definitions of the two separate assessments to be made. Please read them carefully.

Assessment I: The scientist's current scientific performance as compared with other NIH scientists on the same list.

By a person's "scientific performance" we have in mind factors such as:

- Originality and creativeness in locating important problems, or in turning up fruitful leads for attacking the problems

- Wisdom and judgment in deciding which lead is most likely to pay off, or which methods to apply

—continued
Instructions for Assessors

January 28, 1953

—Rigor of thought and precision of methods in carrying out the work and drawing conclusions

—Persistence, industriousness and efficiency in carrying through a project; keeping after the task in hand with a minimum of delays and wasted activity

—Contribution to the work of others, by means of knowledge, insight, and stimulation of ideas. Part of a person's contribution may consist of guiding or stimulating the work of other scientists; or perhaps his "scientific performance" as a whole may depend largely on this factor.

All these factors may contribute to a person's effectiveness. However, some high performers may be outstanding in only one or two of these aspects.

By his "current" performance we mean either work he has completed over the last few years, or work he is currently doing which is likely to "pay off" in the next few years.

To make the assessment more concrete, you may think of the problem this way:

Suppose a sum of money is being made available for research in the area these people are working in. You are being asked to give advice on distributing these funds in research grants or contracts to a small group of scientists — either as individuals or as a team — who, in your estimation, are best equipped to turn out the best possible research within the next few years. (You are not being asked to work with these people, so your personal relations with them should be disregarded.) On the basis of what these scientists have done, are doing, and therefore are likely to do in the next few years, which of them would you recommend first?

Consider what the person has actually done or is now doing, regardless of his age and experience, and regardless of his grade or status. We plan to take the latter factors into consideration in our analysis; and we will compare each person with others of similar age and experience, and status. In your own assessments, therefore, consider only the person's performance as such.

Try not to be influenced by the field in which a person works — as, whether that field is "promising" or "sterile", or whether it falls more in "applied" areas so-called, or more in "basic" areas. Rather, consider the way he performs within his area, in terms of the factors listed above.
Groups within which assessments were made

In all, 72 scientists served as assessors. Of these, 54 gave evaluations of the following laboratory groups:

NIAMDS

L. Biochemistry & Nutrition
L. Pathology & Pharmacology
L. Chemistry
L. Physical Biology

NCI *

L. Biochemistry
L. Biology
L. Biophysics
L. Chemical Pharmacology

NIDR
NHI
NIMH
NMI

L. Biologics Control
L. Infectious Diseases
L. Tropical Diseases

Also, 52 scientists assessed the following disciplines (a few assessors rated more than one field):

Biochemistry & Nutrition
Enzymes & Metabolism
Endocrinology
Organic Chemistry
Pharmacology
Pathology
Physiology
Biophysics
Cellular Functioning
Virology
Microbiology

*/ In NCI, the Laboratory of Pathology and the Endocrinology Branch were entirely included in the corresponding fields of work and were not assessed separately as laboratories.
Obtaining a single score for each scientist

For analysis purposes it was necessary to combine the judgments of various assessors into a single score for each scientist. In doing so, two objectives were sought.

(1) First, it was desirable to utilize only those data in which we could place high confidence. Toward this end the following four steps were taken:

   (a) A few assessors were dropped who provided little information. Included here were those who evaluated fewer than one-fifth of their respective lists, and those who placed nearly all their names (85% or more) in a single "high" or "low" category.

   (b) A few more assessors were dropped whose evaluations were markedly divergent from those of the other judges in the same area. In all, 9 assessors were dropped for either of these reasons.

   (c) Only those scientists who received judgments from two or more of the remaining assessors were used.

   (d) In a few cases a scientist was eliminated if the assessors from his laboratory or discipline group disagreed markedly in their evaluations of him. The criterion adopted for "marked disagreement" is explained below.

(2) After these eliminations the next objective was combining the judgments from assessors within each laboratory and within each discipline. To do this, the qualitative categories of "high" or "medium" or "low" performance had to be given a numerical score, so that scores from several judges could be averaged. Assigning of scores was complicated by the fact that assessors were free to make as many distinctions as they saw fit, and to assign as many persons to each category as they wished. Some judges made one discrimination between "high" and "low" performers, while others preferred to make several distinctions. (For convenience, categories containing fewer than 15% of the names were combined with an adjacent category.)

A numerical score of 1 to 9 was therefore assigned to each category of judgment, depending on the proportion of scientists an assessor assigned to each of his categories. To illustrate this process, we show some hypothetical data from three assessors evaluating the same group of scientists.
Each category was assigned a score corresponding to the decile nearest the middle of the category. For instance, Assessor I gives 25% of the group High ratings. The middle of this category falls at the 87th percentile, or the ninth decile; his High category therefore receives a score of 9.

For Assessor III, the middle of his High category falls at the 70th percentile, so this category receives a score of 7. This is simply another way of saying that it is easier for a scientist to fall in his top category and therefore the score is lower than that of Assessor I.

The combined score of the three assessors on a given scientist is an average of the three category scores. A scientist rated High-High-High (numerically equivalent to 9-8-7) would receive a combined score of 8.

We mentioned above that certain scientists were eliminated if assessors disagreed in their evaluations of him. In the hypothetical example, a scientist rated Low-High-Low would be eliminated. Note that in the
opinion of Assessor I this man has a percentile rank no higher than 20, while in Assessor II's opinion he is no lower than 65 — a discrepancy of 45 percentile points. An arbitrary decision was made that all discrepancies of 30 or more points were "markedly discrepant", and such cases were dropped.

By contrast, however, a rating of High-Low-High in this case would not lead to elimination, since there the discrepancy may be as low as 15 points.

By and large, the ratings were remarkably consistent. Altogether, 577 individual scorings were made in this way (including scorings of the same scientist in two or more of the laboratories and disciplines). Of this number, 64% showed no discrepancies as defined above, while 75% had discrepancies of less than 10 points. Only 8% of the total scorings were eliminated because of "marked discrepancies."

In terms of individuals, however, fewer than 5% were discarded; some individuals lost their discipline score but retained their laboratory score, or vice versa.

(3) For those individuals having both a laboratory and a discipline score, relatively high correlations between the two were found to exist ($r$'s range from .47 to .97 for different pairs of groups; median $r = .84$). Therefore we felt justified in averaging the two scores where both existed, and in using either of the scores if only one existed.

Some data on those scientists used in the research performance analysis

As Table A-1 shows, performance scores were obtained for 266 scientists at NIH. Of these, the large majority received both a laboratory and a discipline score. It is interesting to note that 90% of the scientists received usable laboratory scores, compared with 72% who received discipline scores.

Table A-2 gives a detailed breakdown on those who were assessed and not assessed. Assessments were obtained on 80% of the scientists who filled out questionnaires. Of the total group, 23% scientists were selected for the analysis in Chapter I. Usable assessments were obtained on 87% of this selected group.
### TABLE A-1

**Scientist Having Usable Laboratory and/or Discipline Scores**

<table>
<thead>
<tr>
<th>Scientists having:</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both Laboratory and Discipline scores</td>
<td>166</td>
<td>62%</td>
</tr>
<tr>
<td>Laboratory scores only</td>
<td>73</td>
<td>28</td>
</tr>
<tr>
<td>Discipline scores only</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>266</td>
<td>100%</td>
</tr>
</tbody>
</table>

### TABLE A-2

**Breakdown of Scientists With and Without Assessments**

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Obtained</th>
<th>Not Obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientists selected for the analysis in Chapter I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research scientists having Ph.D and/or M.D., including unit and section chiefs</td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>Same, but having no assessment because:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not assessed (due to unfamiliarity, etc.)</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Only one assessor</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Markedly discrepant assessment pattern</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Remainder excluded from analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laboratory chiefs, Scientific Directors, Institute Directors, etc.</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Grade less than GS-9; non-doctoral;</td>
<td>42</td>
<td>18</td>
</tr>
<tr>
<td>others not included above</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Percent of total institute scientists filling out questionnaire</strong></td>
<td>80%</td>
<td>20% = 100%</td>
</tr>
</tbody>
</table>