1970

The Effects of Group Interaction on the Ratings of Physical Attractiveness

Anthony A. Kopera
Loyola University Chicago

THE EFFECTS OF GROUP INTERACTION ON THE RATINGS OF PHYSICAL ATTRACTIVENESS

by

Anthony A. Kopera

A Thesis Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Master of Arts

June, 1970

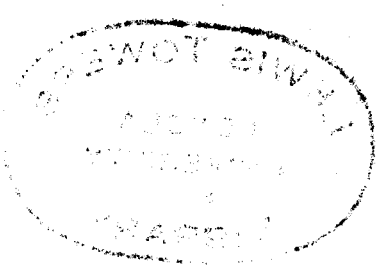

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Recent research has shown that physical attractiveness is an important variable in impression formation. Byrne, London, and Reeves (1968), in studies of initial impressions of strangers, found that interpersonal attraction was greater towards physically attractive strangers, regardless of sex, than towards unattractive ones. In addition they found that if attitudes either similar to or dissimilar from those of the Ss were attributed to the strangers, physical attractiveness still influenced the attraction response.

Walster, Aronson, Abrahams, and Rottman, (1966), conducted a study of dating behavior. Under the guise of getting information for a "computer" dance, the investigators obtained various personality and intellectual measures of their Ss. They also rated them on physical attractiveness. The Ss were then randomly paired. Walster et al., concluded, on the basis of their findings, that "The only important determinant of S's liking for his date was the date's physical attractiveness (p. 508)." This was found to be true for both sexes, even up to six months after the "computer" dance.

These experiments serve to point out that physical attractiveness

is a highly valued attribute. Research on decision making has shown that risk taking is also highly valued, and that such a value demonstrated itself in the "risky shift" phenomenon (e.g., Teger and Pruitt, 1967; Wallach and Wing, 1969). In this risk taking research, individuals are asked to indicate how great a risk they are willing to take in various hypothetical situations. Those same individuals are then randomly grouped, and the groups are presented with the same series of situations. It has been found that individuals working in groups indicate a willingness to take greater risks than when working alone (Wallach, Kogan, and Bem, 1962).

In calling risk taking a valued behavior, Brown (1965), says that being risky is seen as good. Individuals tend to see themselves as being more willing to take risks than the average person. However, in group problem solving situations, they may discover that they are more conservative than they thought they were. In line with the value placed on risk taking, they then become more willing to take risks (i.e., adopt a less conservative position) then they were while working alone.

The present study was conceived as an attempt to discover whether something analogous to the risky shift phenomenon can be demonstrated in the task of rating the physical attractiveness of female strangers. It was thought possible that the effect of working in groups would cause the mean ratings of physical attractiveness to shift in a risky direction in some instances, and a conservative direction in other instances. In this case "risk" is considered a higher rating of physical attractiveness,

"conservatism" a more moderate estimate of the stranger's attractiveness.

It was further considered possible that effects of the group would vary depending on the sex of the group members. Consequently, the use of three types of groups was decided upon: all male, all female, and half male, half female groups.

It was also considered possible that the group reaction would depend on whether the stranger being rated actually was attractive or unattractive. To check this possibility, it was decided that the Ss should be asked to rate strangers of high and low attractiveness.

Finally, it was reasoned that the risky shift phenomenon is but a special instance of the "approval motive", i.e., the motive to gain the approval of others. Marlowe and Gergen (1969), conceptualize the situation in which this motive operates by saying that "the barometer by which we assess our competence and self-worth is largely a result of the reflected appraisals of others (p. 600)." In all male groups with the approval motive operating, the individuals may feel constrained to demonstrate to each other that they have high standards, thus demonstrating their individual good taste. Thus the hypothesis for all male groups was that conservatism in estimates of physical attractiveness of females would be a valued behavior, and specifically that the shift in group consensus ratings relative to individual ratings would be towards the unattractive end of the scale.

The approval motive should also operate in all female groups. The

hypothesis was that female groups would tend to show a regression towards the average rating. The reasoning for this hypothesis was that individual group members would gain approval by not admitting the existence of either extreme--highly attractive or highly unattractive--while working in a group which may be comparing them both to the strangers being rated and the other group members.

Though a male-female group was used to study possible sex interaction effects, no hypothesis was formulated.

CHAPTER II

METHOD

In studies of risk taking Ss first go over a choice dilemma questionaire indicating the degree of risk they are willing to take. Groups are then formed which must reach a consensus opinion on the problems (Wallach et al., 1962). A similar paradigm was used in this study, i.e., individuals first rated a series of pictures on attractiveness, and groups were then randomly formed which gave a consensus rating on the same pictures.

In order to be able to control the level of attractiveness and to check on possible sex differences, data were first gathered on a large group of pictures. Facial photographs of eighty-four female seniors of an Eastern college, Caucasian and without eyeglasses, were projected on a screen for ten seconds and then rated for attractiveness on a one to seven scale, seven being very attractive, one being very unattractive, and four being average. Pictures of Caucasians without glasses were used to control for those factors which clearly might influence ratings. The Ss were one hundred and eight students from an introductory psychology course, fifty-three females and fifty-five males. Only Caucasian Ss were used to again avoid the confound of possible racial differences. The Ss were also single, between the ages of eighteen and twenty-two, and non-

clergy. The aim was to make the raters as similar to the people in the pictures as possible. An analysis of variance with repeated measure across pictures was performed which showed no significant difference between the mean ratings made by males and females, but a difference at the .001 level was found across pictures. This indicates that males and females tend to agree on the ratings they give, and that different pictures get consistently different ratings. Further a Pearson r correlation coefficient of .93 was found for male and female ratings, which indicated substantial agreement between males and females on their ratings.

The eighty-four pictures were rank ordered on the basis of mean ratings of males and females combined. Since the difference between male and female ratings did not reach significance, it was thought that combining their ratings would give a more reliable estimate of the rated attractiveness of the pictures used.

Two groups of eight pictures were selected for use in this study. It was thought that the mean of each group should be approximately one and one-half scale units away from the "average" point. This was done to allow for potentially equal shifts to the mean and to the extremes of the scale in either case. With a seven point scale this meant the means for the high group should be approximately 5.5, and for the low group, 2.5. Because the original set of pictures apparently did not contain pictures of exceptionally attractive girls, the highest mean rating possible was 5.3. It was therefore decided that the mean of the low group

of pictures should be approximately 2.7, or 1.3 scale units below the "average" and to make the high and low groups equivalent in distance from the "average" point of 4.0. To avoid the possibility of a contrast effect, eight pictures in the "average" range (mean equalling approximately 4.0), were also included. The ranges of the three groups were approximately equal, as were the variances (see Table 1).

Table 1

Means and Variances of the Three Groups of Pictures Used

	Picture	Mean Rating	
High			
	1	5.87	
	2	5.62	$\bar{X}_H$ = 5.30
	3	5.41	
	4	5.27	σ^2_H2 = .12
	5	5.25	
	6	5.07	Range = .92
	7	4.98	
	8	4.95	
Middle	9	4.23	
	10	4.11	$\bar{X}_M$ = 3.99
	11	4.11	
	12	4.06	σ^2_M2 = .04
	13	3.88	
	14	3.86	Range = .39
	15	3.85	
	16	3.84	
Low	17	3.06	
	18	2.98	$\bar{X}_L$ = 2.70
	19	2.92	
	20	2.84	σ^2_L2 = .09
	21	2.67	
	22	2.64	Range = .84
	23	2.28	
	24	2.22	

The twenty-four pictures thus selected were arranged in a random order and were presented for ten seconds to the Ss, who were comparable to the Ss used in originally selecting the pictures. The method of presentation was the same as that originally used, i.e., the pictures were projected on a screen. Each S gave an individual rating of each picture without knowledge of the ratings of other raters. Groups of six were then randomly formed, either all male, all female, or made up of three male and three female Ss. Six rather than four were used because Teger and Pruitt (1967) found that larger groups maximized the shift phenomenon. Each group was then presented with the same twenty-four pictures, in the same order, and for the same length of time, and asked to reach a consensus in their ratings within two minutes. Thirty Ss were run in each experimental condition, allowing five groups of six to be formed. A change score was computed based on the difference between the mean ratings of the six individuals in each group and the consensus rating those six people gave.

Thirty males and thirty females were run as controls, with fifteen randomly chosen from each to form a male-female control group. The control condition consisted of individuals who rated the pictures independently once before groups of six were formed, and then rated them independently a second time. The grouping in the control condition was simply for statistical purposes, and no discussion was allowed in the control condition. A change score was computed based on the difference between the mean rating of the six people in each group the first time, and the mean

rating of the six people the second time. Thus each cell in the statistical analysis had an N of five.

The resulting data were analyzed using a three by three by two analysis of variance design, with three levels of attractiveness--high, middle, and low, three levels of sex group--male, female, and male-female, and the experimental and control conditions, with repeated measures across levels of attractiveness.

CHAPTER III

RESULTS

The Ss appeared to find the task of rating the pictures very easy, both as individuals and in groups. Though they were allowed up to two minutes to arrive at a group consensus opinion, no group took longer than a minute and a half, with the majority of the groups finishing within one minute most of the time.

Before any change score in the experimental conditions can be considered significant, it must be determined that no significant change occurred in the control conditions. Therefore, a three by three by two analysis of variance (three levels of pictures, high, middle, and low; three levels of sex of group; and the first versus the second ratings), with repeated measures across levels of pictures was performed as a check to make sure no change had occurred. The results are shown in Table 2. It can be seen that, as expected, there is a significant difference across the levels of pictures ($p < .001$), but that there is no significant difference between the first and the second ratings of the control groups ($F < 1$). There was, however, a significant interaction between levels of pictures (high, middle, low), and the first and second times the control Ss rated the pictures.

Table 3 shows the mean ratings for high, middle, and low level pictures the first and second times the pictures were rated by the control groups. It can be seen that while no change occurred in the highest rated pictures, the middle and low groups tend to be rated lower by the control Ss the second time.

Table 2

Analysis of Variance: Control Groups

First Rating Versus Second Rating

Source	SS	df	MS	F	p
Total	85.06	89	—	—	—
Between	5.74	29	—	—	—
Sex of Group	.33	2	.17	<1	ns
1/2	.19	1	.19	<1	ns
Sex x 1/2	.01	2	.01	<1	ns
$error_b$	5.21	24	.22	—	—
Within	79.32	60	—	—	—
HML	78.65	2	39.33	3,933	.001
HML x Sex	.06	4	.015	1.5	ns
HML x 1/2	.17	2	.085	8.5	.001
HML x sex x 1/2	.01	4	.002	<1	ns
$error_w$	.43	48	.01	—	—

Table 3

Mean Ratings for High, Middle and Low Pictures for Control Groups First and Second Times

	I	II
H	5.21	5.21
M	4.07	3.99
L	3.03	2.83

A second analysis of variance, similar to the first in design, was performed comparing the first ratings of the experimental and control groups. This was done to find out if there was initial agreement between the groups on how the pictures were to be rated (see Table 4). Results show no significant difference between experimental and control groups on initial ratings. There was, again as expected, a difference across levels of attractiveness ($p < .001$). A significant sex of group effect was also found ($p < .05$).

Table 5 shows that the mean of the initial ratings of the all male groups, experimental and control combined, is lower than the mean of the initial ratings of the all female groups, experimental and control combined. The mean for the all female groups, in turn, is lower than the mean of the initial ratings of the male-female groups, experimental and control combined. A Duncan Multiple Range Test shows no significant difference between the mean of the male groups and the mean of the female groups, or between the mean of the female groups and the mean of the male-female groups. The difference between the mean of the male groups and the mean of the male-female groups is significant at the .001 level.

Change scores were computed by subtracting each group's score on the first rating at each level of attractiveness from that group's score on the second rating. Thus if the change was in a negative direction, the direction of change was reflected in the change score. Table 6 shows the

mean change score for each sex group, experimental and control, for each level of attractiveness. It can be seen that all but two changes are negative, that one is positive in direction (control group, all male, high pictures), and one showed no change at all (control group, male-female, high pictures). This, of course, contradicts the hypothesis for female groups on low pictures, which was that the consensus change would be in the positive direction. Note also that the changes in the experimental groups are larger than the changes in the control groups.

Table 4

Analysis of Variance: First Ratings

Experimental Versus Control Groups

Source	SS	df	MS	F	p
Total	80.24	89	—	—	—
Between	4.43	29	—	—	—
Sex	.96	2	.48	3.68	.05
E/C	.02	1	.02	< 1	ns
Sex x E/C	.14	2	.07	< 1	ns
$error_b$	3.31	24	.13	—	—
Within	76.93	60	—	—	—
HML	74.75	2	37.38	934.50	.001
HML x Sex	.06	4	.02	< 1	ns
HML x E/C	.05	2	.03	< 1	ns
HML x Sex x E/C	.10	4	.03	< 1	ns
$error_w$	1.97	48	.04	—	—

Table 5

Mean Ratings by Sex of Group of

Experimental and Control Groups Combined

First Rating Only

M	F	MF
3.96	4.10	4.22

Table 6

Mean Change Score for Each

Group at Each Level of Pictures

	M		F		MF	
	E	C	E	C	E	C
H	-.13	.05	-.28	-.02	-.02	0.0
M	-.46	-.07	-.38	-.06	-.22	- .12
L	-.64	-.19	-.42	-.15	-.39	- .25

A third analysis of variance was performed, this time on change scores, experimental versus control, using the same design used in the first two analyses done. The results are shown in Table 7. Most important is the finding that there is a significant difference between the experimental and the control groups. However, there again is a level of picture difference, indicating that there is a greater change at at least one of the levels of pictures than at the others. In order to further investigate this finding, three two by three analyses of variance were performed, one at each level of pictures with the three levels of sex of group and experimental and control as the factors. Tables 8, 9, and 10 give the results of these analyses. As can be seen, the difference in the change scores of the experimental versus control groups is significant only in the low, and middle level pictures. The difference in change scores is not significant at the high level of attractiveness pictures.

A series of nine t tests was then performed in order to more closely examine the results. Table 11 shows the t ratios and the significance levels they reach for each possible experimental versus control comparison--three levels of pictures, and three types of groups, male, female, and male-female. All significant differences are for changes in the negative direction.

It can be seen that in the all male and all female conditions the difference in change is significant even using two tailed tests. For male-female groups the difference in change is non-significant for any level of

picture, though if a one tailed test is used the difference approaches significance ($p < .10$) for middle and low pictures. It should be noted that the difference in the male high condition is between -.13 and .05. Perhaps if the control condition had not raised its rating the second time the difference would not be significant.

Table 7

Analysis of Variance: Change Scores

Experimental Versus Control

Source	SS	df	MS	F	p
Total	10.61	89	—	—	—
Between	6.69	29	—	—	—
Sex	.07	2	.03	<1	ns
E/C	1.24	1	1.24	5.90	.025
E/C x sex	.29	2	.15	<1	ns
$error_b$	5.09	24	.21	—	—
Within	3.92	60	—	—	—
HML	1.14	2	.57	11.4	.005
HML x sex	.18	4	.045	<1	ns
HML x E/C	.08	2	.04	<1	ns
HML x sex x E/C	.04	4	.01	<1	ns
$error_w$	2.48	48	.05	—	—

Table 8

Analysis of Variance: Experimental Control by Sex of Group

Low Pictures

Source	SS	df	MS	F	p
Total	4.17	29	—	—	—
E/C	.60	1	.60	4.29	.05
Sex	.10	2	.05	<1	ns
E/C x sex	.12	2	.06	<1	ns
error	3.35	24	.14	—	—

Table 9

Analysis of Variance: Experimental Control by Sex of Group

Middle Pictures

Source	SS	df	MS	F	p
Total	3.16	29	—	—	—
E/C	.55	1	.55	5.5	.05
Sex	.04	2	.02	<1	ns
E/C x sex	.12	2	.06	<1	ns
error	2.45	24	—	—	—

Table 10

Analysis of Variance: Experimental Control by Sex of Group

High Pictures

Source	SS	df	MS	F	p
Total	2.17	29	—	—	—
E/C	.18	1	.18	2.57	ns
Sex	.12	2	.06	<1	ns
E/C x Sex	.07	2	.03	<1	ns
error	1.80	24	.07	—	—

Table 11

t* Tests Experimental Versus Control

	t	p				direction of change
Male						
High	2.72	.025	two tailed			-
Middle	5.65	.001				-
Low	7.71	.001				-
Female						
High	5.30	.001	two tailed			-
Middle	5.33	.001				-
Low	3.46	.005				-
Male-Female						
High	0.3	ns	two tailed	ns	two tailed	o
Middle	1.58	ns		.10		o
Low	1.63	ns		.10		o

*df=8 for all comparisons

CHAPTER IV

DISCUSSION

The most important finding of this study is the significant difference in the change scores between the experimental and control groups. This finding indicates that group interaction does have an effect on the task of rating physical attractiveness, and more, indicates a generality of effect one could not assume based solely on studies of risk taking behavior. Specifically, it can be said that people in groups appear to do what they think will make them look good in the eyes of the group, whether this means being critical of someone's physical attributes or demonstrating a willingness to take a risk.

While the difference in change scores for the experimental and control groups is significant, it is not always in the predicted direction. The hypothesis for all male groups was that the change would be in the negative direction, i.e., that the ratings would be lower in groups. This hypothesis was supported. The hypothesis for all female groups was that the females in groups would tend to give all pictures an average rating, i.e., lower ratings on high pictures and higher ratings on low pictures. The results show that females, as well as males, tend to lower their ratings of physical attractiveness in groups. This consistent lowering of ratings in all male and all female groups indicates that the effect of group interaction on

rating physical attractiveness is the same for both sexes. It may be, then, that females, as well as males, enhance their self esteem by being critical of others in groups.

Though no hypothesis was formulated for the male-female groups, the results are of major interest. In male-female groups no significant change occurred, though for middle and low pictures, the change approaches significance. The change in male-female groups for the high pictures is so slight in fact (-0.02), that it makes the over all effect for the high pictures across sex groups non-significant.

The general effect--non-significant change in male-female groups--may be due to inhibiting factors in mixed sex groups. That is, members of one sex may be more reluctant to express their criticality of physical attractiveness in the presence of members of the opposite sex. The fact that the high pictures changed so little merits further discussion. Two other findings help give a plausible explanation for it. First, in the control groups the second ratings for middle and low pictures were lower than the first ratings, while the mean ratings for the high pictures remained the same. Second, the high pictures were most resistant to change across all sex groups. What may have happened is that the Ss took a short while to establish the criteria they used to judge the pictures. When they did establish their criteria, these criteria may have been the high pictures. Using the high pictures as anchors or reference points may have lead to a lowering of ratings on the middle and low pictures by contrast in the

control groups, and smaller changes on high pictures than on middle and low pictures for experimental groups. In the case of male-female groups, where the change for the high, middle and low pictures is non-significant, the anchoring effect and the mixed-sex groups effect combine to keep the change for high pictures minimal.

One further finding bears discussion. There was a significant difference observed among the initial ratings, all levels of pictures combined, across the three sex groups. Since no group interaction had taken place prior to the initial ratings, these ratings should have been approximately equal. In fact, since the pictures were selected after a standardization study, the mean rating for all pictures combined should have been approximately equal to four. However, Ss were allowed to form groups with whomever they chose. Perhaps this means of forming groups allowed subtle nonrandom selection factors to operate, thus leading to the finding of significance.

To summarize the main result: group interaction in either all male or all female groups leads to lower ratings of physical attractiveness than those ratings obtained when Ss rated the stimulus pictures independently. No significant change was noted in the male-female groups, however.

The results of this study suggest some further research. If the high pictures do act as anchors, then perhaps they are better remembered than middle or low pictures. Also the change in the ratings of control Ss suggests that studies of contrast effects in the ratings of physical

attractiveness might yield interesting results, e.g., not using a middle group of pictures could cause the ratings of high and low groups of pictures to be exaggerated towards the extremes of the scale. That male-female interactions in group decision making situations should be studied is certainly suggested by the results of this study. Finally, a major question is what makes some people more critical than others of the attractiveness of strangers, whether they do the rating alone or in groups.

REFERENCES

Brown, R. Social psychology. New York: Free Press, 1965.

Byrne, D., London, O., & Reeves, K. The effects of physical attractiveness, sex, and attitude similarity on interpersonal attraction. Journal of personality, 1968, 36, 259-271.

Marlowe, D., and Gergen, K. J. Personality and social interaction. In G. Lindzey and E. Aronson (Eds.), The handbook of social psychology. (2nd ed.). Vol. 3. Reading, Massachusetts: Addison-Wesley, 1969.

Teger, A. I., and Pruitt, D. G. Components of risk taking. Journal of experimental social psychology, 1967, 3, 189-205.

Wallach, M. A., & Kogan, N. The roles of information, discussion, and consensus in group risk taking. Journal of experimental social psychology, 1967, 1, 1-19.

Wallach, M. A., Kogan, N., & Bem, D. J. Group influence on individual risk taking. Journal of abnormal social psychology, 1962, 65, 75-86.

Wallach, M. A., Kogan, N., & Burt, R. B. Can groupmembers recognize the effects of group discussion upon risk taking? Journal of experimental social psychology, 1965, 1, 379-395.

Wallach, M. A., Kogan, N., & Burt, R. B. Are risk takers more persuasive than conservatives in group discussion? Journal of experimental social psychology, 1968, 4, 76-88.

Wallach, M. A., & Wing, C. W. Is risk a value? Journal of personality social psychology, 1969, 9, 101-106.

Walster, E., Aronson, V., Abrahams, D., & Rottman, L. Importance of physical attractiveness in dating behavior. Journal of personality social psychology, 1966, 4, 508-516.

APPROVAL SHEET

The Thesis submitted by Anthony A. Kopera has been read and approved by members of the Department of Psychology.

The final copies have been examined by the director of the Thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the Thesis is now given final approval with reference to content and form.

The Thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

June 2, 1970
Date

Richard A Maier
Signature of Advisor