DOCUMENT RESUME

ED 199 286 TM 810 196

AUTHOR	Campbell, Patricia B.
TITLE	The Impact of Societal Biases on Research Methods.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
PUB DATE	Feb 81
CONTRACT	NIE-P-79-0120
NOTE	71p.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Racial Bias; Research Design; *Research Methodology; Research Proposals; Sampling; *Sex Bias; Test Bias

ABSTRACT

The effects of racism and sexism can be seen in the various elements of research methods including the selection of topics, design, sampling, measurement, and the generation of conclusions. In selecting topics, potential sources of bias such as funding sources and publishing patterns are covered. How knowledge of appropriate literature, selection of independent variables, and control of sources of invalidity can be sources of bias in design are examined. The effects of these sources in terms of incomplete testing for race or sex similarities and differences, use of biased independent variables, and incomplete blocking or control of confounding variables are discussed. Bias and sampling are discussed covering the sources of bias (composition and selections of samples, the effects of single sex or race samples, organizations, use of different race or sex samples for different content areas), and the results of that bias. The section on bias and measurement investigates the effect of bias on observation, aptitude and achievement tests, vocational tests, and affective tests. The section on bias and the generation of conclusions examines the way that bias can encourage overgeneralization, incorrect attributions of causality, and conclusions based on expectations rather than on data. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED199286

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE IMPACT OF SOCIETAL BIASES ON RESEARCH METHODS

Patricia B. Campbell

U.S. Department of Education
Terrel Bell, Secretary

National Institute of Education
Milt Goldberg, Acting Director

Teaching and Learning
Lois-ellin Datta, Associate Director

Learning and Development
Susan Chipman, Assistant Director

February, 1981

TM 810 196

The work described in this publication was conducted with funding from the National Institute of Education, U.S. Department of Education, under contract number NIE P-79-0120. The opinions expressed in this publication do not necessarily reflect NIE position or policy, and no official endorsement by NIE or ED should be inferred.

THE IMPACT OF SOCIETAL BIASES ON RESEARCH METHODS

ABSTRACT

As members of society, researchers have been affected in both their lives and their work by the racism and sexism that is part of American society. The effects of these biases can be seen in the various elements of research methods, including the selection of topics, design, sampling, measurement, and the generation of conclusions. The purpose of this paper is to present both some of the ways that bias can affect the elements of research and some of the results that can occur from biased research.

Beginning with a discussion on the role of objectivity in science, the paper continues with a summary of how research has, in the past 150 years, been used to preserve the status quo in respect to sex and race equity. Relying heavily on examples, the ways that science has been used, historically, to justify sex and race prejudice and discrimination, are examined and analyzed.

Examined, too, are the ways that bias is currently affecting research. In the area of topic selection, potential sources of bias such as funding sources and publishing patterns are covered as are some of the effects of bias including the existence of holes in research knowledge. A discussion of the sources and effects of bias in design is also given. For example, how the knowledge of appropriate literature, the selection of independent variables, and the control of sources of invalidity can be sources of bias are covered. The effects of these sources in terms of incomplete testing for race and/or sex similarities and differences, use of biased independent variables, and incomplete blocking or control of confounding variables are all discussed, as are the results of this bias.

A similar pattern is used to discuss bias and sampling, covering the sources of bias (composition and selections of samples), the effects of bias (single sex or race samples, organizations, and the use of different race or sex samples for different content areas) and the results of that bias. The section on bias and measurement investigates the effect of bias on a variety of different measures including observations, aptitude and achievement tests, vocational tests, and affective tests. The section on bias and the generation of conclusions examines the way that bias can encourage overgeneralization, incorrect attributions of causality, and conclusions based on expectations rather than on data.

Following the discussions of the negative effects of bias in research methods, guidelines for reducing these effects are presented. The final section of the paper is a summary of the major points covered earlier and a discussion of future directions for the development of new methodologies to overcome and study bias in research.

ACKNOWLEDGMENTS

The author wishes to thank the following people for their reviews of the original manuscript. Their comments did much to improve the work.

Jessie Bernard, U.S. Commission on Civil Rights
Wiley Bolden, Georgia State University
Susan Chipman, National Institute of Education
Carol Dwyer, Educational Testing Service
Hendrik D. Gideonse, University of Cincinnati
Glen Harvey, National Institute of Education
Susan Klein, National Institute of Education
Greta Morine-Dershimer, Syracuse University
Myra Sadker, American University
Charol Shakeshaft, Hofstra University
Bernadine Evans Stake, University of Illinois
Barrie Thorne, Michigan State University
Linda Walter, Pennsylvania Department of Education

and those attending the Sixth Annual Conference on Research on Women and Education's session on bias in research.

CONTENTS

		Page
	Abstract	i
	Acknowledgments	ii
I.	Introduction	1
II.	Historical Use of Research to Strengthen and Preserve the Status Quo	5
III.	Bias and Topic Selection	11
IV.	Bias and Design	20
V.	Bias and Sampling	26
VI.	Bias and Measurement	33
VII.	Bias and the Generation of Conclusions	43
VIII.	Guidelines to Minimize the Effects of Race and Sex Biases on Research Methods	50
IX.	Future Directions	53
X.	Reference Notes	55
XI.	Bibliography	56

1. INTRODUCTION

In the past twenty-five years, there have been major efforts to fight prejudice and discrimination in American society. While these efforts have met, in large part, with success, equity for all members of society is still a goal for the future. Much of society is still biased and those biases affect the members of society in their lives and in their work.

The biases to which we have been exposed throughout our daily lives are also a part of our tradition as researchers. The purpose of this paper is to discuss the ways that these societal biases have affected social science research in the past and continue to affect it in terms of such seemingly objective activities as topic selection, sampling, measurement, design, and the generation of conclusions.

For this paper, bias is defined as a particular tendency or inclination, especially one that prevents reasonable, knowledgeable, thoughtful consideration of a question (Harmon, 1973). There are a number of biases which continue in our society including bias based on sex, race, age, and handicapping condition. This paper focuses on the effects of race and sex bias on research methods. These two areas were selected because, more than other areas of bias, enough work has been on these areas to begin an analysis of their effects on research methods. It is hoped that in the future other areas of bias will be examined as well.

Since bias prevents reasonable consideration of a question, one who is affected by race or sex bias will have difficulty dealing "objectively" with questions covering race and sex. Since objectivity is at the base of the scientific method and empirical research, it would appear that conceptually the presence of bias renders true research impossible.

Researchers are well aware of the negative effects that some sources of error can have on the quality of research and have made great efforts to control for them.

Lists of potential sources of error such as maturation, statistical regression, differential subject mortality, and pretest effect and their relation to research are found in basic research textbooks and in the lecture notes of most research professors. Students are warned of potential problems and designs are created and used to minimize those problems. Yet in spite of this great concern for some sources of error, little attention is paid to the effects of societal biases, particularly sexism and racism, on research. At best the work of people like Rosenthal on the "self-fulfilling prophesy" and the effect of "what the researcher expects to see" on shaping the data is mentioned, but not in detail (Issac, 1975).

Perhaps this lack of attention occurs because the researcher, the follower of the scientific method, is presumed to be "objective", unmoved by bias. Bias is frequently seen as something that can affect the subjects of social science, but not the author. This is unfortunate because, as Thomas and Sillen have concluded, researchers are not immune to the "disease and superstition of American racism" (1972). Too, as Gideonse has suggested, if a society is sexist, then so perhaps is its science (1977) and its scientists. The myth of the objective researcher is one which most of us believe; yet it is a myth.

Writers in the philosophy of science have long commented on the difficulty of attaining scientific objectivity. As Nagel (1961) wrote, "It is not easy in most areas of inquiry to prevent our likes, aversions, hopes, and fears from coloring our conclusions." Russell's (1959) comment is even stronger. "As soon as any strong passion intervenes to warp the experts' judgment, he (sic) becomes unreliable, whatever scientific equipment he (sic) may possess."

There is some empirical data to support the opinions of philosophers about the difficulties of attaining objectivity. Sherwood and Natanpsky (1968) found that scientists' conclusions as to the relative importance of heredity and environment in determining race differences in intelligence could be predicted reasonably well from biographical characteristics of the investigator such as the researcher's age when the research was published, family background, and education.

Even the most respected researchers have had blind spots related to racism and sexism. Former head of psychology at Columbia University, H.E. Garrett was convinced that the "scientific community had been blinded to the truth (of racial intellectual inferiority) by the duplicity of Franz Boas, Communists, Jews, and Sentimentalists" (1961). In the 19th and early 20th century, a whole genre of literature, written by the top scientists of the era, was devoted to treatises on feminine inferiority. The conclusions of these treatises were upheld more by the weight of an eminent name and a sweeping reference to the zoological world than by anything approaching what scientists generally considered evidence (Tolpin, 1972).

The myth of the objective researcher is one which keeps us from an awareness of the role that the researcher's own attitude can play in research. Other myths have also contributed to this lack of awareness, such as the myth that scientists and the public are open to new findings. While one would like to believe this, examples of the tenacity with which people maintain their beliefs must cause it to be questioned.

McDonald (1977) found, for example, that because people select from research only those parts which fit their preconceived notions, little that is new is accepted and frequently all that happens is that already established notions are supported. Once people have made a decision or accepted a judgment, it is difficult to get them to change. In a study of student behavior, half of the subjects in a class were told that a new instructor was a "warm person" while the other half were told that the instructor was "cold". Even though all the subjects saw the same instructor in the same setting, the ones who expected the instructor to be cold described the person as "cold", while the ones expecting the instructor to be warm rated the person as "warm". These results continued even when the subjects were told that the initial information they were given was arbitrary (LaBrecque, 1980).

Most social science, like most people, remains wedded to the status quo. Since much of the task of the social sciences is to describe reality, the probability of describing new realities is very low (Argyris, 1975). Yet, as Parlee (1975) has commented, the very fact that the data of a social science consistently confirm everyday beliefs and practices ought to prompt a critical look at the methods and procedures which produced these "facts".

For example, for over fifty years the theory that there were two kinds of female orgasms, vaginal and clitoral, was accepted by researchers and clinicians. Only fifteen years ago did Masters and Johnson (1966) prove that the theory was wrong. One wonders how an incorrect theory which could have been tested, was perpetuated for so long. How strong the pull toward the established, the status quo, must have been for psychiatrists not to have learned from their patients that there was no "double orgasm" (Weisstein, 1969).

The third myth in the trilogy is that research is value free. Values do however, play an important role in research. Gideonse (1977) has done an excellent analysis of the role of values in research. He explained that social science cannot be separated from the confounding effects of human values, time, and the phenomena of human consciousness. He then goes on to explain that:

1. Everything in social science research is susceptible to one's value premises.
2. The social scientist is always conceptually inside the system or phenomena being studied.
3. Every social scientist must adopt some vantage point for analysis, be it their own values or those of society.

Much of the basis of the objectivity of research is founded on the assumptions of 1) an objective researcher, 2) the openess of researchers and the general public, and 3) the value-free nature of science. Yet, as indicated earlier, these assumptions are myths. A purely objective social science is, as Nage (1961) informs us, a vain hope. Knowledge is socially distributed; what one takes for reality is determined by his/her place in the social structure (Long Laws, 1978).

The myths of research have legitimatized science's role in strengthening and preserving the status quo, particularly in terms of racism and sexism. As section II will indicate, science has traditionally been used as a buttress of racism and sexism. Succeeding sections will focus on specific ways that sex and race bias have influenced research methods including topic selection, design and data analysis, sampling, measurement, and the generation of conclusions. Also included will be guidelines to minimize the effects of these biases.

II. HISTORICAL USE OF RESEARCH TO STRENGTHEN AND PRESERVE THE STATUS QUO

In the 19th century, science was seen by many as "a liberating force against injustice and domination" (Ehrenreich & English, 1979). The evils of humanity were to be banished by science and as a leading engineer of the time explained, "the golden rule will be put into practice through the slide rule of the engineer". The creed for America was science and its high priests were scientific men. Scientists were, by definition, totally objective and above special interests of any kind (Ehrenreich & English, 1979).

Yet is spite of the claims of objectivity and justice, an analysis of the work of scientific men (in reality scientific white men, for few white women and almost no minority men or women entered the halls of science) showed how well science and the results of scientific studies buttressed the prevailing sexism and racism of the era.

Up to the middle of the 20th century, scientists repeatedly "proved" the intellectual inferiority of women and minority men. In 1840, Morton measured a small number of crania and proclaimed that since Blacks had less "gray matter", they were less intelligent than Whites (Stanton, 1960). George Romanes in 1892 tested and found that women's brains were lighter than men's. He then concluded that women and men could not be treated equally until their brain weights were the same. Brain weights, head sizes, and facial proportions were all used by scientists to "prove" the assumption that Anglo-Saxons were highest on the evolutionary ladder, followed by Northern Europeans, Slavs, Jews, and Italians, with Blacks trailing far behind (Ehrenreich & English, 1979; Thomas & Sillen, 1972). This, of course, pertained only to males; the Anglo-Saxon female was considered to be at the level of the Black male and few bothered to categorize women of other groups.

Based on this research, a leading European professor of natural history, Carl Vogt, concluded in the 1860's that..."the grown-up Negro (male) partakes, as regards his intellectual facilities, of the nature of the child, the female, and the senile white" (Haller & Haller, 1974).

Later in the 20th century, intelligence tests began to replace the measurement techniques of the natural scientist; however, the conclusions remained the same. In 1916, Terman concluded that after extensive testing using his revision of the Binet-Simon intelligence test, that a low level of intelligence was "very, very common among Spanish-Indian and Mexican families of the Southwest and also among Negroes. Their dullness seems to be racial" (Terman, 1916). Ignoring Binet and Simon's conclusions regarding the effects of environment and educational opportunity on test scores and their assumption that the children tested and compared should come from closely similar environments, Terman asserted that based on their test scores, children of such Mexican, Spanish-Indian, and Negro parents "are uneducable beyond the merest rudiments of training. No amount of school instruction will ever make them intelligent voters or capable citizens in the true sense of the word. Judged psychologically, they cannot be considered normal" (Terman, 1916). These "scientific conclusions" came not from the Klu Klux Klan, but rather from a well-respected psychologist and educator whose work on gifted children is still in use today.

Racial inferiority was "proven" using other techniques as well. In the 1840 census, data was generated to show that free Blacks were more prone to insanity than were those Blacks in slavery. Data from the census showed that the rate of "insanity and idiocy" among Blacks was eleven times higher in the North than in the South. This data was used by advocates of slavery as scientific confirmation of the natural inferiority of Blacks and the need for slavery to protect Blacks. Thanks to the work of Dr. Edward Jarvis, it was soon discovered that the data were false. Towns with no Black population were reported as having numbers of insane Black residents and, in the case of Worchester, Massachusetts, all the White patients of the local state hospital were described as Black. In spite of the inaccuracies of the data, they remained in the government books and for decades were used to show that Blacks were relatively free of mental illness in the slavery days (Stanton, 1960).

Incidents of purely fictitious data are rare, although after the discovery that the data of Dr. Cyril Burt on the inheritability of intelligence, were also imaginary, some researchers fear that the incidents may be less rare than one would like to think. It is probably safe to say, however, that most scientific conclusions are not based on deliberately falsified data, although they may be based on data which is not empirically verifiable. As Allport suggested in 1961, "it is easy to invent instincts according to one's own needs". Instincts abound in scientific treatises about Blacks. In an article for the Psychoanalytical Review, Evarts (1914) reported, based apparently on Theodore Roosevelt's account of an African safari, that Blacks in spite of being bereft of a moral sense do have a great "compensating gift...they all sing". A 1921 article by Bevis in the first volume of the American Journal of Psychiatry also allowed Blacks a compensating gift, mimicry. The article, "Psychological Traits of the Southern

Negro with Observations as to Some of His Psychoses" concluded that "All Negroes have a fear of darkness...are careless, credulous, childlike, and easily amused". Even as well known a personage as Arnold Gesell fell prey to the lure of developing "instincts" that reinforced the status quo. He explained that it is "well known that among the colored race there are many women who are supremely endowed with almost unique emotional equipment which makes their services ideal for infants and young children." (Gesell & Ilg, 1943.)

The examples can go on and on; however, the preceding seem to be evidence enough to verify that science has been used to support racist ideas. Science has also played a large role in perpetuating stereotypes about minority and majority women and in reinforcing whatever society had defined as women's position. As mentioned earlier, in the 19th century white females, like minority males, were seen as being lower on the evolutionary scale than white males. The white woman was seen as the more primitive version of man, whose body was not only primitive, but deeply pathological. Her mind, body, and soul were all felt to be in the thrall of her all-powerful reproductive organs. While scientists were in agreement about the power of her reproductive organs, there was, however, some discussion as to which of the organs were the most powerful. Holbrook observed in 1870 that it seemed "as if the Almighty in creating the female sex had taken the uterus and built up a woman around it" (Wood, 1973); but in 1883 Austin wrote that ovaries "give woman all her characteristics of body and mind" (Ardetti, 1974). Bliss (1870) added that "women's entire personality was directed by the ovaries and any abnormalities from irritability to insanity could be traced to some ovarian disease". For example, the environment might be the cause of tuberculosis in men; but for women, tuberculosis was a result of reproductive malfunction.

The assumption that a woman had little or no control over her body or her mind was an important one in scientific research up to the middle of the 20th century. In 1903, Weininger found women to have no logical or ethical sensibilities and Spenser (1893) argued that women's innate nature forced them to be either the protected or the prey of men. Woman's control over herself was found to be so minimal that even mothering behavior was thought to be regulated by the pituitary gland (Anthony & Benedek, 1970).

Research on women was such that if society wanted women in the home, not being schooled, then there was a scientific basis for women remaining uneducated and at home. Clarke, for example, in his book, Sex in Education or a Fair Chance for Girls, which went through seventeen editions, concluded that "higher education would cause women's uteruses to atrophy" (1873, 1972). Later, work with a more empirical basis showed that female students were pale, in delicate health, and prey to monstrous deviations from menstrual regularity. One 1902 study showed that 42% of the women admitted to insane asylums were well educated compared to only 16% of men "proving that higher education was driving women crazy" (Bullough & Bullough, 1973). Even G. Stanley Hall, known as the "father of psychology, wrote that the woman who used her brain lost her mammary function first and had little hope to be other than a moral and medical freak (1905).

The concern of Hall and Clarke for the future of "unfeminine" women is also reflected in a much later publication. Sturgis and Menzer-Benaron wrote in the introduction to their 1962 monograph, The Gynecological Patient: A Psycho-Endocrine Study: "We are impressed in particular by the dictum that much of the physical and mental ill health of the individual woman can be properly understood only in the light of her conscious or unconscious acceptance of her feminine role."

While some researchers were examining the affects of using her brain on woman's physical and emotional health, other researchers were looking at the woman at home. In this area, people such as Bowlby were overgeneralizing from research on institutionalized children to warn of the dire consequences of maternal deprivation in such crisis situations as "Death of a parent, Imprisonment of a parent, Social calamity-war, Famine and Fulltime employment of mother" (Bowlby, 1966). Even if the mother was home full-time, there still could be "partial deprivation" from a mother who did not provide "constant attention, day and night, seven days a week, and 365 days in a year" (Bowlby, 1966). The mother who devoted all her time to her child still had to worry specifically about the "psychotoxic diseases of infancy". Spitz (1965) found that there were diseases such as colic, eczema, and fecal play in which the mother's personality acted as the disease-provoking agent, as the "psychological toxin".

This research was, of course, based on middle-class white women. The work on the frailty of women and need for total maternal supervision of children was generalized neither to poor or minority women who needed to work in order to survive, nor to their children.

In the past, research has been used to perpetuate myths that keep people in their places. Too often the role of the scientific expert has not been to seek out what was true, but rather to pronounce on what was appropriate. That which was labeled science has frequently been the ideology of a racist/sexist society dressed up as objective truth. Some of the conclusions about intellectual inferiority and the power of women's reproductive organs mentioned earlier came about as a deliberate effort to preserve the status quo and to preserve privilege. More often they came about because of the ignorance of the researcher as to the effects a society can have on both researchers and their research. Most of these examples are both blatant and antiquated, but the problems and processes that caused them to occur are still with us today.

People, particularly women and minority men, have a number of reasons, some of which have been delineated here, for not trusting research and the proclamations of "science"; however, the process which has been used to justify discriminating against women and minority men is also one of the few ways we have to learn about and perhaps change the world around us. Only by an analysis of the weaknesses of research and research methods and by attempts to strengthen those problem areas can science have any hope of becoming that "liberating force against injustice and domination".

III. BIAS AND TOPIC SELECTION

At the 1980 Annual Meeting of the American Library Association, a publisher commented while schools and libraries influenced what was read, publishers and ultimately authors, had a much greater influence. That which is not written can never be read. Her comments hold as well for the impact of societal biases on research methods: while it is important to discuss the effects of bias in such areas as sampling and measurement, it is even more important to be aware of how bias affects what is published and, even more basically, how it affects what is researched.

Bias has a tremendous affect on what research questions are asked, how they are asked, and if they are studied. As Nage explained, "the things a social scientist selects for study are determined by his (sic) concept of what are socially important values. The student of human affairs deals only with materials to which he (sic) attributes 'cultural significance'." (1961). Traditionally the "cultural significance" of minorities and women, other than in a pathological sense, has been minimal, as has the amount of research being done about them.

There are a number of other influences on the selection of research topics, outside of, but related to, individual interest. The major ones appear to be 1) funding opportunities, 2) publishing opportunities, 3) priorities of professional organizations, and 4) the interests of mentors. Of the above, the greatest influence on topic selection, at least for large-scale studies, is the opportunity for funding.

In today's world, few institutions and even fewer individuals are willing or able to assume the financial burdens of doing research. Most researchers must look to private and public funding sources for support and thus match their research interests to the interests and priorities of the funding agency. In general, research for minorities and women on equity and other issues is less likely to receive funding from either the private or public sector than are many other, more traditional, areas. For example, in fiscal year 1979, the total research budget for the National Institute of Education (NIE), the research arm of the Department of Education, was $80,200,000, of which $5,100,000 or 6.4% was spent on women or sex equity. This figure included $3,200,000 which was spent not on research but on training minority and women researchers (Klein & Goodman, 1980). The figures for FY80 have $3,180,000 being spent by NIE on women or sex equity, including $2,700,000 on training minority and women researchers. For 1980 less than $500,000 was allocated for research on women and sex equity. In the entire Department of Education, only 0.22% of their FY79 resources were allocated for sex equity and sex equity or research on women is not even mentioned as a category or subcategory in the Department's FY81 budget (Klein & Goodman, 1980). In the foundation world, the situation is not much different. In 1976 about 0.6% of foundation grant dollars went to woman's programs or to issues of sex equity (Saario, 1980). The twelve million dollars that foundations spent in this area in 1976 included not only research, but also the development and the provision of services in education and other areas such as health and economics. Much more money has been spent in minority areas. In FY80 the Department of Education, other than NIE, spent $259,300,000 on language and ethnic minorities and $294,500,000 on race desegregation (Klein & Goodman, 1980). It must be remembered that most of this money is spent on development, training, and the provision of services, with very little being left for research. The small amount of funding can contribute to both a lack of research and to a lack of interest in doing research in areas dealing with minorities and women.

Another large influence on topic selection is the opportunity for publishing. Most researchers want or need to publish their work. Through publishing, research is disseminated, feedback is generated, and decisions are made about hiring, firing, and promotion. "Publish or perish" is still very much a reality in the academic and research and development worlds. The unpublished researcher is soon the unemployed researcher. Thus the research foci of journals are an important influence on topic selection. In general, research journals publish little research specifically related to minorities and women.

In an analysis of the articles published in five leading education journals from 1973 to 1978, Lockheed and Stein found that 303 or 13.5% of the 2,239 articles dealt with women, girls, and education (1980). The percentage of articles on women and education ranged from 5.0% of the articles that appeared in the Journal of Educational Measurement to 17.4% of the brief reports appearing in Child Development (Lockheed and Stein, 1980).

An analysis of the American Educational Research Journal (AERJ) produced similar findings. In 1978, three of the forty-four articles published focused on minorities and/or women in education. In 1979, three of the thirty-three articles published by AERJ were on minorities and/or women in education (Note 1).

In response to the apparent lack of interest in minorities and women in education on the part of "mainstream" journals, a number of "special area" journals such as Integrated Education, the Journal of Black Psychology, the Psychology of Women Quarterly and SIGNS: A Journal of Women in Culture and

Society have been established; however, the subscribers and presumably readers of "special area" journals are primarily minorities and/or women. Few majority male researchers read these journals and thus much of the work found in them does not find its way into the mainstream of research or knowledge (Note 2).

Closely tied to the influences of publishing opportunities on topic selection are the research interests and priorities of professional organizations. Since professional organizations publish many of the relevant social science journals, their priorities necessarily influence journal editorial policy and topic selection. Professional organizations' priorities help decide the focus for presentations at the organizations' national conferences. Presentations, like publications, provide opportunities for dissemination and feedback and can increase one's chances for promotion or tenure. The situation has changed somewhat, due in part to the increase of special interest groups and caucuses on minorities and women; however, in general research topics in minority and women's areas have not been high priorities for most professional organization.

The effects of racism and sexism on funding patterns, publishing opportunities and professional organizations have contributed greatly to a deemphasis on and devaluing of research for minorities and women. In part because of this devaluing and lack of interest, experienced people who teach new researchers "the ropes", do not generally guide them toward research dealing with minorities and women. Since professional leadership, recognition, and other rewards less often come to people working in these areas, even experienced researchers with such interests are loath to encourage their students and proteges to work in these areas. Many of the researchers who do work in these areas have found it necessary to "balance" this work with work in more highly valued areas (Campbell, 1980). In order to progress professionally, many researchers, particularly minorities and women, have found that their research cannot be totally or even primarily focused on areas such as racism or sexism. Too often the case is as

Steinem found during her tenure as a Wilson fellow:

> My own work on theories of gender based power was academically suspect as single-factor analysis while my neighbor's work on one man's military acts during one decade was thoughtful, scholarly and basic. (Steinem, 1980, p. 98.)

This devaluation also makes it difficult to develop and implement training in specialized equity research methods. Unlike other training areas, professional organizations are reluctant to arrange the training without outside funds and frequently question the need for, or even the existence of, specialized methods for equity research such as measures of equity in classroom interactions or methods which rely less heavily on statistical analysis. An exception to this was a conference held in 1980 by the National Institute of Education on "Attitudinal and Behavioral Measurement in Social Processes/Women's Research". This conference, the first of a proposed series on equity research methods, had as its focus, ways of expanding upon the range of measurement techniques in the Social Processes/Women's Research area to fill needed gaps and to avoid losing reliability and the capacity for meaningful comparisons with existing instruments (Note 3).

Sexism and racism affect the selection of research topics through the theories and theoretical constructs that form the basis of so much research. Much of the theory that underlies social science research is sexist and racist. For example:

> Arthur Jensen:
>
> The assumption of equal or equivalent intelligence across all human populations is gratuitous and scientifically unwarranted (1980, p. 370).
>
> Joseph Rheingold:
>
> ...woman is nurturance...anatomy decrees the life of a woman (1964, p. 137.
>
> Bruno Bettleheim:
>
> ...as much as women want to be good scientists and engineers, they want, first and foremost, to be womanly companions of men and to be mothers (1965, p. 30).

Research based on such theoretical foundations must necessarily be sexist and/or racist. Thus, as Shakeshaft (1979) concluded from her study of the content of dissertations in educational administration, "we have asked 'Why can't Johnny read not why can't Janey add?' or 'Are women feminizing our schools' not 'Are men polarizing our schools and causing them to become violent places?' ".

Biased questions have also been posed regarding race. We ask "Are Negroes as intelligent as white people?" (Bettleheim and Janowitz, 1964) not "how biased are our tests of intelligence?" or "Are Caucasians as intelligent as Black people?". When studying racism, we ask "What are the effects of racism on Blacks?" and not "What are the effects of racism on whites?"

As a result of these biases on research topic selection, "At a period when the public is demanding information and explanations about social inequity, the academic professions are unable to provide them, in large part because the relevant questions are not on their research agenda" (Long Laws, 1978).

The amount of research being done from the perspective of minorities and women has been minimal; however, the amount of research being done on minorities has been considerable. In 1968, Whitney Young commented in a National Observer article that the "Negro-studying business has become so big that I am afraid that if we just end it quickly, too many folks will be thrown out of work. I'd like to propose a study of white folks. After all, Negroes didn't create the ghetto, white folks did."

Young's comment touches on an important issue in the effects of bias on topic selection. Much of the research done on minority groups has come from a "blame the victim" perspective and focuses solely on the pathological aspects of minority life. As Quarles (1967) explained, "when we pick up a social science book (when) we look in the index under 'Negro', it will read 'see Slavery', 'see Crime', 'see Juvenile Delinquency', perhaps 'see Commission on Civil Disorders', perhaps see anything, except Negro. So when we try to get a perspective on the Negro, we get a distorted perspective."

This distorted perspective, with its emphasis on the negative and the pathological, can be seen in such social science classics as Liebow's (1967) study of the street corner society of unemployed Black males, Talley's Corner, and Lewis' (1961) anthropological study of the "culture of poverty" of a poor Mexican family, Children of Sanchez, as well as other less well-known work. Reviewing the literature on Black families, Billingsley (1968) found that it displays "a selective focus on the negative aspects", while literature on the "psychology of Blacks" fails to accept the fact that Negroes can see themselves in a positive light (Gullattee, 1969). As a result, "seen narrowly as a 'victim', the Black man appears in the learned journals as a patient, a parolee, a petitioner for aid, rarely as a rounded human being" (Thomas and Sillen, 1972).

When Thomas and Sillen used the term "Black man" in the above quote, it is doubtful that they were using it in the generic sense, for while there have been prototype images of Black males in the research literature, distorted or ethnocentric as they may be, there has been little about Black females or indeed about any minority females. As Lightfoot (1978) and Pollard (1976) concluded after surveying the literature, minority females have not been the focus of the research agenda of social science research and, in fact, are presented in the literature minimally, if at all.

An exception to this is Stack's All of our Kin: Strategies for Survival in a Black Community. Stack (1975) examines the world of women and children and the positive as well as negative aspects of life in the Black community.

The absence of information about minority females is not surprising. As Babladelis (1976) explained, "a significant portion of what is considered important to study is determined by male investigators" and male investigators, particularly majority males, rarely study females, particularly minority females. Babladelis is not alone in her conclusions. Researchers from education (Shakeshaft, 1979; McDonald, 1977), sociology (Millman & Kanter, 1975), and psychology (Acker & Van Houton, 1974; Kearney, 1979) have all commented on how the preponderance of male researchers in a male-dominant society has limited the selection of research topics.

The limitation of topics goes beyond a lack of interest in research for or on women to a lack of interest in and devaluating of research topics associated with women's lives and emphasizing those associated with men, or stereotyped as male.

The bulk of work in the social sciences has focused on phenomena and areas in which men dominate: territoriality, aggression, politics, and economics (Shakeshaft, 1979). Work on patterns of research, in early childhood education, has found that the majority of male researchers (who are themselves the majority of researchers) tend to study control of persons and institutions, philosophy, and methodology; while women tend to study the family, the role of women, and the development of young children (McDonald, 1977).

Sex and race stereotypes about what is appropriate behavior for people also affect topic selection. For example, investigating subjects' desires to remain teachers rather than to move from teaching to administration is seen as a study of deviant behavior rather than a study of different levels of aspiration (Shakeshaft, 1979). Similarly, other than some recent work on housework by people like Oakley (1975), unpaid housework and childrearing are not considered "work" and studies on the labor force or on working do not deal with the millions of full-time homemakers.

While not reflecting the dominant male model may be considered deviant, so may acting outside prescribed sex roles. As the Committee on the Status of Women in Sociology (1980) has commented, there is an emphasis in research on the problems of female-headed households and single-parent families, but there is an absence of studies of problems associated with two-parent families because, just as situations outside prescribed sex roles are defined as problematic, situations in which people conform to prescribed sex roles are assumed to be non-problematic. In the same vein, the problems of "working" mothers and the problems of children in child care are studied, not the problems of "non-working" mothers or of children who are at home full-time. The problems of "non-working" fathers are examined, but rarely are, at least in relation to the family, the problems of "working" fathers. Mother-child interactions are studied extensively, father-child interactions, rarely.

When research is done on areas that can challenge traditional gender arrangements, the results are frequently ignored and not followed up. For example, evidence of male hormonal, emotional, and behavioral cycles has been around for decades, but only recently discussed. The need to believe that women are changeable and unreliable while men are consistent and reliable has led psychologists and others to ignore the existing evidence of male cyclicity and refrain from asking themselves the obvious question, "Do men have monthly hormonal and behavioral cycles?" (Silveira, 1972). Other examples include the effects of racism on whites and the use of non-verbal behavior in dominance. On the other hand, research which reinforces stereotypes is encouraged, even when the results are shown to be incorrect. An obvious example is the studies of Black genetic inferiority which seem to follow a cycle of being conducted, being refuted, and then being conducted again (Kamin, 1973).

The result of these biases is that there are significant gaps in our knowledge base. Pettigrew's comment in 1964 that "Many of the most basic and important personality questions about Negro-Americans have not received even tentative answers" still holds true today.

Because of these gaps and the unwillingness or inability of researchers to fill them, decisions are being made, programs are bing developed, and services are being offered without an adequate research base. An examination of government and foundation funding for equity programs indicates that monies are more apt to be spent on direct service activities than on research, development, dissemination or evaluation. (Klein and Goodman, 1980). Thus, much of our response to sexism and racism and our quest for equal opportunity, particularly in education, has become a very expensive "trial and error" process. Without an adequate understanding of the problems being faced and without tested strategies based on non-racist, non-sexist theories to call upon, our educational and social programs have been doomed to failure. As bias in topic selection has contributed to gaps in the knowledge base, it has also allowed us to close our minds to testing alternative hypotheses or explanations for behavior that fall outside of our stereotyped expectations.

To challenge this existing structure is difficult, for, as Kenneth Clark has said:

> To raise serious questions and to doubt established practices, particularly in those disciplines concerned with man's [sic] ego and his [sic] relation with his [sic] fellow man [sic] is to invite personal jeopardy rather than professional reward.
>
> (Clark, 1972, pp. xi-xii)

But challenge this we must.

IV. BIAS AND DESIGN

Race and sex bias can affect research design and cause designs to be developed and implemented under which the data collected answer different questions than the ones the researcher intended to pose. Sources of bias in design include four areas: familiarity with the appropriate literature; selection of independent variables; controlling of sources of invalidity; and data analysis.

Race and sex bias can affect our knowledge of appropriate literature in a number of ways. As indicated in the previous section, one of the results of bias in topic selection is the existence of gaps in research areas, where information pertinent to a research topic simply does not exist. The devaluation of research concerning minorities and women may also mean that the research that is done is published out of the mainstream in sources that may be difficult to find and that are not routinely read by the majority of researchers. Even if the work is readily available, researchers may not see how work concerning, for example, sex roles or racial bias in testing, could affect their work on achievement motivation or self-concept.

Lack of researcher awareness of relevant literature can cause the selection of independent variables to become a source of bias. For example, socio-economic status is often used as an independent variable with a mixed sex sample. This occurs even though common methods of categorizing people according to socio-economic status use the status of the husband of father as tne determinant of the woman's status (Acker, 1973; Edelsky, 1979).

Biased definitions of race can also be used as independent variables. For example, two non-identical concepts, social race and biological race, may be lumped together and a common racial label used for individuals of markedly different backgrounds: Children of a Black parent and a White parent may be identified and studied as Black (Harris, 1968). Jensen's controversial study of race and intelligence used race as an independent variable, categorizing subjects socially identified as Black as biologically Black regardless of their ethnic and racial background (1969).

A third source of bias can be found in the control of potential sources of invalidity. Obviously if one is not aware of, or sensitive to, possible threats to validity, it is impossible to control for them. For example, in many early studies of racial differences, the effects of socio-economic status (SES) were not controlled and the SES of the majority group sample was higher than that of the minority group to which it was being compared (Pettigrew, 1964). Frequently, neither are other variables such as educational level controlled in cross-race studies. Even when differences in educational levels are controlled, differences in the quality of schooling, which are particularly important because of the existence of dual school systems for minorities and majorities, are not.

Finally, the analysis of the data can itself be a source of bias. Data which does not fit the expected pattern can be thrown out and not used. For example, this occurred in studies of sex roles in chimps, work that has often been used as evidence of the genetic basis of sex roles. Yerkes (1943) did not use data from a number of animals when the results were "statistically disappointing". The data not used included the data for two female chimps "because they were highly dominant."

An effect of bias in design is that the presence of sex similarities and differences is frequently not studied. In studies published in 1970 to 1971 in psychology journals, Schwabacker (1972) found that only 50% of the mixed sex samples were checked for sex differences. By 1974, this proportion had increased to 61% (Reardon & Prescott, 1974). A 1972 study of 754 studies in psychology journals found that 35.4% of the samples had been checked for sex differences. It was also found that women researchers (61%) were more apt than men (50%) to check for sex differences (Harris, 1972). This finding has been challenged by a later report that women and men researchers were equally likely to test for sex differences (Etaugh & Spandikow, 1979).

It is suggested by Harris (1972) and Etaugh and Spandikow (1979) that sex differences should be checked and results reported in all mixed sex samples. While there are some excellent reasons for this suggestion, including the possibility of finding new and valuable information about sex similarities and differences, there are also some problems associated with this suggestion. Merely analyzing the data using sex as an independent variable, without previous knowledge of possible confounding variables, can lead to inaccurate conclusions. As Maccoby and Jacklin (1974) have commented, "the appearance of a sex difference often depends upon detailed aspects of the situation in which behavior was studied - details that have so far gone unrecognized."

A similar pattern exists with the testing of racial similarities and differences. Frequently racial differences are neither analyzed nor mentioned.

For example, Wylie's (1961) survey of the literature in self-concept did not mention even one study which considered race as a factor associated with self-concept. When racial differences are examined, the problems with such analysis may be worse than if no analyses had been done. The lack of knowledge of many researchers about minority groups, the subjective nature of the definitions of racial groups, and the frequent lack of controls to insure that the racial comparisons that are being made are made on equivalent groups, have caused many inaccurate conclusions to be drawn in areas such as Black self-concept (Nobles, 1973) and the Black family (Thomas & Sillen, 1972). This generation of inaccurate information has been so prevalent that some minority researchers have called for a moratorium on the analysis of mixed-race samples for race differences (Note 4).

Perhaps the area most rarely tested is the interaction of race and sex. The lack of dual-sex, multi-racial samples and the apparent lack of concern of researchers with race and sex interactions, even where it would seem to be appropriate, is evident. For example, studies of the effects of desegregation and racial isolation rarely tested for sex differences. When sex and race differences are analyzed, it is by comparing the interactions among girls of different racial backgrounds to those among boys, rather than examining race/sex interactions (Campbell, in press). Weinberg (1977) hypothesized that, in this area, sensitivity about interracial romantic liaisons and the widespread fear of racial intermarriage during the first years of school desegregation contribute to the refusal to examine cross-sex, cross-race interactions.

Another major effect of bias on design has been the lack of use of existing information about minorities and women to increase the validity of the study. For example, Caplan (1975) found that the presence of an adult investigator caused boys to become more anti-social, while the absence of an adult was more conducive to finding no sex differences in anti-social behavior. A study of anti-social behavior, designed without knowledge of this information and without necessary controls, could then lead to an inaccurate conclusion about sex differences and a reinforcing of stereotypes about boys' anti-social behavior. Based on her work, Caplan (1975) concluded that the design of a study may play a greater role in producing or abolishing sex differences in behavior than do real sex differences.

Her conclusions have been supported by several other researchers. For example, Serbin and O'Connor found that children were more apt to play with toys considered appropriate for their sex when there were others in the room, than when they were alone (Greenberg, 1978). Thus, studies of children's play behavior carried out with others in the room may be apt to find "sex appropriate" behavior.

In their extensive review of the literature, Maccoby and Jacklin (1974) reported a number of findings that could affect studies of sex differences. For example, they found that college men performed better when observed by peers, while college women's performance did not change. They also found that boys persisted longer in a task with a boy watching than when an adult was watching, while the age and sex of an observer had little affect on girls. Also, different situations were found to affect boys' activity levels and speed of work, while not affecting girls'. Not knowing, or controlling for, this type of information, particularly in studies of sex differences, means that the independent variable being studied may not be the subjects' sex, but rather is the interaction of the situation with the sex of the subject.

The order in which persons are rated can also have an effect on the validity of the study. Intons-Peterson (1980) found that when subjects rated a male first and a female second, they tended to rate them using stereotyped characteristics; however, when they rated the female first, they tended to give her more male-positive and female-positive characteristics than they did the male they rated second.

Studies which have been designed with an eye toward controlling subtle confounding variables have found different results than did studies without these controls. For example, work in child development has shown that middle class White mothers tend to interact with their children in ways that child development experts note as positive. Lower class Black mothers have been viewed as interacting with their children in ways that have been labeled as inadequate. However, a study by Graves and Glick found that when White middle class mothers did not believe that their behavior was being observed, they behaved in ways similar to the lower class Black mothers. In other words, the major difference in parenting behavior between the two groups was not that White middle class mothers had more appropriate parenting behaviors than Black lower class mothers, but rather that the lower class Black mothers did not change their behavior under observation, while the others did (1978).

While more work needs to be done in this area, the Graves and Glick study does indicate that a rethinking of sources of error and a willingness to develop and test alternative hypotheses may cause different results to occur. Another example comes from a study of motor skill development. Wilmore found that prepubescent children of both sexes were about equal in many athletic skills, except the softball throw. Rather than assuming a sex difference, he repeated the tests with the subjects using their non-dominant arms. The results indicated performance was about the same (Kaplan, 1979). Practice made a difference, which, if not controlled for, could cause inaccurate conclusions to be drawn.

Another effect of bias on design can be the lack of "blocking" across racial or sex groups on potentially confounding variables such as socio-economic status. The Graves and Glick study on parental behavior is an example of this. The two groups being compared differ both on race and socio-economic status and, as usually happens, the minority group is from a lower class than the majority group. While not the case in the Graves study, in many studies differences in behavior are then concluded to be racial differences when in reality the differences may be related to socio-economic status, the interaction of race and socio-economic status, or any of a number of variables related to socio-economic status. Geographic area is another variable that is frequently not controlled. For example, Pettigrew(1964b) found a tendency for researchers to compare Northern Whites, Northern Blacks, and Southern Blacks, but not Southern Whites.

Controlling for these variables is not enough if some of the variables are themselves biased. The most obvious of these, as mentioned earlier, is socio-economic status. Nichols describes the ways that socio-economic status can be biased and the way that the bias can affect research in one area, language, in the following passage:

> Unfortunately common methods of categorizing people according to SES use the husband's or father's status as the determinant of the woman's status even though the woman may have more education than the man in question. In addition "stenographer" and "mechanic" may be classified as same status occupations. The result is that women are often misclassified because of a bias in the methodology and are found to use different language than men. What those language differences may actually reflect in some cases is the fact that women are being compared to men of presumably the same, though in reality lower, class status and also to men whose jobs are likely to be less language oriented (1978, p. 6).

The major result of bias in design is the generation of inaccurate information and incorrect conclusions based on that information. Particularly in areas involving sex and race roles and differences, bias in design has influenced research toward stereotypic expectations and away from examinations of complex realities - a dangerous direction for both research and researchers.

V. BIAS AND SAMPLING

The composition and selection of samples and the generalizations made from those samples can all be affected by race and sex bias. In the past several years, a number of researchers from different fields have commented that traditionally males have been the population studied in social science research. Babladelis (1976) wrote that "a significant portion of our knowledge is based on the study of male behavior only." Her conclusion, which related to research in psychology, has been supported by other psychologists (Kearney, 1979; Long Laws, 1978) as well as by researchers in other social science areas, such as education (Shakeshaft, 1979), anthropology (Carey, 1979), and sociology (Millman & Kanter, 1975).

The concerns of these researchers have a basis in fact. Of the 226 studies published by the Journal of Personality and Social Psychology in 1968, 31% studied males, 15% studied females, 44% had mixed sex samples, and 10% did not indicate the sex of the sample (Carlson, 1971). A 1970-71 replication of 199 studies found 27% male samples, 7% female samples, 53% mixed sex samples, and 14% sample sex not indicated (Schwabacker, 1972). A further replication of 99 studies, in 1974, found 12% male samples, 28% female samples, 43% mixed sex samples, and 15% sex unspecified (Reardon & Prescott, 1977).

At least for this journal, a number of single sex samples have been done. Previously, these samples were disproportionately male; however, the most recent data found the single sex samples to be predominantly female. This change does not appear to be reflected in other social sciences where researchers have indicated that in fields such as mobility (Steinmetz, 1974) and management and careers (Jelnick, 1978), studies have focused almost entirely on male subjects.

A somewhat different pattern seems to emerge for minority group members. There have been a large number of studies done on Blacks focusing on areas such as "the Black experience" or "the problems of Blacks in society" (Thomas & Sillen, 1972). Much less frequently, however, have Blacks been found in the samples of studies of "human behavior". The presence of other minority groups, such as Hispanics and Asians, is even more rare (Stivers & Leckie, 1976).

An analysis of the samples studied in articles published in the American Educational Research Journal in 1978, found that while researchers frequently indicate the sex of their samples (69%, or 29 studies), they rarely indicate the race of their samples (23.8%, or 10 studies). (Note 1.) Of the ten studies that did indicate the racial breakdown of their samples, two samples were all

White and two did not use all of the minority data that they collected in their analysis. One of these studies did not use any of the 17% "non-Caucasian subjects" in order to "homogenize the sample" (Smith, Zingale & Coleman, 1978), while the other used data from Black and White subjects, but did not use the data from other minorities because the sample sizes were small and the results would "complicate interpretations" (Peng & Fetters, 1978). Of the remaining six studies, one dealt with desegregation and one with ethnic patterns. Only four of the forty-two studies published in 1978 indicated that they had mixed race samples on studies of general educational interest. The lack of information makes it impossible to draw conclusions about the racial composition of the samples and the degree to which minority group members are included in general studies; however, the very lack of information about racial breakdown is in itself significant.

A number of widely known research studies have been done on single sex and/or single race samples. For example, Bettleheim and Janowitz's (1964) study of Social Change and Prejudice was done on an all-male, all-White sample. Although the conclusions from this study have been generalized extensively, the study itself used no minority subjects in order to, as the authors stated, "simplify the findings".

McClelland's work on achievement motivation also used single sex subjects. There is some indication that McClelland and his co-workers were aware that sex differences probably existed in achievement motivation; however, they neither pursued these possible differences or expanded their samples to include females. Neither did they specify that their theories or conclusions applied only to men (Atkinson, 1958). As a result, knowledge of achievement motivation in women and girls has been spotty and inaccurate. Only in the past ten years has there been an effort to reconcile the realities of women's and girls' achievement motivation with the theory. And indeed, more recently, McClelland (1975) concluded that "sex role turns out to be one of the most important determinants of human behavior."

Moral development is another area in which all-male samples were used as the basis of theory development. Kohlberg's six stages of moral development [illegible] empirically derived from a longitudinal study of 84 American boys. Although the stages were felt to be universal (Kohlberg & Kramer, 1969), in order to reach the higher stages, it was necessary to "enter the arena of male activities (Gilligan, 1980). Women rarely reached the higher stages (Gilligan, 1977).

Gilligan (1977) began developing constructs in moral development from women's own lives, rather than trying to fit them into an existing model. She found an outline of moral conception different from that described by Freud, Piaget, or Kohlberg; one that would never have been found in all-male samples.

Another example can be found in the area of management and careers. Until recently, research on management and careers typically examined White, middle-class male subjects. Surveys on educational levels might cover females and males; however, when occupations and other work-related variables were examined, the subjects were primarily male. For example, in her analysis of the "Relation of Education and Situs of Work in Economic Differences Between Blacks and Whites" Gottfredson (1977) used a sample composed of 20,000 White men, 1,500 Black men, and no Black or White women. Similarly, in 1978 Blau and Duncan published a massive study of occupational structure based on a survey of 20,000 men (race unspecified).

More recently, a major study of occupational changes at midlife, currently being completed, has used a sample of 370 midlife males, but generalizes its results beyond the sample to "individuals" (Osherson, 1981). As Jelinik (1978) concluded, patterns and norms brought to light by this type of research are not necessarily applicable to a wider population.

Although such results are not necessarily applicable to wider samples, authors frequently do try to generalize them to the general population. An analysis of single sex studies found that in 1970 to 1971, 92% of the studies with male subjects and 61% of the studies with female subjects, published in the

Journal of Personality and Social Psychology, were generalized to both women and men (Schwabacker, 1972). By 1974, 97% of the studies with female samples and 92% of the studies with male samples were generalized to both sexes (Reardon & Prescott, 1974). Other work, in 1972, found 92% of the studies with male subjects and 62% of the studies with female subjects were overgeneralized (Dan & Beckman, 1972).

The lack of information of the racial breakdown of samples makes it difficult to assess the amount of overgeneralizing that occurs in single race studies; however, the two studies, in the 1978 American Education Research Journal, that indicated that they used single race samples, did overgeneralize. The larger number of studies (32) that did not indicate the racial breakdown of the sample did not caution the reader about the problems of overgeneralizing their results to those groups not included (Note 1).

Another problem in sampling is the tendency to generalize from skewed samples. For example, Kardner and Oversey wrote The Mark of Oppression as a study of the effects of prejudice and discrimination on Blacks, but it was based on a sample of twenty-five people who were either in therapy or who desired to enter therapy. Black Rage, Grier and Cobbs 1968 book on the effects of discrimination on Blacks, was based on data drawn from a sample of psychiatric patients, but, again, was generalized to the Black community as a whole.

Little work has been done on why researchers chose skewed samples; however, there has been some work on why researchers study single sex subjects. Prescott and Foster (1974) found that the reasons sixty-seven researchers gave for working with single sex samples fell into three major categories: (1) scientific, (2) practical, and (3) extra scientific. "Scientific" reasons given were, for example, "sex differences were known to exist in the phenomena and the investigator did not wish to explore them" and "the theory being studied was restricted to one sex." The sex of the subject pool and the need to keep the number of subjects to a reasonable size were given as the "practical" reasons for using single sex samples while the "extra scientific" reasons were that the use of one sex reduced the variability of the data and that the experimental conditions favored the use of one sex over the other.

No comparable study has been done on those who use single race samples in their research; however, the comments discussed earlier of researchers who did not use minority data in order to "homogenize the sample" or because it would "complicate interpretations" suggest that the categorizations of reasons for using single sex samples may be used as well for single race research. Since it appears that many, if not most, of the single sex and/or single race studies are generalized to other gender and racial groups, the reasons given for limiting the sample seem very weak and, particularly in the case of the extra scientific category, pose a serious threat to the validity of the studies.

Just as there appears to be a tendency to use minority samples in minority topic areas and white samples in more general topics, the use of single sex samples seems to be related to the topic being studied. McKenna and Kessler (1974) found that while in general men were more apt than women to be studied, the sex of the sample depended, to some extent, on the topic. If a research topic was in an interpersonal area, researchers were more apt to include women than they were if the topic were, for example, on aggression or on careers. Also, women have been more apt to be included in samples of research on topics stereotypically considered feminine, such as parent-child interaction (in reality mother-child interaction) (Condrey & Condrey, 1976) and on questions dealing with children, child-rearing, and work around the house (Steinmetz, 1974).

Studies about the family and its economic and social status have, however, been done almost totally on male samples. The male "head of household" has been the basic unit for much data collection on the family. If, for example, the husband was a bricklayer, then the family was categorized as blue collar, regardless of whether the wife was a teacher, a waitress, or unemployed. The Census, source of much of the data for studies on families, has encouraged the use of male only samples. In 1850, the Census dealt only with "male persons over fifteen". By 1970, the Census had changed considerably, but "for ease of tabulation", the male was still considered head (and representative) of the household (Steinmetz, 1974)*. This step for ease of tabulation helped to continue to encourage the use of male only studies on families.

Sampling procedures can also contribute to the selection and use of skewed samples. For example, the Gallop Poll has a very well-developed sampling procedure to gather information on its sample of 1500 "representative" subjects; however, male subjects are interviewed beginning at 6 PM, while female subjects are interviewed beginning at 4 PM (Gallop, 1980). Most of the female interviews are conducted from 4 to 6, thus eliminating the responses of most professional "working women" who would not be home by 6 PM (Note 5).

Except in stereotypically feminine areas dealing with care of children and of home, men and boys have traditionally been the population studied. And, except in the case of research specifically on minorities and on pathological areas such as cultural deprivation, delinquency, and the effects of discrimination, the men and boys who have been studied have been white.

A major result of "ghettoizing" the work done on minority and women samples has been that minorities and women have been viewed as the "deviant", that which confounds the results. As Long Laws (1978) explains, social science "takes for granted that the human we seek to understand is a heterosexual white male". Thus, if white males are the norm, then research on general areas of interest can be done over white male samples. The results can then either be generalized to other groups or the other groups can be studied in terms of the ways that they deviate from [illegible] Minorities and women then become a topic, something to be studied as opposed to a sample to be used to study some general phenomena.

---

* The 1980 [illegible] respondents indicate, themselves, which household member was "head of [illegible] years of assuming that the head of household was male and that fema[illegible] household indicated male absence may cause this step to have little impact.

Another effect of bias on sampling is the inaccurate attribution of characteristics to one sex based solely on the study of the other sex. Sections from studies of one sex frequently conclude, without data, that the other sex is either equivalent or is opposite. For example, Parlee (1975) found, in one article, two instances where studies quoted as "proving" sex differences had only male subjects.

Nowhere is the effect of bias on sampling more evident than in the newly popular and growing field of the life cycles or stages. Beginning with Erikson's work on the "Eight Stages of Man" (1959) through Levinson's Seasons of a Man's Life (1978), the study of life cycles has focused on male subjects. When women are examined, it is in terms of how they fit or don't fit the male model. Without empirical verification, women are said to go through the same cycles as men (Stewart, 1977) or are said to go through cycles that are antithetical to men's (Sheehy, 1977). The samples are overwhelmingly White as well as male, with racial differences rarely even being theorized.

Based on a survey of the literature on life cycles, Sanguilano concluded that "Mostly we (researchers) persist in seeing her (woman) in the reflected light of men" (1978). Her conclusion seems to hold for women and for minority men in a number of other social science areas as well.

VI. BIAS AND MEASUREMENT

While many researchers are not aware of the effects of race and sex bias in such areas as design and sampling, the controversy and debate about race and ability testing has made a number of educators and researchers sensitive to the existence of race bias in testing. The development of tests by White, middle-class authors, which are standardized and normed on White middle-class students and then used to make decisions about minority and poor children is a problem that has been recognized by many test publishers and users. Fewer people are aware of the effects that sexism can also have on ability testing and the negative influences that both race and sex bias can have in other areas of testing such as vocational testing, affective and personality testing, observations, and other unobtrusive measures. Since tests and measures are the basis of data collection for research, bias in measurement has a large effect on research.

One of the largest areas of testing for research and evaluation purposes is in achievement and aptitude. It appears that sex bias in the content of tests in these areas can affect performance. Milton (1957) found that females did better on test problems dealing with stereotypically feminine activities than they did on problems with a more stereotypically masculine orientation. This finding was substantiated by more recent studies which found that girls tended to do better on questions dealing with human relations and that boys tended to do better on questions dealing with science and economics (Coffman, 1961; Donlon, 1971).

The number of females and males appearing in a test item may also have an influence. A 1977 analysis of four achievement tests found a tendency for adolescent girls to be more apt to get an item correct if more girls than boys or if equal numbers of boys and girls were mentioned (Donlon, Ekstrom & Lockheed, 1971). Donlon (1971) has also suggested that changing the content of items on the math section of the Scholastic Aptitude Test to include more subject matter of greater familiarity to females had the potential to reduce the sex difference in test scores by about twenty points.

The influence of item content on performance seems primarily to affect adolescents and adults. To date, research on younger children has been mixed; findings range from effects similar to those found in older students (Montemayor, 1974) to no effect (Plake, Hoover & Loyd, 1978) to a reverse effect, with Black girls doing better on traditionally masculine contents and Black boys doing better on traditionally feminine contents (King & Blount, 1975).

Format has also been shown to affect the performance of females and males. Murphy (1977) found that switching from essay or fill-in-the-blank questions to more objective items such as multiple choice produced higher scores for males. This was found to be true in a wide range of subjects, including those stereotypically considered masculine, such as math, and those stereotypically considered feminine, such as reading comprehension.

"Sex differences" can be created or eliminated through the selection of items to be included in a test (Dwyer, 1976). A test can be slanted toward either sex or balanced to assure sex equity through the format (multiple choice, fill-in-the-blank, essay) or the proportion of stereotypically masculine, feminine, or neutral contents of items selected to make up the test (Campbell & Scott, 1980).

While a number of tests may be unconsciously "balanced" or "unbalanced" by sex, intelligence testing is one of the few testing areas in which some tests have been designed so that females and males will score approximately equally. These "balanced" IQ tests may then be used in research on sex differences in intelligence. The results of this research would not provide accurate data on sex differences in intelligence; results would be biased by one of the original intentions of the test - to equalize female and male scores. Studies using unbalanced tests in other areas can also contribute to inaccurate results. For example, sex differences found in mathematics may be real differences or may just be indicators that the test was slanted by item content or type to favor one sex.

It is interesting to note that IQ tests are not balanced for race equity the way that some are balanced for sex equity. The assumption that minority groups and the majority group have equal intelligence is not made and tests are not developed to reflect an assumption of racial equity.

Racism in tests occurs primarily because tests are the products of the majority culture. The test language, content, illustrations, and even scoring procedures may be drawn from a culture that is foreign to, or at least less familiar to, many minority children (Campbell & Scott, 1980). Cultural background, geographic isolation and low socio-economic status often combine to provide the minority child with a frame of reference very different from that envisioned and expected by test developers (Anastasi, 1976).

Many minority children face the obstacle of unfamiliarity with the settings and concepts used to test a wide variety of skills from reading comprehension to problem solving. In addition, minority children's scores may suffer because of the difference between the language that they speak, be it Spanish, French, or Black English, and the language in which the test is written. Thus, the language structure and syntax of tests are unfamiliar to many test takers (Samuda, 1975). Also, many words have different connotations in different cultures. The hawk, for example, is the wind in Black English and a large bird in standard English; the Spanish word "embrazada" looks very close to the English word "embarrassed", but means "pregnant" (Campbell & Scott, 1980).

Test scores for minority students combine actual skills in an area with a number of other factors and influences. As Tyler and White (1978) have concluded, "unless the material used for these purposes (testing) is equally familiar to all cultural groups, differences in performance are uninterpretable. The difficulties of achieving 'equal familiarity' in this sense are so formidable as to make the idea of culture-fair tests appear unrealizable and perhaps unreasonable." Unfortunately, many researchers do not heed Tyler and White's advice and continue to generate conclusions about racial achievement and aptitude based on tests which may be inequitable.

The effects of bias in measurement go beyond aptitude and achievement tests. In the past, vocational interest inventories severely limited test takers' job choices by sex. Interest inventories had separate questions and there were even separate tests for females and males. Even when the same person took both forms of a vocational interest inventory, the feminine form indicated the person's interests were in low-paying, low-status job areas, while the masculine form indicated the person's interests were in high-paying, high-status job areas. Routinely, females were given job options such as nurse, executive secretary, and stewardess, while males with similar interests were given job options such as physician, business executive, and airline pilot (AMEG, 1973). The implementation of Title IX of the Education Amendments of 1972, which prohibits the use of vocational tests suggesting one set of occupations for males and another for females, helped improve this situation. Now, as Tittle and Denker (1976) conclude, "many of the characterisitics of vocational interest inventories that caused charges of sex bias in interpretive materials and gender-linked items are being eliminated."

Although these changes have been made, much of the research on career aspirations was done using the older biased tests. Results based on the older tests may have been more influenced by the biases of the tests than on the career interests and motivations of the subjects.

Bias can also have an impact on personality tests. For example, social value scales base their scoring on a knowledge of common social values. The correct answer is the common social value, while an incorrect answer is assumed to be based on ignorance of common social values. Answers that are not based on ignorance, but rather on opposition to common social values or to differences in situations, are still considered incorrect and are not differentiated from answers based on ignorance (Jorgensen, 1973).

Cowan, Watkins and Davis (1975) found the Minnesota Multiphasic Personality Inventory seriously overpredicted the incidence of schizophrenia in less well-educated Blacks (less than 12 years of formal education). Nine members of a group of twenty adult Blacks, all functioning members of society who had been examined and found not to be schizophrenic, were classified by the MMPI as schizophrenic. Minority group members, not well-assimilated into the majority culture - either by circumstance or by choice - in general, do not fit the "normal" profile of personality tests, normed on primarily majority populations. In 1973, Brazziel sought a suspension of the use of tests normed on all-White populations, feeling that their use contributed to a further documentation of Black pathology in a "different equals deficient" model.

Bias has affected personality tests in somewhat unusual ways. McWhinni, in his study of the use of figure drawings to assess racial and cultural differences and self-concept, questioned whether:

> Asking a Negro child to draw with a pencil on white or cream paper and rating the resultant drawing with respect to the degree to which Negro features and skin color are portrayed is not exactly the equivalent of asking a white child to draw a person with white chalk on black paper and then rating his [sic] drawings on the degree to which he [sic] portrayed white children, especially skin color (1972, p. 30).

In a pilot study, he found that not one of the twenty-five Black and three White children asked to draw a person on black paper with white chalk, filled in the skin color to make the person white. In a separate study, only two of twenty-six White children asked to draw pictures of their families shaded the skin to make the pictures white (McWhinni, 1972). While these were only pilot studies, the results do indicate the possible existence of hidden biases in tests and measures.

Dress, or differences in dress, may be another hidden bias in measures. For example, if boys and girls are dressed differently, studies of children at play may, in reality, be measuring differences between playing in dresses and playing in pants rather than measuring sex differences in aggression or levels of activity.

Unawareness of different perceptions of reality based on living in a biased society can affect personality test development and interpretation. Personality tests such as the Rotter Internal/External Locus of Control, rather than tapping a subtle and pervasive dynamic dimension on which females and males differ, may be elliciting a realistic appraisal by minority and majority women and men of the extent to which their own efforts determine whether they succeed or fail (Parlee, 1975). This measuring of reality, rather than a personality trait, may also be the case in the use of Black and White dolls to measure children's racial attitudes and self-concept (Pettigrew, 1964b). Giving Black dolls the lesser choices in areas such as jobs or housing may be a reflection of societal reality rather than low Black self-concept or individual White racism.

Reality may also be responsible for racial differences in scores on the femininity /masculinity scale of the MMPI. Pettigrew (1964) suggests that Black males scored more feminine on the scale than White males, because those characteristics necessary for the oppressed to survive were also those characteristics considered feminine. One might extend Pettigrew's argument and conclude that the oppression of women may also have had an influence on the development of some of those characteristics labeled feminine.

The effect of sex bias on masculinity/femininity scales goes beyond the influence of oppression and discrimination on personality characteristics. These primarily bipolar scales more appropriately describe what society thinks about differences between the sexes than reality (Tresemer, 1975). In these scales, masculinity and femininity are seen as totally opposing concepts, with differences between women and men defined as distinctions between masculinity and femininity. Tresemer (1975), carrying this concept to its logical conclusion, asks ironically if since women's buttocks are larger than men's, does this mean that the larger the buttocks, the more feminine the person?

While Tresemer's question is absurd, some of the questions on masculinity/femininity scales are no less absurd. For example, the California Personality Inventory scored girls who don't fear thunderstorms and who don't want to become librarians as low in femininity, while boys with no desire to drive a race car or read Popular Mechanics, were low in masculinity (Vincent, 1966).

The Minnesota Multiphasic Personality Inventory contains one of the most widely used femininity/masculinity scales. The scale was validated by a study of 17 male homosexuals, 108 female workers, and 117 male armed service personnel. Items that discriminated between the armed service personnel as a group and the females and gay men as the other group were used as the basis for the scale (Tresemer, 1975).

Masculinity/femininity is assumed to be a bipolar, unidimensional, continuous, normally distributed variable that is highly important and consistently viewed within the sample population (Tresemer, 1975); yet this is not the case. Women and men are not polar opposites. Definitions of characteristics of women and men are neither as simplistic nor as continuous as most scales assume. The assumptions inherent in most scales of masculinity/femininity appear to be a better example of Archibald's Law that "if the shoe doesn't fit, there is something wrong with the foot" (Caplan & Nelson, 1973) than it does of reality.

Sex role instruments, many of which were developed in response to the biased properties of masculinity/femininity scales also have problems. Thomas (1978) analyzed fifty different measures of sex-typed personality traits used in social research and found "little concern with validity. A new type of validity seemed to be used with sex role measures: declarative validity; if the authors say the measure is valid, it is."

Observation is also affected by bias. Silveria (1972) commented about the work of R.M. Yerkes that "his bias determined his 'observations' and his observations were used to support his bias." Her conclusion holds for a number of other researchers as well. Ratings may reflect the observer's perceptions of what is appropriate behavior for females and males rather than reality. In studies of small children, where one group of observers is told that the subject is female and the other group that it is male, different observations result (Herman & Serbin, 1977; Gerwitz & Dodge, 1975). Condrey and Condrey (1975) found that "the same infant in a particular situation was seen as displaying different emotions and significantly different levels of emotional arousal, depending on the sex attributed to the infant, the sex of the rater, and the rater's experience with young children. Men with little experience with young children were the group most apt to rate the child the same, regardless of attributed sex; while men with a large amount of experience with young children were the most apt to rate the child differently based on the sex attributed to the child."

In a similar study, Meyer and Sobiesgik (1972) determined that observers tended to describe the child with the observer's sex as having more qualities than an opposite-sex child. Rothbart and Maccoby (1966) found that when fathers listened to a tape of an infant, they rated it as being more dependent and aggressive if they thought the tape was of a girl than if they thought it was a boy. Mothers rated the tapes in just the opposite fashion, rating the perceived boy as being more dependent and aggressive than the perceived girl. Maccoby and Jacklin (1974) suggested that observational studies of dependency showed no constant sex differences while ratings by teachers and children of dependency usually showed girls to be more dependent than boys.

The perceived sex of the subject being observed can have an effect on the observation. Sex-related cues to identity should be eliminated in infant studies. That it has not been considered necessary to do this indicates the magnitude of the problem (Levinson, 1972). Eliminating sex-related cues can be done on observational studies of young children, but it grows increasingly difficult for older subjects. Thus, a healthy caution should be exercised in interpreting studies of sex differences obtained by observers who know the sex of the child (or the adult) being studied (Condrey & Condrey, 1976).

It may be possible that a similar phenomenon occurs in observational studies of race. Perhaps as expectations affect the ways females and males are observed and rated, they also affect the ways that subjects from different ethnic and racial backgrounds are observed and rated.

Expectations may also affect ratings on student-teacher interaction studies. In much of the work on student-teacher interaction, it is impossible to determine if the race or sex of the student may be affecting the rater's observations. Sadker and Sadker, in an examination of the interaction analysis systems listed in Mirrors of Behavior, found that only two of the many systems listed could be adapted in such a way to indicate the race and sex of the student participating in the interaction (Note 6).

Bias and stereotyped expectations can influence the experimenter as well as the observer. Kamin's (1973) comment that "the degree to which identical twins resemble one another depends on whether the same psychologist is testing both children or not" holds for a number of testing situations.

While the effect of the race and sex of the experimenter is not fully understood, it appears that these characteristics can affect the results of studies (Pettigrew, 1964b). In an extensive 1970 review of the literature, Sattler concluded that: (1) White experimenters can impede or enhance Black college students' performance, but they are more likely to have an affect on the performance of Black children; (2) the experimenter's race affects subjects' picture and doll preferences, but may not influence their scores on intelligence tests and personality measures; and (3) respondents give socially desirable responses to interviewers of races other than their own.

Later work has also shown the race of the experimenter to have an effect. Shostick (1977) found that the race of the tester or interviewer affected the ability of the interviewer to get candid answers. For example, there were great differences in the answers that high and low militant Blacks gave White interviewers in response to questions about social problems, but few differences in the responses they have to Black interviewers. The interviewers were women and, as so often happens, the interaction of race and sex was not discussed.

While studies have found that performance on a task is very likely to be influenced by the social-psychological characteristics of the testing environment, inconsistency is the rule with studies on the affect of experimenter race (Samuel, 1977). Samuel (1977) surveyed the literature and found, with one exception, that females score higher with a female experimenter. Rikli (1976), however, in her research and survey of the literature found no such consistency. Although she concluded that the sex of the experimenter may be a source of experimental contamination and should be given appropriate consideration in experimentation, she felt that the inconsistency of studies may suggest that the affect of experimenter is related to the type of task and the age of the subjects, with younger subjects being more strongly affected than older.

It does become clear that while the effects of experimenter race and sex on research are complex and inconsistent, they do exist. Systematic use of testers of one race or sex may have a long-range affect on research.

Almost any measure can be influenced by bias and, as a result of that influence, cause inaccurate results to be generated. Results can be influenced by obvious factors as the language of a question, such as "Should Jews be forced to leave the country?" (Bettelheim & Janowitz, 1964) or more subtle factors such as using the SAT to track changes in sex differences in mathematics, but not accounting for the potential affect of a test content change to the SAT, as in removing data insufficiency items on which females did better (Dwyer, 1976).

"In a society that professes educational equity as a goal, equitable non-stereotyped test content is a simple matter of justice" (Campbell & Scott, 1980). It is also a matter of good research.

VII. BIAS AND THE GENERATION OF CONCLUSIONS

After the design has been developed and implemented, the data collected and analyzed, the researcher's final task is to generate the conclusions, the section of a research article that is most often read and quoted. It is to this section that hurried professionals and practitioners turn to find the meaning of a study. It is also from this section that textbook authors, journalists, and politicians find their "research" quotes. Unfortunately, this section, too, can be affected by race and sex bias. The researcher's expectations, colored by stereotypes and by the racism and sexism found in our society, can - and frequently do - have more affect on the generation of conclusions than does the data itself.

An example of this phenomenon can be found in Yerkes' studies of chimpanzee sex role behaviors. Based on his observations, Yerkes concluded that his hypothesis that male chimps were naturally dominant while female chimps were naturally subordinate, was supported (Yerkes, 1943). Herschberger, in her analysis of Yerkes' work, humorously yet effectively, castigated Yerkes' conclusions, when writing from the perspective of Jose, a female chimp in Yerkes' sample, she commented:

> When Jack takes over the food chute, the report calls it his "natural dominance"...While I'm up there lording it over the food chute, the investigator writes down, "the male temporarily defers to her and allows her to act as if dominant over him." Can't I get any satisfaction out of life that isn't allowed me by some male chimp, damn it. (1948, p. 10.)

Yerkes' choice of different, value-laden words to describe similar behaviors by subjects of different sexes is not unusual. Garai and Scheinfeld (1968), in an analysis of sex difference, found females higher on 41 comparisons and males, on 35 comparisons. Yet because of the descriptors used, the reader gets the impression that males' performance was considerably better than was females'. For example, the term "superiority" was used in the descriptions of 45% of the male dominant areas, but only in 27% of the female dominant areas (Parlee, 1975). In both Yerkes' and Garai and Scheinfels's studies, the data may be accurate, but the impressions conveyed by the conclusions are supported more by sex role stereotypes than by data.

Stereotypes can be found in the explanation of data as well. The data are reported, but explanations of results are based on stereotypes rather than using other, equally plausible but non-stereotyped explanations. For example, based on their study of infant behavior when confronted with a barrier keeping them from a desired object, Goldberg and Lewis (1969) concluded that boys attack and girls give up. This is a reasonable explanation of the data showing that girls tend to cry and boys tend to try to push or climb the barrier and is also an explanation that reinforces stereotypes. An equally plausible non-stereotyped explanation that the more verbal girls were crying to attract the attention of those who could remove the barrier was not, however, reported. Thus the reader was left with the impression that one sex sought to solve problems while the other did not, rather than the impression that both sexes sought to solve problems, but used different styles (Parlee, 1975).

The presentation of stereotyped rather than non-stereotyped explanations of behavior occurs in studies of race as well as in studies of sex. As Crain (1976) commented, "the greater family instability of Black families can be interpreted either with Blacks being the helpless victims of racism which destroyed the family or as evidence that Blacks had been successful in retaining elements of African culture". Needless to say, the interpretation that is most often used is the "helpless victim of racism" theory.

There can be a variety of reasons why results "happen", but rarely are a variety of explanations given in research conclusions. Most frequently one explanation, which fits the data and which also supports our stereotypes, is given.

It appears that research conclusions, based on collected [illegible] have not come so far from the 1895 Psychological Review article where White subjects' slower reaction time (compared to Blacks and Indians) was taken as proof that Whites were the superior group (Gossett, 1963) or from Romane's 1887 study of sex differences in reading. Since women, in fact, performed better on the test than men, Romanes assured his readers that the quicker perception of women, as demonstrated by the test, was balanced by the lack in women of the "deeper qualities of mind", as proved by no test at all (Tolbin, 1972).

The preceding have been examples where the researchers fit stereotyped conclusions to the reality of the data, but this is not always the case. As Nobles (1973) asserts, "oftentimes researchers have demonstrated only one insignificant finding and, regardless of their own results, concluded their studies with assertions which were contrary to their own evidence. Brown (1967), for example, found no statistically significant differences between Black and White students on reports of teachers' perceptions of the students in terms of talking, clothes, and intelligence. Yet on each characteristic, he concluded that Black students, more often than White, tended to believe that their teachers perceived them negatively.

Similar inaccuracies have been found in studies of father absence and Black self-concept. Pettigrew (1964a) found no statistical differences between individuals from father-absent and father-present homes; however, he concluded that people from father-absent homes felt more victimized and in less control of the environment than did others. As Thomas and Sillen (1972) indicated, Black self-hatred is so much a part of the stereotype that when it is not found, it is invented. This was the case in a study by McDill, Meyers, and Rigsby (1966) that concluded that the high self-esteem they found in Black subjects was, in reality, a defense mechanism against discrimination.

As bias can cause researchers to misinterpret or deny data, it can also cause them to go far beyond the limitations of the data, or even the topic area of the study, in their conclusions. Nowhere is this more prevalent than in the "blame the victim" conclusions often found in studies of race and sex. Conclusions and interpretations of studies on race and sex frequently focus on person-centered characteristics while ignoring situationally relevant factors. Conclusions as those of Glazer and Moynihan (1970) find the "principal causes of the plight of the poor are found in the internal deficiencies of their own way of life" rather than being found in economic and educational structures. A study of the research published in Psychological Abstracts for the first six months of 1970 found that 82% of the studies dealing with race interpreted the difficulties of Black Americans in terms of personal shortcomings (Caplan and Nelson, 1973).

One of the most blatant and well known examples of this phenomena is Moynihan's famous study of Black families, with its conclusion that:

> The Negro community has been forced into a matriarchal structure which because it is so out of line with the rest of American society, seriously retards the progress of the group as a whole and imposes a crushing burden on the Negro male. Obviously not every instance

> of social pathology affecting the Negro community can be traced to the weakness of the family structure...once or twice removed, it will be found to be the principal source of most of the aberrant, inadequate, or antisocial behavior that did not establish but now serves to perpetuate the cycle of poverty and depression (1965, p. 29).

The emphasis placed on women's alleged fear of success as, at best, a partial explanation of women's and men's inequitable positions in employment (Horner, 1969) is another example of the "blame the victim" phenomena. Later work showing that fear of success affects women and men has not received the press of the original finding and is much less known.

Bias can also affect the generation of conclusions through the selective reporting of research results. For example, Moynihan (1965) concluded in support of his theory of Black matriarchy expressed earlier that "it is clear that Negro females have established a strong position for themselves in white collar and professional employment". He based his conclusion on data indicating that Black males represented 1.2% of all males in white collar occupations while Black females represented 3.1% of all females in white collar occupations. He did not report that at the time employed Black men earned, on the average, 1.5 times as much as employed Black women or that the smaller number of women in white collar occupations and the location of most white collar women in clerical occupations made the comparison, at best, misleading (Wallace, 1978). Employed women earn 59¢ for every dollar earned by employed men. To imply that Black women are doing better than Black men because their earnings are closer to White women's than Black men's are to White men's is ludicrous. Yet these comparisons continue. Gershman's (1980) recent New York Times article still compared the earnings of Black men to White men and Black women to White women. Again the large differences between average female and male earnings was not mentioned, leaving the inaccurate impression that economically Black women are doing better than Black men.

The selective reporting of data was used by Moynihan to substantiate at least one other conclusion. He stated that a greater percentage of Black females than Black males completed one to three years of college in 1963. The reason for this rather unusual comparison appears to be that the greater number of women in one and three-year nursing programs contributed to the female numerical superiority. In the 1960's Black women and Black men varied as to which group had the larger numbers in college. And, in fact, more men than women completed four years of college, but these statistics, which would not support his conclusions, were not reported (Wallace, 1978).

Moynihan is not the only one guilty of selective reporting of data. In general, sex differences, not the lack of differences, are reported in the research literature. Tresemer (1975) found one instance where 442 tests of male/female differences were made, 31 of which were significant. The 31 results (many of which could be explained by cumulative error rates) were reported; the 411 were not. Jensen also used selective reporting of results to substantiate his theories. He quoted one study of the effects of birthweight on intelligence, but did not indicate that two other studies quoted in other contexts in his article found that birthweight did not have an effect on intelligence (Jensen, 1969; Kamin, 1974).

The effects of bias on the generation of conclusions increase when the conclusions are reported in secondary sources, such as textbooks. As Van Tassel (1979) commented, "warnings as to limitations of primary research findings are often dropped as the studies are presented in text materials. This practice leads to an exaggerated list of sex differences and a number of statements which are patently false as written."

Parlee (1975) was also concerned about the inaccuracies in secondary sources. Suggesting that readers should be wary of statements that begin "it is generally accepted that...", she comments that "it is truly astonishing to discover how readily psychological theories about women can be vitiated by noting the existence of an intractable fact or two — usually data from studies misleadingly cited by the investigator as evidence for her or his theory." For example, different patterns of findings within each sex have sometimes been mistaken for sex differences. Sears found preschool girls were more verbally than physically aggressive, while boys were found to be more physically than verbally aggressive. These findings were later interpreted, by others, as showing that girls were higher than boys in verbal aggression. In reality boys showed more verbal and more physical aggression than did girls (Glickman, 1976).

Another area which is a concern in secondary sources is an inaccurate attribution of causality. Since sex and race are not variables that can be manipulated or randomly assigned to subjects, studies using race and sex as independent variables can not conclude that race or sex caused a change in the dependent variable. Race and sex may be related to the dependent variable, but causality may not be assumed. Other variables related to race and sex may be the cause. For example, differences in children's play may be caused by the different sex of the children, their attire,, observer stereotype, or the stereotypes of the children themselves. The study can tell us if differences exist, but not what caused them. Unfortunately this distinction is frequently not made and many reports of research do assume sex or race causality to say that being male or being Black does cause something to happen when, based on the data, all that should be said is that a difference was found (Campbell, 1979).

The major result of bias on the generation of conclusions is that inaccurate conclusions are drawn and incorrect myths, stereotypes, and theories are perpetuated. Paraphrasing Weisstein (1969) it is clear that until social expectations for minority and majority women and men are equal, until we provide equal respect for both sexes of all races, then answers to questions about minority and majority women and men may simply reflect our prejudices.

VII. GUIDELINES TO MINIMIZE THE EFFECTS OF RACE AND SEX BIASES ON RESEARCH METHODS

I. Minimizing Bias in Topic Selection

Issue: Patterns of funding, priorities of professional organizations, publishing opportunities, and the existence of biased theories have all contributed to a deemphasis on and devaluing of research for minorities and women. As a result of this deemphasis and devaluing there are gaps in our research knowledge. The existence of these gaps contributes to the development of programs and services and the making of decisions without an adequate research base.

Standards:

1. Research topics dealing with minorities and women should explore relevant non-traditional areas as well as traditional ones. For example, research on parent-child interaction should include paternal-child interaction as well as maternal-child interaction.

2. Emphasis should be put on studies of minority "non-deviant" behavior as well as on studies of "deviant" behavior. For example, research on minority achievers should be emphasized as is research on minority juvenile delinquency.

3. Research questions on minorities and women should include investigations of causes as well as effects. For example, work on the relative intelligence of minority and majority group members should also include the potential for examining the effects of biased tests.

4. Theories used as the basis of research studies should be examined to insure the groups whose behavior they purport to explain were considered in the development and testing of the theories. For example, theories of moral development, developed on males and generalized to females and males, are now incorporating new theoretical work on female moral development.

II. Minimizing Bias in Design

Issue: The lack of knowledge of literature about minorities and women contributed to the design of studies that are incomplete and inaccurate. The lack of testing for race and sex differences and similarities and the interaction of race and sex contribute to designs that leave gaps in research knowledge.

Standards:

1. Researchers should determine the validity of independent variables dealing with race and sex. For example, classifications of a woman's socio-economic status should be based on her own status rather than that of her father or husband.

2. Researchers should be familiar with the literature that may assist them in gaining a better knowledge of the important characteristics of the potential study subjects

and should use this information in their designs of studies incorporating variables. For example, knowing that the presence of an observer changes boys' anti-social behavior but not that of girls' has implications for the design of studies of sex differences in anti-social behavior.

3. When appropriate, researchers should test for and report race and sex similarities and differences and for interactions between race and sex.

III. Minimizing Bias in Sampling

Issue: The large proportion of studies that have been done on single race and sex samples, the use of different race and sex samples for different content areas, and the overgeneralization of results from these studies have all contributed to bias in sampling. As a result, inaccurate conclusions have been drawn based on overgeneralizations. Since most of the samples studied have been White males, minorities and women must be viewed as the "deviant", that which either conforms or deviates from the norm. Minorities and women have become a special topic to be studied rather than a sample to be used to study some general phenomena.

Standards:

1. Unless there is a demonstrable rationale for restricting the sample to one sex or race, samples should be male and female and multi-racial. For example, studies of aggression or child rearing should include both sexes and be multi-racial. A study on breast feeding, on the other hand, would have a justification for including only one sex.
2. Results from studies should not be generalized to members of racial and gender groups not represented in the sample.
3. Minority and female samples should be used in the development and testing of models, not in investigating post hoc how well they fit existing models devised from White male samples.

IV. Minimizing Bias in Measurement

Issue: Item types, contexts, language, and scoring procedures can cause tests to be developed which are biased for or against a specific racial or gender group. This, as well as the biases of observers, can contribute to inaccurate selection and inaccurate study results and interpretations.

Standards:

1. If observers are used, attempts should be made to control for the effect the observers' race and sex bias can have on the results. For example, observers can be sensitized to be aware of the potential effects stereotyped expectations may have on their ratings.

2. Instruments used in research studies should be examined to insure that they do not systematically favor one racial or gender group. For example, instruments whose item contexts deal primarily with scientific, mathematical, or mechanical areas to the exclusion of contexts in areas such as human relations should not be used to assess general achievement.

3. Researchers should be aware that different value orientations and life experience can affect personality test dedevopment and interpretation. For example, children giving Black dolls the lesser choices in areas such as jobs or housing may be a reflection of reality rather than low Black self-concept or individual White racism.

V. Minimizing Bias in the Generation of Conclusions

Issue: Conclusions incorporating unsupported researcher expectations, unwarranted attributions of causality, attributing unexpected results to unspecified methodological error rather than examining alternate hypotheses and generating conclusions on very small effect sizes, all contribute to inaccurate interpretation of research results. Inaccuracies can be further perpetuated in conclusions in secondary sources without the information about the limitations of the original studies.

Standards:

1. In a *post hoc* or *ex-post facto* study, researchers should not assume that sex or race were causal factors.

2. Researchers should reference their conclusions directly to the results of the study. For example, if non-significant differences are found, then these should not be reported simply as differences.

3. Non-stereotyped as well as stereotyped explanations of results should be explored by researchers. For example, infant girls' crying response when confronted by a barrier can be explained by concluding that girls give up or by concluding that the girls are crying to attract attention and are using an alternative method of problem solving. Both explanations should be given with rationale and supporting evidence.

4. Researchers, using ethnographic techniques, should indicate that the race and sex of the subjects may be changed to protect their identity.

5. Research information in secondary sources should include, where appropriate, the statements of limitations found in the original work. In addition, authors of secondary works should, based on their evaluations of the original works, include cautions and statements of limitations.

IX. FUTURE DIRECTIONS

Bias does affect researchers and the research that they do. As the preceding sections have indicated, because of the racism and sexism present in our society and in researchers who are part of that society, research findings can be limited and inaccurate. As the guidelines indicate, efforts can and are being made to rectify some of the negative effects. More work is being done by people like Gilligan (1977) to involve multi-ethnic samples of women and men in the development and testing of theories. Increasing numbers of researchers such as McClelland (1975) are becoming concerned about the negative effects of single-sex, single-race samples. Research organizations such as the American Educational Research Association are paying increased attention to the need for more and better information about populations that include women and minority men.

These changes and this paper itself are based on the acceptance of a traditional, statistically-oriented model of research. Little has been said about the role that bias has played in the development of this model itself and the ways that use of the "scientific method" of research has limited the generation of knowledge.

Increasing numbers of researchers such as Anton (1979) are suggesting that research needs to be restructured such that the authority of truth rest with the subject, not with the researcher's methodology. Making the subject, rather than the researcher, the source of authority may provide opportunities to give more information about the complex realities that are the basics of social science research. Much of social science research has been based on the isolation and manipulation of variables, tending to simplify complex situations and to generate knowledge about components rather than the whole (Note 7). Even that knowledge may be limited. As Argyris (1975) stated, "if the experiment is designed well, the choice is free in the sense that the subject makes it, but inevitable in that the experimenter designed the experiment to minimize other choices." In a more general way, the social scientist adopts a vantage point for analysis; with that vantage point comes an operating context. "Knowledge gained as a result will fit the parameters of accepted context and will be useful in those terms. Knowledge that might have been accrued as a consequence of adopting a different frame of reference will not be forthcoming nor will the implicit challenge to that context that might otherwise develop be likely to arise"(Gideonse, 1977).

The context that has been most frequently accepted in research is one which Carlson (1979) calls "agentic". Agentic inquiry is based on separating, ordering, quantifying, manipulating, and controlling. She suggests the need for increased attention to communal research with its emphasis on naturalistic observation and the personal participation of the investigator. The human experience is a subjective one and researchers need to develop strategies to help people talk about that subjectivity. Analytic systems are perhaps not the appropriate method for describing experience (Anton, 1979). Also not appropriate may be the social science researchers, ultimately unsuccessful, efforts to be objective, to remove themselves from settings and phenomena that are part of their own experience. Perhaps social science researchers should take greater note of the frank recognition of the historian or the anthropologist that the researcher or scholar is in no way free of the phenomenon of interaction with the data; that in fact interactive effects will occur (Note 7). Also from the historian and the novelist, social scientist researchers can learn about ways to discuss topics for which we have no words, to use metaphors, analogies, and literary passages, which can help illuminate the human experience.

Developing new models of social science research, which combine the best of the old and new, that marry qualitative and quantitative methods is neither easy nor impossible. It may be, however, the best hope for social science research.

REFERENCE NOTES

1. Campbell, P. Equity Content in Education Journals (Unpublished Manuscript).
2. Campbell, P. Readership of Special Interest Journals (Unpublished Manuscript).
3. Wirtenberg, J. Attitudinal and Behavioral Measurement in Social Processes/Women's Research (Conference Agenda, Dec. 1-2, 1980).
4. Morrison, B.M. Personal Communication, August, 1980.
5. Shakeshaft, C. Personal Communication, December, 1980.
6. Sadker, M. Personal Communication, November, 1980.
7. Gideonse, H. Personal Communication, November, 1980.

BIBLIOGRAPHY

Acker, J. Women and social stratification: a case of intellectual sexism. American Journal of Sociology, 72 (4), 1973, 936-945.

Acker, J. & Van Houten, D.R. Differential recruitment and control: the sex structure of organizations. Administrative Sciences Quarterly, 1974, 6, 152-169.

Allport, G.W. Pattern and Growth in Personality. New York: Holt, Rinehart and Winston, 1961.

AMEG Commission report on sex bias in interest measurement. Measurement and Evaluation in Guidance, 1973, 6(3), 171-174.

Anastasi, A. Psychological Testing. New York: Macmillian, 1976.

Anthony, E.J. & Benedek, T. Parenthood: Its Psychology and Psychopathology. Boston: Little, Brown, 1970.

Anton, J. Keynote Address. Fifth Annual Conference on Research on Women and Education, Cleveland, October, 1979.

Arditti, R. Women as objects: science and sexual politics. Science for the People, September, 1974, 6-10.

Argyris, C. Dangers in applying results from experimental social psychology. American Psychologist, 1975, 30 (4), 469-485.

Atkinson, J.W. (Ed). Motives in Fantasy, Action and Society. Princeton, N.J.: Van Nostrand, 1958.

Babladelis, G. Editorial. Psychology of Women Quarterly, 1976, 1 (1), 3-10.

Bettleheim, B. & Janowitz, M. Social Change and Prejudice. New York: Free Press, 1964.

Bettleheim, B. The commitment required of a woman entering a scientific profession in present day American society in Woman and the Scientific Professions. M.I.T. Symposium on Women in Science and Engineering. Cambridge, Mass., 1965.

Bevis, W.M. Psychological traits of the southern Negro with observations as to some of his psychoses. American Journal of Psychiatry. 1921, 1, 69-78.

Blau, P.M. & Duncan, O.D. The American Occupational Structure. Free Press, 1978.

Bliss, W.W. Woman and Her Thirty-Years' Pilgrimmage. Boston: B.B. Russell, 1870.

Billingsley, A. *Black Families in White America*. Englewood Cliffs, NJ: Prentice-Hall, 1968.

Bowlby, J. *Maternal Care and Mental Health*. New York: Schocken Books, 1966.

Brazziel, W.F. White research in black communities. *Journal of Social Issues*, 1973, 29, 41-44.

Brown, B. The assessment of self-concept among four-year-old Negro and white children: A comparative study using the Brown-IDS self-concept referents test, New York: Institute for Developmental Studies, 1967.

Bullough, V.L. & Bullough, B. *The Subordinate Sex: A History of Attitudes Toward Women*, Urbana, Il: The University of Illinois Press, 1973.

Campbell, P.B. Sex, race and student interaction. ERIC Document. In Press.

Campbell, P.B. Things your advisor never told you. Washington: Minority and Women's Program, NIE, 1980.

Campbell, P.B. & Scott, E. Non-biased tests can change the score. *Interracial Books for Children Bulletin*, 1980, 11(6), 7-9.

Caplan, N. & Nelson, S.D. On being useful: The nature and consequences of psychological research on social problems. *American Psychologist*, 1973, 28(3), 119-211.

Caplan, P.J. Sex differences in anti-social behavior: does research methodology produce or abolish them? *Human Development*, 1975, 18(6), 444-460.

Carey C.B. The anthropology of women: An advocacy position. ERIC Document, ED 156543, 1979.

Carlson, R. Where is the person in personality research? *Psychological Bulletin*, 1971, 75, 203-219.

Clarke, E.H. *Sex in Education or a Fair Chance for the Girls*. Boston: James R. Csgood, 1873 (reprint edition by Arno Press, 1972).

Clark, K.B. Forward. In A. Thomas & S. Sillen. *Racism and Psychiatry*. New York: Brunner/Mazel, 1972, xi-xiii.

Coffman, W.E. Sex differences in response to items in an achievement test. *Eighteenth Yearbook, National Council on Measurement in Education*, 1961, 117-124.

Committee on the Status of Women in Sociology. Sexist biases in sociological research: Problems and issues. *ASA Footnotes*, 1980, 1, 8-9.

Condrey, J. & Condrey, S. Sex differences: A study of the eye of the beholder. *Child Development*, 1976, 45, 812-19.

Cowan, M.A., Watkins, B.A., & Davis, W.E. Level of education, diagnosis and race related differences in MMPI performance. Journal of Clinical Psychology, 1975, 31(3), 442-444.

Crain, R.L. Why academic research fails to be helpful. School Review, 1976, 84(30), 337-351.

Dan, A.J. & Beekman, S. Male vs. female representation in psychological research. American Psychologist, 1972, 27(11), 1078.

de Ortego y Gasca, F. Chicano's concepts of culture. ERIC Document ED178246, 1978.

Donlon, T.F. Content factors in sex differences on test questions. A paper presented to the annual meeting of the American Educational Research Association. Boston, 1971.

Donlon. T.F., Ekstrom, R.B., & Lockheed, M.E. Performance consequences of sex bias in the content of major achievement batteries: final report. ERIC Document ED151415, 1977.

Dwyer, C.A. Test content and sex differences in reading. Reading Teacher, 1976, 29, 753-757.

Edelsky, C. Genderlects: A brief review of the literature. ERIC Document ED165187, 1979.

Ehrenreich, B. & English, D. For Her Own Good. Garden City, NY: Anchor Books, 1979.

Erikson, E. Growth and crises of the healthy personality. Psychological Issues, 1959, 1(50), 1-171.

Etaugh, C. & Spandikow, D.B. Attention to sex in psychological research as related to journal policy and author sex. Psychology of Woman Quarterly, 1979, 4(2), 175-184.

Evarts, A.B. Dementia praecos in the colored race. The Psychoanalytic Review, 1914, 1, 388-403.

Gallup, G. The Gallup Poll: Public Opinion, 1979. Princeton, NJ: Scholarly Research, Inc., 1980.

Garai, J.E. & Scheinfeld, A. Sex differences in mental and behavioral traits. Genetic Psychology Monographs, 1968, 77, 169-299.

Garrett, H.E. The equalitarian dogma. Mankind Quarterly, 1961, 1, 253-255.

Gershman, C. The Black plight: A debate. A matter of class. New York Times, 10/5/80, 24, 92-99, 102-104.

Gerwitz, S. & Dodge, K. Adults' evaluation of a child as a function of the sex of the adult and the sex of the child. Journal of Personality and Social Psychology, 1975, 32(5), 822-828.

Gesell, A. & Ilg, F.L. Infant and Child in the Culture of Today. New York and London: Harper, 1943.

Gideonse, H.D. Sex roles, inquiry and social policy. A paper presented to the American Educational Research Association. New York, 1977.

Gilligan, C. Woman's place in man's life cycle. Harvard Educational Review, 1980, 49, 431-446.

Gilligan, C. In a different voice: Women's conceptions of the self and morality. Harvard Educational Review, 1977, 44, 481-517.

Glazer, N. & Monihan, D.P. Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians & Irish of New York. Cambridge, M.I.T. Press, 1970.

Glichman, J. Sex differences in intellectual functioning: Myth or reality. ERIC Document ED20706, 1976.

Goldberg, S. & Lewis, M. Play behavior in the year-old infant: early sex differences. Child Development, 1969, 40, 21-31.

Gossett, T.F. Race: The History of an Idea in America. Dallas: Southern Methodist University Press, 1963.

Gottfredson, L.S. The relation of education and situs of work to economic differences between blacks and whites. ERIC Document, ED150059, 1977.

Graves, Z.R. & Glick, J. The effect of context on mother-child interaction: A progress report. In W.S. Hall and M. Cole (Eds) The Quarterly Newsletter of the Institute for Comparative Human Development. New York: Rockefeller University Press, 1978, 41-46.

Greenberg, S. Right From the Start. Boston: Houghton Mifflin, 1978.

Grier, W.H. & Cobbs, P.M. Black Rage. New York: Basic Books, 1968.

Gullattee, A.C. The Negro psyche: Fact, fiction and fantasy. Journal of the National Medical Association, 1969, 61, 119-129.

Hall, G.S. Adolescence Vol. II. New York: D.D. Appleton, 1905.

Haller, J.G. & Haller, R.M. The Physician & Sexuality in Victorian America. Urbana, Il.: University of Illinois Press, 1974.

Harmon, L.W. Sexual bias in interest measurement. Measurement & Evaluation in Guidance. 1973, 5(4), 496-501.

Harris, S.L. Who studies sex differences. American Psychologist. 1972, 27(11), 1077-1078.

Herman, L.S. & Serbin, L.A. Observer expectations of a child's performance as a function of sex. A paper presented to the Third Annual Midwest Conference on Research on Women in Education, Milwaukee, 1977.

Herschberger, R. *Adam's Rib*. New York: Pellegrini and Cudahy, 1948 (Reprint edition Harper and Row, 1970).

Horner, M. Fail: Bright women. *Psychology Today*, November, 1969, 62, 36-37.

Hunter, D.M. & Babcock, C.G. Some aspects of the intra-psychic structure of certain American Negroes as viewed in the intercultural dynamic. In W. Muensterberg and S. Axelrod (Eds) *The Psychoanalytic Study of Society, Vol. 4.* New York: International Universities Press, 1967,

Inton-Peterson, M.J. Female-male vs. male-female: Does the order affect stereotyped trait attributions. *Psychology of Woman Quarterly*, 1980, 4(4), 460-481.

Isaac, S.I. *Handbook in Research & Evaluation*. San Diego: EDITS, 1975.

Jelnick, M. Career management and women. ERIC Document ED150272, 1978.

Jensen. A.R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 1969, 39, 1-139.

Jensen, A.R. *Bias in Mental Testing*. New York: Free Press, 1980.

Jorgensen, C.C. IQ tests and their educational supporters. *Journal of Social Issues*, 1973, 29(1), 33-40

Kamin, L. The Southern Regional Council's symposium on human intelligence social science, and social policy. ERIC Document ED084318, 1973.

Kaplin, J. *Women and Sports*. New York: Viking Press, 1979.

Kardner, A. & Oversey, L. *The Mark of Oppression*. Cleveland: World Publishing, 1962.

Kearney, H.R. Feminist challenges to the social structure and sex roles. *Psychology of Women Quarterly*, 1979, 4(1), 16-31.

King, P. & Blount, H.P. The influence of sex-typed math test problems on performance of sixth grade students. A paper presented to the National Council on Educational Measurement, Washington, 1975.

Klein, S.S. & Goodman, M.A. *Federal Funding to Promote Sex Equity in Education: 1980*. Washington, DC: Women's Educational Equity Act, U.S. Department of Education, 1980.

Kolhberg, L. & Kramer, R. Continuities and discontinuties in childhood and adult moral development. *Human Development*, 1969, 12, 93-120.

LaBrecque, M. On making sounder judgments. *Psychology Today*, 1980, June, 33-42.

Larsen, M. An exploration of the use of race and sex as variables in federally sponsored educational evaluation. Washington, DC: National Institute of Education, 1980.

Levinson, C.A. Sex differentiation in early infancy: Problems in methodology and interpretation of data. ERIC Document ED06414, 1972.

Levinson, D.J. The Seasons of a Man's Life. New York: Alfred A. Knopf, 1978.

Lewis, O. Children of Sanchez. New York: Random House, 1961.

Liebow, E. Tally's Corner. Boston and Toronto: Little, Brown and Company, 1967.

Lightfoot, S.L. Socialization and education of young Black girls in schools. In Conference on the Educational and Occupational Needs of Black Women, Vol. II. Washington: NIE, 1978, 3-32.

Lockheed, M.E. & Stein, S.L. The status of women's research in educational publications. Educational Researcher, 1980, 9(2), 11-15.

Long Laws, J. Feminism and patriarchy: Competing ways of doing social science. ERIC Document, ED166081, 1978.

Maccoby, E. & Jacklin, C.N. The Psychology of Sex Differences. Stanford: Stanford University Press, 1974.

Marccia, E.S. Prejudice against women and bias in educational research. ERIC Document ED131352, 1977.

Masters, W.H. & Johnson, V. Human Sexual Response. Boston: Little Brown, 1966.

McClelland, D. Power: The Inner Experience. New York: Irvington Publishers, 1975.

McDill, E.L., Meyers, E.D. & Rigsby, L. Sources of Educational Climates in High Schools. Baltimore: John Hopkins, 1966.

McDonald, G. Two windows on research. ERIC Document, ED133053, 1977.

McGrath, J.H. Research Methods & Designs for Education. Scranton, PA: International Textbook Co., 1970.

McKenna, W. & Kessler, S.J. Differential treatment of males and females as a source of bias in social psychology. A paper presented to the annual meeting of the American Psychological Association, New Orleans, 1974.

McWhinni, H. A note on methodology in using children's figure drawing to assess racial and cultural differences. Studies in Art Education, 1972, 13(2), 30-33.

Meyer, J.W. & Sobieszik, B.J. The effect of a child's sex on adult interpretations of its behavior. Developmental Psychology, 1972, 6, 42-48.

Millman, M. & Kantor, R.M. (Eds). Another Voice. Garden City, NY: Anchor Press, 1975.

Milton, G.A. The effects of sex role identification upon problem solving skills. Journal of Abnormal and Social Psychology, 1957, 55, 208-212.

Montemayor, R. Children's performance in a game and their attraction to it as a function of sex typed labels. Child Development, 1974, 45, 152-156.

Moynihan, D.P. The Negro family: The case for national action. Washington, DC: U.S. Department of Labor, 1965.

Murphy, R.C. Sex differences in examination performance. A paper presented to the International Conference on Sex Role Stereotyping, Cardiff, Wales, 1977.

Nage, E. The Structure of Science. New York and Burlingame: Harcourt, Brace and World, 1961.

Nichols, P. Dynamic variation theory as a model for the study of language and sex. A paper presented at the Ninth World Congress of Sociology, Uppsala, Sweden, August, 1978.

Nobles, W.W. Psychological research and the Black self concept: A critical review. The Journal of Social Issues, 1973, 29(1), 11-31.

Oakley, A. The Sociology of Housework. New York: Pantheon Books, 1975.

Osherson, S.D. Adaption to Occupational Changes at Midlife: A Predictive Longitudinal Study. NIE Project G-77-0049, 1981.

Parlee, M.B. Psychology. SIGNS: A Journal of Women in Culture and Society, 1975, 1(1), 119-138.

Peng, S.S. & Fetters, W.B. Variables involved in withdrawal during the first two years of college: preliminary findings from the National Longitudinal study of the high school class of 1972. American Educational Research Journal, 1978, 15(3), 361-373.

Pettigrew, T.F. A Profile of the American Negro. Princeton: Van Nostrand, 1964. (a)

Pettigrew, T.F. Negro-American personality: why isn't more known? Journal of Social Issues, 1964, 20(2), 4-23. (b)

Plake, B.S., Hoover, H.D., & Loyd, B.H. An investigation of differential item performance by sex on the Iowa Tests of Basic Skills. A paper presented to the National Council on Measurement in Education, Toronto, 1978.

Pollard, D. Special needs of Black women in educational research and development. In Doris Leckie (Ed) NIE/AERA Planning Conference to Increase the Participation of Women and Minorities in Educational Research and Development. Washington: NIE, 1976, 31-58.

Prescott, S. & Foster, K. Why researchers don't study women: The responses of 67 researchers. A paper presented to the annual meeting of the American Psychological Association, New Orleans, 1974.

Proshansky, H. & Newton, P. The nature and meaning of Negro self-identity. In M. Deutsch, I. Katz, & A.R. Jensen (Eds) Social Class, Race, and Psychological Development. New York: Holt, Rinehard and Winston, 1968,

Quarles, B. Jet Magazine. 1967, 33(32).

Reardon, P. & Prescott, S. Sex as reported in a recent sample of psychological research. Psychology of Woman Quarterly, 197[illegible], 2(2), 57-61.

Rheingold, J. The Fear of Being a Woman. New York: Grune & Stratton, 1964.

Rikli, R. Physical performance scores as a function of experimenter sex and experimenter bias. The Research Quarterly, 1976, 47(4), 776-82.

Romanes, G. The Darwinian Theory. London, 1892.

Rothbart, M.K. & Maccoby, E.E. Parent's differential reactions to sons and daughters. Journal of Personality and Social Psychology, 1966, 4, 237-243.

Russell, B. The Scientific Outlook. New York: W.W. Norton, 1959.

Saario, T.T. The Ford Foundation. In L. Hunter & J. Marzone (Eds) Funding for Women's Educational Equity. San Francisco: Women's Educational Equity Communications Network, 1980.

Samuda, R.J. Psychological Testing of American Minorities. New York: Dodd, Mead, 1975.

Samuel, W. Observed IQ as a function of test atmosphere, tester expectation, and race of tester: A replication for female subjects. Journal of Experimental Psychology. 1977, 69(5), 593-604.

Sanguiliano, I. In Her Time. New York: William Morrow, 1978.

Sattler, J. Racial experimenter effects in experimentation, testing, interviewing and psychotherapy. Psychological Bulletin, 1970, 72, 137-160.

Schwabacker, S. Male vs. female representation in psychological research: An examination of the Journal of Personality and Social Psychology, 1970, 1971. Journal Supplement Abstracts, 1972.

Shakeshaft, C. Dissertation research on women in academic administration: A synthesis of findings and paradigm for future research. Ph.D. Dissertation, Texas A&M University, 1979.

Sheehy, G. *Passages*. New York: Bantom, 1977.

Sherwood, J.J. & Nataupsky, M. Predicting the conclusions of Negro-White intelligence research from biological characteristics of the investigator. *Journal of Personality and Social Psychology*, 1968, 53-58.

Shosteck, H. Respondent militancy as a control variable for interviewer effect. *Journal of Social Issues*, 1977, 33(4), 36-45.

Silveria, J. The effects of sexism on thought: How male bias hurts psychology and some hopes for a women's psychology. Pittsburgh: KNOW, 1972.

Smith, M.D., Zingale, S.A., & Coleman, J.M. The influence of adult expectancy/child performance discrepancies upon children's self-concepts. *American Educational Research Journal*, 1978, 15(2), 259-265.

Spencer, H. *Domestic Institutions: Principles of Sociology*. New York: D. Appleton & Co., 1893.

Spitz, R. *The First Year of Life: A Psychoanalytic Study of Normal and Deviant Development of Object Relations*. New York: International Universities Press, 1965.

Stack, C.B. *All of Our Kin: Strategies for Survival in the Black Community*. New York: Harper & Row, 1975.

Stanton, W. *The Leopard's Spots: Scientific Attitudes of Race in America, 1815-1959.* Chicago: University of Chicago Press, 1960.

Steinmetz, S.K. The sexual context of social research. *The American Sociologist*, 1974, 9(3), 111-116.

Steinem, G. Feminist notes. *MS*, 1980, 9(3), 98-103.

Stewart, W.A. A psychosocial study of the formation of early adult life structure in women. DRJ 77-08289. Ann Arbor, Mi.: Dissertation Abstracts, 1977.

Stivers, P. & Leckie, D. *NIE/AREA Planning Conference to Increase the Participation of Women and Minorities in Educational Research and Development.* Washington, D.C.: NIE, 1976.

Sturgis, S.H. & Menzer-Benaron, D. *The Gynecological Patient: A Psycho-Endocrine Study*. New York: Grune & Stratton, 1962.

Terman, L.M. *The Measurement of Intelligence*. Boston: Houghton, Mifflin, 1916.

Thomas, A. & Sillen, S. Racism and Psychiatry. New York: Brunner/Mazel, 1972.

Thomas, S. Measurement of sex roles: What are we really measuring? A paper presented to the annual meeting of the National Council of Measurement in Education, Toronto, 1978.

Tittle, C.K. & Denker, E.R. Re-entry women. Review of Educational Research, 1977, 47(4), 531-584.

Tolpin, M. The Darwinian Influence on Psychological Definitions of Femininity in England. Ph.D. Dissertation, Harvard, University, 1972.

Tresemer, D. Assumptions made about gender roles. In Marsha Millman and and Rosebeth Kanter (Eds). Another Voice. New York: Anchor Books, 1975, 308-339.

Tyler, R.W. & White, S.H. Testing, teaching and learning. Chairman's Report of a Conference on Research on Testing, Washington, U.S. Government Printing Office, 1978.

Valentine, C.A. Culture and Poverty: Critique and Counter Proposals. Chicago: University of Chicago Press, 1968.

Van Tassel, E. The science of sexism. Politics and Education, Fall, 1979, 28-30.

Vincent, C. Implications of changes in male-female role expectations for interpreting M-F scores. Journal of Marriage and the Family, 1966, 28, 196-199.

Wallace, M. Black Macho and the Myth of Superwoman. New York: Dial Press, 1978.

Weinberg, M. Minority students: A research approach. Washington, D.C.: U.S. Government Printing Office, 1977.

Weisstein, N. Woman as nigger. Psychology Today, 1969, 10, 20-23.

Wood, A.D. The fashionable diseases, women's complaints and their treatment in nineteenth-century America. Journal of Interdisciplinary History, Summer, 1973, 4, 28-35.

Wylie, R.C. The Self Concept. Lincoln: University of Nebraska Press, 1961.

Yerkes, R.M. Chimpanzees. New Haven: Yale University Press, 1943.

Young, W. National Observer, April 1, 1968.