



Status in Experimentally Produced Groups

Author(s): Muzafer Sherif, B. Jack White and O. J. Harvey

Source: *American Journal of Sociology*, Vol. 60, No. 4 (Jan., 1955), pp. 370-379

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/2772030>

Accessed: 06-04-2015 17:09 UTC

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *American Journal of Sociology*.

<http://www.jstor.org>

## STATUS IN EXPERIMENTALLY PRODUCED GROUPS<sup>1</sup>

MUZAFER SHERIF, B. JACK WHITE, AND O. J. HARVEY

### ABSTRACT

In a study of intra- and intergroup relations utilizing a combination of observational and laboratory methods, two small groups were experimentally produced as a consequence of interaction among individuals. This was done under controlled conditions embodying goals that had a common appeal and that called for interdependent activities for their attainment. It is demonstrated that, when group members face an unstructured task in which they are sufficiently motivated, their estimates of each others' performance are influenced by the status each occupies in the social hierarchy. This finding shows the feasibility of assessing the effects of group interaction by means of laboratory techniques.

This paper reports an experiment carried out as part of a research program on group relations. The problem of this unit concerns judgments of performance as indexes of status relations among members of small groups which are themselves experimentally produced. The focus will be mainly upon the theoretical and methodological considerations which constitute the distinctive features in the research program.

In this series of experiments on intra- and intergroup relations the guiding concern has been to use an interdisciplinary approach—an approach to human relations accepted by an increasing number of recent writers.<sup>2</sup> Lack of the interdisciplinary perspective in studying human relations has resulted, on the whole, either in the formulating of problems and experimental designs which are often artificial, creating serious questions of

validity, or in depriving field studies of the benefit of rigorous techniques developed by experimentalists, thus raising equally serious questions concerning the exact nature and control of factors.

The theoretical and methodological issues involved are illustrated by two major trends in research on the small group. On the one hand, there is the impressive body of literature on small groups from the field work of sociologists. This line of development is rich in content, lifelike in context, and full of suggestive recurrences. Yet the findings need verification by appropriate experiments which make possible specification and manipulation of conditions.

On the other hand, laboratory experiments have been accumulating rapidly during the past decades. Experiments have attempted to measure the effects of *togetherness* on various psychological processes. We have learned a great deal from these studies about the effects on behavior of specific social situations. Yet this strictly experimental approach has proceeded, on the whole, without due regard to suggestions derived from the field work of sociologists for the formulation of significant problems and hypotheses that can be tested.

The aim, therefore, was a program of research which would embody advantages derived from an interdisciplinary approach for an integrated study of group relations. This approach starts with due recognition of the recurrences observed by field workers in intra- and intergroup relations of small

<sup>1</sup> This research program has been continued during the last two years under the direction of Muzafer Sherif with a grant from the Rockefeller Foundation, which is gratefully acknowledged. Other aspects of the present study will be reported in collaboration with various staff members. We are grateful to Professor Carl Hovland and other members of the Yale Communication Research Project for the facilities extended during the carrying-out of the study.

<sup>2</sup> E.g., S. S. Sargent and M. W. Smith (eds.), *Culture and Personality* (New York: Viking Fund, 1949); J. E. Hulett, Jr., and Ross Stagner (eds.), *Problems in Social Psychology* (Urbana: University of Illinois, 1952); R. E. L. Faris, "Development of the Small Group Research Movement," in M. Sherif and M. O. Wilson (eds.), *Group Relations at the Crossroads* (New York: Harper & Bros., 1953), chap. vii.

groups, takes note of the minimum essential properties of groups, and proceeds to the formulation of problems and hypotheses to be tested experimentally on this concrete basis rather than on the basis of hunches that may have little or no relevance to vital social events. At the same time it utilizes psychological principles of validity proven over a period of time and the methods and techniques developed by the experimentalist.<sup>3</sup> This procedure recognizes from the first the structural properties of the setting and the irreducible differential effects produced in the interaction process in intra- and intergroup relations. It does not disregard the unique "traits" or unique contributions of the individual participating but attempts to study their unique characteristics and contributions as they are affected, modified, and even transformed by the reciprocities emerging or standardized in group interaction.

If it can be established that status relations within a group and positive or negative relations between groups can be predicted from judgments obtained from individual group members under specified conditions, we shall be taking a step in the direction of testing laboratory and field findings within one experimental design. Since the judgments are made of performance in tasks chosen for their appeal in the subject's eyes, but unstructured in specified ways, such techniques for the assessment and, finally, prediction of status relations and other products of group interaction will have advantages over existing devices: (1) Such techniques will provide precise measures in terms of perceptions and judgments along specific dimensions, as contrasted with gross behavioral observations. Such indexes are not subject to the experimenter's interpretation of the subjects' reactions, a condition which has occasioned many current controversies over various projective tests. The judgments or perceptions reported in these

<sup>3</sup> Experimental work, primarily of psychologists, is summarized comprehensively up to 1937 in G. Murphy, L. B. Murphy, and T. M. Newcomb, *Experimental Social Psychology* (rev. ed.; New York: Harper & Bros., 1937).

experiments are in numerical form to start with and do not require interpretation or evaluation by the experimenter *prior* to statistical analysis. (2) Since the judgments and perceptions are made in an unstructured situation and do not betray to the subjects the problem being investigated, they are indirect methods of eliciting reactions. As frequently noted recently, direct questioning or other techniques which make the subject aware that he is being observed unavoidably give rise to personal qualms which influence his responses, especially if responding one way or another involves the subject's ego.

It is not suggested here that group attitudes and other group effects be studied solely through judgment or perception. At this early stage in the study of group relations, such indexes obtained experimentally can be advantageously used as a check upon findings reached by other methods, observational, sociometric, and so on.

#### BACKGROUND AND FORMULATION OF HYPOTHESES

The hypotheses in the present investigation, related to both the experimental production of small groups themselves and the assessment of status relations within the groups thus formed, are not postulated on a priori grounds. They are derived from a large amount of evidence, both empirical and experimental, of which only a summary statement will be given here.

The feasibility of producing small groups experimentally is suggested by sociological studies over a period of decades.<sup>4</sup> From these works one can abstract certain minimum characteristics of the rise and functioning of small informal groups. Among these are

<sup>4</sup> The studies of Thrasher, Shaw, Anderson, Zorbaugh, Landesco, Whyte, Roethlisberger, and Dixon and various studies of adolescent cliques are illustrative in this connection. Representative findings are surveyed in M. Sherif and H. Cantril, *The Psychology of Ego-Involvement* (New York: John Wiley & Sons, 1947), chap. x; M. Sherif, *An Outline of Social Psychology* (New York: Harper & Bros., 1948), chaps. v, vi, and xiii; M. Sherif and C. W. Sherif, *Groups in Harmony and Tension* (New York: Harper & Bros., 1953), chap. viii.

some motive or motives shared or endured in common which are conducive to interaction and which impel members toward the attainment of common goals. As members gravitate toward one another in striving for common goals, the interaction process produces differential effects on their behavior. In time, interaction becomes stabilized in a pattern of reciprocities manifested in a group structure consisting of hierarchical statuses and roles for individual members. The established pattern of reciprocities becomes codified in terms of certain norms regulating the expectations, responsibilities, and loyalties of members occupying the respective roles and statuses. Norms are also standardized for other matters of consequence or relevance to the existence, activities, and goals of the group.

Interaction is not made an item in this list of characteristics of informal groups, because interaction is the *sine qua non* of any kind of social relationship, whether interpersonal or group. Nor are common attitudes or sentiments emphasized separately: common attitudes are derived from and formed in relation to social norms which are standardized in a group and which establish for various activities a range of tolerated behavior. If group members do not share in common certain values or norms (at least in matters concerning the identity, continued existence, and practices of the group) and each member does not hold attitudes within the range of tolerated behavior established by these norms, one cannot even speak of a group.

With these features in mind, we define a group as a social unit which consists of a number of individuals who, at a given time, stand in more or less definite interdependent status and role relationships to one another and which explicitly or implicitly possesses a set of values or norms regulating the behavior of members at least in matters of consequence to the group.

If through the introduction of specified conditions we can produce, among individuals with no previously established status or role relations, a structure of statuses and

roles, we can speak of experimental formation of a group. In this report our concern will be mainly with the rise of such a status structure. The hypothesis to be tested here is:

A definite group structure consisting of differentiated status positions and reciprocal roles will be produced when individuals (without previously established interpersonal relationships) interact with one another under conditions which present goals that have a common appeal and which require interdependent activities for their attainment.

Such groups were produced in our study of group relations in 1949,<sup>5</sup> but no attempt was made in that study to assess the status relations among group members through indexes of judgment. The attempt reported here was made in our second large-scale experiment on group relations carried out in the summer of 1953.

Our hypothesis concerning the assessment of status relations in groups which are themselves experimentally produced hinges upon the feasibility of assessing already existing relationships through indexes of judgment. The rationale for doing so was suggested by the principle that the so-called "cognitive" processes (judgment, perception, remembering, etc.) take place within a frame of reference in which both external stimuli and internal factors (including biogenic motives and social attitudes, status expectations, etc.) are jointly operative. Studies revealing variations in judgments as a function of group identification or reference when the checks and restraints of reality are not too compelling give our study an experimental basis. The substantial body of experiments along these lines is too familiar to be repeated here; only a few will be mentioned. For example, in a 1939 study which served as a prototype for many others, Chapman and Volkmann found variations in estimates as affected by the standing of one's own group in relation to other groups perceived

<sup>5</sup> For a full account of this study see M. Sherif and C. W. Sherif, *op. cit.*, esp. chap. ix.

as "higher" or "lower" on a given task.<sup>6</sup> Marks found displacement of judgments of skin color (within limits) in the direction of a desired area of the color continuum.<sup>7</sup>

A series of studies undertaken at the University of Oklahoma during the last five years constitute the more immediate background of the present experiment, which aims at assessment of status relations in experimentally produced groups. Following a demonstration of variation in judgments as a function of positive interpersonal relations,<sup>8</sup> Harvey and Sherif found differential estimates of own future performance and that of a partner in directions predicted on the basis of interpersonal relations (positive or negative) prevailing between the partners before the experiment.<sup>9</sup>

In this work, variations in judgment were chiefly a function of the relationship between subjects, the degree of structure of the experimental situations being held constant. Going a step further, James Thrasher systematically varied gradations of stimulus structure and the nature of interpersonal relationships of subjects participating as partners in the experiment (friends or strangers).<sup>10</sup> He found that, as the stimulus conditions became more unstructured, the correspondence between stimulus and judgment decreased and the effect of social influences increased.

The study most directly related to the present unit is Harvey's experiment on

<sup>6</sup> D. W. Chapman and J. Volkmann, "A Social Determinant of the Level of Aspiration," *Journal of Abnormal and Social Psychology*, XXXIV (1939), 225-38.

<sup>7</sup> E. Marks, "Skin Color Judgments of Negro College Students," *Journal of Abnormal and Social Psychology*, XXXVIII (1943), 370-76.

<sup>8</sup> C. Sherif, reported in M. Sherif, *An Outline of Social Psychology*, pp. 289-92.

<sup>9</sup> O. J. Harvey and M. Sherif, "Level of Aspiration as a Special Case of Judgmental Activity in Which Ego-Involvements Operate as Factors," *Sociometry*, XIV (1951), 121-47.

<sup>10</sup> James D. Thrasher, "Interpersonal Relations and Gradations of Stimulus Structure as Factors in Judgmental Variations: An Experimental Approach," *Sociometry*, XVII (August, 1954), 228-41.

status relations in informal groups.<sup>11</sup> Members of already existing cliques occupying well-differentiated status positions—high, middle, low—participated in the experimental task, three members from each clique representing these positions serving together. Harvey found significant variations in estimates of future performance by one's self and by other members, in line with their respective status positions in the group.

The implication of these experiments is that, if status structure were emerging (verifying our hypothesis concerning the experimental formation of groups stated above), it may be that status relations will be reflected in the members' judgments of each others' actual performance. It is reasonable, therefore, to formulate the following hypothesis:

Variations in judgments made by an experimentally formed group of a member's performance on a task which is of common significance to the group and which provides few external anchorages (i.e., is unstructured) are significantly related to the status of that member.

The higher the status of the member whose performance is judged, the greater the tendency of the others to overestimate his performance. Conversely, the lower the status of the member, the less the tendency of other members to overestimate his performance; they may even underestimate it.

#### EXPERIMENTAL FORMATION OF GROUPS

The procedures and conditions designed to test our hypothesis concerning group formation were essentially similar to those in the 1949 study of group relations.<sup>12</sup> Very briefly, the preconditions for this test involved selecting subjects with no previously established role or status relationships and bringing them together in a situation that was natural and lifelike, that appealed to them, and that presented goals requiring cooperative interaction for their attainment. The experimental site was a summer camp

<sup>11</sup> O. J. Harvey, "An Experimental Approach to the Study of Status Relations in Informal Groups," *American Sociological Review*, XVIII (1953), 357-67.

<sup>12</sup> M. Sherif and C. Sherif, *op. cit.*, chap. ix.

which held possibilities for various activities attractive to boys and was isolated from towns, highways, public amusements, and other diverting influences.

The subjects were boys of about twelve years of age from settled upper-middle-class Protestant families. Through interviews and psychological tests administered by a clinical psychologist, normal boys were selected. They were of similar educational level and somewhat above the average in intelligence. They understood that they were coming to a summer camp planned for the study of camping methods and procedures.

Participant observers were used as counselors and in other staff positions. Experimental personnel was instructed to play their parts as naturally as possible and not to take notes in the boys' presence without good excuse. Further, personnel was instructed, in so far as possible without involving danger or hardship for the boys, to allow freedom to the boys themselves in taking initiative and in planning and executing activities. Activities, in turn, were those for which the boys expressed preference, as weather and other practical conditions permitted.

Upon arriving, all twenty-four boys were placed together in a large bunkhouse, and for two days the activities involved the entire camp. This was done so that the group formation to follow could not be attributed to friendship clusters which might form spontaneously on the basis of personal affinities and common interests and without manipulation of experimental conditions. Following this brief period, the boys were divided into two experimental groupings made up so as to split budding friendship clusters and to be as similar as possible in athletic ability, personality characteristics, and prior acquaintanceship. (It turned out that a number of these boys were not complete strangers to each other but had at least seen each other in school or church.)

Some boys strongly disapproved of the arbitrary division into two groups, which were subsequently located in tents at considerable distance from each other. There-

fore, the first activities after the division were a hike and cook-out, which were very attractive to these boys. Other activities were so planned that the execution of tasks and attainment of goals devolved as much as possible upon the group as a whole. The two groups were kept separate as much as possible during this period. The activities included a treasure hunt for each group separately, the solution of which brought a group award of ten dollars to be spent as the group decided, work on projects initiated in the respective groups (e.g., building a dam, lean-to), campfires, religious services, overnight hikes and camping trips, and other usual camp activities—all engaged in by each group separately.

The scope of the conditions, procedures, and practical matters which have to be manipulated, surmounted, and attended to in an experiment such as this, which attempts to control as completely as possible the total situation and its crucial aspects, is very broad. Even best efforts to maintain the criteria established for each step at times failed. On one occasion rain made it necessary to postpone the overnight hike and camp-out planned to begin that day, and the boys had to spend a good part of the day in tents. By the seventh day two groups could be said to exist, the criterion being that observers could place the various members in their own status positions with relatively little disagreement, especially at the upper and lower levels, on the basis of such behavior as initiative attempted and accepted, choices in activities, etc. Thus our hypothesis concerning group formation was substantiated and the experimental formation of groups reported in the 1949 study essentially verified, including the reversal of spontaneous friendship choices as a by-product of group delineation.

*Experimental groups.*—Before proceeding to the experimental unit on assessment of status relations, a word concerning the two experimental groups is in order.

One group characteristically displayed greater solidarity as well as greater stability of status structure than the other at the end

of the period of formation. This group adopted the name of "Panthers" as they discussed plans for spending their treasure-hunt reward for a group flag with a panther as its emblem. The name is one of the indications of group morale; for example, expressions such as "I'm a Super-Panther" and "We're Super-Panthers" were not uncommon. Of course, the difference in degree of stability between the experimental groups was comparative. The solidarity and stability of status structure which evolved in the Panther group would be exceeded in turn by many informally organized groups functioning spontaneously in actual life.

The second group, which later in the study became known as the "Pythons," had more severe status problems from the beginning, being separated into an "upper-crust" and "a lower-rung" segment. Achievement of an interrelated group structure proceeded with difficulty, and the picture was further complicated by another uncontrollable event—the illness and finally the necessary departure from camp of the acknowledged leader. A second leader developed, a boy who had kept somewhat aloof in taking initiative until this opportunity appeared. In spite of this, the status positions assigned by three observers of this group were in high agreement.

Since the Python name came later in the larger study, the group names will not be used to designate the experimental groups. For present purposes, the experimental groups will be referred to as "Pa group" and "Py group," the Pa group being the boys who had already called themselves "Panthers."

#### ASSESSMENT OF STATUS THROUGH JUDGMENTS OF PERFORMANCE

In order to test our hypothesis concerning the differential effects of status relationships within the two experimentally formed groups, judgments by group members of each others' performance had to be secured on a task which had considerable appeal value common to the group, which seemed "natural" in the camp setting (e.g., a "rea-

sonable" thing to do), and which was sufficiently unstructured or ambiguous for external anchorages (or cues) not to be dominant in determining judgments.

Accordingly, the task chosen was throwing handballs at a target specifically designed to provide few external anchorages for judgments of performance while permitting the experimenter to record actual performance.

*Apparatus.*—The target was a five-foot circle of three-fourths-inch plywood cut into fifteen concentric circles which were attached to a plywood backboard by means of quarter-inch bolts and coil springs. Each of the fifteen concentric circles was cut into segments. When the impact of a handball depressed one of these segments, a bolt was driven against an electrically sensitized contact point. At the back of the board, a small panel containing fifteen radio bulbs was connected with the electrical circuit to correspond with the fifteen concentric circles. Thus the location of a segment of any one of the fifteen concentric circles depressed by a handball could be immediately recorded by the experimenter at the panel.

A cover of blue denim was hung over the front of the target, suspended from the top of the backboard, to hide the concentric circles from the subjects' view when they threw. Since the ball recoiled from the target immediately, no marks were left on the cloth.

*Procedure.*—On the morning of the day of the experiment, both groups were told that a tournament of games between them would begin in the afternoon, the first event to be a softball game. This tournament was to be the activity involving both groups (inter-group relations). The experimental task was presented to the subjects as related to this tournament, in order to make the procedures seem natural and to heighten its appeal; it was suggested that they might like to get a little practice before the game by throwing handballs at a target. The experimenter (who was known to the boys as a member of the camp staff) proposed that the practice be turned into a game with everyone taking

turns in order that all might have equal opportunity to benefit, an idea which they accepted as a good one.

The experiment was carried out in a large recreation hall equipped with concealed recording equipment. The two groups performed separately, the junior counselor of the respective group and the experimenter being the only adults present. Instructions were given orally and informally by the experimenter:

This is a game of skill to see how well you can judge the throwing ability of your friends and yourself. This is the target, covered with blue cloth. The target underneath the cloth has fifteen circles, ranging in score value from 2, the outside circle, to 30, the bull's-eye. (*The denim covering was raised.*) Now look carefully at the target and remember the way the circles run in score values, because the cloth will cover the target all the time you're throwing at it.

Let's let each of you judge your friends' skill and your own skill twenty-five times at throwing the ball at the target. We'll take turns at throwing until each of you has thrown twenty-five times. After each throw, the one who is throwing will judge the score he thinks he made. He will write this down on his score sheet and also call it aloud. The rest of you who are not throwing will write down the score you think the thrower made each time after he throws. You will not call your judgments of his score aloud. (*S's were shown where to record their judgments on the score sheets.*)

Boys, I won't be able to take time to tell you your score. I'll just remind you that 30 is the best possible score you can make and zero is the worst. Be sure to always do your best in throwing and judging.

The experimenter stood behind the target recording actual scores throughout the experiment.

*Subjects.*—Of the twenty-four boys, two in each group are not included in the experiment. Three were in the infirmary at the time. The fourth was rated by observers as an "isolate" in the Pa group. Since an isolate is, by definition, not a functional part of the group structure, variations in his judgments or those of other group members concerning his performance could not legitimately be

attributed to the group structure; rather any relationship between his judgments and the status of other members would have to be attributed to chance variations or possibly to personal factors. Therefore, he was eliminated from the analysis.

*Results.*—In order to test our hypothesis, the status rankings within each group were compared with judgments concerning the performance of the member occupying each position. Status rankings for the members of each group were obtained by averaging the independent ratings of participant observers (two for the Pa group and three for the Py group).

The status ratings made by the participant observers were in significant agreement for each group. For the Pa group a rho of .71 was obtained between the ratings of the two observers, which falls between the .02-.05 probability level when converted to *t*. To test the amount of agreement among the three participant observers for the Py group, the coefficient of concordance<sup>13</sup> was utilized, its value being .912, which is significant at less than the .001 level of confidence.<sup>14</sup>

A variation score was computed for each subject as a measure of the extent his performance was overestimated or underestimated by other members of his group. This variation score was obtained by finding the average difference between judgments of a subject's performance by all other members of his group and his actual performance on the twenty-five trials. It was then possible to rank the members of each group in terms of variation in judgment as well as in terms of status.

Table 1 presents the descending rank order for status and the corresponding rank order of judgment variation scores for the members of each group.

The rank-order correlations between sta-

<sup>13</sup> M. G. Kendall and B. Smith, "The Problem of *m* Rankings," *Annals of Mathematical Statistics*, IX (1938), 133-49.

<sup>14</sup> According to the formula given by P. O. Johnson, *Statistical Methods in Research* (New York: Prentice-Hall, Inc., 1949), p. 175.



tus and judgment variation scores for the two experimental groups are given in Table 2 along with *P* values derived from the conversion of rho to *t*.

As Table 2 indicates, a significant positive relationship was found between group status and judgment variation, that is, between an individual's relative standing in the group and the relative extent to which his performance was overestimated or underestimated by other members.

Since it is possible that this positive relationship between status and variations in

Table 3 gives the corresponding ranks in each group for status, variation in judgment, and mean performance scores (i.e., the average of scores actually made on twenty-five trials). As shown, a higher relationship exists between status and variation in judgment

TABLE 3  
CORRESPONDING RANKS IN STATUS, JUDGMENT VARIATION SCORES (J.V.S.), AND MEAN PERFORMANCE

PA GROUP				PY GROUP			
Subject	Status Rank	J.V.S. Rank	Performance Rank	Subject	Status Rank	J.V.S. Rank	Performance Rank
1...	1	3	9	A...	1	1	4
2...	2	1	2	B...	2	2	6
3...	3	6	8	C...	3	7	8
4...	4.5	4	5	D...	4	5	7
5...	4.5	2	6	E...	5	6	2.5
6...	6	7	4	F...	6	3	1
7...	7	9	1	G...	7	10	10
8...	9	5	3	H...	8	4	5
9...	9	8	10	I...	9	8	9
10...	9	10	7	J...	10	9	2.5

TABLE 1  
STATUS AND JUDGMENT VARIATION SCORES (J.V.S.) IN CORRESPONDING RANK ORDER

PA GROUP			PY GROUP		
Subject	Status Rank	J.V.S. Rank	Subject	Status Rank	J.V.S. Rank
1.....	1	3	A.....	1	1
2.....	2	1	B.....	2	2
3.....	3	6	C.....	3	7
4.....	4.5	4	D.....	4	5
5.....	4.5	2	E.....	5	6
6.....	6	7	F.....	6	3
7.....	7	9	G.....	7	10
8.....	9	5	H.....	8	4
9.....	9	8	I.....	9	8
10.....	9	10	J.....	10	9

TABLE 2  
CORRELATION BETWEEN STATUS AND JUDGMENT VARIATION IN PA AND PY GROUPS

Pa Group		Py Group	
Rho.....	.737	Rho.....	.676
<i>P</i> .....	<.02	<i>P</i> .....	<.05

judgment is traceable simply to a correspondingly high relationship between variations in judgment and actual skill, the subjects' ranks for mean scores actually achieved in the task were compared with their corresponding ranks in judgment variation and status in the group.

TABLE 4

CORRELATION BETWEEN RANK IN PERFORMANCE SCORE (SKILL) AND IN JUDGMENT VARIATION SCORES

Pa Group		Py Group	
Rho.....	.007	Rho.....	.45
<i>P</i> .....	.90-1.00	<i>P</i> .....	.10-.20

ment than between performance and variation in judgment for this particular task, especially in the Pa group.

Table 4 gives the correlations between actual performance (skill) and variation in judgment for the two groups. It shows that the rank-order correlation of .007 between performance level on this task and variation in judgment for the Pa group is clearly not significant (*P* = .90-1.00). While this rela-

tionship falls short of significance for the Py group, the  $P$  value ( $P = .10-.20$ ) may reveal a trend which warrants further investigation.

Further light is thrown on this finding by comparing the correlations between status and variation in judgment (Table 2) and those between performance level (skill) and variation in judgment (Table 4). For the Pa group the difference between the rank-order correlation for status and judgment variation (.737) and that between performance and judgment variation (.007) is .730, which

TABLE 5\*

"U" VALUES FOR DIFFERENCES BETWEEN VARIATION IN JUDGMENT OF UPPER AND LOWER HALVES OF THE TWO STATUS STRUCTURES

Pa Group		Py Group	
$U$ .....	1	$U$ .....	6
$P$ .....	.008	$P$ .....	.11

\* Since the  $U$  values obtained were too large to be evaluated with the tables offered by H. B. Mann and D. R. Whitney, "On a Test Whether One or Two Random Variables Is Stochastically Larger than the Other," *Annals of Mathematical Statistics*, XVIII (1947), 50-60, these values were reduced by a formula presented by L. E. Moses, "Non-parametric Statistics for Psychological Research," *Psychological Bulletin*, XLIX (1952), 122-43.

is significant at the .04 level of confidence. For the Py group the difference between the rank-order correlation for status and judgment variation (.676) and that for performance and judgment variation (.45) is .226, and this difference is not significant ( $P = .26$ ).

Theoretically, these differences in significance of the relationships between status and judgment of performance, on one hand, and between actual performance and judgment, on the other, can be accounted for in terms of differential structures of the Pa and Py groups. As noted, the observers agreed that greater stability of structure and greater solidarity had evolved in the Pa than in the Py group at this time. From a theoretical point of view, it would be expected

that, as group structure becomes better defined and more stabilized, as solidarity increases, the higher the relationship between status in the group and expectations of the individual's performance.

Then, as one increasingly identifies himself with the group, the greater the correspondence between his expectations and shared expectations regulated by the prevailing status hierarchy and the group norms. Further, greater stability of group structure would be reflected in the differences between expectations of the performance of members of higher and lower status, that is, greater stability of structure and greater solidarity are accompanied by higher expectations and thus higher estimates for the performance of higher-ranking members than of lower.

To test the foregoing theoretical considerations in the Pa and Py groups, the members of each were divided into upper and lower halves as to status structure, and the differences between means of variation in judgment for these halves determined.

For this comparison, a statistic appropriate for very small samples that are not normally distributed was needed. The Mann-Whitney  $U$ -Test was used for this purpose.<sup>15</sup> Table 5 gives the  $U$  values and  $P$  for differences between mean variation in judgment of the upper and lower halves of the status structure for each group.

The variations in judgments of those in the upper half of the status structure in the Pa group were significantly higher than those of members in the lower half ( $P = .008$ ). This difference for the Py group was not significant, although the  $U$  value for the difference approached significance ( $P = .11$ ). The comparison (Table 5) corroborates other findings concerning the relationship between stability of status structure and expectations, as revealed in judgments of performance on this experimental task.

Variations in judgment in the Pa group,

<sup>15</sup> H. B. Mann and D. R. Whitney, "On a Test Whether One or Two Random Variables Is Stochastically Larger than the Other," *Annals of Mathematical Statistics*, XVIII (1947), 50-60.

in which stability of structure and solidarity were relatively greater, were closely related to differential status positions and were not significantly related to actual skill. Variations in judgment in the Py group, whose status structure was less stable and where there was less solidarity, were also significantly related to differential status positions, though to a lesser extent. In addition, it seems that variations in judgment in the Py group were influenced to an extent relatively greater than in the Pa group by the skill actually demonstrated.

There is evidence that expectations about an individual member tend toward stability as group structure develops before the individual's expectations of his own performance begin to coincide with the group's. It may be that the degree of coincidence between one's own expectations of one's performance and what other group members expect of him is indicative of the degree to which group structure is stabilized. In this study, a rank-order correlation of .41 was found between judgments of own performance and judgments by other group members in the Pa group, which had the more stable structure. In the less stable and less unified Py group, a rho of .03 was found between judgments of one's own performance and judgments of it by others.

Among established groups of greater stability and solidarity than these experimentally formed groups, we would expect that closer relationships would be found both between status rank and judgments of own performance by the individual member and by other group members. The study on status relations in existing informal groups by Harvey indicates this, although the data of that study, being in terms of expected *future* performance, are not directly comparable to the present research.

It can be concluded on the basis of these findings that variations in judgment of per-

formance are significantly related to status in the group. The performance of members of high status was overestimated; the performance of members of low status was underestimated, the extent of over- or underestimation being positively related to status rankings. Thus our hypothesis was substantiated.

Variations in judgments of performance were not significantly related to actual skill in this task. This is not to be interpreted to mean that skill is irrelevant or that significant positive correlations between variation in judgment and skill would not be found in other tasks with more objective anchorages for judgment.

The finding of note is the differential relationship between variation in judgment and status rankings in the two groups as a function of differential stability of group structure and group solidarity. There is evidence that this relationship is closer in groups of greater stability and that actual skill is given *relatively* greater weight in the judgments of the group with less structural stability. Therefore, in the experimental study of group formation and status relations (including leader-follower relations), attention should be paid to the degree of stability of group structure evolved at the time, since it influences the behavior of members toward one another.

The present unit of our research program was concerned with experimental production of groups and with the development of indexes of status. The technique, already applied to assessment of status relations in existing groups, is being extended to the study of relations between groups. The final step in our research is an experiment in which groups themselves are experimentally produced and indexes devised for assessing relationships within the group and relations between groups.

UNIVERSITY OF OKLAHOMA