

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

- 1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.**
- 2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.**
- 3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again -- beginning below the first row and continuing on until complete.**
- 4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.**
- 5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.**

University Microfilms International

300 North Zeeb Road

Ann Arbor, Michigan 48106 USA

St. John's Road, Tyler's Green

High Wycombe, Bucks, England HP10 8HA

77-13,658

GRADY, Kathleen Elizabeth, 1942-
SEX AS A SOCIAL LABEL: THE ILLUSION OF
SEX DIFFERENCES.

City University of New York, Ph.D., 1977
Psychology, social

Xerox University Microfilms, Ann Arbor, Michigan 48106

© 1977

KATHLEEN ELIZABETH GRADY

ALL RIGHTS RESERVED

SEX AS A SOCIAL LABEL:
THE ILLUSION OF SEX DIFFERENCES

by

KATHLEEN E. GRADY

A dissertation submitted to the Graduate
Faculty in Psychology in partial fulfillment of the
requirements for the degree of Doctor of
Philosophy, The City University of New York

1977

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1/26/77
date

Ellen J. Langer
Chairperson of the Examining
Committee

1/31/77
date

Florence L. Denmark
Executive Officer

Ellen J. Langer

Florence L. Denmark

R. Glen Hass

Supervisory Committee

Abstract

SEX AS A SOCIAL LABEL: THE ILLUSION OF SEX DIFFERENCES

by

Kathleen E. Grady

Advisor: Professor Ellen Langer

The research reported here investigated some of the cognitive and motivational factors that are related to the belief in sex differences. Previous research indicates that many of these beliefs, commonly held in this culture, may be considered illusory in the sense that they lack scientific support; however, they do provide a set of norms or ideals to which people can conform their behavior.

In the first set of studies the discriminability of sex as a characteristic of others was compared to the discriminability of physical characteristics. In the first study, people were interviewed in subway stations and asked to recall the characteristics of token sellers for the purpose of identification. Their descriptions indicated that they considered sex the most descriptive characteristic. Subsequent laboratory studies constructed around a description game supported this finding and indicated that people would use sex as a descriptor even when it was

counter-productive for success in the game. Their behavior was compared to a model of statistical distinctiveness which predicted the most efficient strategy.

The second set of studies investigated expectancy as a mediating variable between the belief in sex differences and behavior. Expectancy was manipulated by giving subjects information about the relationship of sex to ability on a variety of tasks. In one study the effect of sex-related information was compared to similar information about the relationship of eye color and handedness to ability. The results indicated that although subjects held the traditional beliefs about female superiority on verbal tasks and male superiority on math tasks, when given opposite information for specific tasks, their performance conformed to the new information. In contrast, subjects behavior was not influenced by information related to the characteristics of handedness and eye color on motor coordination and perceptual tasks. It was concluded that the expectancy of a sex difference does influence behavior. Further, although beliefs about the relationship of sex to abilities are more widely held than beliefs about the other characteristics studied, they are also easier to manipulate.

In the final study, information that both contradicted and supported existing sex stereotypes was given to different experimental groups. Experimental groups were compared to two control groups. In general, results were

similar to the previous study in that instructional information about the relationship of sex to ability affected expectancy of success, task performance, and evaluation of performance. Subjects were not more influenced by information that supported stereotypes. In addition, there were effects for task, condition, and sex. These results were discussed in terms of the strength of the stereotypes and reactance theory.

Findings from all four studies are discussed in terms of discrimination, the importance of the sex label, the personal cost of stereotyping, and the possibilities for change. The need for a new paradigm for the study of sex as a psychological variable is outlined.

ACKNOWLEDGEMENTS

Over the two years it took to do this work and the several years in which the ideas percolated, there were many people who contributed time, ideas, and later, criticisms and suggestions.

I especially want to thank the members of my committee. My advisor, Ellen Langer, read many drafts and gave helpful suggestions at every stage in the development of the research. It was her unfailing optimism and confidence in me and in the data which helped me through many rough periods in the data analyses. The first set of studies was inspired by a conversation with Glen Hass, who also helped get subjects and gave valuable feedback on both theory and statistics. Florence Denmark supported the development of these ideas over several years. My outside readers, Mary Brown Parlee and Bob Brannon, provided detailed suggestions as well as perspective and encouragement enough to inspire several more research projects.

I am deeply indebted to the Woodrow Wilson Foundation which provided the financial support to enable me to devote a full year to this work. The opportunity to develop a piece of theory and research unhindered by employment pressures was a luxury that I had never known.

In the end it is one's closest friends who live through the daily struggle of a major research effort. Jeff Hayward

and Joyce Borkin provided everything from emotional support to intelligent and detailed criticism; they participated in the early formulation of the ideas and edited final drafts; they provided meals and childcare. They were always there. Other friends and colleagues who gave support during critical periods were Susan Lee, Anne Mulvey, Jeremy Miransky, Buzzy Chanowitz, Linda Small, and John Sabin.

It was through friendly and intelligent arguments with Stanley Milgram that many of the ideas were refined. His insistence on clarity and a certain elegance of style set standards to which I will always aspire.

Most especially I want to thank my wonderful and exceptional children, Paul and Emily, for their understanding, independence, and encouragement of their mother's work.

TABLE OF CONTENTS

Approval page.....	111
Abstract.....	iv
Acknowledgements.....	vii
List of Tables.....	x
List of Figures.....	xiii
List of Appendices.....	xiv
 I. INTRODUCTION.....	 1
Sex as a Subject Variable.....	2
Sex as a Stimulus Variable.....	8
The Illusion of Sex Differences.....	11
 II. SEX AS A SOCIAL LABEL: STIMULUS DISCRIMINATION...	 15
Study 1: A Field Demonstration of Social Stimulus Discrimination.....	18
Study 2: Social Stimulus Discrimination in the Laboratory.....	26
 III. THE EXPECTANCY OF SEX DIFFERENCES: THE INFLUENCE OF OWN SEX LABEL ON BEHAVIOR.....	 52
Study 3: The Expectancy of Characteristic Differences.....	56
Study 4: The Expectancy of Sex Differences.....	85
 IV. CONCLUSIONS AND IMPLICATIONS.....	 121
 V. APPENDIX.....	 127
 VI. BIBLIOGRAPHY.....	 168

LIST OF TABLES

1.	Token Seller Characteristics Reported by Subjects...	22
2.	Rank Order in Which Subjects Gave Descriptive Clues.....	32
3.	Theoretical Distribution of Clues Based on the Statistical Distinctiveness Hypothesis.....	34
4.	Number of Extra Clues Over Trial Blocks.....	40
5.	Rank Order in Which Subjects Gave Descriptive Clues - Study 2a.....	44
6.	Theoretical Distribution of Clues - Study 2a.....	45
7.	Number of Extra Clues Over Trial Blocks - Study 2a..	49
8.	Mean Scores: General Expectancies.....	62
9.	Analysis of Variance Summary Table: General Expectancy Scores by Condition, Characteristic, and Task Ability.....	63
10.	Mean Scores for Expectancy, Performance, and Evaluation on Maze Task.....	68
11.	Analysis of Variance Summary Table: Maze Performance.....	68
12.	Analysis of Variance Summary Table: Maze Evaluation.....	68
13.	Mean Scores for Expectancy, Performance, and Evaluation on Verbal Task.....	70
14.	Analysis of Variance Summary Table: Verbal Performance.....	70
15.	Mean Scores for Expectancy, Performance, and Evaluation on Math Task.....	72
16.	Analysis of Variance Summary Table: Math Expectancy.....	72
17.	Analysis of Variance Summary Table: Math Performance.....	72

18.	Analysis of Variance Summary Table: Math Evaluation.....	73
19.	Analysis of Covariance Summary Table: Math Performance.....	73
20.	Mean Scores for Expectancy, Performance, and Evaluation on Math Task: LOMATH Group.....	73
21.	Analysis of Variance Summary Table: Math Performance: LOMATH Group.....	75
22.	Mean Scores for Expectancy, Performance, and Evaluation on Math Tasks: HIMATH Group.....	75
23.	Analysis of Variance Summary Table: Math Evaluation: HIMATH Group.....	75
24.	Mean Scores for Expectancy, Performance, and Evaluation on Perception Tasks.....	77
25.	Analysis of Variance Summary Table: Little Tommy Tucker Performance.....	78
26.	Analysis of Variance Summary Table: Word Search Performance.....	78
27.	Correlations Among Expectancy, Performance, and Evaluations for Each Task.....	79
28.	Mean Scores: General Expectancies of Control Subjects.....	93
29.	Analysis of Variance Summary Table: General Expectancy by Sex and Task Ability.....	93
30.	Analysis of Covariance Summary Table: Word Knowledge Personal Expectancy.....	96
31.	Analysis of Covariance Summary Table: Word Recognition Personal Expectancy.....	96
32.	Analysis of Covariance Summary Table: Word Knowledge Performance.....	98
33.	Analysis of Covariance Summary Table: Word Recognition Performance.....	98
34.	Analysis of Covariance Summary Table: Word Knowledge Evaluation.....	101

35.	Analysis of Covariance Summary Table: Word Recognition Evaluation.....	i01
36.	Correlations Among Expectancy, Performance, Evaluation, and Experience Scores for Each Task.....	103
37.	Analysis of Covariance Summary Table: Number Series Personal Expectancy.....	108
38.	Analysis of Covariance Summary Table: Shape Recognition Personal Expectancy.....	108
39.	Analysis of Covariance Summary Table: Number Series Performance.....	111
40.	Analysis of Covariance Summary Table: Shape Recognition Performance.....	111
41.	Analysis of Covariance Summary Table: Number Series Evaluation.....	114
42.	Analysis of Covariance Summary Table: Shape Recognition Evaluation.....	114

LIST OF FIGURES

1.	Mean Rank of Clues Given Compared to Theoretical Means Derived from the Statistical Distinctiveness Hypothesis.....	36
2.	Percentage of Each Clue Given and Expected in First Position.....	38
3.	Percentage of Times Sex Used as a First Clue Across Trial Blocks.....	40
4.	Mean Rank of Clues Given Compared to Theoretical Means - Study 2a.....	46
5.	Percentage of Each Clue Given and Expected in First Position - Study 2a.....	48
6.	Percentage of Times Sex Used as a First Clue Across Trial Blocks - Study 2a.....	49
7.	Mean Personal Expectancy Scores on Verbal Tasks.....	95
8.	Percent Correct on Verbal Tasks.....	97
9.	Mean Evaluation on Verbal Tasks.....	100
10.	Mean Personal Expectancy Scores on Math Tasks.....	107
11.	Percent Correct on Math Tasks.....	110
12.	Mean Evaluation on Math Tasks.....	113

LIST OF APPENDICES

A.	Study 1:	Interview Schedule.....	128
B.	Study 2:	Instructions to Subjects.....	130
		Characteristics of Stimulus Pictures.....	131
C.	Study 3:	Characteristics of Subjects.....	133
		Booklet Cover.....	134
		Task Instructions.....	135
		Tasks.....	143
D.	Study 4:	Booklet Cover.....	149
		Instructions to Subjects.....	150
		Number Series Task.....	167

I. INTRODUCTION

The study of sex differences has been and is still an extremely popular scientific endeavor. Routinely, researchers have divided their subjects into males and females and correlated this independent variable with various dependent variables, commonly using sex as a sub-hypothesis to the investigation of interest. Whether researchers have sought the causes of these differences in "nature" or "nurture," the underlying assumption has been the same: there are two types of people, "females" and "males" who differ in terms of specific capacities, abilities, or traits. As long ago as 1909 Franz Boas pointed out that distributions by sex of virtually any characteristic greatly overlap and warned that "types" such as the male type or the female type could be misleading abstractions (cited in Stocking, 1968, pp. 168-9). Nonetheless, these abstractions have been used, and the research task has been to find the causes for the differences that are presumed to exist between these two groups of individuals. Thus sex has been treated as a characteristic of the individual or, in research terms, a subject variable.

Some research has been done which has investigated the significance of sex in a somewhat different way. In these studies, the sex of either the confederate of the experimenter (the stimulus person), or the sex linked to a particular set of stimulus materials (art work, applications, exams) has

been the manipulation of interest rather than the sex of the subject. These studies could be said to have investigated sex as a stimulus variable.

In this section, the research findings of these two approaches will be briefly discussed. A new framework for the interpretation of these results will then be offered, and a research program designed to test the hypotheses generated by this framework and its implications will be outlined.

Sex as a Subject Variable

The classification of two types of people, male and female, is generally assumed to be a biological classification of a more or less fundamental nature. However, in practice, the specific procedure used is usually the self-report of the subject as to sex assignment. The relationship of self-report, or gender identity, to biological measures is not perfect. Recent work, based on treatment of ambiguous or contradictory sex assignment cases (Money & Ehrhardt, 1972; Garfinkel, 1970; Morrison, 1974), indicates that gender identity can be completely independent of chromosomal, gonadal, or hormonal sex. In fact, these factors can be at variance while an individual still has a consistent and clearly defined gender identity.

Cases of this type also point to a more fundamental flaw in the assumption of a bio-genetic foundation for sex. The existence of more than two chromosomal categories, the overlapping of hormonal curves of distribution, and the

identification of ambiguous cases of genitalia indicate that biology, in theory, could be interpreted as dictating more than simply two sexual categories.

Bakan made a strikingly similar point almost a decade ago:

When close and careful investigation of each of these levels of sexual differentiation is conducted, any simple male-female dichotomy breaks down. There are varieties of intersex (hermaphroditic) individuals, based on structural and physiological characteristics. The distributions of secondary sex characteristics, social role, and psychological measures characteristically show overlap of at least some degree, sometimes quite extensive. Culture affects "role" greatly. Correlations among indices.... are far from perfect, so that given individuals may be extremely "masculine" at one level and extremely "feminine" at another.

Yet the distinction between male and female as a simple dichotomy prevails in all cultures and plays an important role in the psychological life of all individuals. (p. 107)

Garfinkel (1970) refers to the dichotomy as a "moral" one that is decided on the basis of motivated compliance with the legitimate social order and not on the basis of biological, medical, urological, sociological, psychiatric, or psychological fact.

What is not being argued here is that there is no relationship of biology to behavior. It would be premature for such a conclusion. What is being argued is that the relationship between biology and psychological or social sex differences is not as clear as is generally believed. There is no clear way to biologically dichotomize sex; there are always some number of "left-overs." Socially, however, there

are just the two categories, male and female, to which individuals are assigned, and, on occasion, reassigned.

Bakan offers and then dismisses the obvious conclusion from this reasoning:

We could, of course, simply reject the notion of male and female principles and limit our considerations to the various types of data.... This would be a very tough-minded approach to the problem of sex differences. We would essentially eliminate male and female as very good scientific categories and limit ourselves to individual differences of various types and the correlations that may prevail among them. But to do this would be to give up something which we know is a very important aspect of human life. With all of the difficulties the various empirical data present, there is a validity to the dichotomy which we must accept. And there is a certain truth associated with the male and female principles, present in both men and women, which we cannot ignore. (p. 108)

What is the "truth associated with the male and female principles" and what is the important role that the dichotomous view plays in human life? Several writers have offered partial answers. Rosenberg's (1974) historical perspective suggests that "scientifically" supported sex differences have provided justification for the existing social structure and the inferior position of women in it. Herschberger (1948) also suggests a social organization rationale but in a more general sense:

The history of mankind has been the search for some difference between individuals and groups of individuals--color, geographical location, sex--that would really count for something, a difference on the basis of which society could get organized, once and for all, and settle down to a peaceful senility. (p. 171)

Levi-Strauss and Rubin suggest economic reasons:

The sexual division of labor is nothing else than a device to institute a reciprocal state of dependency between the sexes (Levi-Strauss, 1971).

The divisions of labor by sex can therefore be seen as a taboo; a taboo against the sameness of men and women, a taboo dividing the sexes into two mutually exclusive categories, a taboo which exacerbates the biological differences between the sexes and thereby creates gender....Gender is a socially imposed division of the sexes. (Rubin, 1975).

Birdwhistle (1974) suggests a biological imperative based on evolutionary necessity; in discussing masculinity and femininity as display, he says that this kind of sex-typed behavior is necessary because we are a weakly dimorphic species. Presumably, without clear signals denoting who are the men and who are the women, reproductive activity would become threateningly inefficient.

Interesting though these different historical and sociological explanations are, they do not address themselves to the psychological questions involved in an assumption of sex differences. There are many such interesting questions. For example, are there sex differences in cognitive or perceptual processes, in sensitivity to reinforcement paradigms, in developmental patterns, in social behavior, in personality characteristics, in psychopathology?

The most recent review of the bulk of the psychological sex difference literature (Maccoby & Jacklin, 1974) discloses that there are surprisingly few reliable findings. Long-believed differences between the sexes are found to be either not empirically supported or explicable by other factors in

the situation, the stimuli, or the data analysis. Hypotheses that girls are more "social" than boys, that girls are more "suggestible" than boys, that girls have lower self-esteem, that girls are better at rote learning and simple repetitive tasks that require high-level cognitive functioning and the inhibition of previously learned responses, that boys are more "analytic," that girls are more affected by heredity and boys by environment, that girls lack achievement motivation, and that girls are auditory and boys visual are all exploded as myths. Several other research areas are judged inconclusive, such as compliance, anxiety, competitiveness, and dominance.

Maccoby and Jacklin emerge from their review with four sex differences they conclude are fairly reliable given the evidence now at hand: an advantage for females in verbal ability, and advantages for males in mathematical ability, spatial visualization, and aggressiveness. The first three differences are not evident until adolescence, and the fourth, physical aggressiveness, is found only in childhood, disappears by adolescence, and permits a far narrower interpretation than generally given: it is unrelated to field independence, activity level, or achievement motivation. Clearly, these limited findings are not sufficient explanation for the pervasive social differences between the sexes. And yet, these social differences are clear: girls get better grades; boys are class presidents; women care for children; men run corporations.

The most recent effort to find the individual psychological difference that will account for social sex differences involves subject sex-role differences. Underlying this approach is the assumption that psychologically there are not two kinds of people, masculine and feminine, but at least three, including androgynous ones who are neither masculine nor feminine but both. In order to allow for this possibility, masculinity and femininity have to be conceptualized not as opposite ends of a single bi-polar dimension but as independent dimensions, an idea which had been suggested by several people (Carlson, 1971; Bem, 1972; Constantinople, 1973; Block, 1973). Scales were developed by Bem (1974) and Spence and her colleagues (1974) to distinguish these various sex-role "types."

Bem (1975) has correlated scores on the Bem Sex Role Inventory with behavior in sex-typed situations. In general, she has found, as hypothesized, that strongly masculine or feminine individuals perform well only in situations consonant with their sex-typing and show a deficit in cross sex-typed situations. Individuals who are less strongly either masculine or feminine perform approximately the same regardless of the task. Bem concludes that these androgynous people who are less restricted by sex-typing are more situationally adaptive, a uniquely positive approach to a lack of sex-role preference.

One of the most interesting facets of Bem's research is the independence of masculinity and femininity from traditional sex categories. The implication is that masculine people

of both sexes are more like one another than males as a category and that feminine people of both sexes are more similar than females in general. Although this generalization seems counter-intuitive, once one substitutes "instrumental" for "masculine" and "expressive" for "feminine," it seems straightforward. The conclusion then is that there are individual differences along a continuum of instrumental-expressive, usually confounded by or interpreted as sex differences, that are more important for predicting behavior than the simple categories "male" and "female."

Sex as a Stimulus Variable

There is a body of literature which demonstrates another kind of sex difference. Money & Ehrhardt (1972) refer to sexually dimorphic responses:

Yet, as soon as the shape of the external genitals is perceived, it sets in motion a chain of communication: It's a daughter! It's a son! This communication itself sets in motion a chain of sexually dimorphic responses, beginning with pink and blue, pronomial use, and name choice, that will be transmitted from person to person to encompass all persons the baby ever encounters day by day, year in and year out, from birth to death. Dimorphism of response on the basis of the shape of the sex organs is one of the most universal and pervasive aspects of human social interactions. It is so ingrained and habitual in most people, that they lose awareness of themselves as shapers of a child's gender-dimorphic behavior, and take for granted their own behavior as a no-options reaction to the signals of the child's behavior which they assume to have preordained by some external verity to be gender-dimorphic. (p. 12)

These responses are based not on the sex of the responder but on stimulus sex differences.

The psychological literature which investigates sex as a stimulus variable offers more information about the discrepancy between social and psychological facts. Often, these studies set out to uncover a difference between subjects. For example, the widely quoted "fear of success" study (Horner, 1972) was at first taken to mean that the female subjects who had given bizarre responses to a cue involving female success were suffering from negative feelings about success that might be able to explain their own approach to achievement situations. Subsequent research (Monahan, Kuhn & Shaver, 1974), however, showed that both males and females gave bizarre responses to the female achievement cue and reasonable responses to the male cue. Therefore, the sex difference was located in the stimulus, not in the subjects.

There is other relevant work indicating that the sexes believe themselves more different than they in fact are. For example, in a study on values, Unger & Siiter (1974) asked subjects to rank order values either (1) for themselves, (2) for members of their same sex, or (3) for members of the opposite sex. They found that the subjects did not differ by sex in the values which they assigned themselves or their own sex but did believe that members of the opposite sex held different sex-stereotypic values. In fact, all the work on sex-role stereotyping indicates that there are firm and reliable beliefs in the differences between the sexes which

include specific ideas about attributes, interests, and abilities. For example, Sherriffs & Jarrett (1953) found that men and women share the same stereotypes about the two sexes. They conclude that "...virtually no behavior or quality escapes inclusion in either a male or female "stereotype," and that these stereotypes are substantially the same whether held by men or women. Nichols (1962) reported that subjects cited sex-stereotyped (but invalid) characteristics for the two sexes. Later, Rosenkrantz et al. (1968) found that there is strong agreement between the sexes about the differences between men and women.

These beliefs presumably underlie different behavior toward others which depends on the sex of the other, the primary example of which is an evaluation bias. For example, Goldberg (1968) found that essays said to be written by women were rated lower by women subjects than the same essays attributed to men. Broverman et al. (1968) found that clinicians used different indicators of mental health for women and men based on sex-stereotypes. Fidell (1970) reported that there are sex-linked discriminations in hiring: highly competent women are hired at lower levels than men with the same credentials; incompetent women are more likely to be hired than equally incompetent men. Pheterson & Kiesler (1971) found that art work attributed to women was ranked lower than the same art work attributed to men; however, if the paintings were said to be "winners" rather than "entries," there was no such difference. Clifford & Walster (1972) concluded from

their research that discrimination is seen more among the average applicants for college and jobs than among the best. Deaux & Taynor (1973) found that men are judged superior to women when applying for a study-abroad program when the merits of each are low. Thus, given identical information about a hypothetical stimulus person but different sex labels, subjects' stereotyped beliefs appear to control their responses. Taynor & Deaux call this phenomenon "perceived sex differences."

The Illusion of Sex Differences

There are three main conclusions to be drawn from the literature outlined thus far. First, differences between "males" and "females" are not great and perhaps can be better explained by other kinds of individual differences. Second, people, at least in this society and probably in others, believe the sexes to be very different. And third, people treat others differently depending on the other's sex and in ways that are predictable from presumed sex differences.

In interpreting these data, the interesting psychological phenomenon seems to be the discrepancy between actual sex similarities and the belief in sex differences. The traditional debate about how much of any given characteristic (aggressiveness, sensitivity, musical ability) is given to or learned by the sexes pales before this robust phenomenon. To whatever extent differences do in fact exist between the sexes, they are like what Unger and Stifter (1974) have called the

"grain of truth" that supports a huge superstructure of beliefs, many of which are illusory.

Scientists may be firmly convinced that men as a group are superior to women as a group on spatial-visualization tasks, for example. But this well-established sex difference plays no part in the everyday belief in sex differences. Most people do not even know about men's superior ability to rotate three dimensional objects mentally. Much more popular are beliefs about women's greater emotionality or sensitivity and men's leadership, assertiveness, or decision-making ability. These beliefs are illusory sex differences in that they are not supported in the psychological literature. They are not illusory in the sense of providing a set of norms or ideals to which people conform their behavior. They set the occasion for a self-fulfilling prophecy. As a "nonconscious ideology" (Bem & Bem, 1971), they may operate as an information processing bias that notices only confirming cases and is insensitive to a statistical notion of correlation.

The "scientific categories" of male and female that Bakan suggests we might consider discarding may not be exhaustive in the sense of discriminating only two "types" of people biologically or in terms of personality, but they are extremely interesting social psychological categories. Considering social categories as the determinant, one then has a picture of individuals of widely varying traits, abilities, interests, and even of varying bio-genetic make-up striving to fit into these categories, being assigned and complying with

the specific social requirement of these labels.

As a first step toward investigating this phenomenon from this perspective, the proposed research will address the way in which sex as a characteristic is treated in relation to other physical characteristics. Is sex more salient than other characteristics? Is it in fact the most salient feature of other individuals, something we notice under all conditions, even when it is not objectively important to any discernable goal attainment? Its centrality to an ideology, its function as a cue to behavioral repertoires, its socially determined consequences would suggest that this is so. The first area of investigation explores the salience of sex relative to other characteristics. A general isomorphism is proposed between statistical distinctiveness and psychological distinctiveness for noticing and remembering physical characteristics. That is, people will tend to remember and report first those things that are more distinctive about another. They will code or communicate those characteristics in an approximation to how much those features distinguish the person from the rest of the population. However, sex as a characteristic is treated somewhat differently. It is hypothesized that sex is so psychologically distinctive that it holds a cognitive place of importance that violates the statistical rule used for other characteristics.

The cognitive importance of sex, demonstrated in an information-processing bias and response bias, does not provide a clear behavioral imperative, however. What mediates

this cognitive importance and behavioral outcomes? What is the process by which sex operates as a psychologically meaningful variable? The second area of investigation studies expectancy as a possible mediating factor. To what extent does the expectancy of a difference produce that difference? If people believe in sex differences, they expect sex to make a difference. Do they also expect other physical characteristics like handedness and eye-color to make a difference? Probably only in a limited way. But given the expectancy of differences based on such physical characteristics, will people behave differently, consonant with such expectations? The focus will be to assess the performance effect of an expectancy of various kinds of differences based on physical characteristics. Some of the expectancies contradict pre-existing notions; their effect demonstrates whether and to what extent long-held stereotypic expectancies are changeable.

It is hoped that an exploration of the discrete steps that lie between beliefs and behavior in terms of established psychological relationships and processes will shed new light on the vast social differences that exist between the sexes and open new possibilities for change.

II. SEX AS A SOCIAL LABEL: STIMULUS DISCRIMINATION

In social psychology, the term "discrimination" is most often used by those interested in prejudice. For example, Allport (1954) defines it as an aspect of "acting out prejudice":

Discrimination: Here the prejudiced person makes detrimental distinctions of an active sort. He undertakes to exclude all members of the groups in question from certain types of employment, from residential housing, political rights, educational or recreational opportunities, churches, hospitals, or from some other social privileges. Segregation is an institutionalized form of discrimination, enforced legally or by custom (p. 15).

Discrimination comes about only when we deny to individuals or groups of people equality of treatment which they may wish (p. 50).

In this general definition there appear to be several key elements:

- (1) distinctions made between people or groups of people
- (2) behavior based on these distinctions
- (3) negative consequences to another as a function of the actor's behavior
- (4) prejudiced personality of the discriminator
- (5) goals or wishes of the discriminated

It is the first element that refers to the more general definition of discrimination in the sense of mere discernment. This kind of psychological discrimination, which here will be called stimulus discrimination, is crucial for the operation of prejudice. Obviously, one must first notice a particular charac-

teristic of another in order to take it into account when behaving.

The literature on social perception offers remarkably little guidance in this basic area of discrimination as mere discernment of physical characteristics or the classification of stimulus persons into social categories. In general social perception has studied two kinds of phenomena: the effects of social and cultural variables on object perception (e.g., Bruner & Goodman, 1959; see Tajfel, 1969 for a summary and review); and the perception of social objects, persons, in terms of their emotions, personality traits, or other qualities (e.g., Heider, 1958; Jones & Davis, 1965; see Taguiri, 1969 for review). Those studies that have used both social variables of the subject and of the stimulus have tended to focus on the attitudes or characteristics of the subject (Horowitz & Horowitz, 1938; Allport & Kramer, 1946; Lindzey & Rogolsky, 1950; Tajfel, 1969; Falbo, 1975).

There is, however, a theoretical thread that runs through this literature that is related to the question of social stimulus discrimination. The theoretical assumption has been that people categorize for specific purposes or to obtain immediate goals. Tajfel (1969) uses the term "functional salience" to describe the learning of minute discriminations for occupational purposes, i.e. wine-tasting or reading X-rays (p. 361).

It is clear that there are times when specific characteristics of another individual are relevant to an instrumental

goal. For example, if one wants help moving furniture, one might notice the muscularity of a potential helper; if one wants to hire a receptionist, one might notice smile, tone of voice, etc., to estimate friendliness of an applicant; if one wants a sexual partner, one usually has a preference for one sex and would therefore notice the sex of another. Immediate goal relevance would appear to be an important and obvious determinant of the salience of various characteristics.

There is some evidence, however, that sex is highly salient regardless of its goal relevance. Norman (1976) reports that in studies of selective attention a curious result emerges. Subjects receive dichotic stimulation, i.e., separate messages are given simultaneously to each ear, but are directed to attend to only one message by repeating it as it is given. Afterward, when they are asked questions about the "rejected" message, the one they were ignoring, they reliably report the sex of the speaker despite the fact that they cannot report the content of the message or even whether it was in English. Such evidence, fragmentary as it is, raises the possibility that the salience of sex as a characteristic is unrelated to any particular goal and cannot be explained by this limited definition of functional salience.

The following studies were designed to explore the salience of the sex of a stimulus person in situations where that characteristic was and was not goal relevant. In both situations the salience of other obvious physical characteristics of the stimulus person was also assessed. The over-

arching question was the judged relevance of sex compared to other characteristics.

Study 1: A Field Demonstration of Social Stimulus Discrimination

There are many brief social interchanges in the course of a day, perhaps particularly in the course of an urban dweller's day. Most of these have a particular instrumental goal, whether it is the repair of an appliance, the transportation from one location to another, the purchase of an item, the service of lunch in a restaurant, or some similar well-defined goal. The general behavioral requirement of the other persons is that they perform their part of the interchange with at least a minimum of competence and efficiency.

The degree to which the physical characteristics of the other person contribute to goal attainment is often quite small. In general, it hardly matters if the taxi driver is relatively short or if the waiter has a beard. One would probably only pay attention to these characteristics if they threatened to interfere with the goal, that is, if the taxi driver could not reach the accelerator or if the waiter's beard trailed in the soup.

For this field demonstration, a situation has been chosen in which no characteristic of the other is of particular importance to goal attainment. Respondents were asked to describe the person from whom they just bought a subway token. This social interaction has several key features: it

is brief; it is instrumental and the goal of obtaining a token is probably paramount; the respondent probably never saw the stimulus person before and probably never expects to see her/him again; the instrumental interchange can be successfully completed despite a wide range of characteristics of either the buyer or the seller; and the job is not sex-stereotyped. The question is what characteristics of the seller are noted? In particular, does the respondent judge sex a critical descriptive characteristic? It was hypothesized that people would notice sex more frequently than other characteristics.

METHOD

Procedure: People waiting for subway trains in midtown Manhattan were approached on the subway platform and asked if they had just bought a token. If they responded in the affirmative, they were asked to participate in a study on eyewitness reporting which would only take a couple of minutes (see Appendix A for the interview schedule).

If respondents agreed they were asked to list things about the token-seller that might help in identification. After the respondent finished speaking, the interviewer probed for other characteristics that might be remembered but not volunteered. Respondents were then thanked for their cooperation, and the study was explained if they asked any questions or showed any interest.

The characteristics of the token-sellers were recorded

by an observer and the interviewer on one occasion and by two interviewers on the second occasion.

Subjects: Sixteen people served as respondents. Approximately twice this many people were approached, but only 16 agreed to the interview. Four of these respondents were women, and twelve were men. They ranged in age from about 16 to 65.

RESULTS

The data are presented in three sections: (1) the spontaneous descriptions offered by the respondents; (2) the characteristics given in response to the probes; and (3) the accuracy of characteristics compared to the judgments of the observers.

Spontaneous descriptions: In their spontaneous descriptions respondents offered a total of 34 characteristics. Most of these characteristics related to sex, age, race, length of hair, weight and clothes, but they varied in frequency (see Table 1). Sex was the most frequent, being offered as a descriptor by every subject who mentioned any characteristic at all in response to the first question. In fact, sex constitutes nearly one-third of all the characteristics volunteered. The importance of sex as a descriptor is also evident in a statement by one of the subjects who gave no spontaneous description: "I don't even know if it was a man or a woman." Interestingly, the next most frequently mentioned characteristic is race which was given by eight subjects. All other

characteristics were only given 2-3 times.

In addition to sheer frequency, the order in which characteristics were listed also provides a measure of their judged importance. Sex is the first descriptor 75% of the time and first or second 100% of the time. In every case when sex was not given first, race was the first descriptor.

Probes: After the spontaneous descriptions, subjects were asked specific questions about characteristics they had omitted.¹ Table 1 indicates the frequency with which subjects remembered characteristics including both spontaneous and probed responses. In general, subjects could identify many more characteristics with these direct questions, but there were differences in recall-rate depending on the characteristic. Every subject but one (99%) knew the sex of the token-seller. Most subjects (81%) also gave an age or weight estimate. Thirty-eight percent of the subjects could not recall the token-seller's height. Height, as the respondents complained, was a difficult judgment because the seller was seated. However, 38% could also not recall the length of her hair, and 31% could not remember whether or not she wore glasses.

Accuracy: Simply because subjects gave responses spontaneously or when probed, one cannot infer that they had noticed or remembered the characteristics of the token-seller. In order to determine how reasonable these responses are,

1. Race and clothing were not specifically probed.

TABLE I

Token Seller Characteristics Reported by Subjects

CHARACTERISTICS	FREQUENCY SPONTANEOUSLY REPORTED	TOTAL FREQUENCY REPORTED INCLUDING PROBES	% CORRECT
sex	12	15	100%
age	3	13	46%
color hair		12	83%
length hair	3	10	58%
glasses		11	45%
height		10	30%
weight	3	13	38%
race	8	9*	89%
clothes	2	3*	33%
other	3	3*	

N=16

*not probed

they must be compared to the conscious judgments made by each interviewer and observer. These two independent judgments varied for some characteristics as noted below:

Stimulus person #1:

sex: woman
 age: 23/40
 color of hair: black
 length of hair: medium(afro)/ short afro
 glasses: no
 height: 5'6"/5'7"
 weight: 135/heavy
 dress: black sweater/red blouse, black sweater
 other: race=Black

Stimulus person #2:

sex: woman
 age: 45/40's
 color of hair: black
 length of hair: short (curled and pulled back)/short-ish, curls, not afro
 glasses: yes
 height: sitting down (about 5'2")/sitting down
 weight: overweight (150)/slightly overweight, 150
 dress: white jacket, blue shirt/white sweater, blue shirt, chain on glasses
 other: race=Black

In cases where the judges disagreed, respondents' judgments are said to be correct if they agree with either judge. The last column of Table 1 indicates the percentage of responses that agreed with the judges. Not surprisingly, height and weight judgments show considerable disagreement. The one characteristic on which there is no disagreement at all is sex: all respondents and judges said the token-sellers were women. The most interesting contrast to this finding is the other dichotomous judgment, presence or absence of glasses. Although 69% of the subjects said they remembered whether or not the token-seller was wearing glasses, only 45% of the

time were their memories accurate.

DISCUSSION

The responses of these subjects demonstrate the importance of sex as a characteristic. The sex of the target person is remembered if any characteristics are; furthermore, it is always mentioned and usually given first. Mere attentiveness cannot explain these results since subjects were quite inaccurate in remembering the presence or absence of glasses. Clearly, subjects were attending quite selectively. The phenomenon is very like Norman's selective attention cited previously in that subjects coded the sex of the stimulus person apparently effortlessly and despite what their main focus of attention was.

The frequency with which subjects could recall the sex of the token-seller and the unanimity of their judgment would indicate that sex is indeed a highly salient category, in fact, the most salient of all physical characteristics. However, there is an additional conclusion to be drawn. The primary task here was to describe a target person to someone else. Therefore, subjects were basically making a judgment of descriptiveness. They were, in effect, saying that not only is sex salient for them but that it would be salient for others and a helpful characteristic to know in order to select this person from a crowd.

What makes these findings interesting is their irration-

ality in a statistical sense. Research (Chapman, 1967; Hamilton & Gifford, 1976) would suggest that a rational model for a judgment of descriptiveness is the statistical distinctiveness of a characteristic. For example, if someone has one leg, that is a relatively unique characteristic; although one might not notice this feature immediately (if the person were behind a counter for example), one would certainly report it first when describing that person for general identification. Sex is a moderately distinctive characteristic in this sense; by naming sex as an identifying characteristic, one rules out about half the population. However, it is probably not as distinctive as wearing glasses, having blonde hair, or being bald in the general population. It would appear that sex has a psychological distinctiveness which is far greater than its statistical distinctiveness.

In the subway study when respondents cited sex as a descriptor in preference to other kinds of characteristics, it is unclear whether they were using a rule of statistical distinctiveness. Was this one of the few token-sellers who are female? Do people assume token-sellers are male and notice departures from this expectation?

In order to determine the relationship between psychological distinctiveness and statistical distinctiveness for sex and other characteristics, a measure of statistical distinctiveness was necessary. One difficulty in arriving at a hard measure of statistical distinctiveness in this field setting was in determining the reference population the

respondent was using. Was the token-seller being compared to (1) all people, (2) people in the geographical area, (3) other token sellers, (4) other Transit Authority workers, or (5) other clerks in general? We could only know how statistically distinctive the characteristics mentioned were if we knew the range and distribution of characteristics in the reference population. Further research was needed to systematically investigate the relationship of psychological distinctiveness to statistical distinctiveness.

Study 2: Social Stimulus Discrimination in the Laboratory

This study attempted to test the relationship of statistical distinctiveness to psychological distinctiveness in the more controlled setting of the laboratory. Subjects were given specific reference groups by means of a limited number of pictures of people and asked to describe targets within that group and in relation to it. The distribution of specific characteristics within these groups was varied. For example, sex was 50-50 male/female, glasses/not glasses was 60-40, and only one person was blonde. If the target person was that one blonde, that characteristic was statistically most distinctive and should have been cited as the most descriptive feature. The order in which subjects in fact cited descriptors was considered the measure of psychological distinctiveness and was compared to the statistical distinctiveness within that group. It was hypothesized that psychological and statistical measures of distinctiveness would

be highly similar except for sex which would be judge descriptive despite its statistical distribution in the population.¹

METHOD

Procedure: Two subjects at a time entered the laboratory and were seated opposite one another. The instructions given by the experimenter (see Appendix B) emphasized that this was a simple exercise, rather "like a game," the object of which was to describe pictures to one another as accurately and rapidly as possible. Subjects each had the same ten pictures with corresponding numbers 1-10 mounted on cardboard in front of them.

Subjects assumed one of two roles, "clue-giver" or "clue-receiver," by drawing a numbered ball from a can. The one with the lower number was the first clue-giver. Clues were limited to five dichotomous characteristics, e.g., blonde hair/dark hair, and were typed on small slips of paper so that subjects could array them in any order they chose, move them aside as they gave them, etc. Target pictures were selected by the clue-giver from the can of numbered balls with

1. Race was not used as a characteristic since like sex it is undoubtedly a socially consequential feature. One might assume that it would be considered psychologically distinctive in the same way as sex except that there are strong social pressures not to "notice" race. Since the data is being obtained from what is basically a social interaction, the inhibitions against mentioning race might confound the results. It is precisely because race prejudice is such an important topic that it is being excluded here as an incidental feature to the main object of inquiry.

replacement after every round. Clue-givers were instructed to give all five clues and not to talk except for the giving of clues. Thus, the basic task of the clue-giver was to rank five clues for descriptiveness.

The clue-receivers were given sheets on which to record their guesses and to state a level of confidence (1-11) for each guess. They were asked not to indicate when they knew which picture was correct but to record a response for each of the five clues and to say "next" when they were ready for the next clue. The giving of all five clues for each target picture constituted a trial. Feedback about the correct target was given at the end of each trial.

After each block of 5 trials, subjects changed roles for a total of 10 trials each. The experimenter recorded the order of the clues given for each trial. After all 20 trials, subjects changed roles for a total of 10 trials each. The experimenter recorded the order of the clues given for each trial. After all 20 trials, subjects filled out a questionnaire of demographic data. They were also asked the characteristics of each picture to determine whether they had "seen" the characteristics as intended. They were then debriefed.

Subjects: Twenty-one subjects were recruited through a classified advertisement in a newspaper. The ad was rather neutrally worded calling for "participants for a psychological study" involving a "simple, fun exercise" that took less than an hour and paid four dollars. They were 13 males and 8 females, ranging in age from 17 to 45 with most subjects in

their late 20's and early 30's.

Subjects were scheduled in pairs. In only one case did one of the pair not show up. That single subject was discarded from the analysis.

Stimulus pictures: The 10 stimulus pictures were 3 X 5 snapshots of undergraduates photographed at a large urban college. They were all full-face head and shoulder shots. The critical characteristics which constituted the clues and their frequency were as follows:

blonde hair (1)	dark hair (9)
head covering (2)	no head covering (8)
cap or scarf	
light clothes (3)	dark clothes (7)
glasses (4)	no glasses (6)
female (5)	male (5)

Appendix B includes a table summarizing the characteristics of each of the ten stimulus pictures in their statistically distinctive order. For example, blondeness is the most distinctive feature of picture #10; light clothes is the most distinctive feature of #1, etc.

Treatment of the data: The data, which consisted of the ranking of five clues over ten trials per subject, was examined in light of three questions:

(1) How did subjects order the clues? Did they give them randomly? Did they follow a rule of statistical distinctiveness?

(2) Did subjects treat the sex clues differently from other clues, specifically were they more likely to use "male"/"female" as first clue?

(3) Was there a change over trial blocks?
Did subjects generally become more efficient
at describing pictures in the fewest clues?
Did subjects become less likely to use sex
as a first clue?

RESULTS

Fixed-order subjects: One immediately apparent feature of the data was that four subjects used a fixed order of clue presentation. The first two of these subjects adopted the following order for all 20 trials: sex; color of clothing; head covering or not; glasses or not; and hair color. Surprisingly, the next subject who adopted a fixed order used exactly the same order except that head covering and glasses were reversed. For the second block of 5 trials, hair color was promoted in the list to the following order: sex; color of clothing; hair color; glasses or not; head covering or not. The partner followed suit.

The similarity of these fixed orders for these four subjects is quite startling. In every case they used sex and color of clothing as the first two clues. One of these subjects, when questioned by the experimenter about the fixed order, said, "But it's obvious that the most important thing in describing a person is to tell their sex, then what they're wearing, then whether or not they have glasses and the color of their hair. It's only logical."

It is clear that these subjects were not giving clues randomly, nor were they following a rule of statistical distinctiveness in terms of this group of pictures. For this

reference group of pictures, color of hair is most discriminating when it is "blonde" and least discriminating when it is "dark." Invariably using hair color in the same rank order indicates that these subjects were insensitive to that distinction. Although they all may well have had a rationale about descriptiveness that is somewhat interesting in terms of the overall study, one must conclude that they did not understand the nature of this particular task. They were therefore omitted from the data analysis.

Overall order of clues: Table 2 summarizes the order in which the remaining 16 subjects ranked the ten clues. The availability of clues varied a great deal since clues like "blonde hair" and "head covering" occurred infrequently and their counterparts, "dark hair" and "no head covering" occurred correspondingly frequently. The ranking of these clues is therefore also given in percentage of availability for comparison. It is evident that subjects most frequently used sex as the first clue. They also placed the highly discriminating clues of "blonde hair" and "head covering" more frequently near the beginning of the list and their much less discriminating counterparts near the end of the list.

A contingency coefficient was computed to determine the overall degree of relationship between all ten clues and their ranking ($C = .58$). This contingency coefficient is highly significant based on the Chi-square from which it is computed ($\chi^2 = 412, 36 \text{ df}, p < .001$). It is clear that subjects did not randomly order their clues.

TABLE 2

Rank Order in Which Subjects Gave
Descriptive Clues

CLUES												
		blonde hair	head covering	light clothes	glasses	male	female	no glasses	dark clothes	no head covering	dark hair	
RANK	I	11 34%	10 28%	4 6%	14 28%	53 64%	43 56%	7 6%	7 7%	4 3%	7 5%	160
	II	2 6%	16 44%	10 16%	23 43%	19 23%	26 34%	31 29%	11 12%	8 6%	14 11%	160
	III	8 25%	2 6%	19 29%	11 21%	4 5%	2 3%	33 31%	29 31%	29 23%	23 18%	160
	IV	7 22%	4 11%	17 27%	2 4%	2 2%	2 3%	23 21%	16 17%	38 31%	49 38%	160
	V	4 13%	4 11%	15 21%	2 4%	5 6%	4 5%	14 13%	32 34%	45 36%	35 27%	160
N of clues	32 100%	36 100%	65 99%	52 100%	83 100%	77 101%	108 100%	95 101%	124 99%	128 99%		800
Mean rank	2.7	2.3	3.5	2.1	1.6	1.7	3.1	3.4	3.9	3.7		

N of subjects=16

Derivation of theoretical distribution: In order to determine whether subjects were using a rule of statistical distinctiveness to order their clues, specific predictions had to be derived from the distinctiveness hypothesis. A theoretical distribution was constructed indicating the frequency with which each clue would have occurred in each position if subjects were rigorously ordering the clues according to their statistical frequency in the reference group. In most cases it was a simple matter to determine the expected frequency. For example, #10 was drawn approximately 25 times. This number corresponded to the only picture with light blonde hair. Therefore, the clue "blonde hair" should have occurred 25 times as the first clue according to the statistical distinctiveness hypothesis. Similarly the other characteristics of that picture could receive theoretically derived rankings.

It was possible, however, that subjects had "seen" the pictures differently from the experimenter and differently from one another. Some of the pictures were rather ambiguous in that, for example, some light hair could have been described as blonde in a general way. No mention had been made of the distribution of characteristics before the task since it might have made the probabilities too apparent. However, after the task, subjects were asked to describe the pictures and indeed eight subjects saw the pictures in idiosyncratic ways. For these subjects special orders of the characteristics had to be taken into account in order to make the ranking predictions from the discrimination hypothesis.

TABLE 3

Theoretical Distribution of Clues Based on the
Statistical Distinctiveness Hypothesis

		CLUES										
		blonde hair	head covering	light clothes	glasses	male	female	no glasses	dark clothes	no head covering		dark hair
RANK	I	32 100%	36 100%	18.5 28%	22 42%	32.5 39%	15 19%		4 4%			160
	II			29 45%	29.5 57%	11 13%	42 55%	39 36%	9.5 10%			160
	III			10.5 16%	.5 1%	36.5 44%	20 26%	38.5 36%	54 57%			160
	IV			7 11%		3 4%		30.5 28%	27.5 29%	84 68%	8 6%	160
	V									40 32%	120 94%	160
N of clues		32 100%	36 100%	65 100%	52 100%	83 100%	77 100%	108 100%	95 100%	124 100%	128 100%	800

Table 3 indicates the theoretical distribution generated by this method. It would be ideal to use a goodness-of-fit test to compare the actual distribution of clues (Table 2) to the theoretical distribution of clues (Table 3), but the theoretical distribution predicts too many zero cells to make this kind of comparison possible. However, an approximation of the fit between these distributions may be visualized. Figure 1 plots the mean ranking of each clue as it was actually given and the theoretical ranking predicted from the statistical distinctiveness hypothesis. This graph illustrates the degree of linearity predicted and the rough linearity achieved by the subjects' rankings of the clues. It can be seen that the three most discriminating and the three least discriminating clues are falling on opposite sides of the median (rank 3), and thus subjects may be said to be roughly approximating the theoretical distribution.

Use of the sex clues: The mean ranking of the sex clues, "male" and "female," is quite interesting in Figure 1. They are both displaced toward the first position. In a general sense this displacement supports the hypothesis that sex will be considered more psychologically distinctive than its statistical distinctiveness would warrant.

Stronger support is offered, however, by the frequency with which sex is used as a first clue. It should be noted that the first clue is extremely important given the instructions to have the clue-receiver guess correctly as early as possible. There were in all 160 first clues given. It is

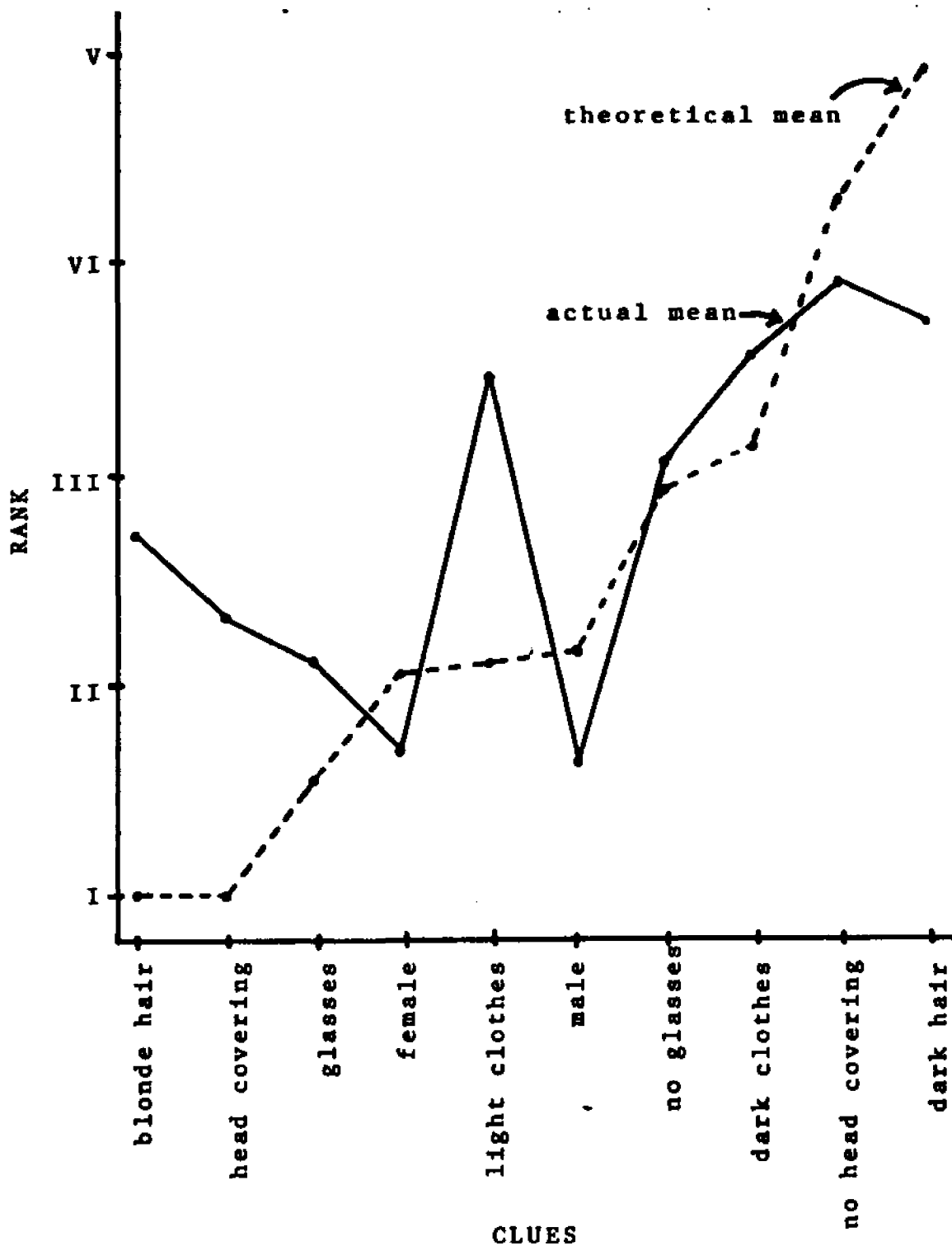


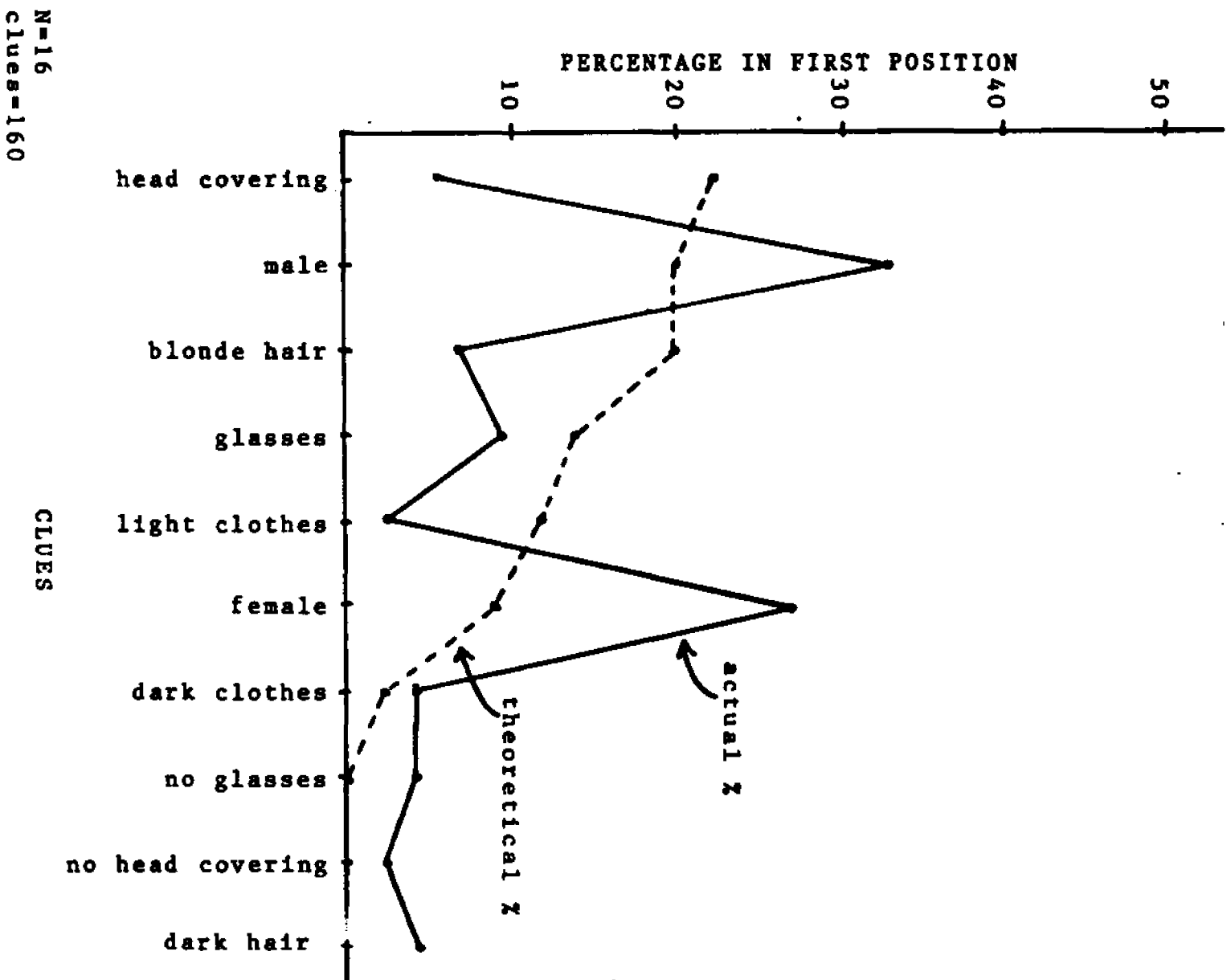
Figure 1

Mean Rank of Clues Given Compared to
Theoretical Mean Derived from Statistical
Distinctiveness Hypothesis

possible to compare the percentage of times each clue was given in the first position to the theoretically-predicted percentage using the frequencies for the first position in Tables 2 and 3. Chance, of course, would predict that each clue be given 10% of the time. Figure 2 displays the percentage of first clues of each type compared to a distribution predicted by the statistical distinctiveness hypothesis. It is clear from this graph that the clues "male" and "female" constitute the largest percentage of first clues and are different from all other clues in this regard.

Trial effects: One final interesting question is whether there was task learning over the four blocks of trials. Did subjects become more efficient in describing the target picture to their partners? In order to determine an efficiency criterion against which subjects' performances could be judged, the minimum number of clues necessary to discriminate each picture from all others was calculated. Although there were in many cases several ways to describe the picture, it was relatively simple to determine the number of clues that were critical for the most efficient description.

Table 4 summarizes the number of clues subjects used, the minimum number of clues necessary, and the discrepancy between them. The eight subjects in Trial Block I used 15 extra clues for their 40 trials. Their partners, who might be expected to have developed a more efficient strategy during their turn as clue-receivers, then used a total of 24 extra clues in Trial Block II. In Trial Block III when the initial



subjects had their second turn, they again used 15 extra clues, and when the partners tried again in Trial Block IV, they used 26 extra clues. If learning were taking place, these number of extra clues should have decreased across trial blocks. Instead, it seems as though one group of subjects was simply more efficient, and practice did not improve performance.

It was still possible, however, that subjects did improve with respect to the placement of the sex clues. Overall, subjects had used either "male" or "female" as a first clue 60% of the time (Fig. 2). The statistical distinctiveness hypothesis would have predicted the combined use of these two clues to be 30% of first clues, and chance would have predicted 20%. Did subjects give sex less precedence over trial blocks?

Figure 3 compares the percentage of times sex was given as a first clue over each of the four trial blocks. It is clear that the largest discrepancy between the actual and predicted percentages is in the first trial block. There is a dramatic reduction in the second and third trial blocks and then a substantial discrepancy again in the fourth trial block.

Overall it appears that there was less use of sex as a first clue over time. However, there seems to be a difference between the two groups of subjects. The first group, the clue-givers in Trial Blocks I and III, improved greatly from their first set of trials to their second while their partners actually got worse. This result is similar to that

TABLE 4
Number of Extra Clues Over Trial Blocks

Trial Block	Number of clues actually used	Minimum number of clues necessary	Discrepancy
I	126	111	15
II	153	129	24
III	139	124	15
IV	148	122	26

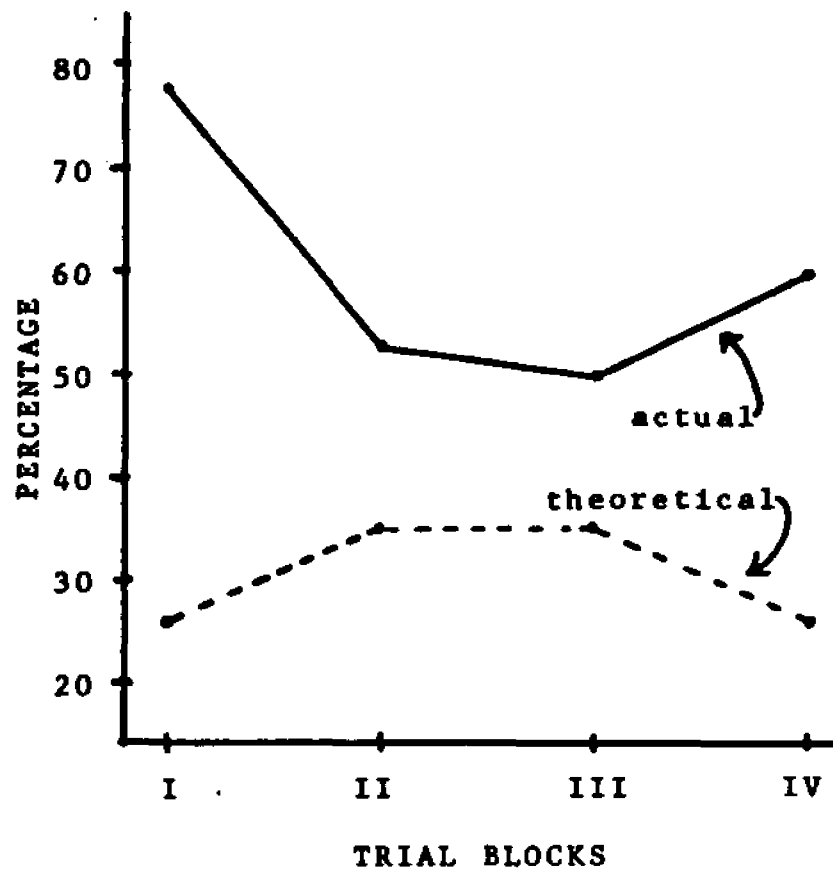


Figure 3
Percentage of Times Sex Used as a First Clue
Across Trial Blocks

obtained by the "minimum-clue" measure and indicates the important relationship between the placement of sex as a first clue and overall efficiency at the task.

DISCUSSION

Taken together, these findings generally supported the hypothesis that subjects use a rule of statistical distinctiveness to rank order descriptive characteristics but violate that rule for the characteristic of sex. Subjects did consider sex highly descriptive, so much so that they almost always gave it as a first or second clue. In a statistical sense and according to the rules of this "game," this precedence was an inappropriate strategy. By giving sex as a first clue, the partner had to make a guess from five pictures and therefore only had a 20% chance of correctness on the first clue. For most stimulus pictures there was a different clue available that narrowed the field to fewer pictures for partner to guess from.

Instead of narrowing the field for partner, subjects in general preferred to give the sex of the stimulus person. The fixed order subject who said it was only "logical" was referring to a "psycho-logic" rather than a mathematical or statistical logic. With a small set of pictures in plain view, given instructions to use the fewest clues, and given feedback at the end of every trial, subjects ignored the logic inherent in this situation.

The possibility that subjects did not understand the goal of this task is contradicted by the rankings given to the other clues. In general, they did place the most distinctive clues near the beginning of the list and the least distinctive clues near the end. They were to some degree following a rule of statistical distinctiveness for the other clues. The decrease in the use of sex as a first clue over trial blocks also indicates some understanding of the goal of the game.

It is true, however, that subjects did not understand the task very well. This limited understanding is clear from the overall lack of improvement over time. Although the task was by no means as easy as it sounds, practice should have had an effect. It was thought possible that a change in the instructions to strengthen the feedback would lead to increased learning. The next study was undertaken with this in mind.

Study 2a

METHOD

Procedure: The procedure was basically the same as in Study 2 with the following modifications to make clearer the nature of the task:

- (1) The clue-receiver was instructed to announce when (s)he knew which picture was correct, and to stop the trial at that point.
- (2) Each subject was given a practice trial.
- (3) After the practice trial, the characteristics of each picture were reviewed to verify

that there were indeed only three persons in the stimulus pictures wearing very light clothes and that there was only one blonde.

(4) To eliminate some of the trials on which sex would be appropriate as a first clue, ball #9 was taken out of the can without informing the subjects.

RESULTS

Order of clues: No subjects in this group used fixed orders for presentation of the clues. Table 5 indicates the actual order in which the subjects gave the clues. Again, subjects used sex frequently as the first clue, but, when available, they gave the clues of "blonde hair" and "head covering" an even larger percentage of the time and very rarely used "no head covering" or "dark hair" as first clues.

The contingency coefficient for Table 5 is high and significant ($C = .50, \chi^2 = 189, 36df, p < .001$) indicating that subjects were far from random in ordering their clues. Because the frequencies in positions 4 and 5 were low, these categories were combined for this analysis.

Table 6 presents the theoretical distribution of clues predicted from the statistical distinctiveness hypothesis and Figure 4 compares the actual to the theoretical distribution. Again, the approximation of the mean order in which subjects actually gave clues to the theoretical prediction is apparent. In fact, the two distributions are much closer than in the prior study.

However, despite this effect, the clues "male" and "female" are once again given precedence as a first clue. Figure

TABLE 5

Order in Which Subjects Gave Clues - Study 2a

CLUES												
CLUES		RANK										
I	14 74%	23 61%	17 33%	31 40%	40 48%	42 57%	3 7%	3 4%	2 4%	5 9%	180	
blonde hair												
II	2 10.5%	10 26%	18 35%	25 32%	17 20%	20 27%	15 37%	36 48%	8 16%	19 34.5%	170	
head covering												
III	1 5%	3 8%	11 21%	16 21%	22 26%	10 14%	12 29%	20 27%	15 30%	19 34.5%	129	
light clothes												
IV	2 10.5%	1 3%	4 8%	3 4%	3 4%	2 3%	7 17%	12 16%	13 26%	8 15%	55	
glasses												
V	1 3%	2 4%	2 3%	2 2%	2 2%		4 10%	4 5%	12 24%	4 7%	31	
male												
female												
no glasses												
dark clothes												
no head covering												
dark hair												

N of Ss=18

Theoretical Distribution of Clues - Study 2a

blonde hair

head covering

light clothes

glasses

male

female

no glasses

dark clothes

no head covering

dark hair

I	19 100%	39 100%	42 79%	42 62%	20 25%	18 29%				180
II			11 21%	13 19%	35.5 44%	43.5 71%	38 100%	20 28%		161
III				13 19%	24.5 31%			37.2 53%	15.7 41%	106.1
IV								6.7 9.5%	15.7 41%	38.1
V								6.7 9.5%	6.7 18%	20.1
N of clues	19 100%	39 100%	53 100%	68 100%	80 100%	61.5 100%	38 100%	70.6 100%	38.1 100%	505.3

N of Sa=18

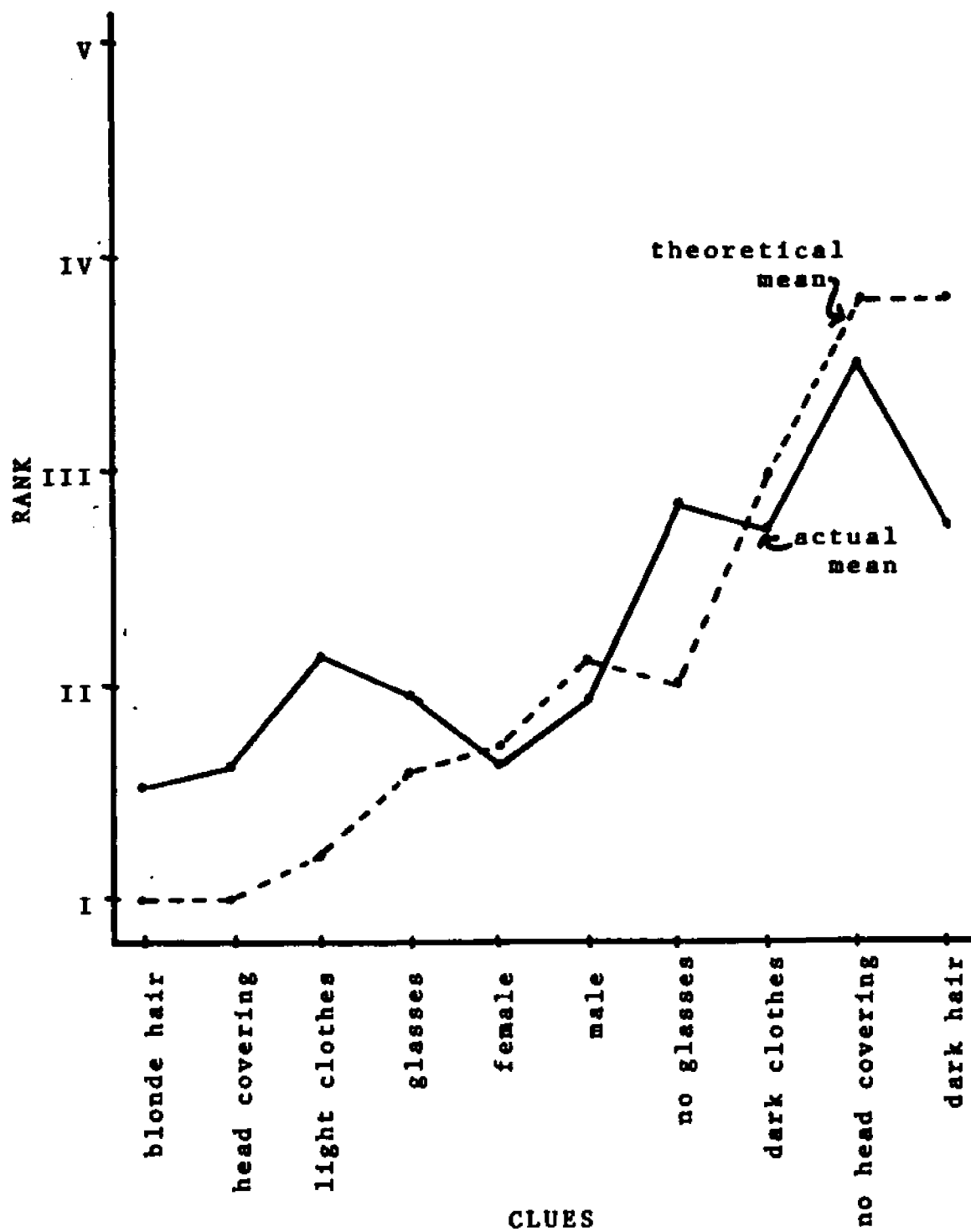


Figure 4

Mean Rank of Clues Given Compared to
Theoretical Means - Study 2a

5 illustrates the large percentage of first clues that are "male" or "female."

Trial effects: The efficiency of this group of subjects in light of the changes in procedure is particularly interesting. With increased feedback, it was thought that subjects' learning might improve. Using the same comparison of actual number of clues used to minimum number of clues necessary, subjects' performance over trial blocks was examined (Table 7). In the first block of 5 trials, subjects used 25 more clues than the minimum. Their partners showed a dramatic improvement in the second block of trials, using only 11 extra clues. The initial group showed a similar improvement with only 12 extra clues in the third trial block. In the last trial block, their partners used 14 extra clues. This change across trial blocks does represent significant improvement ($\chi^2 = 8.06, 3 \text{ df}, p < .05$), although it is clear that it is the first block of trials that is worse than all others. Subjects improve immediately on the second block of trials and performance remains relatively stable.

Figure 6 indicates the percentage of times sex was used as a first clue over trial blocks. Again, clearly the greatest percentage occurred in the first trial block, as well as the greatest discrepancy between actual and predicted percentages. In general there appears to be improvement over time. Although the subjects of Trial Blocks II and IV used sex as a first clue more often in their second block of trials, it was also more distinctive; the discrepancy between the

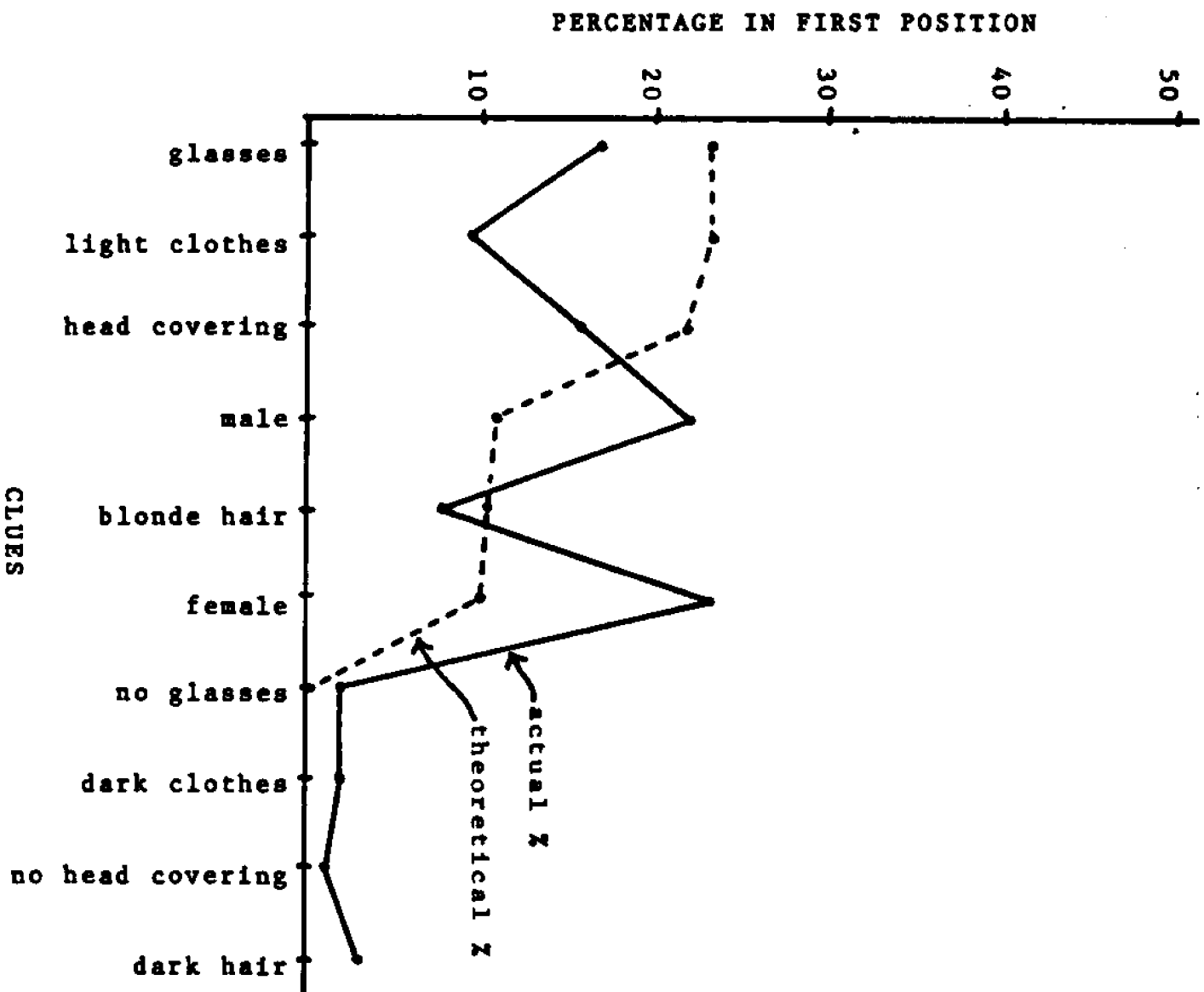


Figure 5

Percentage of Each Clue Given and Expected
in First Position - Study 2a

TABLE 7
Number of Extra Clues Over Trial Blocks
Study 2a

Trial Block	Number of Clues actually used	Minimum number of clues necessary	Discrepancy
I	155	130	25
II	130	119	11
III	141	129	12
IV	141	127	14

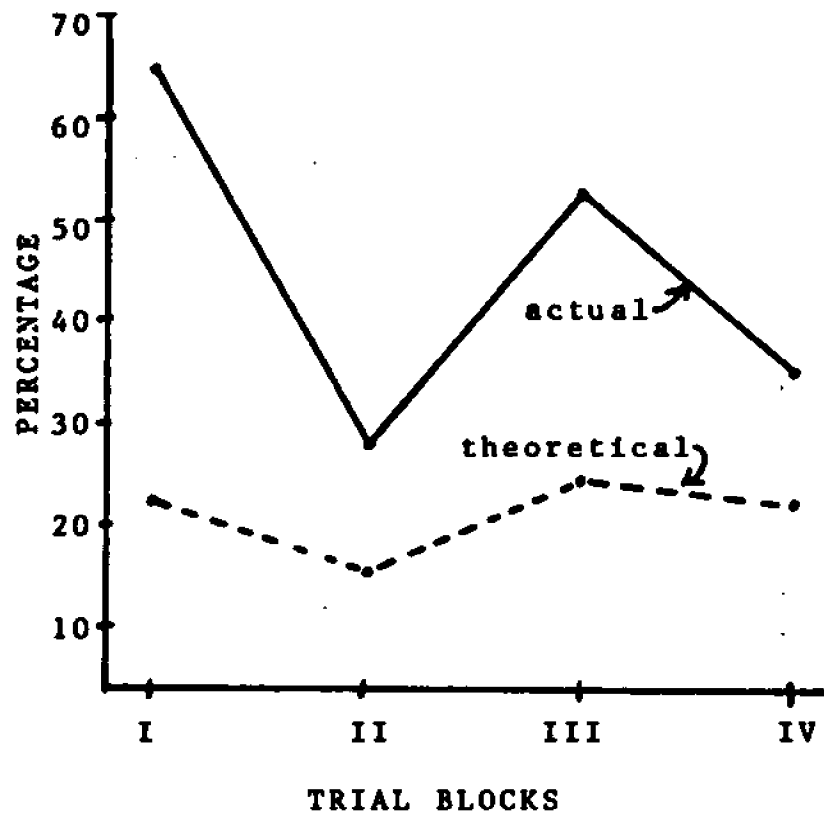


Figure 6
Percentage of Times Sex Used as a First Clue
Across Trial Blocks

actual use and the frequency predicted by the statistical distinctiveness hypothesis is just about equal for those two trial blocks.

DISCUSSION

Overall, these findings support and strengthen those of Study 2. With the procedure of immediate feedback used in this study, subjects were much better at using the rule of statistical distinctiveness as is apparent in the closer fit between the actual and the theoretical rankings of all the clues. Despite this more efficient use of that rule, subjects still overused sex as a first clue to approximately the same degree. Thus, the violation of the rule is even clearer.

Another piece of information gained in this variation of the laboratory game is the finding of some improvement across trial blocks. Undoubtedly this improvement was also because of the change in procedure which provided more immediate feedback by stopping the trial when the partner knew which picture was being described. Thus, subjects did understand the rules of the game and adopted a relevant strategy which practice improved.

CONCLUSIONS

The irrationality of subjects' using sex as a first clue in both the field and laboratory studies stands in marked contrast to the rationality of much of their strategy.

Clearly, sex as a category is considered extremely important. This importance is not functional in any immediate sense in the subway situation and is actually counter-productive in the laboratory game. It appears that sex has acquired a cognitive salience that is to a large extent independent of the immediate situation.

These studies raise many interesting questions. For example, is there a developmental trend in the acquisition of this salience? Would children be more likely to use a rule of statistical distinctiveness equally for all characteristics and not treat sex as an especially important characteristic? Tracing a developmental trend would be helpful in determining the factors that lead to this acquisition.

What contingencies would be necessary to reverse this salience or even to bring it under the control of the immediate situation? In most job-hiring situations employers are being constrained by the law to ignore the sex of the applicant. Can they do this? What would make compliance with that requirement more likely? Can the cognitive set to treat sex differently from other characteristics be broken?

Most important, if individuals use sex as a critical social label for categorizing others, how does their own sex label affect their self-perception and behavior? To what extent does sex as an important social label not only shape stimulus input but also influence an individual's personal aspirations, motivations, and achievement-related behaviors? The next set of studies was designed to explore these questions.

III. THE EXPECTANCY OF SEX DIFFERENCES: THE INFLUENCE OF OWN SEX LABEL ON BEHAVIOR

The previous studies investigated the possibility that sex is cognitively processed differently from other characteristics. They do not indicate why this might be so or what other kinds of behavior might be influenced by such a process. A reasonable reaction to these findings might be that on an important level, perhaps sex is a meaningful category because it implies not only certain physical characteristics but also kinds of behavior, attitudes, etc. There was a time, of course, when handedness also implied a great deal; however, being left-handed is no longer taken to mean that someone is sinister, possessed by the devil, or intellectually retarded. If one held all those beliefs related to handedness, it would be important to know that characteristic of others, and handedness would certainly be an important social category. Knowing someone's handedness would provide one with useful information for one's own behavior in response to that person.

In the same way, the category of male or female implies numerous specific beliefs about abilities, interests and traits; it is an extremely informative category if these beliefs, or stereotypes,¹ are based on fact. The critical

1. The word stereotype is often reserved for false beliefs. Here, the more general definition of stereotype will be adopted: "generalizations about the characteristics of people in some designated social category which are widely shared within a given culture" (Brannon, 1975).

research issue is to untangle the facts about sex differences from the beliefs about those differences. To what extent are there real differences between the sexes and to what extent are apparent differences a result of the belief itself? Does the belief in sex differences occasion its own fulfillment?

The "self-fulfilling prophecy" and the "Rosenthal effect" are both used to describe the phenomenon that people tend toward behavior which confirms their beliefs. This behavior usually represents such a subtle change that the behavior is not aware that (s)he is in effect making the belief come true. Hass (1969) has elaborated on the basic notion and outlined three processes for confirmation of one's prior beliefs: (1) one may perceive confirmation by selectively attending to information; (2) the environment may be changed or restricted so that only certain outcomes are possible; and (3) the individual who is the object of the prophecy or of others' beliefs may internalize that prophecy and change behavior to fulfill it.

The present set of studies investigates this third process. The overall goal of the studies is threefold: (1) to verify the everyday belief in sex differences as related to the specific abilities being tested; (2) to manipulate those beliefs by giving new ones about sex differences and assess the impact of the new beliefs on behavior; and (3) to compare beliefs related to sex to ones related to other characteristics like handedness and eye-color.

In addition to sex, handedness and eye-color have been chosen because, like sex, they are given, non-changeable characteristics of the subject which may plausibly be taken to affect performance. At the same time they are not emotionally-laden characteristics like race, religion, or country of origin that might offend subjects or at least elicit a complicated response. At its simplest, sex is merely a physical characteristic like eye-color; it is important to treat it as such to determine to what extent expectancies linked to it are similar to or different from expectancies linked to other characteristics.

The most parsimonious way to describe the intervening process being investigated here is one of expectancy. Motivation theorists have long treated expectancy as a crucial variable that influences performance. The assumption has been that if someone expects to succeed, (s)he will try harder; if (s)he expects to fail, (s)he will not bother. Many studies have shown that the expectancy of success influences performance (Tyler, 1958; Crandall, Katkovsky, & Preston, 1962; Feather, 1963; Battle, 1966). In fact, Crandall (1975) concludes from years of such research that expectancy is even more important than actual competence: "skill or ability itself does not constitute a dynamic, activating variable, but lies dormant until and unless the subject is motivated to approach the goal (p. 654)." Like Crandall, the assumption in this investigation will be that performance differences based on sex label are greater than

any competence differences that may exist and that these performance differences are mediated by expectancies.

Four kinds of tasks have been chosen. Three of them are conceptually related to the subject sex differences Maccoby & Jacklin (1974) say are fairly reliable: (1) verbal ability; (2) mathematical ability; (3) spatial relations¹; and (4) motor coordination. These four tasks provide an interesting range of beliefs about sex differences. It is probable that in verbal and mathematical ability psychologists and non-psychologists alike believe the sexes to be different with a female advantage for verbal ability and a male advantage for math ability. In spatial relations or embedded figures, however, it is quite likely that only psychologists believe there is a difference. Most people, it would seem, are not aware of the finding of a male advantage for rotating three-dimensional objects mentally. Motor coordination is probably not considered sex-related by anyone. Therefore the potential emergence of sex differences across tasks becomes quite interesting in terms of differences suggested in the literature and the everyday belief in sex differences.

In terms of everyday beliefs about differences related

1. The particular task chosen is not a traditional spatial visualization or embedded figures test. The male advantage on such tests has been hotly debated (see esp., Parlee, 1974) and appears to be task specific and affected by the number of trials among other variables. For the purposes of this study, a task was needed that was plausibly related to this ability for the naive subject but allowed for a range of scores and ambiguity as to success or failure. Most traditional embedded-figures tasks are a series of all-or-none insight problems.

to the other characteristics being studied, it is probably quite commonly believed that left-handed people are more ambidextrous and have superior motor coordination with their dominant hand. However, handedness is not usually linked with the other three abilities. Eye-color is probably not commonly linked to any of the abilities being tested, although apparently some psychologists believe there is a relationship of eye pigmentation and academic performance (Gary & Glover, 1975).

Study 3: The Expectancy of Characteristic Differences

The overall design of Study 3 was quite simple. There were just two groups, an experimental and a control group. Both groups were given expectancies of differential performances based on handedness, sex, or eye-color. However, the experimental group was told which characteristic was related to success, i.e. the direction of the differences.

The sex-linked expectancies given were opposite to the reported findings, that is, men were said to be better at this specific kind of verbal task and women at the mathematical one. The motor coordination expectancy was also assumed to go against the generally held expectancy, that is, experimental subjects were told that right-handed people were more ambidextrous. The expectancy linked to eye-color was assumed to be in the same direction as any expectancy the subject may have had already although it was thought more likely that subjects would have no expectancy of

ability differences based on eye-color. Previously held expectancies were tested by asking all subjects at the completion of the tasks to judge the ability of each of the groups on each of the tasks.

It was hypothesized that these instructional expectancies would differentially influence the performance of the experimental group as compared to the control group. Whether or not subjects in fact differ along these lines (e.g., females have greater verbal ability; left-handed people are more ambidextrous), it was hypothesized that expectancy would be the more important predictor of performance. Such a finding would have implications for the extent to which the belief in sex differences (or similar kinds of differences) is a self-fulfilling prophecy. It would support the view that competence on these kinds of tasks is equally distributed over people and that relevant performance is differentially affected by expectations. Most important, it would indicate that these kinds of expectancies are changeable.

METHOD

Subjects: Ninety-eight introductory psychology students at Pratt Institute, a college of Fine Arts, served as subjects. Of these, twelve were dropped from all analyses because of a failure to understand the directions. The specific criterion for elimination was a score of four or lower on the verbal task. Most, but not all, of these

subjects were students for whom English is a second language.

The distribution of characteristics in question (sex, handedness, and eye-color) for the remaining 86 subjects is summarized in Appendix C. It should be noted that the critical characteristics were not evenly distributed. Only 15.1% of the group were left-handed, and only 32% were light-eyed. More surprising was the fact that in these introductory psychology classes, 72.1% of the students were male and only 27.9% were female.

The age of the subjects ranged from 17 to 44, but the average was 21.8 and the modal age only 19. Subjects were in general in the second year of college and had a mean grade point average of 2.97, a B-. For 14% of the students, English was a second language. There were no apparent differences between the sexes on any of these characteristics (age, year in college, grade point average, language, handedness, eye-color) as tested by a Chi-square analysis. However, an analysis of their math experience, which ranged from less than algebra to more than calculus, indicated that the men had taken significantly more math courses ($\chi^2 = 13.67, 4df, p < .008$).

Procedure: Subjects were tested during class time in groups of about 20. All materials were in booklet form (see Appendix C). The first page of the booklet gave the rationale for the study and was the same for all subjects: it explained that it is important to have information about group differences so that academic subjects can be taught

in such a way that all groups can perform their best rather than having one at a disadvantage. The booklet was similar in format to other standardized tests. Experimental and control booklets were randomly assigned and appeared so similar that subjects never mentioned noticing any differences.

Expectancy manipulation: The task instructions for the subjects in the experimental group included extra sentences which conveyed an expectancy as to who would perform better on each task, e.g., right-handed people, men, women, light-eyed people.

Tasks: Each task was timed for four minutes and the order was the same for both groups:

- (1) motor coordination: using their non-dominant hand, subjects traced their way out of a maze.
- (2) verbal ability: subjects made as many shorter words as possible out of the word "information."*
- (3) mathematical ability: twenty problems were presented; each problem consisted of a series of numbers with the last one missing. In order to solve the problems, subjects had to grasp the relationship between the given number series.
- (4) perceptual ability: subjects were asked (a) to find specific items in a detailed picture and (b) to find words within a matrix of letters.

Dependent variables: Subjects were asked to estimate

* An anagrams task included in the booklet was omitted during the first and all subsequent experimental sessions when the experimenter noticed the difficulty some of the subjects were having with the first verbal task which was far simpler.

on a scale of 1-7 how well they personally expected to do on each task (personal expectancy measure). After completion of task, subjects were asked to judge how well they thought they had performed (evaluation measure). After all tasks, subjects were asked to make judgments on 7-point scales as to the relationship of the characteristics of handedness, sex, and eye-color to specific abilities (general expectancy measure). Performance was determined by scoring each task for correctness.

RESULTS

The results of the general expectancy measure will be presented first in order to determine the background expectancies related to the abilities tested. Then the task results, which include the other three dependent variables (personal expectancy, performance, and evaluation), will be reported.

General expectancies: After the task were completed, subjects in both the experimental and control conditions were asked to indicate what kind of people in general (1) were more ambidextrous, (2) had more verbal ability, (3) had more mathematical ability, and (4) were more sensitive to detail, that is, had more perceptual ability. For each of these kinds of ability, there were separate 7-point scales for handedness, sex, and eye-color. Thus, subjects were asked to make 12 judgments as to the effect of characteristics on ability in fairly rapid succession. These responses

provide a measure of expectancies or stereotypes based on characteristics, that is, general ideas about group differences that subjects held.

Table 8 summarizes the mean general expectancies of ability on each kind of task for each group characteristic (handedness, sex, and eye-color). A repeated measures analysis of variance indicated that there were no differences between the experimental and control groups in general expectancies (Table 9). There was, however, a significant interaction between perceived task ability and characteristics. A post-hoc comparison (Tukey(a), Winer, 1972) indicated that it was the math-sex mean that was significantly different from most of the others. Subjects in general judged males to be superior in math ability. They also tended to judge females superior verbally and perceptually and left-handed more ambidextrous.

It should be remembered that 4.0 is the midpoint of the scale. In fact, in every case except math and sex, the midpoint was the modal answer. One might be tempted to conclude that except for math and sex these subjects were relatively free of stereotypes. However, the particular format of the questions, that is, their proximity to one another, their rapid succession, the total lack of disguise as to content, their implausability in many cases (e.g., who ever thought that light-eyed people were more verbal?), the fact that only the endpoints and the midpoint were labelled, all seemed to make the midpoint an extremely attractive answer.

TABLE 8
Mean Scores: General Expectancies

	CHARACTERISTIC		
	¹ Handedness	² Sex	² Eye-color
Ambidexterity	3.76	3.84	4.00
Verbal Ability	4.05	4.29	4.05
Math Ability	4.19	3.38	4.05
Perceptual Ability	3.98	4.24	3.79

1. Scale: 1=left-handed most ability...7=right-handed most ability
2. Scale: 1=males most ability...7= females most ability
3. Scale: 1=dark-eyed most ability...7=light-eyed most ability

Midpoint of Scales=4.0

TABLE 9

Analysis of Variance Summary Table:
General Expectancy Scores by Condition,
Characteristic, and Task Ability

Source	df	Mean Square	F	Significance F
Condition	1	.486	.474	n.s.
*Unit	75	1.024		
Ability	3	3.585	4.574	.004**
Condition x Ability	3	.556	.710	n.s.
*Ability x Unit	225	.784		
Characteristic	2	.236	.372	n.s.
Condition by Characteristic	2	.474	.746	n.s.
*Characteristic x Unit	150	.635		
Ability x Char- acteristic	6	6.896	10.982	.001***
Cond. x Ability x Characteristic	6	.637	1.015	n.s.
*Ability x Char. x Unit	450	.628		

*marks the effect used in testing the preceding effects.

** $p < .01$

*** $p < .001$

Under these conditions, one might wonder why subjects ever chose anything other than the midpoint.

In summary, then, it appeared that subjects most strongly held the familiar stereotype of males' superior math ability. In addition, they tended to believe that left-handed people were more ambidextrous and that females have superior verbal and perceptual ability. The first three of these stereotypes were assumed to exist when this study was designed. The information chosen for the experimental manipulation was anticipated to be in direct opposition to these prior expectancies. From these data, it appears they were. The fourth trend toward a stereotype that emerged from this analysis is rather surprising. It had been assumed that subjects would have no expectancy related to eye-color and perceptual ability. That assumption was justified. However, they did think that females had more perceptual ability than males.

These data provide the background of pre-existing expectancies against which the effects of the experimental manipulation may be evaluated. On the maze task, subjects in the experimental condition were given "scientific" information that right-handed people would be more ambidextrous, information that was counter to their prior beliefs. Similarly, on the verbal and math tasks, they read that males and females (respectively) would perform better, two counter-stereotyped expectancies. On the perceptual tasks, the information that light-eyed people would perform better can

be viewed as irrelevant to their prior beliefs, which involved a tendency to believe that females would be superior on these tasks.

Overview of the task results: The overall strategy for the analysis of the task results was to compare the scores on the expectancy, performance, and evaluation measures for each of the five tasks by means of a two-way analysis of variance, using two levels of condition (experimental/control) and two levels of characteristic (right-handed/left-handed, male/female, light-eyed/dark-eyed). If the expectancy given in the instructions affected the experimental group, an interaction would be predicted such that subjects who had the "success-linked" characteristic would perform better than their control counterparts, and subjects who lacked that characteristic would do worse. Expectancy and evaluation measures were hypothesized to vary in the same way. Main effects, at least for sex, were also distinct possibilities since the literature seems so clear that females have more verbal ability and males have more mathematical ability. Overall correlations were also computed for expectancy, performance, and evaluation scores on each task.

Maze task: Subjects in the experimental condition were told that right-handed people would perform better using their non-dominant hand than left-handed people. Performance on the maze was scored by placing a grid of 1" squares over it and counting how many squares each subject

had traversed; accuracy was not taken into account. This measure would seem mainly to be one of persistence and could reasonably be expected to reflect motivational differences: right-handed people in the experimental condition were predicted to have a higher score than their control counterparts, and left-handed people in the experimental condition were predicted to have lower scores.

Table 10 summarizes the mean scores on the expectancy, performance, and evaluation measures for this task, and Tables 11 and 12 summarize the analyses of variance that indicated significant results. There were only two significant results.

There was a strong ($p < .02$) effect for left-handed people to perform better regardless of condition. Inspection of the means indicates that although right-handed subjects in the experimental condition did perform slightly better than those in the control condition as predicted, this slight difference was overwhelmed by a huge difference between experimental and control left-handed subjects. Left-handed subjects in the experimental group, who were implicitly given the information that they would not perform as well as right-handed subjects, performed far better than all right-handed subjects and better than left-handed subjects in the control group. Left-handed subjects seemed "inspired" by the counter-stereotype information that right-handed subjects are more ambidextrous.

On the evaluation scores there was a significant ($p < .03$) interaction, but an inspection of the means shows that it was in a direction opposed to the prediction: right-handed people in the experimental condition had a lower evaluation of their performance than those in the control group, and left-handed people in the experimental condition had a higher evaluation of their performance. Comparing these evaluations scores to the performance scores, however, indicates that it is the right-handed people in the control condition whose evaluations do not seem to correspond to actual performance. They seem to overrate a rather poor performance.

Verbal Task: The verbal task was an extremely simple one; it called for making as many words as possible out of the word "information" within the time limit of four minutes. Subjects generated from 5 to 38 words with an average of 19.15. A strict count of the number of words constituted the performance score. Although some subjects came up with startlingly sophisticated words (e.g., minion), the most intelligent strategy, given the instructions and time limit, was to generate many short, everyday words (e.g., for, nor, a, I, etc.). Subjects in the experimental condition were told that males would perform better on this kind of "analytic" verbal task.

Table 13 summarizes the mean scores for males and females on the dependent measures of expectancy, performance, and evaluation. The analysis of variance indicated no

TABLE 10

Mean Scores for Expectancy, Performance
and Evaluation Measures on Maze Task

CONDITION	CHARACTERISTIC	MEAN EXPECTANCY	MEAN PERFORMANCE	MEAN EVALUATION
Control	Right-handed	4.12	17.35	3.38
	Left-handed	3.86	19.86	2.43
Experi- mental	Right-handed ^a	4.32	19.06	2.88
	Left-handed	5.00	28.83	3.83

^a
manipulated expectancy of success

TABLE 11

Analysis of Variance Summary Table: Maze Performance

Source	df	Mean Square	F	Significance of F
Condition	1	182.816	2.61	.11
Handedness	1	392.1	5.61	.02*
Condition x Handedness	1	136.5	1.95	.16
Residual	76	69.953		

TABLE 12

Analysis of Variance Summary Table: Maze Evaluation

Source	df	Mean Square	F	Significance of F
Condition	1	.812	.412	n.s.
Handedness	1	.048	.024	n.s.
Condition x Handedness	1	9.946	5.041	.03*
Residual	76	1.973		

* $p < .05$

difference on the expectancy or evaluation measures. However, the analysis of the performance scores (Table 14) showed an interaction between condition and sex that tended toward significance ($p < .06$) in the predicted direction. An inspection of the performance means indicates that most of the variance appears to be accounted for by the experimental females' poor performance. These women seem to have accepted the implicit negative expectancy in the instructions and to have performed much worse than their controls.

It is interesting that as with the maze task, the group apparently most affected by the experimental manipulation was the group that was implicitly given information that they would fail. The direction of the influence, however, was markedly different. On the maze task, that group did much better than all other groups; on the verbal task, they did much worse.

Math task: The number series task was quite difficult. Out of twenty problems eleven was the highest score of anyone in either group, and the average was only about five. On this task experimental subjects were told that females would perform better because the task reflected "basic arithmetic ability." Again, the prediction was an interaction between sex and condition.

The results summarized in Tables 15-18 indicate that this interaction was non-significant for the group as a whole. There was, however, a strong and consistent main effect for sex with men solving significantly more problems

TABLE 13

Mean Scores for Expectancy, Performance,
and Evaluation Measures on Verbal Task

CONDITION	CHARACTERISTIC	MEAN EXPECTANCY	MEAN PERFORMANCE	MEAN EVALUATION
Control	Male	4.38	18.97	4.41
	Female	4.67	22.00	4.42
Experimental	Male*	4.64	19.15	4.26
	Female	4.73	16.55	4.00

*Manipulated expectancy of success

TABLE 14

Analysis of Variance Summary Table: Verbal Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	38.18	1.01	n.s.
Sex	1	1.08	.03	n.s.
Condition x Sex	1	133.10	3.53	.06+
Residual	81	37.75		

+p < .10

($p < .04$), stating an expectancy to do better ($p < .01$), and evaluating their performance higher ($p < .01$).

This superior performance of the males was not very surprising since the men in these classes had much more experience with math than the women. Responses to one of the final questions ("What was the content of the most advanced math course you have had?") indicated that 81% of the men had taken trigonometry, calculus, or other advanced math courses whereas only 54% of the women had done so. Since this difference in math experience was so marked, an analysis of covariance was done using math experience as the covariate (Table 19). This analysis washed out the main effect for sex and indicated a strong relationship between math experience and performance ($p < .001$). However, the predicted interaction between sex and condition was still nonsignificant.

It seemed possible that the effect of math experience, which was stronger than the expectancy manipulation for the group as a whole, might be different for those with more or less math experience. Subjects who had not taken many math courses would have less concrete information about their own ability to counteract the information given in the instructions and might therefore be more susceptible to the expectancy information. Thus, separate analyses of variance were performed on the scores of subjects who had trigonometry and calculus (HIMATH group) and subjects who had algebra, geometry, or less (LOMATH group).

TABLE 15

Mean Scores for Expectancy, Performance,
and Evaluation Measures on Math Task

CONDITION	CHARACTERISTIC	MEAN EXPECTANCY	MEAN PERFORMANCE	MEAN EVALUATION
Control	Male	4.64	5.83	3.41
	Female	4.15	4.23	2.69
Experimental	Male ^a	4.97	5.97	3.49
	Female	3.55	4.82	2.46

^amanipulated expectancy of success

TABLE 16

Analysis of Variance Summary Table: Math Expectancy

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	.08	.04	n.s.
Sex	1	15.12	6.45	.01**
Condition x Sex	1	3.74	1.60	n.s.
Residual	81	2.34		

TABLE 17

Analysis of Variance Summary Table: Math Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	1.52	.20	n.s.
Sex	1	32.98	4.35	.04*
Condition x Sex	1	.85	.113	n.s.
Residual	82	7.57		

*p < .05
**p < .01

TABLE 18

Analysis of Variance Summary Table: Math Evaluation

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	.13	.06	n.s.
Sex	1	14.34	6.22	.01**
Condition x Sex	1	.21	.09	n.s.
Residual	81	2.31		

TABLE 19

Analysis of Covariance Summary Table: Math Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Covariate Math Ex- perience	1	177.68	30.77	.001***
Condition	1	1.28	.28	n.s.
Sex	1	1.73	.30	n.s.
Condition x Sex	1	8.16	1.41	n.s.
Residual	80	5.78		

TABLE 20

Mean Scores for Expectancy, Performance,
and Evaluation Measures on Math Task: LOMATH Group

CONDITION	CHARACTERISTIC	MEAN EXPECTANCY	MEAN PERFORMANCE	MEAN EVALUATION
Control	Male	4.00	5.17	3.17
	Female	3.25	2.00	3.00
Experimental	Male	3.83	3.00	2.33
	Female ^a	2.71	4.43	2.29

^a

manipulated expectancy of success

**p < .01

***p < .001

The results of these analyses (Tables 20-23) show distinctly different patterns for the two groups. In the LOMATH group, the interaction of sex by condition was significant in the predicted direction on the performance scores ($p < .03$). In the HIMATH group no such effect was found. There were no main effects on performance or expectancy for either group. However, in the HIMATH group, males and females tended to evaluate their performances differently: men rated their performances much higher than females although in fact they were not significantly different ($p < .03$).

Perception tasks: There were two separate tasks for the "perception" part of the study. "Little Tommy Tucker" was a picture in which subjects had to find 20 hidden items, such as a fish, a cup, etc. The "Word Search" involved finding specific words embedded in a matrix of letters. Thus both could in a sense be considered disembedding tasks, although one involved pictures and one involved words. Subjects were told that light-eyed people were more sensitive to light and to detail and thus should perform better on the two tasks.

Table 24 summarizes the mean scores on the dependent variables broken down by eye-color. An analysis of variance indicated no significant differences on the expectancy measure. The performance scores, although not showing the predicted interaction effects, did indicate a main effect for eye-color on the Little Tommy Tucker task: dark-eyed

TABLE 21

Analysis of Variance Summary Table:
Math Performance LOMATH Group

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	.02	.15	n.s.
Sex	1	1.58	.003	n.s.
Condition x Sex	1	29.08	5.19	.03*
Residual	19	5.61		

*p < .05

TABLE 22

Mean Scores for Expectancy, Performance,
and Evaluation Measures on Math Task: HIMATH Group

CONDITION	CHARACTERISTIC	MEAN EXPECTANCY	MEAN PERFORMANCE	MEAN EVALUATION
Control	Male	4.91	6.14	3.55
	Female	4.56	5.22	2.56
Experimental	Male ^a	5.22	6.63	3.74
	Female	5.00	5.50	2.75

^a

manipulated expectancy of success

TABLE 23

Analysis of Variance Summary Table:
Math Evaluation HIMATH Group

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	.14	.06	n.s.
Sex	1	11.17	4.65	.03*
Condition x Sex	1	.03	.01	n.s.
Residual	57	2.40		

*p < .05

people did better regardless of condition ($p < .05$; Table 25).

This finding, of course, is exactly opposite to the given expectancy. There were no differences in evaluation despite this performance difference. Thus, it appears that the information given to the subjects that light-eyed people should do better had no impact on their expectancy, performance, or evaluation. Rather, dark-eyed people did better regardless of information. There was also a trend for a main effect for sex on the "Word Search" with females performing somewhat better in both conditions ($p < .07$; Table 26). Although sex was not the characteristic associated with the manipulation in this case, it should be remembered that in their general expectancies subjects had stated a belief that women would perform better on perceptual tasks.

Correlations: Pearson product-moment correlations, computed for expectancy, performance, and evaluation within each task, are summarized in Table 27. The most interesting pattern to emerge was the virtual lack of a relationship between expectancy and performance for three of the five tasks, the maze and the two perceptual tasks ($r = -.01$; $r = .09$; $r = -.03$). Expectancy was slightly related to performance on the verbal task ($r = .21$, $p < .05$) and very strongly related on the math task ($r = .50$, $p < .001$). Thus, except for the math task, it would seem that subjects were not very accurate in predicting their performance.

TABLE 24

Mean Scores for Expectancy, Performance,
and Evaluation Measures on Perception Tasks

	CONDITION	CHARACTERISTIC	MEAN		
			EXPECTANCY	PERFORMANCE	EVALUATION
"Little Tommy Tucker"	Control	Dark-eyed	4.75	13.46	4.96
		Light-eyed	5.00	12.42	5.25
	Experimental	Dark-eyed	4.93	14.00	5.42
		Light-eyed*	4.63	12.69	5.06
"Word Search"	Control	Dark-eyed		7.96	3.75
		Light-eyed		7.75	3.67
	Experimental	Dark-eyed		7.70	3.73
		Light-eyed*		6.81	3.56
"Word Search"	Control	Male	4.38	7.21	3.45
		Female	4.67	9.39	4.39
	Experimental	Male	4.64	7.15	3.56
		Female	4.73	7.91	3.91

*manipulated expectancy of success

TABLE 25

Analysis of Variance: Little Tommy Tucker Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	10.75	1.11	n.s.
Eyecolor	1	36.90	3.82	
Condition x Eyecolor	1	2.15	.22	.05* n.s.
Residual	78	9.66		

*p < .05

TABLE 26

Analysis of Variance: Word Search Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Condition	1	4.71	.40	n.s.
Sex	1	39.50	3.34	.07+
Condition x Sex	1	9.86	.83	n.s.
Residual	81	11.84		

+p < .10

TABLE 27

Correlations Among Expectancy, Performance, and Evaluation
Scores for Each Task

	TASK				
	Maze	Word Power	Number Series	Little Tommy Tucker	Word Search
Expectancy & Performance	-.01	.21*	.50***	.09	-.03
Performance & Evaluation	.28**	.59***	.50***	.56***	.53***
Expectancy & Evaluation	.22*	.23*	.57***	.45***	.37***
Math Experience & Expectancy			.53***		
Math Experience & Performance			.52***		
Math Experience & Evaluation			.38***		

* $p < .05$
 ** $p < .01$
 *** $p < .001$

They were, however, quite accurate in evaluating it. The performance-evaluation correlations were strong and significant on every task, ranging from .50 to .59 ($p < .001$) for all tasks except the maze ($r = .28$, $p < .01$).

All the correlations for the math task, the Number Series, were very strong. Since an experience measure, the number of math courses taken, was available, it was also correlated with the dependent measures. Math experience was shown to be significantly related to expectancy, performance, and evaluation ($r = .53$; $r = .52$; $r = .38$, $p < .001$).

SUMMARY AND DISCUSSION

Summary of the Performance Scores: Of the predicted interaction effects, two are supported in a limited way by these data. (1) On verbal performance the experimental manipulation affected the females so that their scores were lower when given the information that males should do better. (2) The performance of both males and females who had limited experience with math was affected by the information that females should do better; the result of this information was that females did better and males did worse to approximately the same degree.

The other performance results were four main effects for various characteristics. (1) Left-handed people scored higher on the maze task. (2) Without taking experience into account, males solved more number series problems. (3) Dark-eyed people disembedded more items on the Little Tommy

Tucker task. (4) Females disembedded more words on the word search. Three of these four main effects are in the direction of stereotypic expectancies.

Summary of the Expectancy and Evaluation Scores: The expectancy measure, the 7-point scale on which subjects predicted their performance after reading the instructions, yielded no significant results except for the strong sex difference on the math task. The most interesting fact about the expectancy measure is that it bears little or no relationship to actual performance as evidenced by the generally low correlations. These low expectancy-performance correlations for the maze task, Little Tommy Tucker, and the word search could be because of the novelty of the tasks and the lack of perceived correspondence between the tasks and the general ability they were said to be measuring. The math expectancy, the one case in which performance is strongly related to expectancy, might indicate that this task was in a sense a "true test" of math ability. This idea is also supported by the relationship of math experience and expectancy and performance. Giving subjects a clearer idea of the nature of each task and controlling for each kind of experience might strengthen the correlations and thus the predictive value of the expectancy measure.

In contrast to the expectancy measure, the evaluation scores are generally strongly related to performance. These correlations are somewhat surprising considering the novelty of the tasks and the fact that subjects could not have had

a very clear idea of the "norms" for their performance. In some cases the group differences in the evaluation measures were not as strong as performance differences (e.g., on the verbal task) and occasionally they were stronger (e.g., the HIMATH group on the math task) but overall their pattern followed the performance pattern.

General conclusions: In addition to the effects of the manipulation and prior expectancies, there seem to be at least two other important factors that influenced these results, that is, task experience and the plausability of the new information.

The effects of task experience are clearest on the math performance results. The critical factor in the predicted interaction observed in the LOMATH group may have been the limited prior information about their abilities available to the subjects. This lack of prior information may be a requirement for vulnerability to new information such as was contained in the manipulation. In contrast, the resistance of the HIMATH group to the information contained in the manipulation may have been because of their experience with math in general.

Supporting this interpretation are the results of the maze task. The main effect for handedness on the maze task was very strong and straight-forward. It appears that left-handed people are better at using their non-dominant hand and thus are more ambidextrous. Undoubtedly, left-handed people are also forced more often to use their non-dominant

hand and are aware of their own resultant ambidexterity. Given all this experience and the possibility of real and conscious ability differences, left-handed subjects were, if anything, aroused to perform even better when given information contrary to their beliefs. A possible conclusion based on these maze and to some extent the math data is that a person who has a great deal of information and practice on a task cannot be influenced easily by a new expectancy conveying opposite information.

The inability to create an expectancy "whole cloth," i.e., when subjects had no prior beliefs about a characteristic-performance relationship such as that asserted for eye-color and perception, may be because of the implausibility of the information. In the discussion and debriefing at the end of the experimental session, subjects said that the stated relationship of ability to eye-color was laughable. They also indicated that they found the counterstereotypic handedness information implausible. It is interesting that no one dismissed the information about the relationship of sex to ability even though it was quite counter to what they believed to be true. In the subsequent discussion, it appeared that subjects were even willing to offer rationales supportive of the new information. Thus, it seems that performance cannot be affected by linking any characteristic to an expectancy; rather, of the ones studied, only sex is plausible to the subjects as a characteristic that can influence a number of abilities in either direction.

Overall, this study suggests that the belief in sex differences is different from beliefs, whether previously held or newly created, about other characteristics. Prior beliefs about sex differences relate to a number of abilities, three out of the four which were tested, and are more widely shared but at the same time easier to manipulate, that is, more malleable.

The manipulation, however, was only in a direction opposite to the stereotypes. What would happen if the manipulation were in the same direction as the stereotypes? Intuitively, one would assume that changes would be even easier to produce, that similar information would make a bigger difference if it added to a prior belief. In order to measure such changes, it is clear that a good baseline control is needed, one without any salient sex-related expectancies.

What can be said from this study about the facts about sex differences? There were two main effects for sex found. The male advantage on math was clearly related to experience, but it should be noted that experience is a bad confound since it is probably high ability people who continue to take math courses. However, if males are tracked into math courses or at least encouraged to take them and females are discouraged, then the educational system is providing a restricted environment in which only certain outcomes are possible, outcomes which fit in with widespread beliefs.

Similarly, the female advantage on the Word Search could also be experiential; they may have had more practice on this specific task. However, if this is true, one can't tell whether they do these puzzles because they're good at them or whether they're good at them because they do them. Minimally, however, there must be experience measures for all tasks.

Although the effect of the manipulation was limited in the ways described, there seemed to be enough effect to justify further investigation of manipulated sex expectancies. The manipulation, after all, was merely a few extra words embedded in each task instruction. Most interesting and encouraging was the fact that the impact was stronger on performance than on the prediction of performance. One might have anticipated exactly the opposite finding, that it would be easier to influence a mere statement than actual performance. The next study was undertaken to explore some of these findings further with a larger sample and factors like experience and plausability more carefully controlled.

Study 4: The Expectancy of Sex Differences

This study was conceptually similar to Study 3. The overall hypothesis was identical, i.e., that instructional expectancies of success linked to characteristics would influence performance in a direction consistent with those expectancies. The same dependent measures of personal expectancy, performance, and evaluation, and the same tasks

were used. The format of booklet presentation of experimental materials administered to groups of subjects was also used.

There were, however, some important modifications. First, since expectancies linked to characteristics other than sex did not seem plausible to subjects and did not effect their performance in the same way, they were omitted and all expectancies were related only to sex. In addition, in order to assess more fully the relationship of beliefs in sex differences to behavior a more elaborate design was needed.

The direction of the differences specified for the experimental group in Study 3 was opposite to existing stereotypes, that is, subjects read that males were superior on the verbal task and females were superior on the math task. The question was raised whether information supporting existing stereotypes would enhance performance sex differences. In order to assess the impact of particular kinds of information, experimental groups were needed that received both stereotypic and counter-stereotypic information. In addition, a control group similar to that of the last study was created, one which was asked about sex differences but received no specific information about the direction of such differences. Most important a baseline control was needed to provide data independent of salient sex differences. Therefore, four conditions were created for this study: two experimental groups received different expectancies about which sex would be successful on each task;

in one control group, existing sex-linked expectancies were assessed without manipulation; and in the other control group, the baseline condition, sex and sex differences were not mentioned as relevant variables.

There were also three other minor changes incorporated into this study. Specific task experience was assessed for every task in order to control for the possibility of sex differences in experience. Also, because the low expectancy-performance correlations of Study 3 could have been due to the unexpected nature of the tasks, subjects were given "sample tests" which conveyed some idea of the task before they were asked to predict their performance. In addition, experimental subjects were encouraged to pay more attention to the instructional expectancies by giving them a ranking task for reasons why the particular sex difference was true. These reasons were chosen primarily for plausability, and the main function of this task was to strengthen the manipulation.

In summary, the main hypotheses may be stated as follows:

- (1) instructional information about the relationship of sex to ability will affect estimates of personal expectancy of success, performance, and evaluation of performance in a direction consonant with that information and the subject's own sex;
- (2) instructional information that supports sex-role stereotypes will be more effective in influencing behavior than information that is counter-stereotypic.

METHOD

Procedure: All materials were in booklet form (see Appendix D). All subjects performed the same four tasks in the same order and within a 4-minute time limit for each task. Booklets representing all four conditions were distributed to each group of subjects.

Tasks: The first two tasks were presented as related to verbal ability and the second two to mathematical ability.¹ The tasks were as follows:

- (1) word knowledge: subjects made as many words as they could out of the word "information."
- (2) word recognition: subjects had to circle words embedded in a matrix of letters.
- (3) arithmetic ability: subjects chose which of two numbers was correct to complete a number series. There were 20 problems.
- (4) shape recognition: subjects found specific items embedded in the picture "Little Tommy Tucker."

Experimental and Control Conditions: There were four conditions, two experimental and two control.

Conditions 1 & 2: In these two experimental conditions the booklet was labelled "The New York Test of Sex Differences," and subjects read that either males or females were better on each of the four tasks. Subjects in Condition 1 read that females were better at the first task, males on the second, females on the third, and males on the fourth. In

1. The same tasks were used as in Study 3 for comparability. The stated relationships to general verbal and mathematical abilities are based on plausability. The actual relationship of a task such as "Little Tommy Tucker" to mathematical ability is unknown.

Condition 2, subjects read that males were better on the first task, females on the second, etc. Thus each group was told that each sex was superior on one kind of verbal and math task and inferior on the other.

Condition 3: In this control condition, subjects were also given booklets labelled "The New York Test of Sex Differences." However, rather than being told which sex had an advantage on each task, they were asked to judge on a 23-point scale whether males or females would be better and to give a reason for their judgment.

Condition 4: This control condition provided a baseline measure of performance given no information about sex differences. The test was simply labelled "The New York Test of Cognitive Abilities."

Dependent variables: Subjects in Condition 3 stated a general expectancy about the relationship of sex to performance on each task. Subjects in all conditions estimated on an 11-point scale their personal expectancy of success and made an evaluation of each performance after completion. In addition, subjects were asked to indicate their specific experience with each task. Performance scores were determined independent of the experimental situation.

Subjects: One hundred forty-two subjects from Brooklyn College served as the total sample for this study. Of these, 47 were recruited from the "subject pool," made up of introductory psychology students who had to participate in two experiments as part of their course requirement. The other 95 subjects were run during three class meetings of Introductory Social Psychology courses.

Overall, there were 72 males and 70 females in the total sample. The ages ranged from 16 to 56 with a mean of

20.7 and a modal age of only 18. Seventy-two percent were in their first two years of college. For 10.7% of the subjects English was a second language.

Comparisons of these demographic characteristics between the two groups of subjects, subject pool vs. classroom participants, by chi-square indicated several differences between the groups. There was a difference in the distribution by age ($\chi^2 = 29.93$, 16 df, $p < .02$). The subject pool participants were somewhat older with a mean age of 22.5 compared to a mean age of 19.8 for the classroom participants although more of the classroom participants were advanced students ($\chi^2 = 26.23$, 4 df, $p < .001$). For a larger percentage of those in the subject pool (19.6% vs. 6.4%) English was a second language ($\chi^2 = 4.32$, 1 df, $p < .04$). However, despite these demographic differences, no significant differences on dependent variables were evident between the two groups as tested by chi-square. Consequently, the two groups were combined for all data analyses.

Overview of the analyses: The first analysis task was to determine what beliefs about sex differences were already held by subjects in this population. In order to determine these beliefs, the general expectancy scores of the subjects in Condition 3 were analyzed by a repeated measures analysis of variance. The manipulation can then be evaluated against these background expectancies.

The overall strategy for the analysis of the task results was to do 2(sex) by 4(condition) analyses of covariance

on the other dependent measures, personal expectancy, performance, and evaluation, using the specific task experience as the covariate. The purpose of these general analyses of covariance was to determine any main effects for sex that might exist across conditions. They were also used to determine whether the differences in condition affected the subjects' behavior.

It was most important to determine whether there were significant differences between male and female subjects given the manipulated expectancies of success and failure. Therefore, planned comparisons were made among the four means of subjects in the experimental condition: males in Condition 1; females in Condition 1; males in Condition 2; females in Condition 2 (Hayes, 1963, pp. 471-477).

In addition, Pearson product moment correlations were computed to assess the degree of relationship between experience, personal expectancy, performance, and evaluation for each task.

RESULTS

General Expectancies: The general expectancies, or stereotypes, that subjects held with regard to sex differences in task-related abilities were assessed by analysis of the responses of the subjects in Condition 3, who were asked to indicate on a 23-point scale whether males or females would be better on each task. The mean scale numbers chosen by these subjects are summarized in Table 28.

Subjects indicated that they believed females would do better on all tasks except the Number Series; they judged males to be superior on that task. In addition, it appears that the female subjects were more extreme in their responses.

A repeated measures analysis of variance (Table 29) indicated that both of these effects were significant. Male and female subjects differed from one another in their responses ($p < .023$). Most important, there was a strong effect for task ability ($p < .001$): subjects judged that males and females would perform differently depending on the task. In general, then, these data indicate that subjects held the familiar stereotypes of female superiority for verbal tasks and male superiority for a math task like the Number Series.

Verbal Tasks

Experience: When asked whether they had ever done these kinds of tasks before, the overwhelming majority of subjects (81% on the Word Knowledge task and 77% on the Word Recognition task) chose the responses "occasionally" and "not often, but I have done them." There was no sex difference on these experience reports ($\chi^2 = 2.30, 3 \text{ df}, p < .51$; $\chi^2 = 2.58, 3 \text{ df}, p < .46$).

Personal Expectancy: The pattern of results for the personal expectancy measure is remarkably similar for both verbal tasks. Figure 7 displays the mean personal expectancy scores for males and females in all four conditions on both tasks. In general, the males set a much higher level of expectancy of success, regardless of condition.

TABLE 28

Mean Scores: General Expectancy

TASK	MALES	FEMALES
Word Knowledge	12.88	15.33
Word Recognition	12.65	14.80
Number Series	10.88	9.80
Shape Recognition "Little Tommy Tucker"	13.06	13.60

SCALE: 1=males a great deal better...23=females a great deal better

TABLE 29

Analysis of Variance Summary Table: General Expectancy

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Sex	1	32.88	5.75	.02*
Unit	30	5.72		
Task Ability	3	94.21	9.88	.001***
Sex x Task Ability	3	21.20	2.22	.09
^a Task ability x Unit	90	9.54		

^amarks the effect used in testing the preceding effects

* $p < .05$

*** $p < .001$

The analysis of covariance (Tables 30 & 31) indicates that this sex difference was significant ($p < .01$; $p < .05$). In addition, subjects in the control condition stated higher levels of expectancy than the subjects in the experimental conditions. This trend is particularly apparent for female subjects. This main effect for condition was significant for both tasks ($p < .05$; $p < .05$).

The effect of the information about sex differences is apparent in the differences among the means of the male and female subjects in the experimental conditions (Conditions 1 & 2). On each task the men who had received the information that males would do better set their expectancies higher than anyone else. On the Word Knowledge task these were the males in Condition 2. The planned comparisons indicated that they tended to be different from the other three groups $F(1, 121) = 3.52$, $p < .10$. On the Word Recognition task, this effect was significant, $F(1, 126) = 7.17$, $p < .01$; the males in Condition 1 had a significantly higher expectancy of success.

Performance: The performance scores also show a sex difference, but one that is opposite to the personal expectancy estimates: women performed better than men on both verbal tasks across all conditions (Figure 8). This difference was only a trend on the Word Knowledge task ($p < .10$; Table 32) but was significant on the Word Recognition task ($p < .05$; Table 33). Unlike the expectancy scores, there were no main effects for condition.

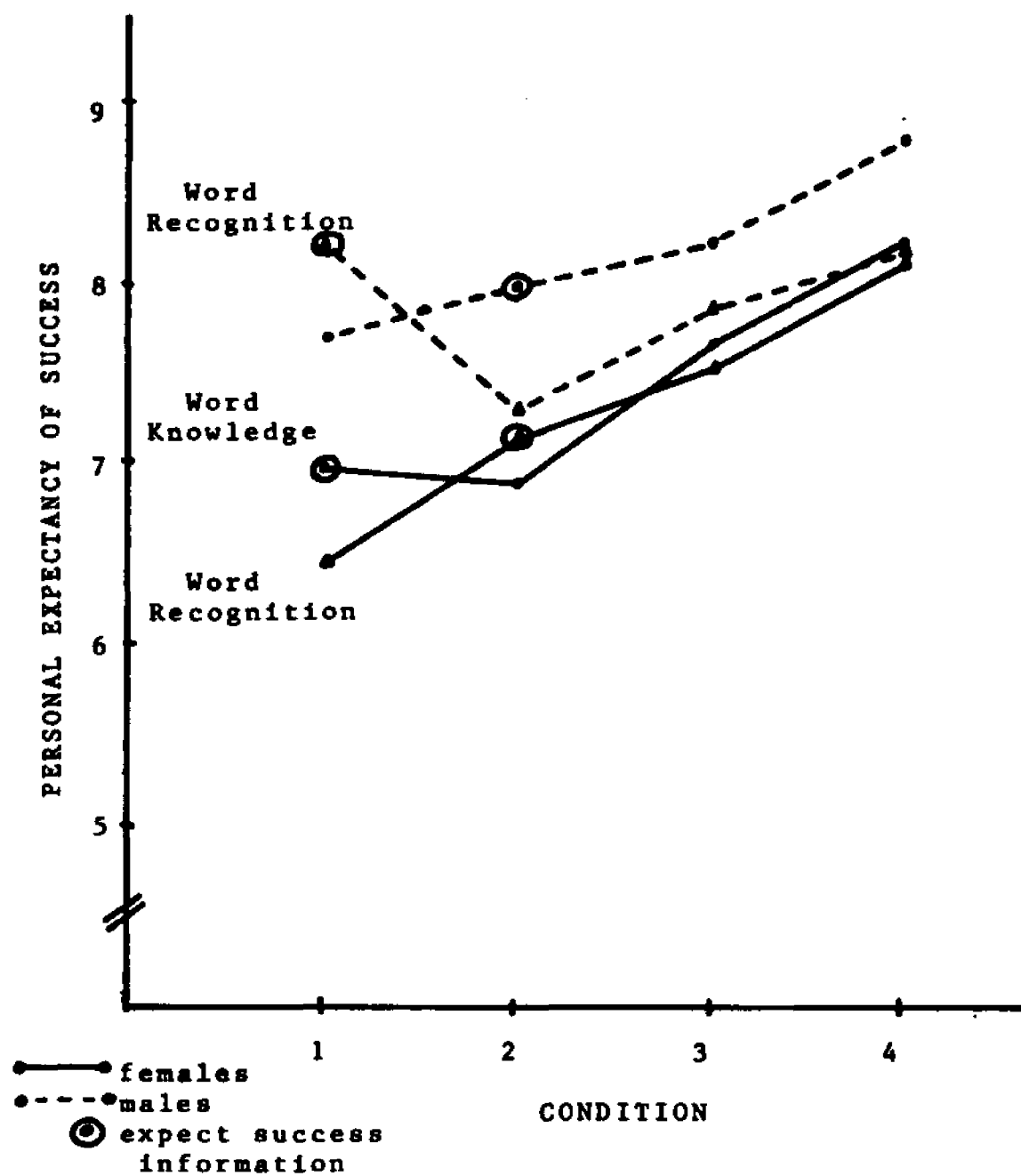


Figure 7

Mean Personal Expectancy Scores
on Verbal Tasks

TABLE 30

Analysis of Covariance Summary Table:
Word Knowledge Personal Expectancy

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	13.408	4.842	.03*
Condition	3	9.970	3.600	.02*
Sex	1	20.390	7.363	.01**
Condition x Sex	3	.641	.232	n.s.
Residual	121	2.769		

TABLE 31

Analysis of Covariance Summary Table:
Word Recognition Personal Expectancy

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	40.745	13.293	.001***
Condition	3	8.260	2.695	.05*
Sex	1	12.884	4.204	.04*
Condition x Sex	3	2.286	.746	n.s.
Residual	126	3.065		

*p < .05

**p < .01

***p < .001

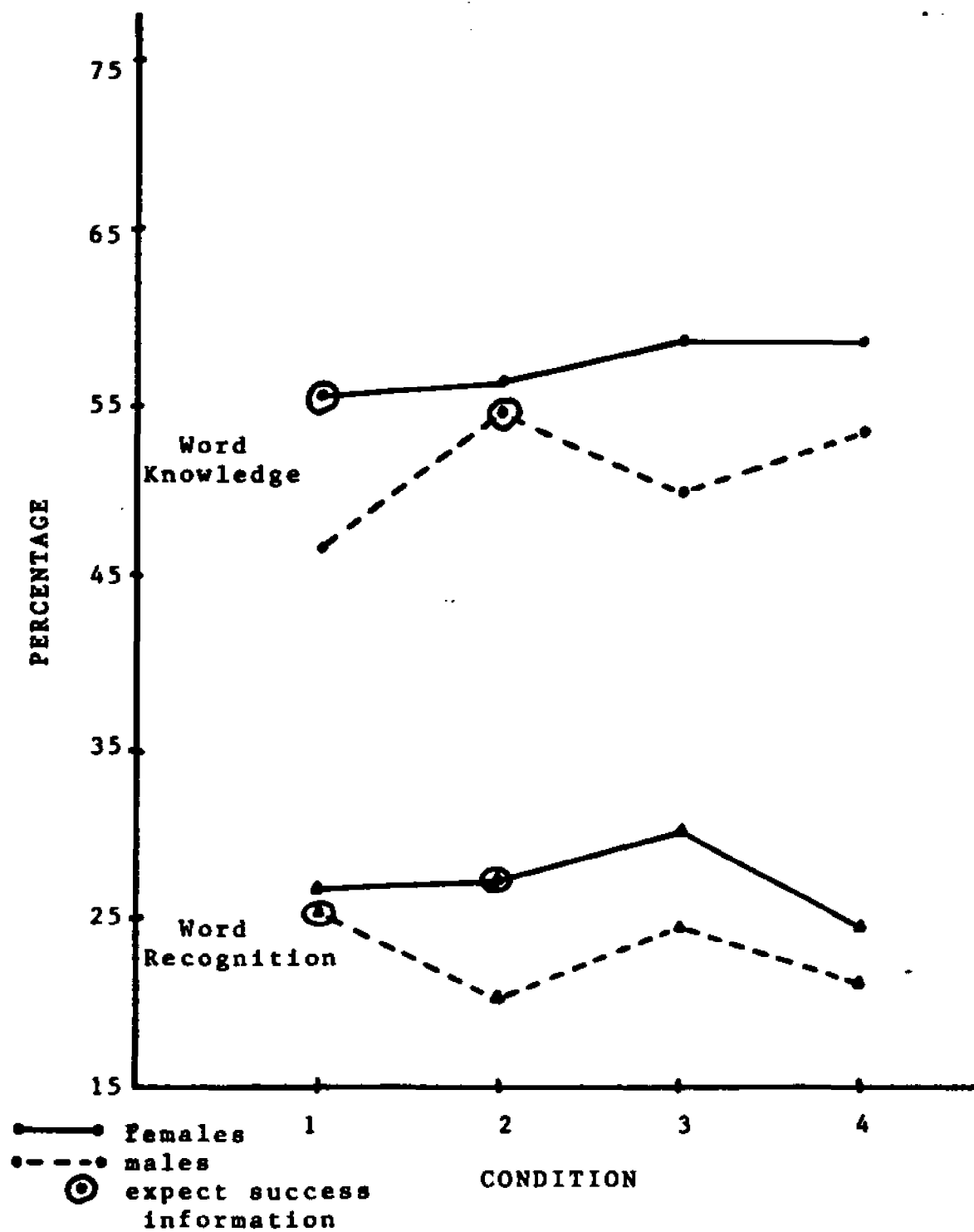


Figure 8

Percent Correct on Verbal Tasks

TABLE 32

Analysis of Covariance Summary Table:
Word Knowledge Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	153.007	4.299	.04*
Condition	3	15.567	.437	n.s.
Sex	1	93.005	2.613	.10+
Condition x Sex	3	30.238	.850	n.s.
Residual	121	35.592		

TABLE 33

Analysis of Covariance Summary Table:
Word Recognition Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	362.408	38.819	.001***
Condition	3	18.852	1.915	n.s.
Sex	1	38.321	3.893	.05*
Condition x Sex	3	4.605	.468	n.s.
Residual	126	9.843		

+p < .10
 *p < .05
 **p < .01
 ***p < .001

The pattern of means for men and women in the experimental conditions is also different from the expectancy scores but is consistent for both verbal tasks. The males who were given the information that females would perform better actually did significantly worse than anyone else. The planned comparisons among the experimental means indicated that on the Word Knowledge task the mean performance of the males in Condition 1 was significantly lower, $F(1, 126)=4.86, p < .05$, and on the Word Recognition task, the males in Condition 2 were also significantly lower, $F(1, 126)=4.86, p < .05$.

Evaluation: Inspection of the evaluation means for both tasks revealed certain regularities (Figure 9). The trend across conditions for the female scores was quite similar in both cases, and, as with the expectancy measure, the evaluation scores were higher in the control conditions. In general, men evaluated their performance somewhat higher than women. These effects did not reach significance for the Word Knowledge task (Table 34), but the analysis of covariance for the Word Recognition task (Table 35) indicated significant effects for both sex ($p < .01$) and condition ($p < .02$) as well as experience ($p < .001$).

The planned comparisons showed that for Word Knowledge, the women in Condition 1 evaluated their performance significantly lower than others in the experimental groups, $F(1, 121)=6.35, p < .05$. It is surprising that these women who had stated a moderately high expectancy of success and

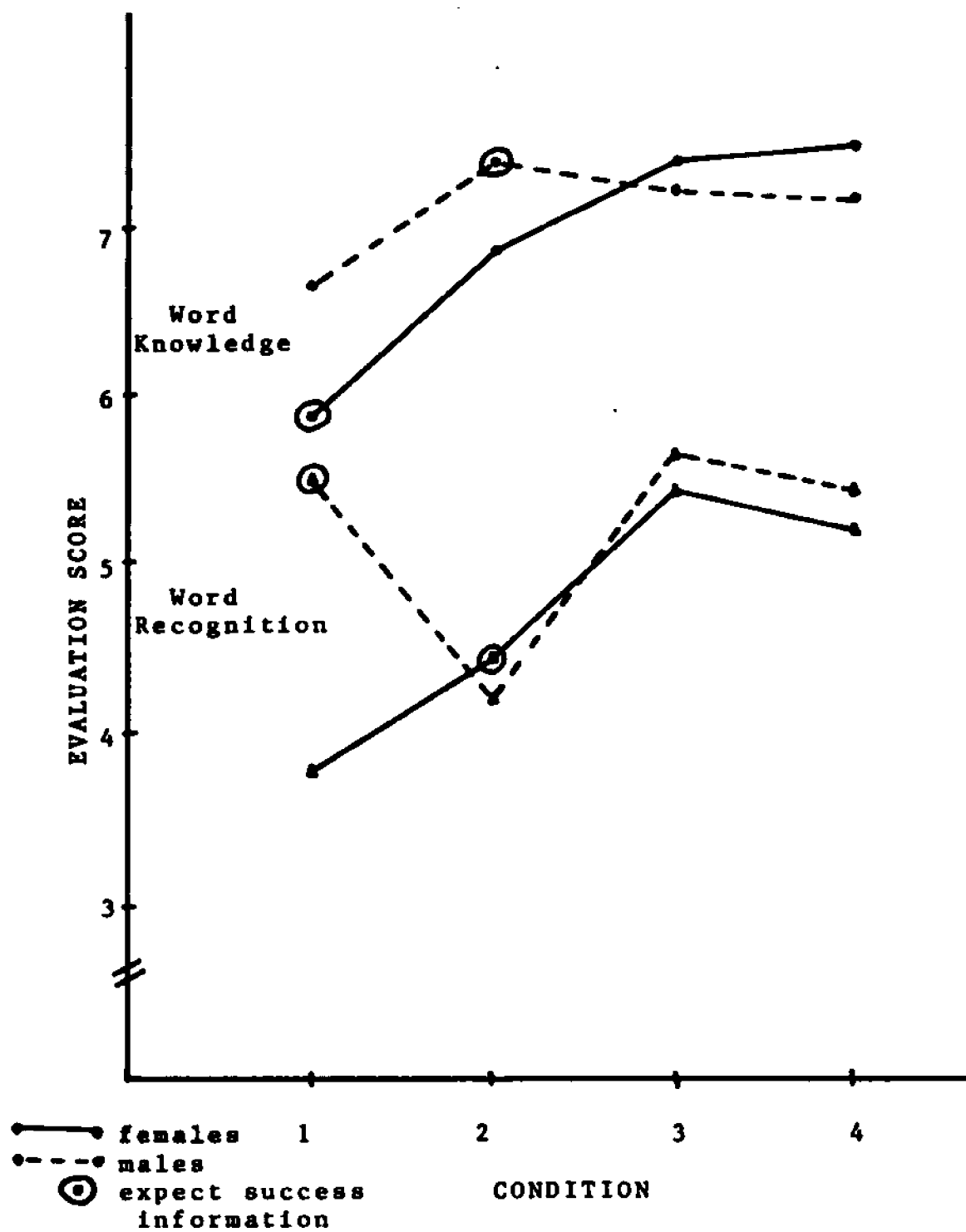


Figure 9

Mean Evaluation on Verbal Tasks

TABLE 34

Analysis of Covariance Summary Table:
Word Knowledge Evaluation

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	4.083	1.595	.21
Condition	3	5.729	2.237	.09+
Sex	1	2.471	.965	n.s.
Condition x Sex	3	2.528	.987	n.s.
Residual	121	2.560		

TABLE 35

Analysis of Covariance Summary Table:
Word Recognition Evaluation

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	99.356	31.932	.001***
Condition	3	10.383	3.337	.02*
Sex	1	21.114	6.786	.01**
Condition x Sex	3	2.183	.701	n.s.
Residual	126	3.111		

+p < .10
 *p < .05
 **p < .01
 ***p < .001

had performed relatively well, should have evaluated that performance so poorly. This discrepancy is in a direction opposite to the manipulation. For Word Recognition the planned comparisons indicated that the males in Condition 1 were significantly higher than all other experimental subjects, $F(1, 126) = 7.97$, $p < .01$. As with the expectancy measure, these were the males who were given an expectancy of success.

Correlations: Table 36 summarizes the correlations among the various dependent measures. As in Study 3, the relationship between expectancy and performance was not significant ($r = .11$; $r = .15$). Separate correlations computed for male and female subjects indicated that males were slightly, though not significantly better at predicting their performance ($r = .23$ for males on both tasks; $r = .09$ and $.15$ for females).

As in Study 3, the correlations between performance and evaluation were quite high. For all subjects, the correlation for Word Knowledge was $.51$, and for Word Recognition, it was $.45$. Females were slightly though not significantly better at evaluating their performance ($r = .60$, $.52$ for females; $r = .48$ for males).

Overall, experience was positively related for both tasks to personal expectancy, performance, and evaluation. However, the relationships were generally stronger for the Word Recognition task than for the Word Knowledge task, and stronger for performance than for the other measures.

TABLE 36

Correlations Among Expectancy, Performance, Evaluation,
and Experience Scores for Each Task

	WORD KNOWLEDGE	WORD RECOGNITION	NUMBER SERIES	SHAPE RECOGNITION
Expectancy & Performance	.11	.15	.25**	.33***
Performance & Evaluation	.51***	.45***	.55***	.70***
Evaluation & Expectancy	.49***	.55***	.68***	.55***
Experience & Expectancy	.19*	.29***	.10 ¹ .39 ² ***	.20*
Experience & Performance	.21*	.47***	.05 ¹ .21 ² *	.32***
Experience & Evaluation	.16	.43***	.06 ¹ .41 ² ***	.23**

1. Experience measure was specific number series experience.
2. Experience measure was number of math courses taken.

* $p < .05$
 ** $p < .01$
 *** $p < .001$

Summary of findings for the verbal task: For both of these verbal tasks, male subjects stated a higher expectancy of success and in general evaluated their performance higher. Female subjects, however, actually performed better. The differences in the correlations suggest that the women underestimated on the personal expectancy measure how well they would do and the men overestimated on the evaluation measure how well they actually did.

In general, it appears that the male subjects were more susceptible to the manipulation. On the personal expectancy and evaluation scales, it was the men who were given expectancies that they would do well who selected the highest scale number of the experimental groups. On performance, it was the men who were told that women would do well who did much worse.

The main effects for condition, particularly apparent in the personal expectancy and evaluation scale scores, were quite surprising. The control groups were generally higher than the experimental groups. Thus, the control group scores indicated a strong effect for the manipulation itself rather than simply its content. In other words, giving subjects, particularly female subjects, any information about the expected direction of differences depressed their prediction and evaluation of performance.

Math Tasks

Experience: There were two experience measures for the Number Series task: (1) as with the other tasks,

subjects were asked about their specific task experience, and (2) subjects were asked to indicate the highest level of math courses taken. These measures were only slightly related to each other ($r = .15$, $p < .09$), and although number series experience was not significantly related to any of the dependent variables, math experience was (see Table 35). Therefore the number of math courses taken was used as the control for experience, that is, as the covariate for the Number Series task analyses.

When asked whether they had ever done shape recognition tasks like "Little Tommy Tucker," the majority of both male and female subjects (79% and 70% respectively) chose the moderate responses "every now and then" or "I did them a few times a long time ago," but there were differences in the more extreme responses. Twenty-one percent of the women compared to only 7% of the men chose the response "I used to do them a lot, some time ago," and 14% of the men compared to only 9% of the women indicated "I have never done them." In general, then, there was a sex difference with women reporting more experience with shape recognition ($\chi^2 = 9.07$, 3 df, $p < .03$). In addition, a greater percentage of males (77%) than females (66%) said they had taken trigonometry or more advanced math courses ($\chi^2 = 14.45$, 5 df, $p < .01$). Thus, unlike the verbal task experience measures, both of these experience measures showed strong sex differences.

Personal Expectancy: The results of this measure were basically the same as for the verbal tasks: for both math tasks men stated a much higher expectancy of success (Figure 10). The analyses of covariance (Tables 37 & 38) showed that although experience was an important factor, this sex difference was significant for both tasks ($p < .001$; $p < .01$). It is also evident in Figure 10 that the females in the control conditions generally stated a higher expectancy of success than those in the experimental conditions although the overall condition effects did not reach significance in the statistical analyses.

It is interesting that, as with the verbal tasks, the manipulation of information in the experimental conditions appeared to affect the men's personal expectancies more than the women's. Given the information that males should perform better, men stated higher expectancies of success than men with opposite information, as well as higher expectancies than women with both kinds of information. In contrast, the women's scores appeared to be unaffected by the manipulation. On the Number Series the men in Condition 2, those who were given the information that men would perform better, stated significantly higher expectancies than the other groups in the experimental conditions $F(1, 125)=15.44$, $p < .001$. On the shape recognition task, those in Condition 1 were told that men would perform better, and the men in that condition also clearly stated higher expectancies of success, $F(1, 127)=3.89$, $p < .10$.

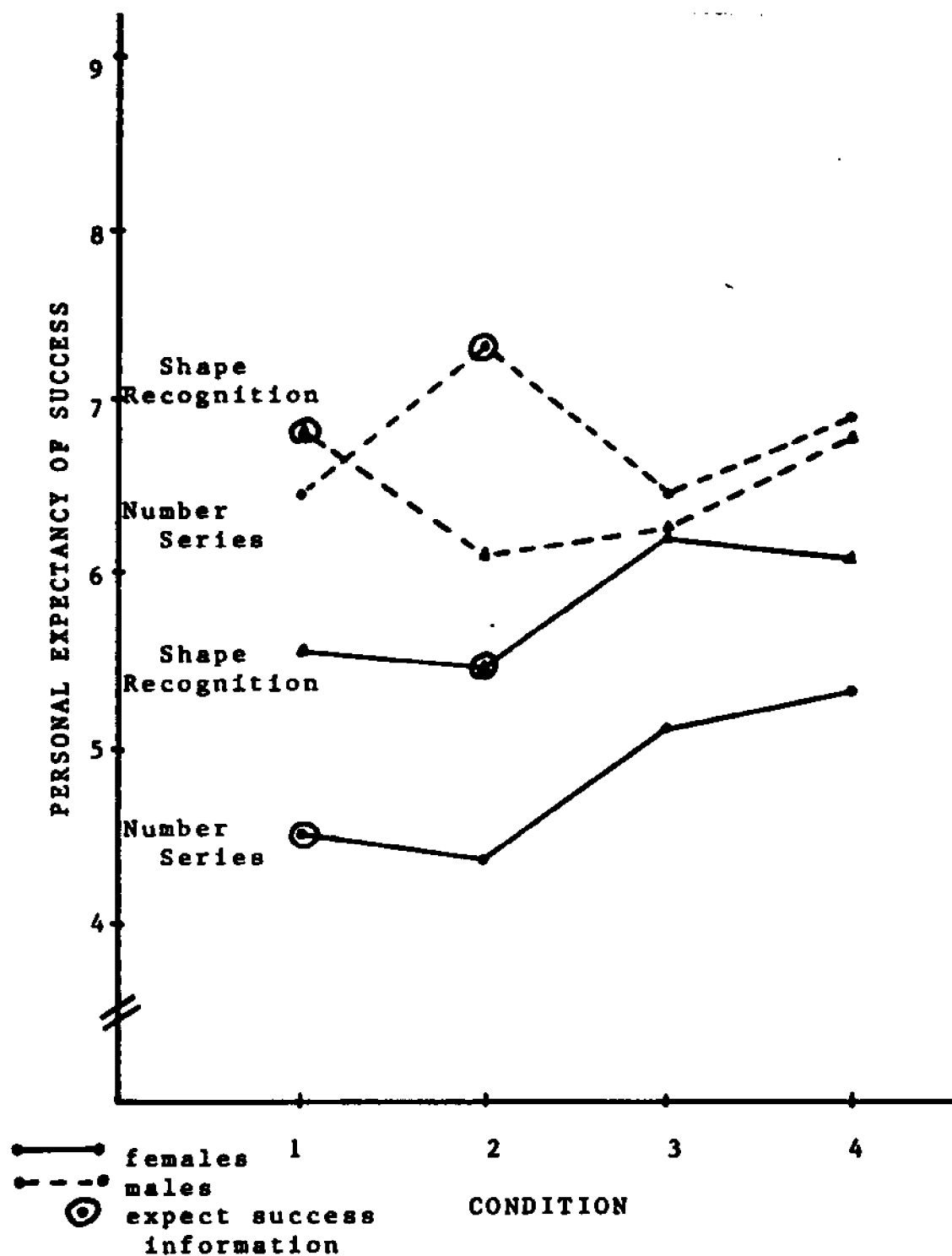


Figure 10

Mean Personal Expectancy Scores
on Math Tasks

TABLE 37

Analysis of Covariance Summary Table:
Number Series Personal Expectancy

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	112.118	25.177	.001***
Condition	3	.852	.191	n.s.
Sex	1	67.963	15.262	.001***
Condition x Sex	3	5.779	1.298	n.s.
Residual	125	4.453		

TABLE 38

Analysis of Covariance Summary Table:
Shape Recognition Personal Expectancy

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	21.603	5.596	.02*
Condition	3	3.020	.782	n.s.
Sex	1	24.703	6.399	.01**
Condition x Sex	3	1.891	.490	n.s.
Residual	127	3.861		

* $p < .05$
 ** $p < .01$
 *** $p < .001$

Performance: The performance means for these two tasks reflect opposite but consistent sex differences: on Number Series, men solved more problems in every condition; on shape recognition, women disembedded more figures in every condition (Figure 11). The analyses of covariance (Tables 39 & 40) indicate that experience was an important factor for both tasks ($p < .02$; $p < .001$). When experience was controlled, however, there was still a sex difference on the Number Series ($p < .01$). On shape recognition, when the superior experience of females was controlled, the performance sex difference did not reach significance ($p < .20$).

The overall condition effects were not significant, but females on shape recognition particularly and in Condition 4 on Number Series, seemed once again to perform better in the control conditions. Males, on the other hand, appeared to do better in experimental conditions generally.

Within the experimental groups, the patterns of the performance means were opposite to the predictions, opposite to the effects demonstrated on the verbal tasks, and opposite to the statements of personal expectancy. The most startling discrepancy was the average performance of the men in Condition 1, those who were told that females would do better: these men performed significantly better than all other experimental subjects, $F(1, 125) = 13.60$, $p < .001$, as well as better than subjects in the control conditions. The effects for shape recognition and for the females on the Number Series were weak and non-significant but appeared

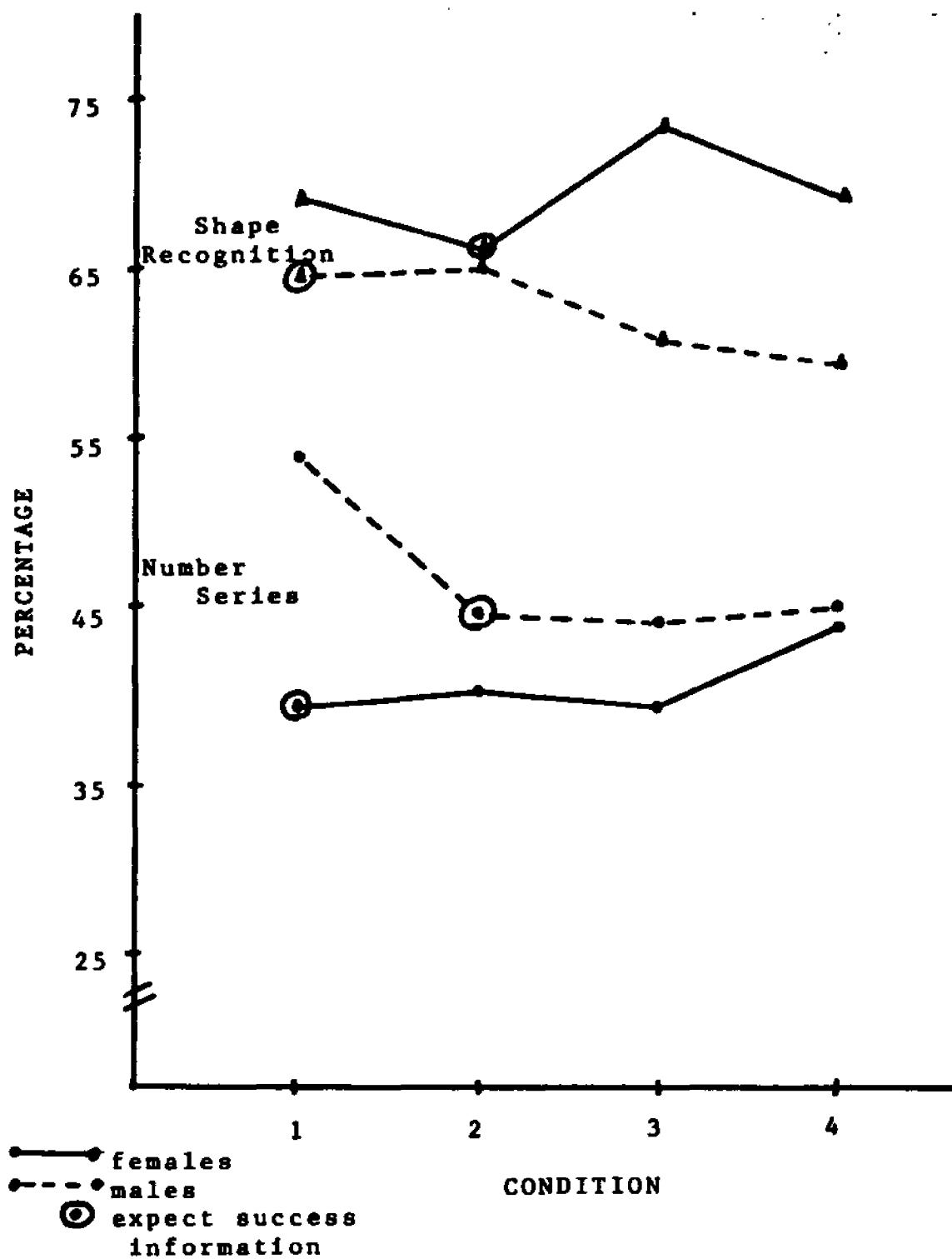


Figure 11

Percent Correct on Math tasks

TABLE 39

Analysis of Covariance Summary Table:
Number Series Performance

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	35.092	5.690	.02*
Condition	3	7.127	1.190	n.s.
Sex	1	36.604	6.109	.01**
Condition x Sex	3	4.728	.789	n.s.
Residual	125	5.992		

TABLE 40

Analysis of Covariance Summary Table:
Shape Recognition

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	164.385	13.929	.001***
Condition	3	4.806	.407	n.s.
Sex	1	21.495	1.821	.18
Condition x Sex	3	8.200	.695	n.s.
Residual	127	11.801		

* $p < .05$
 ** $p < .01$
 *** $p < .001$

to follow this pattern rather than that of the verbal tasks, that is, the groups who were given information that they would not do well appeared to do better than the groups who were given an expectancy of success.

Evaluation: The clearest effect for the evaluation scores was the sex difference on the Number Series evaluation: males stated a much higher evaluation in all conditions (Figure 12). This effect was highly significant ($p < .001$; Table 41). In general, the males' evaluations appear to follow the performance means rather than their stated personal expectancy or the manipulation.

There were no significant effects for condition or sex on the shape recognition evaluations. One can see in Figure 12, however, that the pattern of results for females on both tasks was much the same as it was on the expectancy and performance measures. That is, within the experimental conditions there was a slight trend in the direction opposite to the manipulation, and there were generally superior evaluations in the control conditions.

Correlations: Table 36 summarizes the correlations among the various dependent measures. In contrast to the relatively weak relationships on the verbal tasks, the correlations between personal expectancy and performance were moderately strong for both math tasks ($r = .25$, $p < .01$ on the Number Series; $r = .33$, $p < .001$ on shape recognition).

Overall, as in Study 3 and on the verbal tasks, these subjects demonstrated realistic assessments of their own

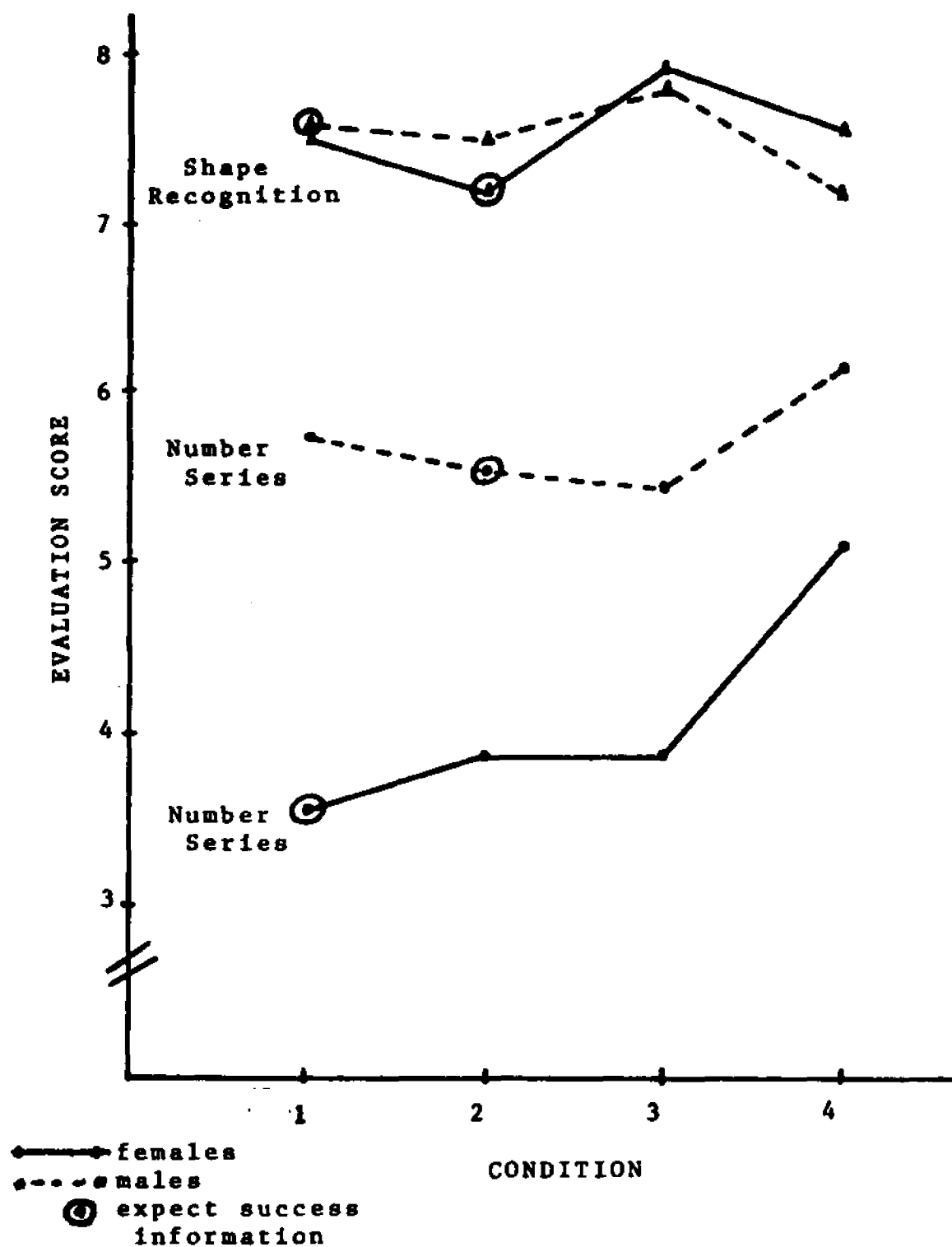


Figure 12

Mean Evaluation on Math Tasks

TABLE 41

Analysis of Covariance Summary Table:
Number Series Evaluation

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	127.487	29.544	.001***
Condition	3	4.999	1.158	n.s.
Sex	1	52.536	12.175	.001***
Condition x Sex	3	1.227	.284	n.s.
Residual	125	4.315		

TABLE 42

Analysis of Covariance Summary Table:
Shape Recognition Evaluation

SOURCE	DF	MEAN SQUARE	F	SIGNIFICANCE OF F
Experience	1	28.493	7.222	.008**
Condition	3	.942	.239	n.s.
Sex	1	2.807	.712	n.s.
Condition x Sex	3	1.330	.337	n.s.
Residual	127	3.945		

**p < .01
***p < .001

performance. Evaluation and performance were strongly correlated for both tasks ($\underline{r}=.55$, $p < .001$ for Number Series; $\underline{r}=.70$, $p < .001$ for shape recognition).

As mentioned previously, the two experience measures for the Number Series task, number of math courses taken and specific task experience, showed a different degree of relationship to the dependent variables. Number of math courses taken was rather strongly related to personal expectancy ($\underline{r}=.39$, $p < .001$) and evaluation ($\underline{r}=.41$, $p < .001$) and moderately related to performance ($\underline{r}=.21$, $p < .05$). Specific task experience, on the other hand, was not significantly related to either expectancy ($\underline{r}=.10$), performance ($\underline{r}=.05$), or evaluation ($\underline{r}=.06$). The specific task experience measure for shape recognition was moderately related to personal expectancy ($\underline{r}=.20$, $p < .05$), evaluation ($\underline{r}=.23$, $p < .01$), and performance ($\underline{r}=.32$, $p < .001$).

Summary of findings for the math tasks: As on the verbal tasks men stated a higher expectancy of success than women. They also evaluated their performance on the Number Series higher, but, unlike the verbal tasks, this evaluation was realistic since they actually had performed better.

In terms of the impact of the manipulation, again men seemed more susceptible. In the experimental conditions, their scores were much more variable than the women's. The personal expectancy scores were like those on the verbal tasks, that is, the men who were told they would do well, stated very high predictions. The effect of the manipulation

was quite different for performance, however, particularly for the Number Series. It was the men who were told women would do better who performed much better.

The condition effects never reached significance on these tasks. However, females seemed to be following the trend noted in the verbal tasks: they had lower scores on all measures in the experimental conditions than in the control conditions. Again, there seemed to be an effect for any specific information regardless of its content.

DISCUSSION

An examination of the results of both the verbal and math tasks in this study indicates certain consistent findings. Men always state a higher personal expectancy and evaluation than women. In addition, control subjects generally predict more success and evaluate their performance higher than experimental subjects. This trend is particularly true for females. In terms of actual performance, women subjects score higher on the verbal tasks and men on the Number Series task. There are no sex differences on the shape recognition task, "Little Tommy Tucker."

This superior confidence of men has been demonstrated numerous times by other researchers, but there still exists no satisfactory theoretical explanation for the phenomenon (see especially, Crandall, 1962). It is interesting to note, however, that in this study these statements of personal expectancy bear virtually no relationship to actual

performance. The main effects for performance reported here are always in the direction of stereotypes.

In terms of the hypotheses of this study, overall the data support the first hypothesis, that instructional information about the relationship of sex to ability affects estimates of personal expectancy, task performance, and evaluation of performance in a direction consonant with that information and the subject's own sex. There are two important qualifications to this conclusion, however.

First, the instructional expectancies affect male and female subjects differently. Males seem much more sensitive to these expectancies with more of the predicted variability in their scores in the experimental conditions. Their personal expectancy estimates and actual task performance differ according to the content of the information given in the instructions. Females, on the other hand, do not show much variability in the experimental conditions. In general, their scores seem depressed and unresponsive to the manipulation. In fact, these experimental scores are consistently lower than those of females in the control condition.

Second, there were noticeable task effects. Verbal performance scores of males in the experimental conditions reflect the direction of the manipulated expectancies and are consonant with them. Number Series performance scores, on the other hand, are opposite to the information contained in the manipulation.

The second hypothesis, that instructional information that supports sex-role stereotypes is more effective in influencing behavior, is not supported because of these task effects. For the verbal tasks the most influenced subjects in terms of performance were the men who read the stereotypic information that females would do better: these men performed worse than the other experimental groups. They appear to have given up. This finding is consistent with the hypothesis. For the Number Series task, however, the most influenced subjects were the men who read the counter-stereotypic information that females would do better: these men outperformed all other groups by a considerable margin. They appear to have been motivated by this counter-stereotypic information. This finding is exactly opposite to that hypothesized.

These results may be interpreted by a closer inspection of the existing stereotypes. The responses of the control subjects who stated their beliefs about whether men or women would perform better on each task suggest two conclusions: (1) women subjects hold stronger stereotypes than men; and (2) the firmest belief of all subjects is a male superiority on math tasks like the Number Series.

If the women subjects believe more strongly in the stereotypes, then it makes sense that they are less influenced by the particular information contained in the manipulation. One might conclude that they are so clear in their beliefs that a little information pro or con cannot

affect them. However, the mere attempt to influence them did depress their scores, particularly their personal expectancy of success. Although this effect is also evident for the men on the verbal tasks, the pattern is more consistent for women across all tasks and all measures.

The strength of the particular stereotype is probably not the only factor accounting for the behavior of the men who were given the counter-stereotypic information on the Number Series task. Their superior performance may be interpretable as reactance, that is, when a person's freedom is eliminated or threatened, (s)he will act to exercise that freedom (Brehm, 1972). In this case, the specific freedom that was threatened is not clear. It cannot be simply the freedom to perform maximally since the reactance effect was not evident on the verbal tasks. Possibly it was the freedom to accept a positive stereotype or to reject a negative one although again such an interpretation should apply to the verbal tasks. Clearly, some aspect of the task differences affected the responses of the male subjects.

This effect is very similar to the results of the handedness expectancy given in Study 3. When the left-handed people in that study read that right-handed were more ambidextrous, they also outperformed all other groups. A possible uncontrolled factor in both cases is the value of the ability to the particular group. Left-handed people might well value their ambidexterity, one of the few positive outcomes of living in a right-handed culture. Although not for

the same reasons, men may be willing to concede verbal superiority, at least on these kinds of simple tasks, to females, but react strongly against the implication that females are superior on something as highly-valued as basic math ability. In fact, Brehm predicted that the importance or value of a threatened freedom will increase reactance.

In summary, these findings suggest caution in attempting to influence a stereotype or stereotypic behavior. A direct presentation of related information can lead to lessened personal expectancy of success, inferior task performance, and a subsequently lowered evaluation for subjects who firmly believe in the stereotype. The only exception is the reactance phenomenon observed in the performance only of subjects who held a personally-relevant strong belief about superior ability on a highly-valued task. For most subjects an emphasis on sex-dependent success for a number of tasks had negative effects.

Within these overall lower scores, however, it does seem that stereotypic behavior can be influenced by the direct presentation of information. It seems far easier, however, to depress scores than to elevate them. These studies indicate that it is easier to convey, even implicitly, an expectancy of failure than an expectancy of success.

IV. CONCLUSIONS AND IMPLICATIONS

A conceptualization of discrimination in the everyday sense is attention to ascriptive characteristics despite the availability of more relevant information. The implications of non-discrimination are that for most things, information other than sex should be considered. This is the thrust of the Federal Government's insistence on BFOQ's (bona fide occupational qualifications) as hiring guidelines: people should be judged on the basis of their job-relevant qualifications and not on the basis of sex, race, country of national origin, previous condition of servitude, or other such factors.

Allport (1954) argued from a kind of cognitive complexity viewpoint that forming and using stereotypes was a natural, normal phenomenon that made interaction easier. He suggested that categorising the complex stimuli of the social world is a necessary simplifying strategy. In recent times, however, we have witnessed a conscious rejection of race as a meaningful category. As blacks are allowed into different arenas of activities, as depictions of black home-life are portrayed on television, additional information, much of it conflicting with old stereotyped notions, is made available. The stereotype cannot withstand the strain of such information. As people become aware that blacks and whites share child-rearing problems and economic and job-related dilemmas, the similarities are emerging to the

detriment of the traditional emphasis on differences.

The research reported here is relevant to a consideration of whether male-female stereotypes could change through such a process of recategorization. This process does not require ignorance of the categories male and female; it requires a subtraction of meaning from those categories so that attention can be directed to personal characteristics relevant to particular goals.

THE IMPORTANCE OF SEX

The subway study demonstrated that sex is such an overwhelmingly important category that it is coded and remembered for no apparent reason. If respondents remembered anything at all it was the sex of the token-seller. One respondent who in fact could not describe the seller at all said he didn't even notice whether it was a man or a woman, implying that everyone "knows" what an important characteristic sex is. The laboratory game indicated that everyone also "knows" that the most efficient way to describe people is to name the most statistically distinctive features first. On the face of it these are clearly contradictory pieces of knowledge.

Yet anyone who has ever tried to read fiction where the sex of the characters is ignored (e.g., Arnold, 1973) or ambiguous (LeGuin, 1969) also knows how extremely difficult it is to form a mental picture of the characters. Descriptions may be provided that are quite detailed as to height,

weight, complexion, musculature, hair color, and so forth, and still it is difficult to conjure up an image of the person. Sex appears to be an important cognitive category that serves to organize subsequent information; perhaps a stereotypic image is formed at first and details merely modify it in specific ways.

The stimulus discrimination studies as well as the first expectancy study (Study 3) indicate that sex is treated differently from other characteristics. Not only does sex perhaps serve this organizing function, but people are then more prepared to accept more information as relevant to that category. Subjects rejected the counter-stereotypic information about handedness and totally ignored the information about eye-color. In contrast, the information about abilities related to sex, even though it was counter-stereotypic, was accepted as plausible and affected task performance.

THE PERSONAL COST OF DISCRIMINATING/STEREOTYPING

Much of the theorizing and research on discrimination and stereotyping has focused on the cost to the person who is the object of the discrimination. Although this is certainly a valuable approach, the research reported here describes some of the costs to the person who does the stereotyping. For example, in the laboratory game (Studies 2 & 2a) the persistence of treating sex as an important descriptive category interfered with the subjects' ability to

maximize outcomes. Although the negative consequences of this strategy were not great, participants did seem sufficiently involved in the game to desire to do their best and seemed disappointed that they were not more successful in describing pictures in fewer clues.

In the second expectancy study (Study 4) an emphasis on sex-dependent abilities generally depressed subjects' personal expectancy and evaluation and interfered with their task performance compared to subjects for whom sex was not salient, the baseline control. This effect was even stronger for subjects who were most sex-stereotyped in their beliefs, the women in this case. The only result that resembled a positive outcome from the subjects' point of view was the energizing of performance, the reactance, of the males on Number Series performance and the left-handed on maze performance. With these exceptions, making salient a relationship between sex and ability had negative consequences for most subjects.

POSSIBILITIES FOR CHANGE

These studies present some evidence for possibilities for change. In the laboratory game, subjects did improve in their ability to attend to characteristics other than sex. This improvement depended on standard principles of learning: knowledge of outcomes, practice, and immediate feedback. It is not a trivial conclusion that the cognitive importance of sex as a category, which one might be tempted

to accept as a given, whether innate or learned/socialized, is amenable to modification through such standard procedures. Such a conclusion suggests approaches for re-training or sensitizing key people such as personnel managers using a more directly relevant learning paradigm.

The expectancy studies suggest that a de-emphasis on sex and ability connections would be more effective in approaching equality than a reversal of the accepted stereotypes. To whatever small extent new information supportive of counter-stereotypic behavior is effective, the implicit negative expectancy, the implication that one group will not be successful, is more effective in influencing behavior in a negative direction.

GENERAL CONCLUSIONS

Kuhn (1962) has proposed that the nature of scientific revolutions is that an old paradigm disintegrates or is abandoned as its usefulness for explaining phenomena breaks down. A new paradigm emerges and science can then proceed with its "normal" problem-solving activities. The study of sex differences seems to be in a transition state between paradigms. The traditional search for subject sex differences has yielded meagre results. Yet many researchers are eager to say something relevant about the social changes occurring around men's and women's roles. Competing paradigms are beginning to emerge. Bem's and Spence's androgyny scales represent a strong challenge to the old masculinity-

femininity formulation. The number of studies on stereotyping or sex as a stimulus variable are multiplying. Interest in the unique experiences of women, like menstruation, pregnancy, and childbirth, is being expressed in major research efforts. Variables like power which are thought to be implicated in the relations between the sexes are being revived and reinvigorated.

Similarly, the research reported here attempts to proceed with the normal problem-solving activities of scientific inquiry. There is not yet a sophisticated or detailed new paradigm; however, the research does take as its starting place the social construction of sex. The perspective itself is not entirely novel, but the emphasis is. As is clear from the literature review, the ideas have been around for a while, particularly in sociology, but not integrated theoretically or practically with the theory and methodology of experimental psychology. The effort here is to raise fundamental questions about the cognitive importance of sex as a social category in such a way that relates it to other variables and other questions of general interest. Hopefully, the perspective and the experimental investigations which inform it will raise new questions and inspire new approaches.

APPENDIX A

R # _____

EYE-WITNESS ACCOUNTS

Excuse me. Did you just buy a token? I'm doing a study on eye-witness reports. I would just like you to describe the token-seller for me. Would you be willing to do this? It will only take a couple of minutes. (NOTE RESPONSE RATE ON 1 SHEET)

1. First, could you just list things about that person that you think might help in identification? (PLEASE NUMBER)

2. (IF OMITTED)

a. Was the seller a man or a woman? M _____ W _____

b. How old would you say (s)he is? _____

c. Did you happen to notice: the color of his/her hair? _____

the length of his/her hair? _____

whether or not (s)he wore glasses? _____

d. Could you estimate his/her height? _____ (IF NO) taller or shorter than me? _____

weight? _____ (IF NO) heavier than me? _____

3. Is there anything else you can remember that would help to identify this person?

Thank you very much. You've been very helpful.

R characteristics:

M _____ W _____

Age _____

Color of hair _____

Length of hair _____

Glasses _____

Ht. _____

Wt. _____

Interviewer _____

S.P.# _____

I sheet # _____

APPENDIX B

STUDY 2
INSTRUCTIONS TO SUBJECTS

This is a very simple exercise--like a game, really. You each have the same 10 pictures in front of you. The idea is for one of you to give clues to the other person and for that person to guess which picture you are describing.

First we will decide who is to go first. Please pick a ball from this can and whoever has the lower number will be the first clue-giver. (DO THAT.)

Okay. You will give the clues. To make it just a little harder, you will only be able to use 5 of these 10 clues. (GIVE CLUE CARDS: PUT BACK BALLS.)

(TO THE OTHER PERSON:) You will write down your guesses on this sheet. After each guess, indicate how confident you are that that is the right picture on the scale: #1 is very confident and #11 is not at all confident. Then tell you partner you are ready for the next clue.

Please do not talk during the "game." For example, do not say when you are certain you know the correct answer. Just mark it on your sheet. At the end of 5 clues, we will see what the correct picture is.

The object is to guess the picture on as few clues as possible.

Are there any questions?

You will each have 2 turns as a clue-giver and 2 turns as a receiver--5 pictures at a time.

CHARACTERISTICS OF STIMULUS PICTURES

#1	#2	#3	#4
light clothes glasses male no head cover dark hair	glasses male dark clothes no head cover dark hair	light clothes glasses female no head cover dark hair	male no glasses dark clothes no head cover dark hair
#5	#6	#7	#8
glasses female dark clothes no head cover dark hair	head cover light clothes male no glasses dark hair	female no glasses dark clothes no head cover dark hair	head cover female no glasses dark clothes dark hair
#9	#10		
male no glasses dark clothes no head cover dark hair	blonde female no glasses dark clothes no head cover		

APPENDIX C

Characteristics of Subjects

Study 3

	Control Group	Experimental Group	% Total
Males	29	33	72.1
Females	13	11	27.9
Right-handed	34	34	79.1
Left-handed	7	6	15.1
Both	1	2	3.5
No Answer		2	2.3
Light-eyed	12	16	32.6
Dark-eyed	28	27	64.0
Can't decide	2	1	3.5
Total N=	42	44	

NEW YORK TEST OF CHARACTERISTIC ABILITIES

Psychologists have long been interested in the effects of personal characteristics on success in certain kinds of tasks. For example, it is well known that women do better in some academic subjects than men, and that men do better in others. Recent work on brain "lateralization" shows that people with left-hemisphere dominance do better on some tasks and those with right dominance do better on other tasks. Similarly, perception experiments have shown that people with light and dark eye pigmentation (color) actually see things differently.

In all of these cases, it is not that one group is overall better than the other but that each can do certain kinds of things better. There is no reason why academic subjects can't be taught so that both groups are performing their very best, rather than having certain methods give an advantage to one group or the other.

In order to reach this goal, however, more basic information on group differences needs to be gathered. That is the purpose of this test. Please try to do your very best on each of the four tasks in this booklet.

Each problem or set of problems has been timed. Please work as rapidly as possible on each page and do not turn to the next page until the signal is given unless it says GO ON at the bottom of the page. And please do not turn back once you have begun the next test.

WAIT TO BEGIN

Part I: Hand Dominance

(Experimental Group)

Hand dominance is one way to measure brain lateralization. People tend to think that those who are left-handed are more able to use both hands well, that is, are more ambidextrous. However, scientific studies have shown that this is not so. In fact, it is usually the right-handed person who is able to transfer control to the other hand when necessary. Scientists now think that it may be because right-handed people are more lateralized, and thus more truly ambidextrous.

In this section of the test you will be given a maze to find your way through. Please use a PEN and hold it in your NON-DOMINANT HAND: that is, if you usually write with your right hand, use your left hand; if you usually write with your left hand, use your right hand; if you can write with either hand, use the one that you don't prefer.

Before beginning, please answer these few questions:

1. In general, are you

right-handed _____ left-handed _____ both _____?

2. To what degree would you say that you are right-handed or left-handed?
(circle one number)

1	2	3	4	5	6	7
completely left-handed			about equal			completely right-handed

3. How well do you expect to do on this maze?

1	2	3	4	5	6	7
very poorly			about average			very well

WAIT TO BEGIN

Part II: Analytic Verbal Ability

(Experimental Group)

Sex/gender often seems to have an effect on intellectual tasks like verbal ability. For example, females usually do well on verbal kinds of material. This is not always true, however. If the material involves a certain kind of analytic ability, then males do better. The following two tests have been selected to test that kind of analytic ability. Males should perform better; the question is how much better.

Before beginning this series, please answer the following questions:

1. Are you

female _____

male _____?

PLEASE ANSWER BOTH QUESTIONS #2 & #3:

2. To what degree would you say that you are masculine?

1	2	3	4	5	6	7
not at all			about			completely
masculine			average			masculine

3. To what degree would you say that you are feminine?

1	2	3	4	5	6	7
not at all			about			completely
feminine			average			feminine

4. How well do you expect to do on this test of verbal ability?

1	2	3	4	5	6	7
very poorly			about			very well
			average			

WAIT TO BEGIN

(Experimental Group)

Part III: Basic Mathematical Ability

Sex/gender also seems to be related to how well a person does on intellectual tasks like mathematics. Many assume that males are better at mathematics because men are more likely to take advanced math courses. However, if one compares males and females who have taken the same number of math courses and thus have the same experience with mathematics, then females as a group always do better.

The following items have been selected to test basic ability (regardless of the number of math courses taken). Again, please answer the following question before beginning:

1. How well do you expect to do on this test of basic mathematical ability?



WAIT TO BEGIN

(Experimental Group)

Part IV: Perception of Detail

Eye pigmentation (color) is known to affect perception in the sense of how sensitive one is to light and detail. It has long been known that light-eyed people, those with blue, light gray, or light hazel eyes, are more sensitive to light than those with dark eyes. There is now good reason to believe that light-eyed people are more sensitive to detail as well. The following tests involve seeing fine detail in pictures. Before turning the page, please answer these questions:

1. Do you have

light eyes _____ dark eyes _____ can't decide _____?

2. How light or dark would you say the color of your eyes is?

1	2	3	4	5	6	7
very			can't			very
light			decide			dark

3. How well do you expect to do on this task?

1	2	3	4	5	6	7
very			about			very
poorly			average			well

WAIT TO BEGIN

Part I: Hand Dominance

(Control Group)

Hand dominance is one way to measure brain lateralization. In this section of the test you will be given a maze to find your way through. Please use a PEN and hold it in your NON-DOMINANT HAND: that is, if you usually write with your right hand, use your left hand; if you usually write with your left hand, use your right hand; if you can write with either hand, use the one that you don't prefer.

Before beginning, please answer these few questions:

1. In general, are you

right-handed _____ left-handed _____ both _____?

2. To what degree would you say that you are right-handed or left-handed? (circle one number)

1 2 3 4 5 6 7
completely about completely
left-handed equal right-handed

3. How well do you expect to do on this maze?

1 2 3 4 5 6 7
very poorly about very well
average

WAIT TO BEGIN

Part II: Analytic Verbal Ability

(Control Group)

Sex/gender often seems to have an effect on intellectual tasks like verbal ability. Before beginning this series, please answer the following questions:

1. Are you

female_____

male_____?

PLEASE ANSWER BOTH QUESTIONS #2 & #3:

2. To what degree would you say that you are masculine?

1	2	3	4	5	6	7
not at all			about			completely
masculine			average			masculine

3. To what degree would you say that you are feminine?

1	2	3	4	5	6	7
not at all			about			completely
feminine			average			feminine

4. How well do you expect to do on this test of verbal ability?

1	2	3	4	5	6	7
very poorly			about			very well
			average			

WAIT TO BEGIN

Part III: Basic Mathematical Ability

(Control Group)

Sex/gender also seems to be related to how well a person does on intellectual tasks like mathematics. The following items have been selected to test basic ability(regardless of the number of math courses taken). Again, please answer the following question before beginning:

1. How well do you expect to do on this test of basic mathematical ability?



WAIT TO BEGIN

Part IV: Perception of Detail

(Control Group)

Eye pigmentation (color) is known to affect perception in the sense of how sensitive one is to light and detail. The following tests involve seeing fine detail in pictures. Before turning the page, please answer these questions:

1. Do you have

light eyes_____ dark eyes_____ can't decide_____?

2. How light or dark would you say the color of your eyes is?

1 2 3 4 5 6 7
very can't very
light decide dark

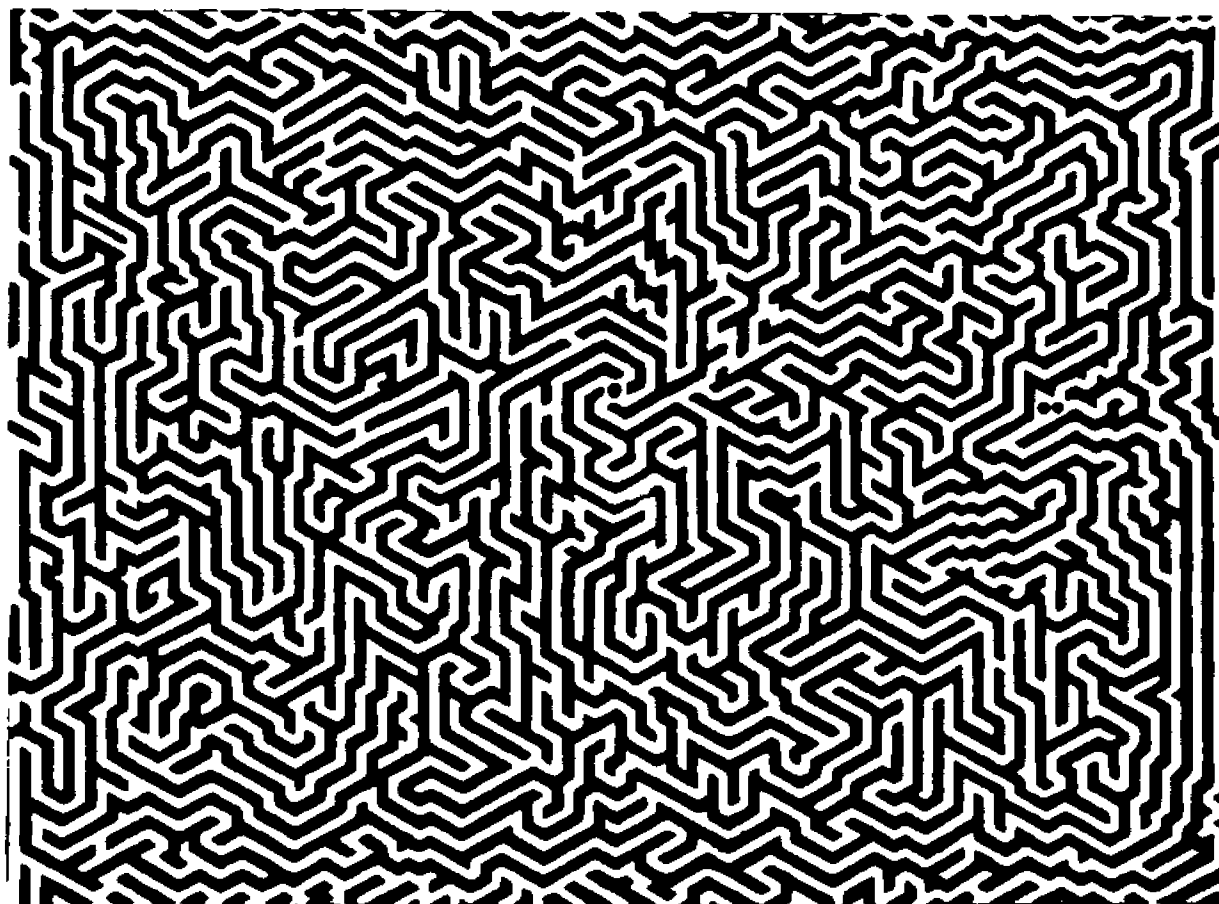
3. How well do you expect to do on this task?

1 2 3 4 5 6 7
very poorly about very well
average

WAIT TO BEGIN

Maze

Remember to use your PEN and your NON-DOMINANT HAND. You will be scored on how far you get in any direction. But you will lose points whenever you cross a line. Therefore, work RAPIDLY but CAREFULLY. Start in the middle of the maze at the single dot and travel to the double dots at the right. You may start over as many times as you wish, and you do not have to retrace lines.

**GO ON**

Word Power

Make as many words as you can using the letters in the word **INFORMATION**. The words may be of any length, but must be recognized English words.

INFORMATION

- | | |
|-----------|-------|
| 1. _____ | _____ |
| 2. _____ | _____ |
| 3. _____ | _____ |
| 4. _____ | _____ |
| 5. _____ | _____ |
| 6. _____ | _____ |
| 7. _____ | _____ |
| 8. _____ | _____ |
| 9. _____ | _____ |
| 10. _____ | _____ |
| 11. _____ | _____ |
| 12. _____ | _____ |
| 13. _____ | _____ |
| 14. _____ | _____ |
| 15. _____ | _____ |
| 16. _____ | _____ |
| 17. _____ | _____ |
| 18. _____ | _____ |
| 19. _____ | _____ |
| 20. _____ | _____ |
| 21. _____ | _____ |

WAIT TO GO ON



Hidden Pictures

Little Tommy Tucker

Can you find these objects in the large picture above?

cup	book	spool	spoon	spider	padlock	letter E	boy's cap	2 cupcakes	crochet hook
key	fish	heart	pencil	rowboat	feather	umbrella	toothbrush	toothpaste	fireman's hat

WAIT TO GO ON

Number Series

The numbers in each line below are in a certain order. In each line, find the number that should come next. You may do them in any order.

1. 5 8 11 14 ____
2. 3 12 48 ____
3. 2 3 5 9 ____
4. 1 7 3 9 ____
5. 1 4 9 ____
6. 51 43 35 27 ____
7. 2 6 18 54 ____
8. 1 2 3 2 4 6 3 6 ____
9. 1 1 4 2 9 3 ____
10. 4 7 8 2 5 6 0 ____
11. 13 13 17 11 25 7 37 1 ____
12. 23 17 12 8 5 3 ____
13. 9 27 13 39 17 51 ____
14. 16 20 14 20 24 18 24 ____
15. 3 $\frac{1}{3}$ 9 4 $\frac{1}{4}$ 16 5 ____
16. 3 16 5 36 7 64 9 ____
17. 11 12 15 20 27 32 ____
18. 1 1 2 3 5 8 ____
19. 2 3 5 7 11 ____
20. 1 10 11 100 ____

BREAKFAST FOODS

Look in the diagram of letters for the words given in the word list. Find them by reading FORWARDS, BACKWARDS, UP, DOWN, and DIAGONALLY. The words are always in a straight line and are never formed by skipping over any letters. Each word is spelled in the diagram exactly as it is in the word list. Each time you find a word, circle it in the diagram.

BACON
BISCUITS
BUTTER
EUN
BUTTERED ROLL
CINNAMON TOAST
COCOA
COFFEE
CORNFLAKES
CORN MUFFIN
CREAM
DATE NUT BREAD
EGGS
ENGLISH MUFFIN
FARINA
FRENCH TOAST
FRUIT
GRAPEFRUIT
HAM
HASH BROWNS
HOMINY GRITS
JAM
MARMALADE
MILK
OATMEAL
ORANGE JUICE
PANCAKES
PRUNE DANISH
RAISIN BREAD
SAUSAGE
SCONE
SUGAR
TEA
TOAST
WAFFLES

C I N N A M O N T O A S T O K A
O O S E K A C N A P E N A M L K
R L R D C O O M O L S T H L I P
N R A A E O A C F C M K O H M L
P A I L N H R F O E A R M A A D
L G S A G G A N A C D B I S E A
A U I M L W E L M E F C N H R E
K S N R I B B J R U J O Y B C R
E T B A S O U E U H F F G R S B
S I R M H F T N G I M F R O F T
G U E D N T T N P A C E I W R U
G C A E U G E O J I S E T N U N
E S D E F A R I N A R U S S I E
T I U R F E P A R G T O A S T T
E B H S I N A D E N U R P S E A
A F R E N C H T O A S T T G U D

GO ON

APPENDIX D

NEW YORK TEST OF SEX DIFFERENCES

Psychologists have long been interested in ability differences between the sexes. For example, it is well known that women do better in some academic subjects than men, and that men do better in others. It cannot be concluded, however, that one group is overall more intelligent than the other; rather each can do certain kinds of things better.

One recent theory is that there are not fundamental ability differences between males and females, but there are differences in style, that is, men and women have characteristic but different ways of approaching problems. In terms of academic success, then, perhaps it is not the content of the courses but the style in which they are taught that causes one group to do better than the other. There is no reason why these courses can't be taught so that both groups are performing their very best, rather than having certain methods give an advantage to one group or the other.

In order to reach this goal, however, more basic information on sex differences needs to be gathered. That is the primary purpose of this test. We have selected basic verbal and mathematical problems that reflect different styles. We are also interested in your judgment as to why males or females are better on certain tasks. Please answer all questions that come before and after each of the four subtests.

Each subtest has been timed for four minutes. It is very important that everyone spend the same amount of time on each test. Therefore, please wait for the signal before beginning each subtest. Please work as rapidly as possible and try to do your very best on each of the problems or set of problems. Remember, do not turn to the next page until the signal is given unless it says GO ON at the bottom of the page. And please do not turn back once you have begun the next test.

WAIT TO BEGIN

II. WORD RECOGNITION

This test, which involves recognizing words in a complicated design, reflects the kind of verbal ability that males usually have. Again, there are many reasons suggested as to why males would be better at this kind of test. How important do you think the following reasons are in explaining this sex difference? Please rank them 1 to 4 (1=most important reason; 2=next most important reason; 3=next most important; 4=least important):

- _____ Boys learn to understand and follow instructions when working on mechanical projects like building models.
- _____ Little boys are more interested in adventure stories with complicated plots.
- _____ Males are more logical in their thinking in general.
- _____ This ability has to do with biological factors like brain structures or hormones that males have.
- _____ Other (please explain): _____

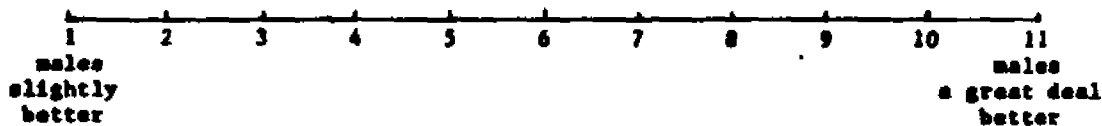
SAMPLE TEST

Look in the diagram of letters for the words given in the word list and circle them. Find them by reading **FORWARDS**, **BACKWARDS**, **UP**, **DOWN**, and **DIAGONALLY**.

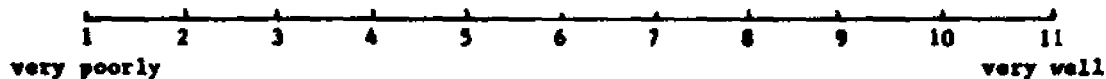
paper ✓
pen
test ✓
art
now

G B R I N
K A C E J
M H P D L
O T S E T
W Y F G R

How much better than females do you think males at Brooklyn College will do on this test? (Circle one number)



How well do you expect to do on this test of word recognition?



WAIT TO BEGIN

MATH ABILITY

I. ARITHMETIC CONCEPTS

On this kind of test involving basic arithmetic operations, females usually do better than males. Again, please rank the following reasons as to their importance (1-most important reason, etc.):

____ Girls are motivated to learn these skills because they wish to please their teachers.

____ Females are more patient.

____ Little girls learn practical mathematical concepts by watching their mothers shop, cut food portions, etc.

____ This ability has to do with biological differences between the sexes.

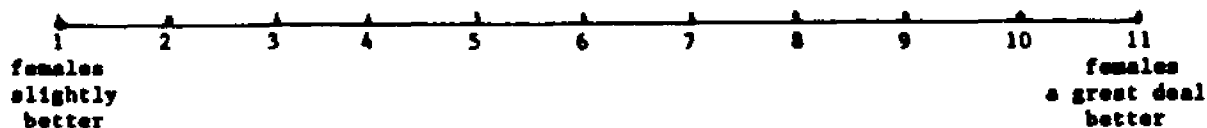
____ Other (please explain): _____

SAMPLE TEST

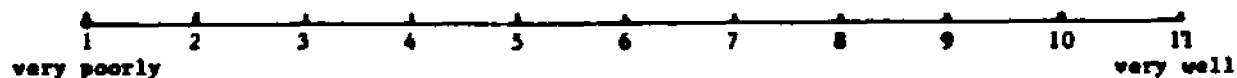
The numbers in each line below are in a certain order. In each line, circle the number that should come next.

1. 2 4 6 ____ 9 (8)
2. 1 3 6 10 ____ (15) 30
3. .75 .625 .5 .375 ____ .125 .25
4. 8.2 9 9.8 ____ 12.8 10.6

How much better do you think females at Brooklyn College will do than males on this kind of test?



How well do you expect to do on this test?

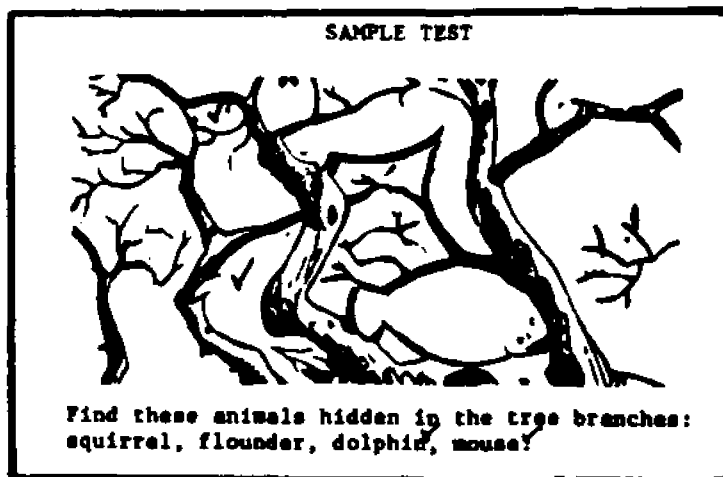


WAIT TO BEGIN

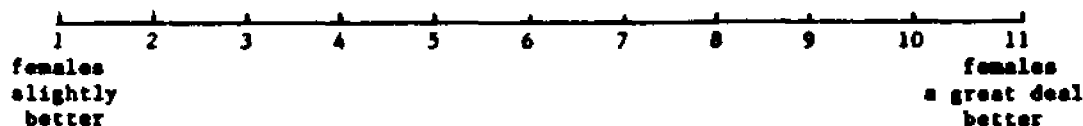
II. SHAPE RECOGNITION

An important part of mathematical aptitude is the ability to recognize and conceptualize shapes. This is not only important in geometry but in the most basic ability to get a mental picture of sets or groups of things. Males of all ages are very good at this kind of task. Again, please rank these possible reasons for their importance in this ability (1=most important reason, etc.):

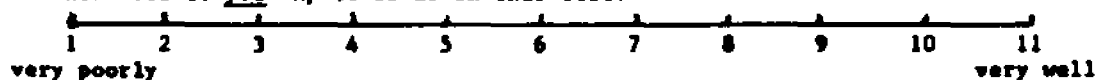
- _____ Boys are motivated to learn these skills because they want to grow up to be engineers, scientists, etc.
- _____ Males are able to think more abstractly.
- _____ Little boys play with toys that encourage mathematical thinking.
- _____ This ability has to do with biological differences between the sexes.
- _____ Other (please explain): _____



How much better than females do you think males at Brooklyn College will do on this test?



How well do you expect to do on this test?



WAIT TO BEGIN

Condition 2

VERBAL ABILITY

I. WORD KNOWLEDGE

Research has demonstrated that males usually do much better on this kind of verbal test. Although no reason fully explains this sex difference, many theories have been suggested that could partly account for such differences. Several of these suggestions are listed below. How important do you think the following reasons are? Please rank them from 1 to 4 (1=most important reason; 2=next most important reason; 3=next most important reason; 4=least important reason):

Boys learn to understand and follow instructions when working on mechanical projects like building models.

 Little boys are more interested in adventure stories with complicated plots.

Males are more logical in their thinking in general.

_____ This ability has to do with biological factors like brain structures or hormones that males have.

Other (please explain):

SAMPLE TEST

Make as many words as you can using the letters in the word ASHTRAY.

1. as 4. a 7. eat
2. ash 5. ray 8. _____
3. trey 6. say 9. _____

How much better than females do you think males at Brooklyn College will do on this test? (Circle one number)

1 2 3 4 5 6 7 8 9 10 11

males males
slightly a great deal
better better

How well do you expect to do on this test of word knowledge?

1 2 3 4 5 6 7 8 9 10 11
very poorly very well

WAIT TO BEGIN

II. WORD RECOGNITION

This test, which involves recognizing words in a complicated design, reflects the kind of verbal ability that females usually have. Again, there are many reasons suggested as to why females would be better at this kind of test. How important do you think the following reasons are in explaining this sex difference? Please rank them 1 to 4 (1=most important reason; 2=next most important reason; 3=next most important reason; 4=least important):

- _____ Little girls spend more time with their mothers and other adults and therefore learn better language skills.
- _____ Girls are more interested in quiet activities like reading and writing.
- _____ Verbal ability is a social skill, and females are more sociable.
- _____ This ability has to do with biological factors like brain structures or hormones that females have.
- _____ Other (please explain): _____

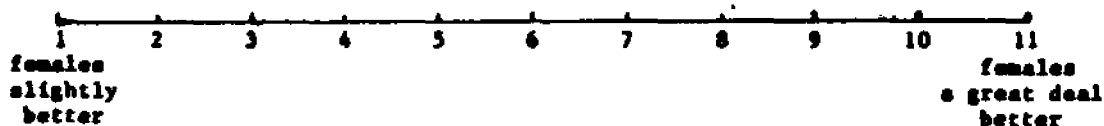
SAMPLE TEST

Look in the diagram of letters for the words given in the word list and circle them. Find them by reading FORWARDS, BACKWARDS, UP, DOWN, and DIAGONALLY.

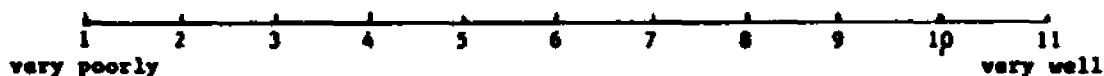
paper ✓
pen
test ✓
art
now

Q E H I N
K A C E J
M K P D L
O T S P T
W Y F G R

How much better than males do you think females at Brooklyn College will do on this test? (Circle one number)



How well do you expect to do on this test of word recognition?



WAIT TO BEGIN

MATH ABILITY

1. ARITHMETIC CONCEPTS

On this kind of test involving basic arithmetic operations, males usually do better than females. Again, please rank the following reasons as to their importance (1=most important reason, etc.):

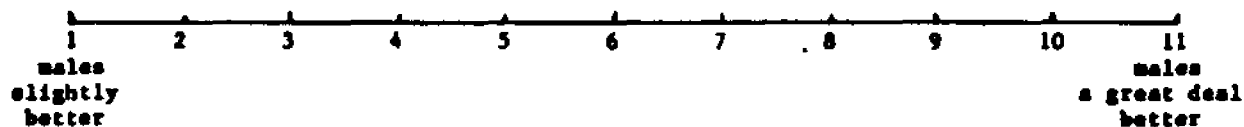
- _____ Boys are motivated to learn these skills because they want to grow up to be engineers, scientists, etc.
- _____ Males are able to think more abstractly.
- _____ Little boys play with toys that encourage mathematical thinking.
- _____ This ability has to do with biological differences between the sexes.
- _____ Other (please explain): _____

SAMPLE TEST

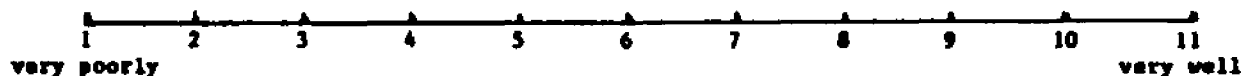
The numbers in each line below are in a certain order. In each line, circle the number that should come next.

1. 2 4 6 ____ 9 **(8)**
2. 1 3 6 10 ____ **(15)** 30
3. .75 .625 .5 .375 ____ .125 .25
4. 8.2 9 9.8 ____ 12.8 10.6

How much better do you think males at Brooklyn College will do than females on this kind of test?



How well do you expect to do on this test?



WAIT TO BEGIN

II. SHAPE RECOGNITION

An important part of mathematical aptitude is the ability to recognize and conceptualize shapes. This is not only important in geometry but in the most basic ability to get a mental picture of sets or groups of things. Females of all ages are very good at this kind of task. Again, please rank these possible reasons for their importance in this ability (1=most important reason, etc.):

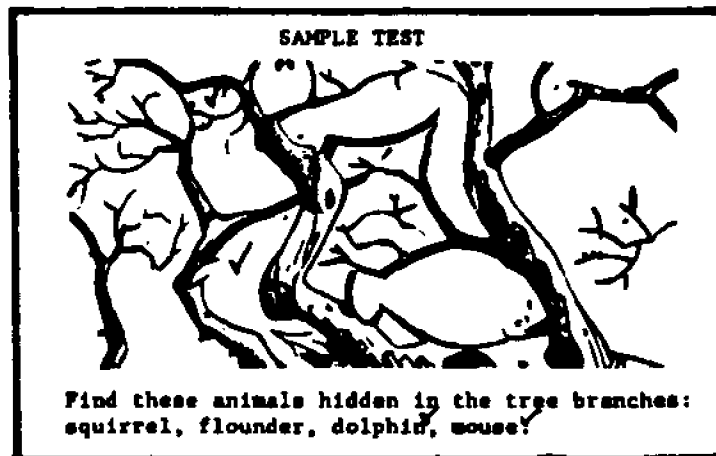
_____ Girls are motivated to learn these skills because they wish to please their teachers.

_____ Females are more patient.

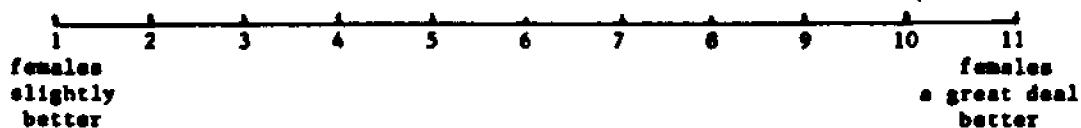
_____ Little girls learn practical mathematical concepts by watching their mothers shop, cut food portions, etc.

_____ This ability has to do with biological differences between the sexes.

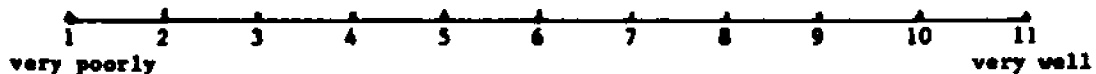
_____ Other (please explain): _____



How much better than males do you think females at Brooklyn College will do on this test?



How well do you expect to do on this test?



WAIT TO BEGIN

(CONDITION 3)

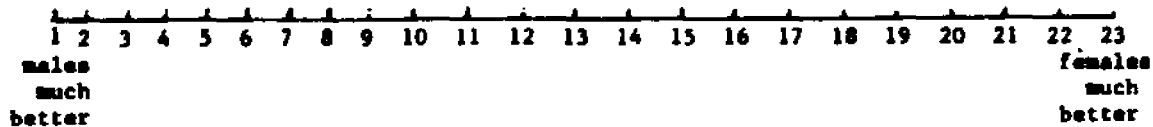
VERBAL ABILITY

1. WORD KNOWLEDGE

Research has demonstrated a sex difference on this kind of test. Although no reason fully explains this sex difference, there are probably factors that could partly account for such differences. Look at the sample test and then answer the questions below.

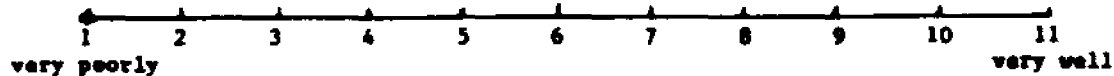
SAMPLE TEST		
Make as many words as you can using the letters in the word ASHTRAY.		
1. <u>as</u>	4. <u>a</u>	7. <u>sat</u>
2. <u>ash</u>	5. <u>ray</u>	8. _____
3. <u>tray</u>	6. <u>say</u>	9. _____

Do you think males or females will do better on this particular test? (Circle one number.)



Why do you think that group will do better?

How well do you expect to do on this test?



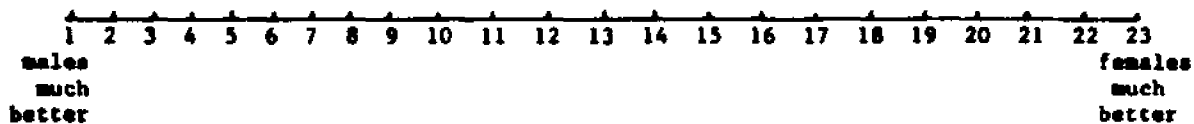
WAIT TO BEGIN

II. WORD RECOGNITION

This test, which involves recognizing words in a complicated design, also usually reflects a sex difference. Again, please explain why you think one group would be better than the other.

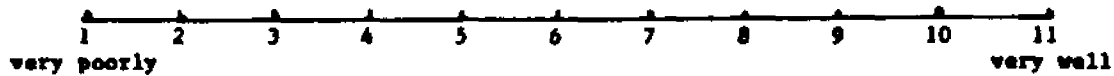
SAMPLE TEST	
Look in the diagram of letters for the words given in the word list and circle them. Find them by reading FORWARDS, BACKWARDS, UP, DOWN, and DIAGONALLY.	
paper✓ pen test✓ art now	

Do you think males or females will do better on this test?



Why do you think that group will do better?

How well do you expect to do on this test?



WAIT TO BEGIN

MATH ABILITY

1. ARITHMETIC CONCEPTS

This kind of test involves basic arithmetic operations and also shows a sex difference. Please answer the questions below, giving your opinion and possible reasons for the sex difference.

SAMPLE TEST	
The numbers in each line below are in a certain order. In each line, circle the number that should come next.	
1. 2 4 6 ____	9 ⑧
2. 1 3 6 10 ____	⑮ 30
3. .75 .625 .5 .375 ____	.125 .25
4. 8.2 9 9.8 ____	12.8 10.6

Do you think males or females will do better on this test?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Why do you think that group will do better?

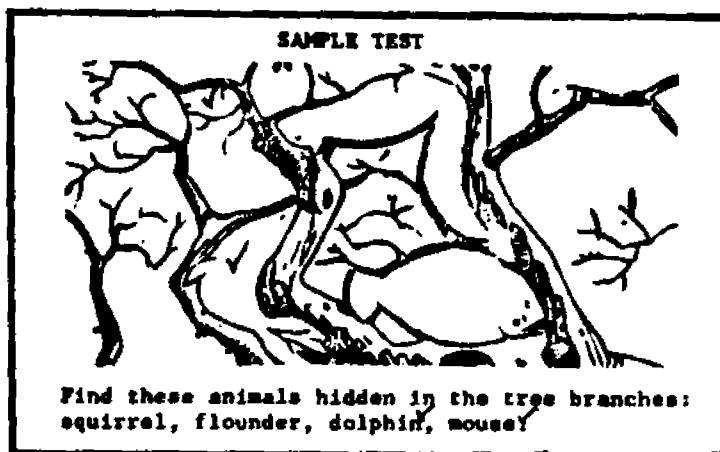
How well do you expect to do on this test?

1 2 3 4 5 6 7 8 9 10 11
very poorly very well

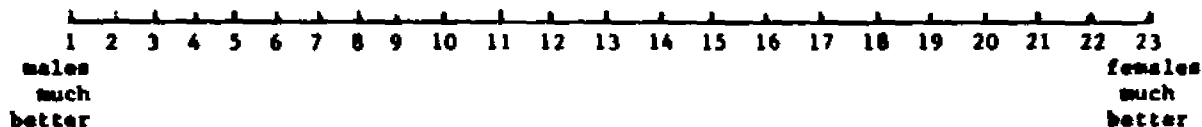
WAIT TO BEGIN

11. SHAPE RECOGNITION

An important part of mathematical aptitude is the ability to recognize and conceptualize shapes. This is not only important in geometry but in the most basic ability to get a mental picture of sets or groups of things. Again, please indicate whether you think males or females are better at this kind of test and why.



Do you think males or females will do better on this test?



Why do you think that group will do better?

How well do you expect to do on this test?



WAIT TO BEGIN

(CONDITION 4)

NEW YORK TEST OF COGNITIVE ABILITIES

This test consists of four short subtests. Each subtest has been timed for four minutes. It is very important that everyone spend the same amount of time on each part. Therefore, please wait for the signal before beginning each subtest.

Please work as rapidly as possible and try to do your very best on each of the problems or set of problems. Remember, do not turn to the next page until the signal is given unless it says GO ON at the bottom of the page. And please do not turn back once you have begun the next test.

WAIT TO BEGIN

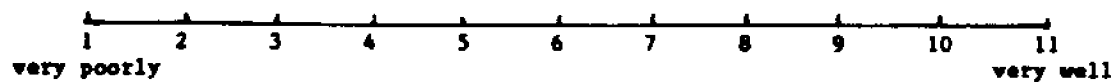
TEST 1

SAMPLE TEST

Make as many words as you can using the letters in the word ASHTRAY.

1. as4. a7. eat2. ash5. ray8. 3. tray6. say9.

How well do you expect to do on this particular test? (Circle one number.)



WAIT TO BEGIN

TEST 2

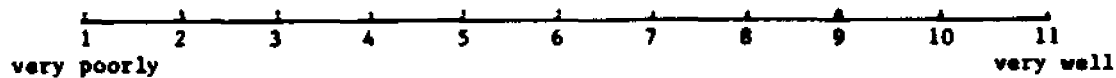
SAMPLE TEST

Look in the diagram of letters for the words given in the word list and circle them. Find them by reading FORWARDS, BACKWARDS, UP, DOWN, and DIAGONALLY.

paper✓
pen
test✓
art
now

SHIN
KACEJ
MFDL
OTS
WYFC

How well do you expect to do on this test?



WAIT TO BEGIN

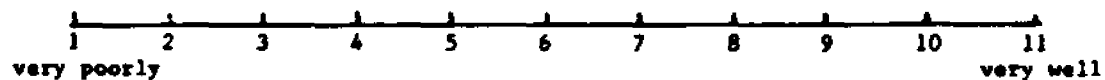
TEST 3

SAMPLE TEST

The numbers in each line below are in a certain order. In each line, circle the number that should come next.

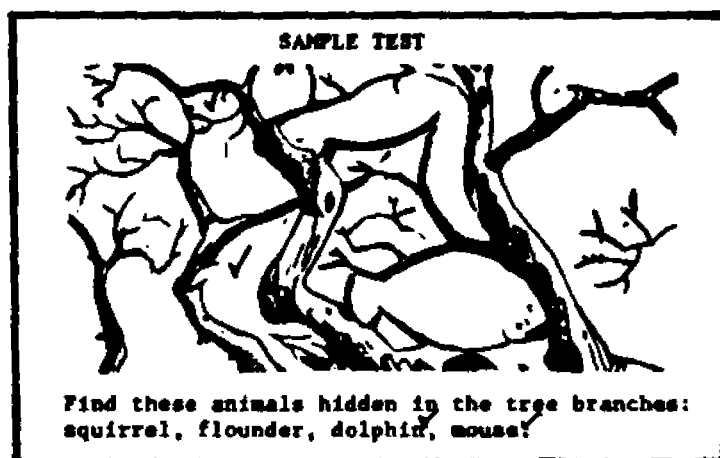
1. 2 4 6 ____ 9 **(8)**
2. 1 3 6 10 ____ **(15)** 30
3. .75 .625 .5 .375 ____ .125 .25
4. 8.2 9 9.8 ____ 12.8 10.6

How well do you expect to do on this test?

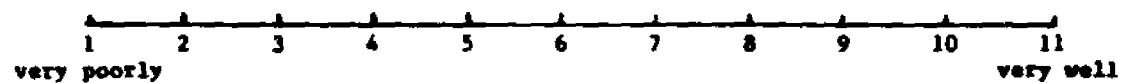


WAIT TO BEGIN

TEST 4



How well do you expect to do on this test?



WAIT TO BEGIN

Number Series

The numbers in each line below are in a certain order. In each line, find which of the two numbers at the right should come next. You may do the problems in any order.

1. 5 8 11 14 ____	17 20
2. 3 12 48 ____	192 36
3. 2 3 5 9 ____	17 14
4. 1 7 3 9 ____	5 11
5. 1 4 9 ____	25 16
6. 51 43 35 27 ____	23 19
7. 2 6 18 54 ____	162 144
8. 1 2 3 2 4 6 3 6 ____	5 9
9. 1 1 4 2 9 3 ____	6 16
10. 4 7 8 2 5 6 0 ____	3 4
11. 13 13 17 11 25 7 37 1 ____	46 53
12. 23 17 12 8 5 3 ____	1 2
13. 9 27 13 39 17 51 ____	21 19
14. 16 20 14 20 24 18 24 ____	28 18
15. 3 $\frac{1}{3}$ 9 4 $\frac{1}{4}$ 16 5 ____	$\frac{1}{5}$ $\frac{1}{2}$
16. 3 16 5 36 7 64 9 ____	100 88
17. 11 12 15 20 27 32 ____	41 35
18. 1 1 2 3 5 8 ____	8 13
19. 2 3 5 7 11 ____	13 16
20. 1 10 11 100 ____	110 101

GO ON

VI. BIBLIOGRAPHY

- Allport, G. The Nature of Prejudice. New York: Doubleday, 1958.
- Allport, G. and Kramer, B. Some roots of prejudice. Journal of Psychology, 1946, 22, 9-39.
- Arnold, J. The Cook and the Carpenter. Plainfield, Vermont: Daughters, Inc., 1973.
- Bakan, D. The Duality of Human Existence. Chicago: Rand McNally Co., 1966.
- Battle, E. Motivational determinants of academic competence. Journal of Personality and Social Psychology, 1966, 4, 634-642.
- Bem, D. and Bem, S. Case study of a nonconscious ideology: Training the woman to know her place. In D. Bem, Beliefs, Attitudes, and Human Affairs. Belmont, Calif.: Brooks/Cole, 1970.
- Bem, S. Psychology looks at sex roles: Where have all the androgynous people gone? Paper presented at UCLA Symposium on Women, May, 1972.
- The measurement of psychological androgyny. Journal of Consulting and Clinical Psychology, 1974, 42, 155-162.
- Sex-role adaptability: One consequence of psychological androgyny. Journal of Personality and Social Psychology, 1975, 31(4), 634-643.
- Birdwhistle, R. Masculinity and femininity as display. In S. Weitz (Ed.), Nonverbal Communication. New York: Oxford University Press, 1974.
- Block, J. Conceptions of sex-role: Some cross-cultural and longitudinal perspectives. American Psychologist, 1973, 28, 512-527.
- Brannon, R. Measuring attitudes toward women (and otherwise): A methodological critique. The Futures of Women: Issues in Psychology. New York: Psychological Dimensions, 1976.
- Brehm, J. Responses to Loss of Freedom: A Theory of Psychological Reactance. New York: General Learning Press, 1972.
- Broverman, I., Broverman, D., Clarkson, F., Rosenkrantz, P., and Vogel, S. Sex-role stereotypes and clinical judgments of mental health. Journal of Consulting and Clinical Psychology, 1970, 34 (1), 1-7.

- Broverman, I., Vogel, S., Broverman, D., Clarkson, F., and Rosenkrantz, P. Sex-role stereotypes: A current appraisal. Journal of Social Issues, 1972, 28, 59-78.
- Bruner, J. and Goodman, C. Value and need as organizing factors in perception. Journal of Abnormal and Social Psychology, 1947, 42, 33-44.
- Carey, G. Sex differences in problem-solving performance as a function of attitude differences. Journal of Abnormal and Social Psychology, 1958, 56, 256-260.
- Carlson, R. Sex differences in ego functioning. Journal of Consulting and Clinical Psychology, 1971, 37, 267-277.
- Chapman, L. Illusory correlation in observational report. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 151-155.
- Clifford, M. and Walster, E. The effect of sex on college admission, work evaluation and job interviews. Journal of Experimental Education, 1972, 41(2), 1-5.
- Constantinople, A. Masculinity-femininity: An exception to a famous dictum. Psychological Bulletin, 1973, 80, 389-407.
- Crandall, V.C. Sex differences in expectancy of intellectual and academic reinforcement. In R. Unger and F. Denmark (Eds.) Woman: Independent or Dependent Variable? New York: Psychological Dimensions, 1975.
- Crandall, V.J., Katkovsky, W., and Preston, A. Motivational and ability determinants of young children's intellectual achievement behaviors. Child Development, 1962, 33, 643-661.
- Deaux, K. and Taynor, J. Evaluation of male and female ability: Bias works both ways. Psychological Reports, 1973, 32, 261-262.
- Falbo, T. The effects of sex and "masculinity" on person and space-related perception. Sex Roles, 1975, 1(3), 283-295.
- Feather, N. Effects of prior success and failure on expectations of success and subsequent performance. Journal of Personality and Social Psychology, 1966, 3, 287-98.
- Fidell, L. Empirical verification of sex discrimination in hiring practices in psychology. American Psychologist, 1970, 25, 1094-1098.

- Garfinkel, H. "Passing" as a woman: A study of sex change. In A. Skolnick and J. Skolnick (Eds.) Family in Transition. Boston: Little, Brown, & Co., 1971.
- Gary, A. and Glover, J. Eye Color, Sex and Children's Behavior. New York: Nelson-Hall, 1975.
- Goldberg, P. Are women prejudiced against women? Transaction, 1968, 5, 28-30.
- Grady, K. Androgyny reconsidered. Paper presented at Easter Psychological Association Meetings, New York, 1975.
- Hampson, J. Determinants of psychosexual orientation. In F. Beach (Ed.) Sex and Behavior. New York: Wiley, 1965.
- Hass, R. Self-fulfilling prophecy as a determinant of social class and racial differences in learning ability. Unpublished manuscript, Duke University, 1969.
- Heider, F. The Psychology of Interpersonal Relations. New York: Wiley, 1958.
- Herschberger, R. Adam's Rib. New York: Harper and Row, 1948.
- Horner, M. Femininity and successful achievement: A basic inconsistency. In M. Garskoff (Ed.) Roles Women Play. Belmont, California: Brooks/Cole, 1971.
- Jones, E. and Davis, K. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.) Advances in Experimental Social Psychology, vol. 2. New York: Academic Press, 1965.
- Kuhn, T. The Structure of Scientific Revolution. Chicago: University of Chicago Press, 1962.
- LeGuin, U. The Left Hand of Darkness. New York: Ace Books, 1969.
- Levi-Strauss, C. The family. In H. Shapiro (Ed.) Man, Culture, and Society. London: Oxford University Press, 1971.
- Lindzey, G. and Rogolsky, S. Prejudice and identification of minority group membership. Journal of Abnormal and Social Psychology, 1950, 45, 37-53.
- Maccoby, E. and Jacklin, C. The Psychology of Sex Differences. Stanford, California: Stanford University Press, 1974.

- Milton, G. Sex differences in problem solving as a function of role appropriateness of the problem content. Psychological Reports, 1959, 5, 705-708.
- Monahan, L., Kuhn, D., and Shaver, P. Intra-psychic versus cultural explanations of the "fear of success" motive. Journal of Personality and Social Psychology, 1974, 29, 60-64.
- Money, J. and Ehrhardt, A. Man & Woman, Boy & Girl. Baltimore: Johns Hopkins University Press, 1972.
- Nichols, R. Subtle, obvious, and stereotype measures of masculinity-femininity. Education and Psychological Measurement, 1962, 22, 449-461.
- Norman, D. Memory and Attention. New York: Wiley, 1976.
- Parlee, M. Sex differences in perceptual field dependence: A look at some data embedded in theory. Unpublished manuscript, University of South Carolina, 1974.
- Pheterson, G. and Kiesler, S. Evaluation of the performance of women as a function of their sex, achievement, and personal history. Journal of Personality and Social Psychology, 1971, 19, 114-118.
- Rosenberg, R. The dissent from Darwin 1890-1930: The new view of women among American social scientists. Doctoral dissertation, Stanford University, 1974.
- Rosenkrantz, P., Vogel, S., Bee, H., Broverman, I., and Broverman, D. Sex-role stereotypes and self-concepts in college students. Journal of Consulting and Clinical Psychology, 1968, 32, 287-295.
- Rubin, G. The traffic in women: Notes on the political economy of sex. In R. Reiter (Ed.) Toward an Anthropology of Women. New York: Monthly Review Press, 1975.
- Sherriffs, A. and Jarrett, R. Sex differences in attitudes about sex differences. Journal of Psychology, 1953, 35, 161-168.
- Spence, J., Helmreich, R., and Stapp, J. The personal attributes questionnaire: A measure of sex role stereotypes and masculinity-femininity. Journal Supplement Abstract Service Catalog of Selected Documents in Psychology, 1974, 4, 43 (ms. no. 617).
- Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. Journal of Personality and Social Psychology, 1975, 32(1), 29-39.

- Stocking, G. Race, Culture, and Evolution. Chicago: University of Chicago Press, 1968.
- Tagiuri, R. Person perception. In G. Lindzey and E. Aronson (Eds.) The Handbook of Social Psychology, vol. III, New York: Addison-Wesley, 1969.
- Tajfel, H. Social and cultural factors in perception. In G. Lindzey and E. Aronson (Eds.) The Handbook of Social Psychology, vol. III. New York: Addison-Wesley, 1969.
- Taynor, J. and Deaux, K. When women are more deserving than men: Equity, attribution, and perceived sex differences. Journal of Personality and Social Psychology, 1973, 28(3), 360-367.
- Tyler, B. Expectancy for eventual success as a factor in problem solving behavior. Journal of Educational Psychology, 1958, 49, 166-172.
- Unger, R. and Siiter, R. Sex-role stereotypes: The weight of a "grain of truth." Paper presented at the Eastern Psychological Association Meetings, Philadelphia, April, 1974.