This is to certify that the

thesis entitled

PEER RATINGS OF MENTAL HEALTH PROFESSIONALS:
SEX BIAS OR REALITY

presented by

Kathleen Sue Long

has been accepted towards fulfillment
of the requirements for

M.A. degree in Psychology

Major professor
John B. Hurley

Date 7/19/83

# PEER RATINGS OF MENTAL HEALTH PROFESSIONALS: SEX BIAS OR REALITY.

By

Kathleen Sue Long

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1983

ABSTRACT

PEER RATINGS OF MENTAL HEALTH PROFESSIONALS:
SEX BIAS OR REALITY

By

Kathleen Sue Long

Experienced mental health professional's perceptions were explored among 97 men and 97 women participants in 1974 and 1977 Annual Institutes of the American Group Psychotherapy Association. Following 11½ hours of psychodynamic group participation, all members rated each other's within-group behavior for Acceptance versus Rejection of Self (ARS) and Others (ARO). Calculated from variance analysis of two independent samples, joint probabilities showed main effects for sex of rater, sex of ratee, group, and two interactions. Most groups showed sharply polarized ratings of the sexes (G x R $p < .0001$). Women gave higher ratings (ARS $p < .01$; ARO $p < .05$), supporting other research that women therapists are slightly more liberal and accepting of non-stereotypic sex role behaviors and attitudes. Women received higher ratings than men (ARS $p < .05$; ARO $p < .09$). A pro-successful-woman phenomenon was postulated; perhaps these women were more motivated and had greater coping skills and abilities than their modal male peer. Further research is needed to clarify why women were rated higher, and rated others higher than men.

ACKNOWLEDGEMENTS

I would like to express my thanks and appreciation to the members of my thesis committee for the time and suggestions and support they gave me:

To Dr. John Hurley, who served as my chairperson, my thanks for his time and help with revisions of this research.

To Dr. Elaine Donelson, who lent an empathetic ear when I needed it, and whose confidence in me helped me get through some difficult times, my thanks also for her intellectual challenges which stimulated me to try to find better solutions, and for the determination and hope that her humane intelligence has inspired in me.

To Herr Professor Bertram Karon, who gave me support and encouragement, my thanks for teaching me the ways of knighthood, and for helping me maintain my courage to meet and slay the dragon.

TABLE OF CONTENTS

LIST OF TABLES

# INTRODUCTION

## Purpose of Study

The purpose of this study is to investigate sex biases among mental health professionals through the use of peer ratings. It focuses on whether or not women are biased toward women, and it also explores how men rate other men. In addition, the study examines how both men and women rate members of their own sex in comparison with their ratings of the opposite sex.

## Theoretical and Clinical Postulates

Several psychologists (Adler, 1954; Allport, 1958; Horney, 1967; Thompson, 1964) have observed and theorized about prejudice against women in Western culture. They suggested that not only were men prejudiced against women, but also that women were prejudiced against women. Allport (1958) noted that members of several "out" groups (e.g., women, Jews, Blacks) often perpetuated the biases of dominant societal groups against themselves. Called identification with the aggressor, this phenomenon seems especially likely when the "in" group is powerful with respect to the "out" group's welfare. Adler, Horney, and Thompson as a result of such an identification with men by women, separately observed that many women believed that women, and their traditional roles, values, and achievements, were inferior to those of males. The internalization of male values by women also facilitated an acceptance of women's inferior status in society and

the devaluation of women's achievements. Horney stated that women had adapted to the male view of themselves to such a degree that they viewed it as their true nature. She wrote:

> "If we are clear about the extent to which all our being, thinking, and doing conform to these masculine standards, we can also see how difficult it is for the individual man and also for the individual woman really to shake off this mode of thought" (Horney, 1967, p. 57).

As a result, women were often seen as unhappy with their feminine roles, and if they wanted to achieve outside of these traditional roles, they adopted "male" values and behaviors. Adler similarly described the process and its effects on the self-esteem of women:

> "A girl comes into the world with a prejudice sounding in her ears which is designed to rob her of her belief in her own value, to shatter her self-confidence and destroy her hope of ever doing anything worthwhile."

He added:

> "The obvious advantages of being a man (lead to) an almost universal dissatisfaction with the feminine role" (Adler, 1954, pp. 110-111).

Thus, according to Allport's theory of prejudice and the clinical observations of Adler, Horney, and others, it would seem that women as well as men are often perpetrators of prejudicial reactions to women, and of prejudicial evaluations of their achievements. Still, although such theories and observations are cogent to some, others argue that if such conceptions once were valid, they no longer are.

In 1968, Philip Goldberg ignited a renewed interest in the topic of women's bias against women. Goldberg's study revealed that even today's modern women (i.e., college coeds) were still prejudiced against women. His original experiment constituted some of the first

research to emerge in recent times on the subject. In his somewhat dramatically written article, Goldberg began with a quote from Aristotle. It read:

> "Woman," advised Aristotle, "may be said to be inferior to man" (Goldberg, 1968, p. 28).

Goldberg then concluded his report by stating:

> "Whatever lip service these girls pay to modern ideas of equality between men and women, their beliefs are staunchly traditional. Their real coach in the battle of the sexes is not Simone de Beauvoir or Betty Friedan. There coach is Aristotle" (1968, p. 30).

Though Goldberg's findings have often been cited as evidence of bias by women against other women, studies designed to replicate and/or expand upon his original work have yielded mixed results. Thus, a closer examination of this literature is necessary before his conclusions can be accepted, refuted, or modified.

## REVIEW OF LITERATURE

### Sex Bias in Performance and Achievement Ratings

Goldberg's 1968 study has often served as the model for later research. In his experiment, forty undergraduate women rated articles from six professional fields to see if, by changing only the apparent sex of the author, the articles would be evaluated differently. Preliminary work with 100 other coeds had established that law and city planning were commonly viewed as "masculine" fields, that elementary teaching and dietetics were similarly regarded as "feminine" fields, and that linguistics and art history seemed relatively "neutral". The participants were given one of two versions of a booklet of six articles, one from each of the aforementioned fields. Thus, for example, the article on law was authored by "Joan T. McKay" half of the time, and by "John T. McKay" the other half. Goldberg found that the undergraduate women revealed the most adverse bias against female-authored articles in the masculine fields, and that they also rated the male-authored articles in the feminine fields more favorably than those female-authored. In addition, male-authored articles were more valued overall, and male authors were seen as more competent than their female counterparts. Goldberg thus claimed to have shown that women in the nineteen-sixties were prejudiced against women.

In an attempt to verify Goldberg's findings, Pheterson (1969, cited in Pheterson, Kiesler, & Goldberg, 1971) followed Goldberg's procedures using a different population ("middle-aged, uneducated

women"), and selected professional articles on marriage, child discipline, and special education. In effect, Goldberg's findings were not confirmed. No significant differences were found between male-authored and female-authored articles. Pheterson did not, however, view this outcome as a refutation of Goldberg's (1968) findings. She suggested that an additional factor was present, namely "success", which had contributed to the null findings. She thought that the "uneducated" women had considered the authoring of any article a sign of success, and Pheterson further postulated that women should not be biased against other women who had achieved outward success and recognition. Pheterson (1969) never defined in the study what constituted "uneducated" women, and in addition to the presumed success factor, it is noted that her selection of target articles were all from traditionally feminine fields, thus perhaps further confounding the results.

Pheterson, Kiesler, and Goldberg (1971) then conducted a second study in which they hypothesized that women who were attempting to achieve would be rated lower than men who were attempting to achieve, while women who had achieved success would be rated equal to or higher than men who had achieved the same success. The format was also shifted from the evaluation of professional articles to the evaluation of paintings. Undergraduate women were asked to rate eight modern paintings on the technical competence of the artist, the artistic future of the artist, the creativity of the artist, the quality of the painting, and its emotional impact. They were also asked to read brief biographical sketches of the artists in conjunction with viewing slides of the paintings. These brief biographies were

identical except for manipulations for the sex of artists, status of the painting (seen as tantamount to the status of the artist), and the personal odds faced by the artist. The status condition was accomplished by having paintings which were "winners" of a contest (thus denoting a successful and recognized artist), or paintings which were "entries" in contests (thus denoting attempts to achieve success and recognition by the artist). Overall, the 120 undergraduate women rated "males" significantly higher than "females" on technical competence and artistic future. When the ratings were broken down by status of painting/artist condition, there were no significant differences between the ratings given to "males" and "females" in the "winner" category. However, "male" artists were rated significantly higher than "female" artists in the "entry" condition. Thus, Pheterson et al. (1971) concluded that two hypotheses had been supported: (a) women will be biased against women who are attempting to succeed, and (b) women will not be biased against women who have already achieved success and recognition.

Three subsequent studies have been unable to replicate Goldberg's 1968 findings. Soto and Cole (1975) conducted two separate studies, using college undergraduates in the first and high school students in the second. The subjects ranked four applicants for a business manager's position by reading their letters of recommendation. These subjects had previously ranked a list of ten attributes of a successful business manager. The attributes were then varied in the letters so that they would have been the determining factors in the difference in rankings of the candidates. Were this not the case, the sex of the candidate would be the determining factor for that

subject. In both studies, the subjects' first choices were based on qualifications rather than upon the sex of the applicant. This was interpreted by Soto and Cole (1975) as evidence against sex bias.

Another study (Chobot, Goldberg, Abramson, and Abramson, 1974) followed Goldberg's procedures, but used articles from only four of the professional fields, art history, education, dietetics, and linguistics, those previously viewed as feminine or neutral. The subjects were again undergraduate female and male students. In this study, the articles written by women were not rated less competent than the articles written by men. The results appeared to refute Goldberg's original findings. The conclusion by Chobot et al. (1974) suggested that perhaps a historical phenomenon had occurred, namely that awareness of feminist issues and more current feminist gains had accounted for the change. The Chobot et al. (1974) research was conducted five to six years after that of Goldberg.

Levenson, Burford, Bonno, and Davis (1975) also conducted two separate studies using both male and female subjects. In the first study they attempted to replicate Goldberg's work using professional articles from the same fields. They asked the subjects to rate the articles on the author's writing style, professional competence, persuasive ability, overall ability, and the value of the specific article. The "author's" sex was not found to be an important variable, nor were any significant differences found between the evaluations of the female and male subjects. In the second study, 33 female and 112 male undergraduates evaluated a student's essay answer to a political science quiz. The subjects read the same "average" essay and graded it from "A" to "F". Half of the subjects

read "Joan" Thompson's answer while the other half read "John" Thompson's answer. Overall, no significant differences were found for sex of author or sex of subject. However, when the evaluations by the 33 female subjects were examined separately, they were found to have given "Joan" significantly higher grades than "John".

At this juncture it is clear that the previously reviewed research of Soto and Cole (1975), Chobot et al. (1974), and Levenson et al. (1975) all refuted Goldberg's findings in some way. Levenson et al. (1975) in fact found evidence even suggestive of a positive bias toward women by other women. Further, a study by Panek, Deitchman, Burkholder, Speroff, and Haude (1976) challenged the Pheterson et al. (1971) study which had shown that women were biased against women who were attempting to achieve success, but not against those who had achieved it. Panek et al. (1976) followed the original Goldberg format with an additional manipulation for level of accomplishment (i.e., success). The success manipulation was accomplished by stating that the article had been published in a professional journal (thus denoting success), or that the article was submitted to a journal for publication (thus denoting attempting-to-succeed). The subjects were 57 female and 39 male undergraduates from a private university. Manipulations were sex of author, sex of subject, level of accomplishment, and the sex-appropriateness of the field. No overall effects were found for sex of subject, sex of author, or level of accomplishment. There was a main effect for the sex-appropriateness of the field. However, upon examination of the ratings assigned to "female" authors alone, Panek et al. (1976) found there was a significant difference between females who had "accomplished" and those who were "attempting-to-

accomplish". Both female and male subjects rated the women who were "attempting-to-accomplish" significantly higher than the women who had "accomplished", the antithesis of the Pheterson et al. (1971) finding.

Further research on the bias by women against women had revealed a complex pattern of prejudice which is at one at odds with the Panek et al. (1976) finding, and which apparently supports the earlier view of Pheterson et al. (1971). In an interesting study by Peck (1978), to test interactions among sex of achievers, sex of subject, and status of achiever, the following four hypotheses were generated:

(1) Both sexes would rate the professional ability of the women in the high status condition significantly higher than the women in the low status condition.

(2) Both sexes would rate the professional ability of the men in the low status condition significantly higher than the women in the low status condition.

(3) Only female subjects would rate high status women significantly higher than the high status men.

(4) Only female subjects would rate low status women significantly lower than low status men.

The subjects, 71 female and 65 male undergraduate education majors, rated two articles on the teaching of reading. Peck had chosen the teaching-of-reading field in an attempt to control for the sex-appropriateness of the profession. Her rationale was that "Instruction of reading was chosen...because teaching and instruction were shown by Feather (1975) to be neutral in sex-appropriateness" (Peck, 1978, p. 207). The two articles and biographical sketches of the authors were given to

the students. The authors of the high status (success) condition were given associate professor titles, with a long accompanying list of publication and research achievements. Lower status (attempting-to-succeed) authors were described as doctoral candidates in education who had written the paper for a term project. Peck found that overall, the subjects rated the high status author significantly above the low status author. Her hypotheses were also supported by several significant interactions. Both sexes rated high status women significantly higher than low status women, and they also rated low status women significantly lower than low status men. In addition, the female students rated high status women significantly more favorably than did the male students. The female students also rated the low status women significantly lower than did males. This positive bias (i.e., prejudice) of the female subjects for the high status, "successful" woman was labeled "overvaluing" by Peck. She noted that Feldman-Summers and Kiesler (1974) had earlier reported that professional status is perceived to result from internal control and stability. Peck then speculated that the subjects may have been attributing greater and stronger internal motivation to the successful women for their being able to manage both their professional demands and the more traditional demands of the feminine role (e.g., marriage, pregnancy, child-rearing, homemaking). Since successful women would then be viewed as having successfully reconciled the two seemingly competing roles, and as having overcome the additional obstacles of being female in the work world, they would be seen as more competent and more internally motivated than their male counterparts. Peck used the converse rationale to explain why the female subjects had

devalued the low status women. Since the low status women had not yet reached a higher professional status, they had not yet shown that they possessed the necessary internal control and motivation to succeed. If the low status women were viewed as having less internal control, they would then also be seen as more vulnerable to the external demands of the traditional feminine role, and therefore eventually less likely to succeed than their male counterparts. Thus, Peck concluded that women's "attempting-to-achieve" efforts may well be devalued by other women, hence revealing a real bias against women by other women. Once women are actually successful, however, this bias may be reversed.

In a format different from those studies cited above, Spence and Helmreich (1972) investigated how 343 women and 275 men (undergraduates) viewed women in four different roles. The four roles depicted an upperclass women interviewing for a job as an orientation advisor, and all roles were played by the same woman. The "applicant" was asked to tell about her academic background, and to respond to a series of questions about her interests and achievements. The woman was depicted as either: (a) competent, with masculine interests; (b) competent, with feminine interests; (c) incompetent, with masculine interests, or; (d) incompetent, with feminine interests. The subjects viewed one of these four videotapes and immediately rated that "applicant" for both "work likability" and "social likability". Although not quite the equivalent, the competent vs. incompetent conditions are analogous to the previously discussed "success" vs. "attempting success" variables, and the masculine interests vs. feminine interests are similar to the "sex-appropriateness of profession" variable. The researchers here wanted to see how women would rate a

competent woman in a masculine field, how men would rate the similarly portrayed competent woman, and so forth. Of particular note, the competent masculine stimulus portrayed an applicant who was an A-average physics major who sought a professional career. She was described as:

> "A successful tournament tennis player (who) liked reading history and biography, was program chairman of a campus group on ecology, had had Red Cross experience teaching life-saving, and directed a campus program for tutoring disadvantaged highschoolers in mathematics and science" (Spence & Helmreich, 1972, p. 202).

Further:

> "Parallel scripts were used in the two Feminine conditions to convey Competence or Incompetence, but interests varied. The applicant in each case was portrayed as a fine arts major, specializing in interior design; she designed and made her own clothes, cooked, and played bridge as hobbies, was not particularly interested in athletics; and was involved in (or had thought about) a campus musical club and tutoring disadvantaged students in English composition" (p. 202).

Still, Spence and Helmreich (1972) found that the "competent woman" stimuli (masculine or feminine interests) were always rated higher than the "incompetent woman" stimuli. In addition, in the competent conditions, the woman with "masculine" interests received higher ratings than the one with "feminine" interests. Further, mean scores for the two "competent" stimulus conditions yielded overall rankings by the male and female subjects as follows:

1st -- competent, masculine interests (ranked by females)

2nd -- competent, masculine interests (ranked by males)

3rd -- competent, feminine interests (ranked by females)

4th -- competent, feminine interests (ranked by males)

Thus, in the "competent" conditions, females and males agreed on the ordering by ranking the applicant with masculine interests above the applicant with feminine interests, but they did not agree on the absolute degree of favorableness. Since there were no comparable, competent male stimuli, it is yet unclear whether the male subjects were expressing a bias against women, or whether the female subjects were expressing a bias for women. Because the woman with masculine interests was rated highest by both females and males, the authors concluded that to gain superior status, the new, more liberal women were taking on more "masculine" values and characteristics. The authors also posed the question of whether or not their results would have been the same under more realistic conditions, to wit: How would men feel in closer relationships (e.g., peer relationships) with competent women, especially in work (though also in social) settings? Spence and Helmreich suggested:

> "When viewing them at a distance, men may truly like and admire women who are competent in masculine activities...because these are achievements they would like for themselves...Many men, particularly perhaps those who are young, may therefore feel threatened and uncertain about their own masculinity when close to women who are successful" (1972, p. 212).

It is to this general arena of closer, more realistic relationships that the present study will shortly return.

## Sex Bias Among Mental Health Professionals

While the analogue studies of Goldberg and others were being conducted, another important line of research on sex biases was occurring. Researchers were studying therapists' attitudes toward, and expectations of, their female and male clients. Broverman,

Broverman, Clarkson, Rosenkrantz, and Vogel (1970) found that practicing clinicians (46 men, 33 women) held the traditional sex role stereotypes of our society. Both the female and male clinicians agreed in their descriptions of a healthy adult, a healthy woman, and a healthy man. What was important in the authors' findings was that there was no significant difference between the descriptions of the "healthy, mature, socially competent" adult and the "healthy, mature, socially competent" man, but there was a significant difference between the "healthy" adult and the "healthy" woman. This revealed not only that clinicians have different standards of mental health for men and women, but also that their standards of mental health for women were less healthy than their standards of mental health for "adults" and men. The Broverman et al. (1970) study has often been used to criticize psychotherapists for having sex-role stereotypic attitudes and acting upon these attitudes in therapy, thus revealing a therapeutic sex bias that would ultimately encourage sex-role stereotypic behavior in clients.

Bernard E. Whitley in "Sex Roles and Psychotherapy: A Current Appraisal" (1979) reviewed much of the research about psychotherapists and their views of sex-role stereotypes and their impact on therapy. He concluded that even though there is a great deal of evidence to suggest that psychotherapists do hold their culture's sex-role stereotypes, "there is little evidence that these stereotypes affect professional judgment or treatment goals" (Whitley, 1979, p. 1318). Whether or not Whitley's conclusion is correct needs to be examined. Most of the literature in this area utilized sex-role questionnaires and/or were analogue studies.

A more accurate view of sex biases could be obtained by examining the therapists' interactions with each other. As Spence and Helmreich (1972) suggested, the closer men are to successful, competent women, the more likely they might be to feel threatened and to react with sex bias. Putting it more generally, the more real the situation or interaction, the more likely we are to find real biases. Thus, more authentic evidence of sex bias may be found by studying successful mental health professionals in peer relationships.

## Peer Ratings

In such contexts, a viable measurement method is that of peer rating, and several studies have shown that peer ratings have good reliability and predictive validity (Booker & Miller, 1966; Carroll, 1952; Helfer, 1972; Smith, 1967; Wiggins, Blackburn, & Hackman, 1969). Peer rating

> "consists of having each group member rate each other group member on a given set of performance or personal characteristics using any one of several kinds of rating scales" (Kane & Lawler, 1978, p. 557).

Peer ratings provide data on all group members that may be subsequently used to give feedback to the participants. In particular, mental health professionals have reported that feedback from peers enables them to further their own interpersonal growth and competency with their clients (Folkins & Spensley, 1977).

A peer rating system appropriate for use with mental health professionals is that devised by Hurley (1976a, 1976b, 1978, 1980). He has researched and reviewed interpersonal behaviors of individuals within groups and has identified two prepotent dimensions of interpersonal behavior which other authors have often examined,

though under other rubrics. Hurley (1980) offered much evidence that these two dimensions are clusters of behaviors that can be described as "Acceptance vs. Rejection of Others" (ARO), and "Acceptance vs. Rejection of Self" (ARS). The ARO dimension has been researched under such names as affiliation-hostility, positive-negative, loving-rejecting, and love-hate. The ARS dimension is a synthesis of such variables as dominance-submission, activity-passivity, and autonomy-control. These interpersonal variables have been found to be indicators of the "helpfulness" of participants and leaders in psychodynamic groups (Hurley, 1975, 1976a, 1976b, 1978, 1979).

## Statement of the Problem

The reviewed sex bias findings are puzzling. Some research has supported Goldberg's (1968) initial claim that women are prejudiced against women (Pheterson et al., 1971), whereas other research has found no such bias (Pheterson, 1969; Chobot et al., 1974; Levenson et al., 1975; Soto & Cole, 1975). One study suggested a pro-women bias by other women (Levenson et al., 1975), and other findings also support a pro-women bias when the women are considered successful (Spence & Helmreich, 1972; Peck, 1978). In fact, Spence and Helmreich (1972) found that the female and male student participants gave the highest ratings of "work likability" and "social likability" to the woman who was portrayed on videotape as "competent" (i.e., successful) with "masculine interests" over the same woman when she was portrayed as "competent" with "feminine interests". Especially noteworthy is the fact that the female student participants rated the competent-masculine interest woman higher than the male students rated her, though she was rated highest of all by both female and

male raters, and would therefore presumably have received the at by proffered "job" as an orientation advisor. Peck's (1978) rese·timuli, suggested a negative bias toward women who were attempting to su. Panek et al. (1976) found just the opposite, a pro-women bias for ci- women attempting to succeed and a negative bias toward successful women. There may indeed be a sex bias dichotomy between "successful" and "attempting-to-be-successful" women, but the direction of bias remains undetermined. In the face of such mixed results, clarifying studies are needed.

As well as finding no consistent evidence of biases toward women by other women, some of the studies included male raters while others did not. One study (Spence & Helmreich, 1972) utilized both female and male raters, but used exclusively female stimuli. To investigate sex bias more adequately, studies need to include stimuli and raters of both sexes. In the present study, therefore, sex bias is examined within sex and between sex by using female and male raters, and female and male stimuli.

Another issue of sex bias research is that the subjects were usually undergraduate students who were emotionally and physically (time, generation, etc.) distant from the professional stimulus persons they were asked to evaluate. This "distance" may have induced less bias than they might have shown were they in face-to-face contacts with such individuals. On the other hand, their own age and identity strivings may have led to even greater bias, as they in effect perhaps clung to safer and more stereotypical conceptions of male and female roles. In order to approximate a more natural situation, this study utilizes adult participants.

A fourth problem with much of the current research is that by having students rate articles, paintings, or even videotaped stimuli, creates an artificial atmosphere with possible strong demand characteristics. In this study the adult professionals were participants of psychodynamic groups, and the data collection was incidental to these groups' primary function, namely the gaining of professional experience and personal growth. These professionals, too, were in the public work world as peers and potential competitors for jobs, promotions, or simply influence at work. Peer ratings have no sex-role or stereotypic demand characteristics. In effect there is no sex-stereotypic face validity, and thus such peer ratings may yield a more accurate account of extant biases. How mental health professionals view each other with respect to sex biases in a more adult, real-life, interactive setting seems likely to yield a more accurate estimate of how these professionals would tend to view and treat their own male and female clients.

Hypotheses

For present purposes, these adult professional participants are viewed as having achieved "success" in their fields. "Success" is defined here as having reached a reasonably high level of professional status. These participants are established professionals who have attained at least master's degrees with substantial work experience. Since most, if not all, attended the American Group Psychotherapy Association (AGPA) conferences voluntarily, they are viewed as competent professionals invested in enhancing their skills.

One hypothesis is that female ratees will be rated significantly higher on ARO than male ratees by both female and male raters. As

suggested by Peck (1978), successful women may be viewed as having had to overcome greater obstacles to success than men, and are thus rated higher. Others may think that these women had to be extra smart, competent, or to have greater internal controls than males in order to have achieved success.

A second hypothesis is that female raters will give higher ARO ratings to female ratees than to male ratees. Levenson et al. (1975) and Spence and Helmreich (1972) found that females rated other females higher than males rated females. Peck (1978) also found that successful females were rated significantly higher by females than by males. It was suggested by Spence and Helmreich (1972) that males in more realistic or in closer interactions with competent and successful females may feel more threatened by their competence, and thus such male ratings might reveal a greater sex bias. In the Spence and Helmreich (1972) study, males rated competent women lower than did female raters. It was suggested by Chobot et al. (1974) that "no bias" in ratings of females may suggest an increased awareness of feminist issues and gains.

A third hypothesis is that male ratees will be given higher overall ARS scores than female ratees by both female and male raters. Though the feminist movement has made inroads into the societal view of women, the women participants in these studies did grow up in a largely male-dominated society. Their formative years were essentially in the 1940's and 1950's, times when males were even more openly valued than females. A successful woman tended to fit the more traditional roles of wife, mother, and housekeeper, and the professional working woman was a greater rarity. Thus, the present participants

likely grew to maturity with a stronger sense that males were valued more, and were therefore of more value. This may well have left them with a non-conscious, lesser degree of ARS than comparable men of this time period. Although these female participants may indeed have a greater acceptance and positive regard toward women in a conscious feminist sense, their own ARS may not be as accessible, nor as susceptible to change.

A fourth hypothesis is that female raters will rate male significantly higher on ARS than will male raters. Following the rationale for the third hypothesis, hypothesis four is similarly suggested. Female raters may automatically attribute ARS qualities to men.

# METHOD

## Subjects

The persons were mental health professionals who attended the American Group Psychotherapy Association's Annual Institute Groups in 1974 and 1977. The typical participant has a professional degree, was generally over 35 years of age, and has had several years of professional experience as either an individual or group psychotherapist or both. At the 1974 AGPA Institute, the participants were composed of M.S.W. social workers (44.5%), M.D. psychiatrists (18%), Ph.D. clinical psychologists (11%), others (21%), educators (4%), and R.N. nurses (1%). The 1977 AGPA Institute was composed of M.S.W. social workers (34%), M.D. psychiatrists (23%), Ph.D. clinical psychologists (15.5%), others (21%), and R.N. nurses (6%). Having achieved a reasonably high professional level of success and/or competence, it seems reasonable to regard these participants as accomplished.

## Measures

Hurley's (1976a, 1980) peer rating scales of the ARO and ARS dimensions were used. The measures were presented in a mini-booklet composed of bipolar semantic differential scales. Each of the nine scales was presented on a separate page, with check points between the anchoring phrases that were translated into scores ranging from 0 to 9. In 1974, the four scales that comprised the ARO dimension were:

Warm-Cold; Helps Others-Harms Others; Involved-Detached; and Accepts Others-Rejects Others. The ARS dimension was tapped by scales: Shows Feelings-Hides Feelings; Expressive-Guarded; Active-Passive; and Independent-Dependent. In 1977, two slight changes were made in the measure. Involved-Detached was replaced by Gentle-Harsh, and Independent-Dependent was superceded by Dominant-Submissive. In both 1974 and 1977 an additional scale, Liked-Disliked was included in the first position in order to "facilitate the ventilation of strong feelings and elicit more cognitive and less emotionally-laden ratings on the subsequent scales" (Hurley, 1979, p. 6). The instructions requested the group participants to place an "X" in the space that reflected their "personal impression" of each member's "within-group behavior". The emphasis was on interpersonal behaviors except for the Liked-Disliked scale, which asked for a "reaction" to each group member. The peer ratings ranged from 0 to 36 for each dimension, with 36 (9's on all four scales) being the highest possible rating, and zero the minimum. Each person was rated individually by all other group members, and the ARO and ARS scores were intended to reflect the interpersonal competence as perceived by peers.

Environment

The data were collected at the Annual Institutes of the American Group Psychotherapy Association in 1974 and 1977. Participants were given the opportunity to become involved in psychodynamic groups for both personal and professional growth. In 1974 these groups were labeled the "General Experience Section", and in 1977 the groups were called the "Psychodynamic Group Process Section". The

psychodynamic groups in which the participants were involved:

> "offered to help persons increase their effectiveness as group therapists and their awareness of the dimensions of the group process.
>
> The group will stress the essential elements of analytic group psychotherapy, including problems of transference, individual and group resistances and defense mechanisms that are intrinsic in group dynamic processes. The Institute group provides a unique opportunity for fulfilling its educational purposes by experiencing many of these same group processes discussed...Both the structure and the dynamics of the group will accordingly by reflected in this experience and will involve cognitive as well as emotional aspects of learning" (Twenty-First Annual Institute, AGPA, Inc., 1977, p. 9).

Procedure

At the 1974 AGPA Institute, 15 groups were randomly chosen from 30 possible "Psychodynamic Process" groups. Seven of these were solo-led, while eight were co-led. Data were collected on two occasions. Phase I data were collected on Day 1, after the groups had met for only 2½ hours. Phase II data were collected after the groups had met twice on Day 2, after a total of 11½ hours of interaction, and just before the groups terminated. Due to a short supply of rating booklets, Day 2 data were solicited from only eight of the fifteen Day 1 groups. Prior permission to conduct this research was granted by the AGPA.

At the 1977 AGPA Institute, 19 randomly selected groups constituting one-half of all the participating in the "Psychodynamic Group Process Section" were used. At the end of the groups' initial 2½ hours of interaction, and at the end of the second day (following a total of 11½ hours of interaction), data were collected. Two groups were excluded from the 19, one because it was the only co-led group, and

the second because it held only two women participants, yielding one solo-based rating from the same-sex peer.

Data Analysis

Separate three-way analyses of variance were applied respectively to 2 x 2 x 8 (1974 study) and 2 x 2 x 17 (1977 study) factorial designs. The main effects of Rater's Sex, Ratee's Sex, and Groups were investigated using Coyle and Frankman's (1977) Balanced Design (BALANOVA) program. Since the data sets are from two independent studies done three years apart, each with substantial N's, joint probabilities were also calculated for main effects and interactions in keeping with Hays (1973, pp. 150-151). The BALANOVA permitted data analyses with different sized groups, but it required equal numbers of female and male raters and ratees per cell. To balance the number of opposite-sex persons within each group, data from excess numbers of the largest groups were randomly dropped. Such persons were randomly selected by numbering the largest group from 1 through 5, and then consulting a random table for the ordering of the participants. For example, if there were four women and three men in a group, the women were numbered from one through four, and then the random table would determine the order of the numbers. If number two was the last of the four numbers to appear on the random table, then person two's data were excluded. Only persons who both rated, and were rated on ARS and ARO were used. Table 1 shows the number of excluded persons per group.

All participants rated themselves on each variable, but self-ratings were excluded from the mean peer ratings. To have a balanced number of peer ratings from each sex, one opposite-sex peer rating was randomly

TABLE 1

Data Sets Excluded from Analyses: 1974 and 1977 Groups

| Year | Group | Persons in Balanced Group | | Excluded Data Sets | |
|---|---|---|---|---|---|
| | | Women | Men | Women | Men |
| 1974 | 1 | 5 | 5 | - | - |
| | 2 | 4 | 4 | 1 | - |
| | 3 | 3 | 3 | 2 | - |
| | 4 | 4 | 4 | 2 | - |
| | 5 | 5 | 5 | - | - |
| | 6 | 4 | 4 | 1 | - |
| | 7 | 5 | 5 | 1 | - |
| | 8 | 3 | 3 | - | 1 |
| 1977 | 1 | 4 | 4 | - | - |
| | 2 | 4 | 4 | 2 | - |
| | 3 | 3 | 3 | - | - |
| | 4 | 5 | 5 | - | - |
| | 5 | 5 | 5 | - | - |
| | 6 | 5 | 5 | - | - |
| | 7 | 6 | 6 | - | - |
| | 8 | 3 | 3 | 2 | - |
| | 9 | 3 | 3 | 1 | - |
| | 10 | 3 | 3 | - | 2 |
| | 11 | 3 | 3 | - | 1 |
| | 12 | 3 | 3 | - | 1 |
| | 13 | 4 | 4 | - | 1 |
| | 14 | 3 | 3 | - | - |
| | 15 | 3 | 3 | 2 | - |
| | 16 | 4 | 4 | 1 | - |
| | 17 | 3 | 3 | - | - |

dropped before calculation of the mean peer ratings. Each group was set up in a table (see Table 2) so that all female and male participants were randomly assigned within their sex and then identified by letter. The first woman "A"'s self-rating would be in the A x A column, and the first man "E"'s rating of "A" was be dropped to yield an equal number of peer ratings by same-sex and opposite-sex raters. Thus, the opposite-sex rater who held the same position within his/her sex group as the person being rated, would have his/her rating randomly excluded.

On rare occasions, one or two ratings were missing from an otherwise complete set of data, possibly due to rater errors such as accidentally skipping someone. Replacement values were determined for these omissions. These values were calculated by totaling the row and column means of the missing values, then dividing it by two. For example (Table 2), if person "B" had not been rated by person "C", "B"'s row mean of 25 plus "C"'s column mean of 29.5 totalled 54.5 which, when divided by two, established a replacement value of 27.25. No replacement values were needed in the 1974 data and only eight among a total of 1426 ratings in 1977.

TABLE 2

Raw Group Data: An Example of How Scores were Randomly Excluded to Insure Equal Cell Size

| | | RATERS | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Women | | | | | Men | | | |
| | | A | B | C | D | Mean Rating | E | F | G | H | Mean Rating |
| RATEES | | | | | | | | | | | |
| Women | A | ○ | 24 | 30 | 16 | 23.33 | (20) | 32 | 19 | 29 | 26.67 |
| | B | 23 | ○ | 27 | 27 | 25.67 | 22 | (23) | 17 | 35 | 24.67 |
| | C | 30 | 22 | ○ | 26 | 26.00 | 25 | 32 | (23) | 34 | 30.33 |
| | D | 23 | 20 | 29 | ○ | 24.00 | 20 | 19 | 18 | (32) | 19.00 |
| Men | E | (25) | 12 | 24 | 5 | 13.67 | ○ | 16 | 8 | 16 | 13.33 |
| | F | 22 | (15) | 26 | 30 | 26.00 | 21 | ○ | 17 | 31 | 23.00 |
| | G | 21 | 14 | (26) | 17 | 17.33 | 18 | 23 | ○ | 30 | 23.67 |
| | H | 29 | 14 | 26 | (27) | 23.00 | 17 | 30 | 16 | ○ | 21.00 |

Note. Rater A and Ratee A are the same person. Therefore, the A x A box is a self-rating.

○ = denotes self-rating

(10) = denotes excluded rating

## RESULTS

There were four major trends for both the 1974 and 1977 studies. Women were consistently rated above men on both ARO and ARS. Women assigned higher ratings to others than did men raters, and women rated other women higher than they did men. Men also rated women higher on both variables than they rated other men. Table 3 displays these major trends by showing the group means for Sex-of-Rater by Sex-of-Ratee for ARO and ARS for both studies.

### Study I - AGPA Institute, 1974

There was a main effect for Groups. The eight groups differed significantly on both variables for ARS ($F = 2.98$, $p < .01$) and for ARO ($F = 3.11$, $p < .01$). See Table 4 for ANOVA summary.

There was also a significant Sex-of-Rater effect for both ARS ($F = 6.50$, $p < .01$) and ARO ($F = 9.70$, $p < .01$). Examination of these means revealed that women raters generally gave higher ratings than did men.

A significant interaction between Group and Sex-of-Rater (G x R) was found for each measure. The $F$-ratio for ARS was 3.74 ($p < .01$) and ARO's corresponding value was 6.02 ($p < .001$). The women gave higher mean ARS and ARO ratings than did men in six of the eight groups.

Finally, Sex-of-Ratee approached significance ($p = .097$) for ARS, but not for ARO ($p = .26$). Women received higher ratings on both measures.

TABLE 3

Mean ARO and ARS Ratings by Sex of Rater and Ratees

| | | | | RATEES | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | ARO | | | ARS | | |
| | | | N | Women | Men | Both | Women | Men | Both |
| RATERS | Women | 1974 | 33 | 25.44 | 23.99 | 24.72 | 23.88 | 21.69 | 22.79 |
| | | 1977 | 64 | 24.90 | 24.48 | 24.69 | 21.32 | 20.45 | 20.88 |
| | Men | 1974 | 33 | 23.72 | 22.60 | 23.16 | 22.39 | 20.74 | 21.57 |
| | | 1977 | 64 | 24.56 | 23.52 | 24.04 | 20.96 | 19.68 | 20.32 |
| | Both | 1974 | 66 | 24.58 | 23.30 | 23.94 | 23.14 | 21.21 | 22.18 |
| | | 1977 | 128 | 24.73 | 24.00 | 24.37 | 21.14 | 20.06 | 20.60 |

TABLE 4

The Analyses of Variance for 1974 AGPA Groups

| Source of Variation | | Acceptance vs. Rejection of Self | | Acceptance vs. Rejection of Others | |
|---|---|---|---|---|---|
| | df | MS | F | MS | F |
| Group (G) | 7 | 127.39 | 2.98[b] | 132.56 | 3.11[b] |
| Sex of ratee (X) | 1 | 122.42 | 2.87[c] | 54.62 | 1.28 |
| Individual ratee (S) | 50 | 42.73 | | 42.67 | |
| Sex of Rater (R) | 1 | 49.09 | 6.50[b] | 79.95 | 9.71[b] |
| G x X | 7 | 44.35 | 1.04 | 17.08 | 0.40 |
| G x R | 7 | 23.29 | 3.74[b] | 49.60 | 6.02[a] |
| X x R | 1 | 2.41 | 0.32 | 0.85 | 0.10 |
| S x R | 50 | 7.56 | | 8.23 | |
| G x X x R | 7 | 7.41 | 0.98 | 11.28 | 1.37 |
| Total | 131 | | | | |

[a] $p < .001$
[b] $p < .01$
[c] $p = .097$ ns

Study II - AGPA Institute, 1977

There were no significant main effects, although two interactions reached statistical significance. As before, the Group by Sex-of-Rater interaction (G x R) was significant for both ARS and ARO (Table 5). The ARS F-ratio was 4.18 ($p < .001$), and the ARO F-value was 3.44 ($p < .001$). Women again gave higher mean ratings than did men for 10 of the 17 groups by both measures.

The other significant interaction was between Group and Sex-of-Ratee (G x X). This reached significance only with the ARS analysis ($F = 2.27$, $p < .01$). In 12 of the 17 groups, women received higher mean ARS ratings than did men. For ARO, the same interaction was not significant, although women's mean scores were higher in 10 of the 17 groups, congruent with the ARS trend.

Joint Probabilities for 1974 and 1977 AGPA Groups

Groups and Rater's sex were significant main effects for both ARO and ARS when the joint probabilities were calculated. The G x R interaction was also significant over both variables. The ARS variable yielded three other significant findings. For ARS, Sex-of-Ratee was significant as were the G x X and G x X x R interactions (see Table 6).

The Sex-of-Rater effect for ARS p was < .01, and for ARO p was < .05. Women raters gave significantly higher ratings than men.

The main effect of Sex-of-Ratee for ARS p was < .05. Women ratees received significantly higher ratings than men from both women and men raters. ARO p was < .09.

Hypothesis I

It was predicted that women would be rated significantly higher on ARO than men by raters of both sexes. Although the ARO scores for

TABLE 5

The Analyses of Variance for 1977 AGPA Groups

| Source of Variation | | Acceptance vs. Rejection of Self | | Acceptance vs. Rejection of Others | |
|---|---|---|---|---|---|
| | df | MS | F | MS | F |
| Group (G) | 16 | 58.37 | 0.89 | 44.92 | 1.28 |
| Sex of ratee (X) | 1 | 73.90 | 1.13 | 34.19 | 0.96 |
| Individual ratee (S) | 94 | 65.42 | | 35.07 | |
| Sex of rater (R) | 1 | 20.14 | 2.48 | 27.29 | 2.64 |
| G x X | 16 | 148.19 | 2.27[b] | 37.96 | 1.06 |
| G x R | 16 | 33.97 | 4.18[a] | 35.49 | 3.44[a] |
| X x R | 1 | 2.57 | 0.32 | 6.09 | 0.59 |
| S x R | 94 | 8.12 | | 10.32 | |
| G x X x R | 16 | 12.63 | 1.56 | 8.38 | 0.81 |
| Total | 255 | | | | |

[a] $p < .001$

[b] $p < .01$

TABLE 6

Joint Probabilities of Independent 1974 and 1977 Studies

| | ARS | | | ARO | | |
|---|---|---|---|---|---|---|
| | 1974 | 1977 | P(1974 x 1977) | 1974 | 1977 | P(1974 x 1977) |
| Groups (G) | .011 | .580 | .0063[b] | .008 | .226 | .0018[b] |
| Sex of ratee (X) | .097 | .291 | .0282[c] | .263 | .326 | .0857[d] |
| Sex of rater (R) | .014 | .119 | .0016[b] | .003 | .107 | .0003[a] |
| G x X | .417 | .008 | .0033[b, e] | .898 | .383 | .3439 |
| G x R | .002 | <.0005 | <.0001[a, f] | <.0005 | <.0005 | <.0001[a] |
| X x R | .575 | .575 | .3306[g] | .750 | .444 | .3330 |
| G x X x R | .456 | .097 | .0442[c] | .239 | .669 | .1599 |

[a] $p < .001$

[b] $p < .01$

[c] $p < .05$

[d] $p < .09$

[e] The sex of the ratee was even more critical within the groups in accounting for the ratee's score. Most of the groups had a large difference between the way women and men were rated.

[f] Raters polarized within groups by sex regardless of whether women were rating higher or men were rating higher.

[g] There was no significant interaction for Sex of Rater x Sex of Ratee.

Sex-of-Ratee were not significant in either study, the joint probabilities for the studies ($p < .09$) revealed a weak ~~strong~~ trend for a ratee's sex effect (Table 6). This trend is supported as the women had higher mean ARO scores than men (Table 3).

Hypothesis II

Hypothesis II stated that women would give significantly higher ARO ratings to women than to men. The (X x R) interaction was not statistically significant. In 1974, women gave a mean ARO score of 25.44 to women and 23.99 to men. Again in 1977, women gave higher mean ARO scores to women (24.90) than to men (24.48). While this interaction did not reach significance, the trend was consistent toward support of this hypothesis.

Hypothesis III

It was predicted that men would be given higher overall ARS ratings than women by both sexes. There was a significant main effect for Sex-of-Ratee; the joint probability was $p < .05$. However, the mean ARS scores were consistently contrary to the direction predicted. The ARS mean in 1974 for women ratees were 23.14, nearly two points above the men's mean of 21.21. The 1977 data also followed the same pattern. ARS scores for women were 21.14, higher than those for men (20.06).

Hypothesis IV

Hypothesis IV stated that women would rate men significantly higher on ARS than would men raters. Neither the 1974 nor the 1977 data supported this view. There was no significant interaction between Sex-of-Rater and Sex-of-Ratee (X x R). In both studies,

women rated men higher on ARS than did male raters, but women rated other women even higher on ARS than they rated men (Table 3).

## DISCUSSION

Both the (1974 and 1977) findings refute Goldberg's (1968) original claim that women are prejudiced against women. Before discussing the trends and possible implications of these studies, however, the fact that it was done on a special sample needs to be reiterated. This research utilized adult participants from the work arena who were involved in small interpersonal groups for their own professional development, not to generate research data. In terms of the pertinent literature, these participants would have been placed in the "successful" categories. In most of the studies reviewed (e.g., Peck, 1978; Panek et al., 1976; Pheterson et al.,1971), "success" was denoted by already having achieved one's goal and public recognition for it vs. attempting to achieve one's goal but no indication of whether or not the achievement would be accomplished and acknowledge (e.g., professor vs. doctoral students, winner of contest vs. contestant in contest, published article vs. submitted article). The participants of this research had achieved professional degrees, had significant work experience, and were engaged in furthering their professional competency, all indicators which place them in the realm of having achieved "success".

One of the major trends of these studies was that women received higher mean ratings on ARO and ARS than men. Conducted three years apart, this evidence consistently revealed that neither women nor

men were biased against women. The joint probabilities of these two studies found that women were rated significantly higher on ARS than men by both sexes $p$ was < .05, while for ARO $p$ = .0857 was not quite significant. More importantly, Hypothesis II, which predicted that women would rate other women higher on ARO than men raters was tenuously supported. Women raters gave women ARO scores in 1974 (N=33) and 1977 (N=64) of 25.44 and 24.90, respectively, whereas they gave men comparable means of 23.99 and 24.48. Though below statistical significance, the respectable sample size consistently showed a tendency for women to perceive other women as more accepting of others in actual interpersonal/professional interactions than were men.

Why were women rated higher on both ARO and ARS by both sexes, but also rated even higher on these measures by women raters than men raters? One of the tenents of this research was that the participants were successful, mature, established professionals. Thus, Peck's (1978) finding that "successful"women were rated higher by women than men was confirmed. Peck (1978) labeled this "overvaluing" of women by women, and speculated that these successful women were thought to have had to be more competent and internally motivated than their male peers, as they had to overcome more obstacles in the work world to achieve comparable career success. Peck's speculations followed Feldman-Summers and Kiesler (1974) in that they found that subjects attributed more motivation to females than to males. But even more importantly, whereas male subjects had attributed more ability to male physicians than to comparable female physicians, women subjects did not think that the male physicians had more ability than female physicians.

> "Instead, they (the female subjects) attributed the male's success to his having an easier task than the female. Female success was then explained in terms of the female physician's motivation to succeed being greater than the male's..." (Feldman-Summers & Kiesler, 1974, p. 853).

Successful women may be rated higher than successful men especially by women because they are viewed as having to overcome greater obstacles and thus may be more motivated than their male counterparts.

Abramson, Goldberg, Greenberg, and Abramson (1977) found that both women and men, when rating biographies that were identical except for assigned sex, gave higher competency ratings to female than to male "lawyers". Like Peck (1978), these authors labeled the ratings assigned to the woman lawyer "inflated" and suggested that their female subjects had rated the women lawyer the highest due to their "greater sensitivity" to the obstacles to the woman's success. These findings seemed reported in a jaundiced, sexist, and rather demeaning style. Having first stated that their results showed that "credit is given where credit is due", Abramson et al. (1977) wrote:

> "However, it may also be said that the finding concerning vocational competence exemplifies the talking platypus phenomenon, that is, when an individual achieves a level of success not anticipated, his/her achievement tends to be magnified rather than diminished. After all, it matters little what the platypus says, the wonder is that it can say anything at all" (Abramson et al., 1977, p. 123).

The last sentence suggests that no matter what women do, if they are seen as competent and successful it is only because no one expected anything from them. One of these researchers was Goldberg, who had earlier (1968) asserted that:

> "Women--at least these young college women--are prejudiced against female professionals and, regardless of the actual accomplishments of these professionals, will firmly refuse to recognize them as the equals of their male colleagues" (Goldberg, 1968, p. 30).

These authors seem unable to take seriously either their own research results or women. Their findings suggest that there is a pro-woman phenomenon for successful women extant in both women and men. Why phrase these findings so derogatorily? Another possible explanation may be that successful women as a group <u>are</u> more competent than their male peers as a group because they do indeed require something extra to attain the same level of success. As Denmark (1980) suggested in her presidential address to the American Psychological Association:

> "To be recognized as influential requires, in addition to ability and perseverance, access to professional networks, admittance to research laboratories, and affiliation with prestigious universities. On each of of these fronts women have been placed at a disadvantage. Although the exceptional person probably rises to prominence regardless of her gender, female psychologists of average or even above average capability and motivation are more likely to be relegated to anonymity than are equivalent males" (Denmark, 1980, p. 1059).

Denmark's address seems to describe rather accurately the obstacles to achievement for women professionals. This might be difficult to prove through research, however, though not necessarily impossible.

A well known study entitled "Empirical evidence of sex discrimination in hiring practices in psychology" (Fidell, 1970) sent identical biographies for "Patricia" or "Patrick" to psychology departments at several universities to apply for jobs. "Patrick" was offered more positions than "Patricia", and he was offered

positions of higher rank than she was also, thus revealing that being equal was not enough to get "Patricia" the job, and possibly suggesting further that she would have to be better in order to be offered the same higher level positions.

The pro-successful-woman phenomenon may not be a bias at all. It may be a reality. Successful women may be brighter, have a wider variety of coping skills, and more motivation and perseverance to overcome the obstacles for women to achieve the same level of success as their male peers. Gullahorn (1977) noted that the evidence reveals:

> "there is a greater selectivity operating for women than for men in higher education. Thus only the intellectually "fittest" of the women tend to survive, as demonstrated by the fact that graduate students and professionals rather consistently score higher on intelligence tests than do their male peers" (Bernard, 1964; Astin, 1969; cited in Gullahorn, 1977, p. 253).

Though the above studies are in no way exhaustive, they do suggest that some research has found at least a tendency for successful professional women to have more abilities and acquired skills than their professional male counterparts or they would not be where they are.

Another important significant finding was that women raters gave higher ratings than men to participants of both sexes. The joint probability for ARS was $\underline{p} < .01$ while for ARO, $\underline{p}$ was $< .05$, suggesting that these women seemed to be more accepting of others in their groups and that they see their peers in a more positive light than men. Other studies support this explanation. Whitley (1979), in his review on "Sex Roles and Psychotherapy: A Current Appraisal", stated that:

> "women therapists tend to be slightly more accepting of counterstereotypic behavior than men (Chasen, 1975; Chasen & Weinberg, 1975) and more lenient in their ratings of all stimulus persons (Abramowitz et al., 1975; Harris & Lucas, 1976; Maxfield, 1976)" (Whitley, 1979, p. 1315).

Sherman, Koufacos, and Kentworthy (1978) studies therapists and their attitudes and information about women. In their study, 184 therapists (social workers, psychiatrists, and psychologists) responded to a questionnaire which consisted of demographic and background information, some short answer questions, the Therapists' Information About Women Scale (TIWS) and the Therapists' Attitudes Toward Women Scale (TAWS). The study found significant differences between the female and male therapists' responses on both the TIWS and TAWS. The means revealed that women therapists were better informed, more liberal, and held less stereotypic views than the men. Two interesting facts about this study (Sherman et al., 1978) are: (a) that the average ages of the therapists were 42 for both men and women, and (b) that there was no significant difference in attitudes among professions. Thus, this sample seems quite similar to the AGPA samples.

Denmark (1980) discussed a study (Kelly & Kiersky, 1979) which is pertinent for the understanding of the AGPA data. The study suggested that females demonstrated more tolerance for "out-of-role" behavior in both men and women than males did. Possibly in the AGPA groups, the women therapists, as in the previously discussed research, were less critical and more accepting of a variety of behaviors and responses from their female and male peers. For example, women who exhibited high agentic traits would not be criticized, just as men therapists who exhibited more nurturant traits would also be accepted.

The AGPA women raters may have been more accepting of androgynous behaviors by both sexes than were the men raters.

The 1974 ANOVA revealed a main effect for Groups for both ARO and ARS $p$ was < .01. The 1977 ANOVA did not find a main effect for Groups (ARS $p$ = .58 while ARO $p$ = .226). However when the joint probabilities were calculated there was a main effect for Groups for both ARO and ARS. According to the joint probabilities for Groups $p$ was < .01, thus indicating that the groups differed from each other. Within each group the interactions between participants were unique and different. Some groups may have had more positive climate and interpersonal interactions while other groups may have had a more confrontive or negative nature. This is speculation for it is not known why the groups differed except to say they were composed of different individuals.

Women were rated higher on ARS than men by both sexes, and women rated women higher on ARS than men, both the opposite of what was predicted. These surprising results can be understood in terms of sex-role stereotypes and the concept of androgyny.

Bem's (1974) new sex-role inventory dealt with femininity and masculinity as two separate dimensions of personality rather than the traditional view of them as bipolar opposites. Bem stated that:

> "The concept of psychological androgyny implies that it is possible for an individual to be both assertive and compassionate, both instrumental and expressive, both masculine and feminine, depending upon the situational appropriateness of these various modalities; and it further implies that an individual may even blend complementary modalities in a single act" (Bem, 1977, p. 196).

The concept of a healthy, androgynous individual, whose many-faceted personality sharply juxtaposes the traditionally sex-typed masculine male and female, is crucial to the understanding of the earlier mentioned results. An androgynous person, for purposes of this discussion, is an individual who scores high on both the femininity and masculinity scales of Bem's Sex Role Inventory Scale (Bem, 1974). Several researchers and scholars studying the psychology of women, sex roles, androgyny, and related self-esteem have stated that this is the definition of androgyny they feel is most useful (Bem, 1977; Heilbrun, 1976; Spence, Helmreich & Stapp, 1975).

It seems that the two measures ARO and ARS strongly tap androgynous behaviors which, as stated by Spence et al. (1975), is "the most desirable state of affairs" (p. 38). In some preliminary results from research in progress (L.S. Blank, personal communication, June 20, 1983), significant correlations have been found between high scores on ARO and Bem's Femininity scale, and between Bem's Masculinity scale and ARS.

Spence and Helmreich (1972) also found that the competent, successful female stimulus with masculine interests (i.e., androgynous) was rated highest by both sexes. Following Spence and Helmreich's findings, the women in the 1974 and 1977 AGPA studies probably exhibited several agentic or traditionally masculine behaviors as well as feminine behaviors, and therefore the women were rated high on the masculine ARS.

In their study "Psychological androgyny and interpersonal behavior", Wiggins and Holzmuller (1978) found that androgynous women were more dominant-ambitious and significantly less introverted than

androgynous men. The authors further found that androgynous men and women presented quite different personality profiles:

> "In comparison with the androgynous male, the androgynous female presents herself as extraverted, warm, excitable, emotional, aggressive, and vivacious. In contrast, the androgynous male is introverted, cold, calm, unemotional, passive and undramatic" (Wiggins & Holzmuller, 1978, p. 46).

This being the case, the components of ARS, namely Shows feelings-Hides feelings, Expressive-Guarded, Active-Passive, Independent-Dependent (or Dominant-Submissive), would help explain why women were rated higher on ARS than men. Further, since the components of ARO are Warm-Cold, Helps others-Harms others, Involved-Detached (or Gentle-Harsh), and Accepts others-Rejects others, it can also be seen why women were rated higher on ARO also. It seems that ARS taps agentic (masculine) behaviors while ARO taps communal (feminine) behaviors.

Since the women in the studies were probably more androgynous than the men, they were rated for their desirable feminine and masculine attributes. Women were higher on ARO because they were more accepting of others and high on many feminine items from Bem's Sex Role Inventory Scale, which are also associated with Bakan's (1966) description of a communion orientation. Women were higher on ARS because they displayed higher masculine traits of Bem's scale which are more "agentic".

It seems that the women in these studies have adapted some traditional male values and traits (i.e., agentic orientation), but have not rejected traditionally feminine attributes and values (i.e., communal orientation). Thus, these women do not reject the

female role as Allport (1958), Horney (1967), and Adler (1954) observed, but have added to it many attributes and behaviors of the male role. It seems that they have combined both orientations, allowing themselves greater flexibility and individuality. They seem to be truly androgynous.

# LIST OF REFERENCES

LIST OF REFERENCES


Abramson, P.R., Goldberg, P.A., Greenberg, J.H., and Abramson, L.M. The talking platypus phenomenon: Competency ratings as a function of sex and professional status. Psychology of Women Quarterly, 1977, 2, 114-124.

Adler, A. Understanding human nature, (W.B. Wolfe, Trans.). Greenwich, CT.: Fawcett Publications, 1954. (Original work published 1927).

Allport, G.W. The nature of prejudice. New York: Addison-Wesley, 1958.

Bakan, D. The duality of human existance. Boston: Beacon Press, 1966.

Bem, S.L. The measurement of psychological androgyny. Journal of Consulting and Clinical Psychology, 1974, 42, 155-162.

Bem, S.L. On the utility of alternative procedures for assessing psychological androgyny. Journal of Consulting and Clinical Psychology, 1977, 45, 196-205.

Booker, G.S., and Miller, R.W. A closer look at peer ratings. Personnel, 1966, 43, 42-47.

Broverman, I.K., Broverman, D.M., Clarkson, F.E., Rosenkrantz, P.S., and Vogel, S.R. Sex-role stereotypes: A current appraisal. Journal of Consulting and Clinical Psychology, 1970, 34, 1-7.

Carroll, J.B. Ratings on traits measured by a factored personality inventory. Journal of Abnormal and Social Psychology, 1952, 47, 626-632.

Chobot, D.S., Goldberg, P.A., Abramson, L.M., and Abramson, P.R. Prejudice against women: A replication and extension. Psychological Reports, 1974, 35, 478.

Coyle, B.W., and Frankmann, R.W. Balanced designs analysis of variance programs: User's manual. Department of Psychology, Michigan State University, October, 1977.

Denmark, F. Psyche: From rocking the cradle to rocking the boat. American Psychologist, 1980, 35, 1057-1065.

Donelson, E. and Gullahorn, J.E. Women: A psychological perspective. New York: John Wiley and Sons, 1977.

Feldman-Summers, S. and Kiesler, S.B. Those who are number two try harder: The effect of sex on attributions of causality. Journal of Personality and Social Psychology, 1974, 30, 846-855.

Fidell, L. Empirical verification of sex discrimination in hiring practices in psychology. American Psychologist, 1970, 25, 1094-1097.

Folkens, C. and Spensley, J. Peer ratings by a community mental health team: A positive approach to accountability. American Journal of Orthopsychiatry, 1977, 47, 331-334.

Goldberg, P. Are women prejudiced against women? Transaction, 1968, 4, 28-30.

Hays, W.L. Statistics for the social sciences (2nd Ed.). New York: Holt, Rinehart, and Winston, 1973.

Heilbrun, A.B., Jr. Measurement of masculine and feminine sex-role identities as independent dimensions. Journal of Consulting and Clinical Psychology, 1976, 44, 183-190.

Helfer, R. Peer evaluation: Its potential usefulness in medical education. British Journal of Medical Education, 1972, 6, 224-231.

Horney, K. Feminine psychology (H. Kelman, Ed.). New York: Norton and Company, 1967.

Hurley, J.R. "Some effects of trainers on their T-groups" reconsidered. Journal of Applied Behavioral Science, 1975, 11, 191-196.

Hurley, J.R. Helpful behaviors in groups of mental health professionals and undergraduates. International Journal of Group Psychotherapy, 1976a, 26, 173-189.

Hurley, J.R. Two prepotent interpersonal dimensions and the effects of trainers on T-groups. Small Group Behavior, 1976b, 7, 77-98.

Hurley, J.R. Toward seeing ourselves as others see us. IN L.R. Wolberg, M.L. Aronson, and A.R. Wolberg (Eds.), Group therapy. New York: Stratton Medical Books, 1978.

Hurley, J.R. Group members' gains and leader behaviors. Unpublished manuscript, 1979.

Hurley, J.R. Two interpersonal dimensions relevant to group and family therapy. In L.R. Wolberg and M.L. Aronson (Eds.), Group and family therapy. New York: Brunner/Mazel, 1980.

Kane, J.S. and Lawler, E.E., III. Methods of peer assessment. Psychological Bulletin, 1978, 85, 555-586.

Levenson, H., Burford, B., Bonno, B., and Davis, L. Are women still prejudiced against women? A replication and extension of Goldberg study. The Journal of Psychology, 1975, 89, 67-71.

Panek, P.E., Deitchman, R., Burkholder, J.H., Speroff, T., and Haude, R.H. Evaluation of feminine professional competence as a function of level of accomplishment. Psychological Reports, 1976, 38, 875-880.

Peck, T. When women evaluate women, nothing succeeds like success: The differential effects of status upon evaluations of male and female professional ability. Sex-Roles, 1978, 4, 205-213.

Pheterson, G.I., Kiesler, S.B., and Goldberg, P.A. Evaluation of the performance of women as a function of their sex, achievement and personal history. Journal of Personality and Social Psychology, 1971, 19, 114-118.

Sherman, J., Koufacos, C., and Kenworthy, J.A. Therapists; Their attitudes and information about women. Psychology of Women Quarterly, 1978, 2, 299-313.

Smith, G. Usefulness of peer ratings of personality in educational research. Educational and Psychological Measurement, 1967, 27, 267-284.

Soto, D.H. and Cole, C. Prejudice against women: A new perspective. Sex-Roles, 1975, 1, 385-393.

Spence, J.T. and Helmreich, R. Who likes competent women? Competence, sex-role congruence of interests, and subjects' attitudes toward women as determinants of interpersonal attraction. Journal of Applied Psychology, 1972, 2, 197-213.

Spence, J.T., Helmreich, R., and Stapp, J. Ratings of self and peers on sex-role attributes and their relation to self-esteem and conceptions of masculinity and femininity. Journal of Personality and Social Psychology, 1975, 32, 29-39.

Thompson, C.M. Interpersonal psychoanalysis: The selected papers of Clara Thompson (M.R. Green, Ed.). New York: Basic Books, 1964.

Twenty-First Annual Institute Program. American Psychotherapy Association, Inc., 1995 Broadway, New York.

Whitley, B.E., Jr., Sex-roles and psychotherapy: A current appraisal. Psychological Bulletin, 1979, 86, 1309-1321.

Wiggins, J.S. and Holzmuller, A. Psychological androgyny and interpersonal behavior. Journal of Consulting and Clinical Psychology, 1978, 46, 40-52.

Wiggins, N., Blackburn, M., and Hackman, J.R. Prediction of first-year graduate success in psychology: Peer ratings. Journal of Educational Research, 1969, 63, 81-85.