

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

NOTE TO USERS

The original manuscript received by UMI contains pages with slanted print. Pages were microfilmed as received.

This reproduction is the best copy available

UMI



Université d'Ottawa • University of Ottawa

Doctoral Dissertation

The Effects of Sex and Sex-Typing on Motivation, Delay, and Distortion in
Negative Feedback Situations

Gretchen L. Conrad

University of Ottawa

1996

This document is being submitted to the School of Graduate Studies and Research in partial fulfilment of the requirements for the degree Doctor of Philosophy, School of Psychology.



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-28332-1

"We must focus on the interaction between employee and boss, because it is only through honest, personal communication, task assessment and direction, that superior performance will develop."

- Cohen & Jaffee, 1982, p. 98

Dedication

To my husband, Mark Shulist

Without your support and your belief in me,
I would never have started upon this journey.

To Katarina Spencer
October 13, 1963 - April 9, 1996.

A woman who taught me much about
courage and perseverance
in the face of seemingly unsurmountable odds.

Acknowledgements

First and foremost, I have to thank my supervisor, Dr. Michael McCarrey. You have been a wonderful friend and role model. Your support, encouragement, gentle prodding, persistence, and sense of humour have carried me through these many years. It has been my privilege to work with you on both my Honours and Doctoral theses, and I leave now with many lessons learned from you, some dealing with academia, but many on just how to deal with life in general.

Thank you to the members of my thesis committee -- Dr. Gail Crombie, Dr. Pierre Mercier, and Dr. Luc Pelletier -- for your guidance, suggestions, expertise, and cooperation. To Dr. Dwayne Schindler and Dr. Keith Busby, for your statistical knowledge and patience.

Thank you to my many research assistants: Lisa Lynk, Dominic Rockall, Stephanie Bennett, Doreen Deveen, Sylvie Coté, Lisa Blanchard, Steve Boisclair, Victor Gallant, Paul Freeman, Gary Goldstein, Eran Talitman, and Serge Lessard. You were wonderfully cooperative -- often in the early hours of the morning! Without you, this project would never have gotten off the ground...

To my many friends, in particular, Margaret Wood, Elecia Elliott, Jennifer Totten, Roberta Anderson, Nancy DelCellier, and Jane Gayton. Your friendships have been a true blessing. We have shared many experiences in the past, and I look forward to what our futures hold. Elecia and Jane, I hope you realize just how invaluable your computer skills were in the completion of this project.

To my parents, Jacqueline and Walter Conrad, for the many lessons I have learned. Thank you.

And to the rest of my family. To Mark -- my husband, my partner, my lover and my best friend -- what can I say? Your love, support, encouragement, and perseverance have kept me going on many an occasion. That we have finally reached this stage is a testament to your belief in me and in our future together. To my daughter, Merina, for helping me set my priorities and for reminding me of the wonder in the world. And to my *soon-to-arrive* sons, "Humpty and Dumpty", thank you for giving me the final push I needed to complete this document.

TABLE OF CONTENTS

ABSTRACT	1
Reluctance to Deliver Bad News	2
Sex-Typing	6
Hypotheses and Rationale	9
Supervisor Appraisal and Effectance Motivation	12
Hypotheses and Rationale	20
PILOT RESEARCH	23
<u>Pilot #1</u> -- Category Correlation between the Bem Sex Role Inventory and the Personal Attributes Questionnaire	23
Method	23
Results	23
Discussion	25
<u>Pilot #2</u> -- Factor Analytic Examination of the Bem Sex Role Inventory - Short Form	26
Method	29
Results and Discussion	30
<u>Pilot #3</u> -- The Impact of Sex-Typing and Motivational Expectancies on the Delay and Distortion of Negative Performance Appraisal Feedback	36
Hypotheses	36
Method	37
Subjects and Recruitment	37
Measures	37
Procedure	40
Results	41
Discussion	49

TABLE OF CONTENTS (Cont'd)

On Site Research

The Effects of Sex and Sex-Typing on Motivation, Delay, and Distortion in Negative Feedback Situations	52
---	----

Method	52
Subjects and Recruitment	52
Measures	53
Procedure	57
Results	59
Discussion	81
Conclusions	94

References	96
------------	----

Appendix 1

The Bem Sex Role Inventory - Short Form (BSRI-S)

Appendix 2

Supervisory Control Worksheet
Computer Science Aptitude Test -- Norms
Computer Science Aptitude Test -- Test Booklet
Computer Science Aptitude Test -- Answer Sheets 1 & 2
Supervisor Qualifications Sheet -- Pilot Study
Supervisor Qualifications Sheet -- On Site Study

Appendix 3

Pilot Study Questionnaires

Supervisor Performance Evaluation Sheet
Appraisal Effectance Motivation Questionnaire
Post-Evaluation Questionnaire

TABLE OF CONTENTS (Cont'd)

Appendix 4

On Site Study Questionnaires

- Supervisor Performance Evaluation Sheet
- Volunteer Performance Factors
- Supervisor Evaluation Factors
- Post-Evaluation Questionnaire
- Appraisal Effectance Motivation Questionnaire

Appendix 5

Power Calculations

TABLE LIST

1.	Correlation and Comparison of Classifications Using the <u>BSRI</u> With Those Using the PAQ	24
2.	Items on the BEM Sex Role Inventory - Short Form	31
3.	Items Loading from the Pattern Matrix for the Three-Factor Solution	32
4.	Summary of Indices Used to Assess Goodness-Of-Fit	35
5.	Pilot Study - Delay and Distortion Values for Male and Female Volunteer Supervisors	43
6.	Pilot Study - Delay and Distortion Value for Instrumental Male and Instrumental Female Volunteer Supervisors	44
7.	Pilot Study - Delay and Distortion Values for Instrumental vs Expressive Female Volunteer Supervisors	45
8.	Pilot Study - Motivation Values for Male and Female Volunteer Supervisors	46
9.	Pilot Study - Motivation Values for Instrumental vs Expressive Sex-Typed Female Volunteer Supervisors	47
10.	Pilot Study - Motivation Values for Instrumental Male and Instrumental Female Volunteer Supervisors	47
11.	Pilot Study - Delay, Distortion and Effectance Motivation Values for Trained vs Untrained Volunteer Supervisors	48
12.	Demographic Data for <u>BSRI-S</u> Manager Normative Sample(N - 137)	60
13.	Norms for <u>BSRI</u> and <u>BSRI-S</u>	60
14.	Mean Delay and Distortion Values for Supervisor Sample (2 X 4 Factorial MANOVA)	63
15.	Delay and Distortion Values for Male Versus Female Volunteer Supervisors	64

TABLE LIST (Cont'd)

16.	Mean Delay and Distortion Values for Instrumental and Expressive Groups (2 X 2 Factorial MANOVA)	65
17.	Mean Delay and Distortion Values for Instrumental Versus Expressive Supervisors	66
18.	Mean Values for Composite Distortion Scores and for Delay Scores (2 X 4 Factorial MANOVA)	68
19.	Composite Distortion Values (Means) and # of Test Trials (Delay) for Gender Groupings of Volunteer Supervisors	69
20.	Mean Motivation Values -- Self-Efficacy Expectancy, Outcome Expectancy, and Outcome Value -- for Supervisor Sample (2 X 4 Factorial MANOVA)	73
21.	Mean Appraisal Motivation Scores for Male Versus Female Supervisors	74
22.	Mean Appraisal Motivation Scores for Supervisors With and Without Previous Performance Appraisal Training	74
23.	Mean Performance Attribution Values for Supervisory Sample (2 X 4 Factorial MANOVA)	75
24.	Mean Performance Attribution Variables for Male Versus Female Supervisors	76
25.	Mean Supervisor Variable Scores (Confidence, Concern and Anxiety) for Gender Groupings of Volunteer Supervisors	78
26.	Mean Supervisor Variables (Confidence, Concern and Anxiety) for Male Versus Female Volunteer Supervisors	80

FIGURE LIST

1.	Standard Estimates for the Baseline Model of the Feminine and Masculine Subscales of the Bem Sex Role Inventory - Short Form	34
2.	Percentage Breakdown of Biological Sex by Sex-Type Classification	61
3.	Composite Distortion Means or Gender Groups of Volunteer Supervisors	70
4.	Mean Delay (# of Test Trials Allowed) for Gender Groupings of Volunteer Supervisors	71
5.	Mean Supervisor Concern Scores for Gender Groups of Volunteer Supervisors	79

ABSTRACT

Personnel evaluation reports are an important tool in the work place. Unfortunately, much research suggests not only that performance evaluations themselves are often not accurate, but also that the results may be distorted and feedback delayed.

Research has focused on many possible determinants of accuracy in performance ratings. Landy and Farr (1980) have suggested that focusing on the techniques of evaluation is not sufficient; we need to have a better understanding of the processes involved in appraisal in order to advance in this area. If the factors that influence inaccurate evaluations could be isolated, then proactive measures could be taken to circumvent or reduce the likelihood of their occurring.

Benedict and Levine (1988), examined a group of business students and found that females have more difficulty with negative feedback than do males, positively distorting results more and further delaying feedback. Building upon this work, the purpose of this research is to further examine the relation between sex, gender or sex-typing, effectance motivation, and delay and distortion in personnel evaluations. It was hypothesized that it is the sex-typing or gender of the supervisors, and not their biological sex per se, that influences the extent to which they delay and distort. It was expected that in comparison with instrumental supervisors, expressive supervisors would positively distort and delay more on the feedback of poor performance results, and that female supervisors would also have lower levels of self-efficacy expectancy, outcome expectancy, and outcome value for the delivery of prompt and accurate negative feedback. Attribution factors, supervisory concern factors, and the impact of training were also considered.

For the most, results did not support the hypotheses, with almost no significant differences between sexes or between sex-types on the variables examined. The findings are significant because they set limits on the results of Benedict and Levine (1988). Using a sample of supervisors from the Federal Service, no significant sex differences were found, suggesting that, within their work context, these individuals act as supervisors first and as males and females second. Interestingly, there was nonetheless a high level of distortion of negative performance results on the part of all supervisors. Training appeared to have little impact on the extent to which supervisors delayed or distorted, and only minimal impact on motivational factors. There was some suggestion that sex-type, as opposed to biological sex, could account for some previously measured sex differences in performance evaluations.

The Effect of Sex and Sex-Typing on Motivation, Delay, and Distortion in
Negative Feedback Situations

Reluctance to Deliver Bad News

Considerable research indicated that, in general, people are reluctant to deliver bad news to others (e.g., Conlee & Tesser, 1973; Deaux, 1974; Tesser, Rosen & Tesser, 1971; Tesser, Rosen & Batchelor, 1972). For example, Tesser et al. (1971) wanted to see if this reluctance occurred in situations where the conveyance of bad news is routine (e.g., an employment counsellor telling a client s/he did not get the job). They examined twenty-seven applications for assistance from a local social service agency in Georgia, USA. They discovered that all of the clients who were granted aid were informed on the same day, however only 73% of those not granted aid were informed as promptly. It is interesting to note that this reluctance occurred even though notification was by form letter, there was no face-to-face interaction or verbal contact with the rejected party. Despite the small sample size of this study, the results showed a trend in the predicted direction -- not to convey bad news.

Tesser and Rosen (1975) summarized the research (both theirs and that of others) on components of the MUM effect -- the alleged tendency to keep mum about unpleasant messages (p. 193). They proposed a multicausal model: a) the communicator's self-concern, b) his/her concern with the recipient, and c) his/her concern with societal norms. Several factors may be influencing the communicator's self-concern: guilt, fear of negative self-evaluation, and mood. The guilt hypothesis assumes that one would avoid communicating bad news because of guilt over the fate of the recipient. Research in this area is ambiguous, with some supporting results and others suggesting caution. The fear-of-negative-evaluation hypothesis proposes that communicators wish to avoid delivering bad news for fear they will receive a corresponding negative evaluation (i.e., that anyone associated with a negative stimulus acquires similar properties). Again, although many studies are supportive, Tesser and Rosen (1975) suggest there is sufficient contradictory evidence to render this hypothesis tentative. According to the mood hypothesis, individuals are reluctant to assume the negative mood state they feel is required when delivering bad news. Although there is less direct research on this suggestion

than on the previous two, the available data has been supportive.

The second umbrella hypothesis, concern for the recipient, has been examined from two perspectives: a) the recipient's emotionality, and b) the recipient's desire-to-hear. The former holds that good news is delivered more than bad because the latter would emotionally disturb the recipient. The desire-to-hear hypothesis proposes that individuals assume others wish to be told good, and not bad, news. Most research is supportive of both proposals.

Concern with norms has also been investigated using two models: a) that norms proscribe the transmission of bad news, and b) that norms concerning such transmission are relatively ambiguous. Although there was no support for the former hypothesis, the norm ambiguity hypothesis was upheld.

Tesser and Rosen (1975) discuss several possible implications of the MUM effect, particularly in a work place environment. Admirable performance is more likely to be brought to the attention of the individual concerned than is less satisfactory performance. Although this could be likened to using positive reinforcement to promote desirable actions and behaviours, withholding of bad news can also have negative consequences. Individuals who are experiencing difficulty may be further disadvantaged as they do not receive full information about their situations. Moreover, they may remain unaware of their shortcomings, and, thus, they may not take the steps necessary to change. As well, although bad news may not be mentioned to the person directly involved, it may be discussed with others, in this case, co-workers. This could promote a negative work environment through oblique negative feedback (e.g., rumours) and promoting distrust, particularly with regards to superiors.

Bond and Anderson (1978) examined whether or not the reluctance to transmit bad news was due to intrapsychic discomfort or to self-presentational concerns. They hypothesized that such reluctance is a public display. "By affecting the reluctance, people regulate a situated image, avoid an unfavourable impression and pay homage to a social norm" (p. 177).

Subjects ($N=38$) were required to administer intelligence tests to confederate participants and to give feedback in one of two conditions: visible to the testee, not visible to the testee. Half of the test results placed the testee in the upper 20% of the university population, the other half were in the lower 20%. Subjects who were visible took twice as long to deliver the

unfavourable results as they did to deliver the successful results. There were no significant differences when the subjects believed they were not visible to those taking the tests. Video cameras recorded the subjects for behavioural evidence of the discomfort (e.g., gaze aversion, self-manipulations) at the failure of the test takers. Subjects in the visible category filled in the feedback delay with active self-presentation (e.g., rechecking the scores). Subjects in the non-visible category did not display such behaviours. However, subjects in the non-visible group did show discomfort (e.g. making excuse on the test-taker's behalf, claimed displeasure) after the transaction, when they were asked to fill out a questionnaire for the experimenter. Bond and Anderson (1987) suggest that had the delay and discomfort been due to conflicting self-image, and not to the maintenance of a public image, all participants would have hesitated before returning negative results. That reluctance to transmit bad news is due to maintaining a public image is also consistent with the results of Deaux (1974), who found that under conditions of anonymity, the positive or negative aspect of the message is not as crucial as it is in face-to-face situations.

Ilgen & Knowlton Jr. (1980) examined distortion of poor performance results, both how evaluation results were communicated to subordinates and how performance attributions (e.g., ability and effort) affected this feedback. They hypothesized that performance evaluations of low performers would be significantly higher when the rater is aware of his/her need to give feedback, than when evaluations are made prior to being aware of this requirement. They further hypothesized that feedback would vary more as a function of attributions based on effort rather than on ability. Thus, the extent to which performance evaluations are slanted, may not only be a function of a face-to-face interaction, but also the result of attributions made for the cause of poor performance.

Forty undergraduate psychology students acted as supervisors of three-person work groups. These volunteers directed the performance activity of the groups, and they were required to provide feedback to one of their three subordinates. All the subordinates were confederates, and performance and attributionally relevant information about performance were manipulated.

The results showed that supervisors significantly distorted feedback in a positive direction

for low performers. These results were particularly pronounced if the supervisors had been lead to believe that poor performance was due to lack of ability. Additionally, supervisors made lower ratings prior to being aware of the need to provide feedback; the evaluations were raised when they became aware of this requirement. Ilgen & Knowlton Jr. (1980) suggest there is an additive effect between the attributions for poor performance and the need to give feedback. Thus, when performance is attributed to lack of ability rather than to lack of effort, supervisors provide even less appropriate feedback to the subordinate.

Benedict and Levine (1988) also examined the impact of several variables thought to influence the delivery of Performance Evaluation Results (PERs): supervisor gender, level of subordinate performance, and whether or not face-to-face feedback was provided. The subjects ($N=64$; half male, half female) were white management and business students from west Texas universities. All had at least six months in a paid supervisory position. Each supervisor was introduced to a male confederate subordinate who attempted to establish rapport, expressing concern about his upcoming test performance as he was looking for job opportunities and wished to use these test results. The supervisor's responsibilities were then explained. After they had permitted up to ten subordinate performance trials, the supervisors judged the quality and capability of subordinate performance (measures of distortion), and made appointments for a face-to-face feedback session (measure of delay). All confederate subordinates performed at a fixed substandard level -- the 25th percentile.

All three independent variables influenced the subsequent performance appraisals. All supervisors delayed and distorted performance ratings, and delayed reviews of sub-standard performances. Female supervisors distorted performance ratings more than did males. In addition, the female supervisors delayed longer than did the males when setting appointments to provide feedback. This tendency for female supervisors to both delay and distort was exacerbated when face-to-face feedback was required. Benedict and Levine (1988) suggested several factors which might be at work: 1. lower confidence of female supervisors in their evaluations of male subordinates, 2. greater concern of female supervisors about overall accuracy of their own ratings, and 3. greater concern of the female supervisors about the effects the ratings would have on the subordinates.

Other research has also shown differences between evaluations completed by male versus females. London and Poplawski (1976) found that evaluations along individual performance dimensions in appraisal and interview situations were higher when made by female college students than when made by male college students. This trend seems to suggest that women have a tendency to evaluate others more positively, or with more leniency, than do men.

Sex-Typing

Historically, masculinity and femininity have been viewed as opposite ends of a bipolar continuum. Bem (1974, 1981) and Spence, Helmreich and Stapp (1975a, 1975b) adopted a different view; they maintain that these are distinct dimensions of personality and should be measured separately. Further, they postulate that both masculinity and femininity could be integrated, to varying degrees, within an individual. Bem (1974) developed the Bem Sex Role Inventory (BSRI), and Spence et al. (1975a) the Personal Attributes Questionnaire (PAQ) to facilitate empirical research in this area. These tests measure socially desirable instrumental and expressive characteristics in an individual, and have their greatest predictive utility when relating to behaviours requiring expressive or instrumental capabilities (Spence & Helmreich, 1980).

The BSRI and the PAQ categorize individuals according to their gender or sex-type: sex-typed (feminine or masculine), cross-sex-typed, androgynous, and undifferentiated. (For the purposes of this paper, and in keeping with APA guidelines (1995), the term "sex" will be used to describe an individual's biological sex; the terms "gender" or "sex-type" will be used to describe the traits or characteristics that are considered sex-related.) Sex-typed individuals are those exhibiting traditionally appropriate male or female characteristics. Cross-sex-typed persons exhibit characteristics that are stereotypically gender inappropriate. Thus, a person who is high feminine and low masculine would be called sex-typed if she were female, and cross-sex-typed if he were male. An individual with low feminine characteristics but high masculine ones would be cross-sex-typed for females, and sex-typed for males. Androgynous individuals are those who incorporate traits typically belonging to both sexes, high feminine and high masculine characteristics. Undifferentiated persons appear to possess personality characteristics that are

not linked to sex-role stereotypes; they rate low masculine and low feminine.

"The BSRI is based on a conception of the traditionally sex-typed person as someone who is highly attuned to cultural definitions as the ideal standard against which her or his own behaviour is to be evaluated" (Bem, 1981, p. 4). Such individuals would be motivated to initiate and maintain behaviours and actions consistent with those of a stereotypic image of masculinity or femininity. Bem (1981) suggests androgynous individuals have more flexibility in their behavioural repertoire enabling them to draw on either stereotypically male or stereotypically female behaviours depending upon the situation. Thus, it would be more the requirement of a situation, in combination with both gender based schemas, that would dictate the androgynous individual's behaviour. Those individuals who are categorized as undifferentiated may determine their behaviours without any reference to stereotypic gender-based schema, possibly using situational requirements alone.

Orlofsky (1977) examined the relation between sex-role orientation, identity formation (as per Erickson, 1959), and self-esteem in college men and women ($N = 111$; 55 males, 56 females). Results showed that possessing a high level of both masculine and feminine characteristics (i.e., androgyny) is conducive to identity achievement and higher self-esteem. Low levels of both masculine and feminine traits (e.g., undifferentiation) is associated with low self-esteem and a lack of personal integration. Sex-typing was most frequently associated with premature identity commitment and lack of personal differentiation, and with high self-esteem in males but with low self-esteem in females. The results for cross-sex-typing were almost the exact reverse. Cross-sex-typing was associated with high levels of self-esteem and identity achievement in females, but with somewhat lower self-esteem in males. Overall these results suggest that high levels of masculinity, with or without high levels of femininity, are associated with higher levels of self-esteem.

Waterman and Witbourne (1982) examined the relations between sex-typing and psychological development among college students and adults. They found the highest scores on psychosocial development were obtained by those with a androgynous orientation followed by masculine, feminine and undifferentiated individuals, in that order. The findings applied equally to both sexes and to both age groups examined (Group 1: 299 undergraduate students,

144 male, 155 female, average age=19.7 years; Group 2: 155 Alumni, 79 male, 76 female, average age=30.7 years). These results support the view that an endorsement of both expressive* and instrumental* characteristics represents an added adaptive capacity. As in other studies (e.g., O'Connor, Mann & Bardwick, 1978; Spence, Helmreich & Stapp, 1975b), the high psychosocial effectiveness scores obtained by both androgynous and masculine individuals, regardless of sex, suggest the relative importance of masculine over feminine characteristics for effective functioning. However, the significant differences between androgynous and masculine participants suggest that the endorsement of feminine qualities in combination with masculine ones makes an important contribution beyond that of masculine ones alone.

Cano, Solomon and Homes (1984) examined sex-role identity, and fear of success. Two hundred and four (80 female, 124 male) subjects were administered the Sadd Fear of Success Scale (SFOS; Sadd, Lenauer, Shaver & Dunivant, 1978), the BSRI (Bem, 1974) and the PAQ (Spence, Helmreich & Stapp, 1975a, 1975b). Regardless of sex, androgynous and masculine individuals reported less fear of success than did feminine or undifferentiated individuals. The fear of success seemed to be related more to the absence of masculine traits rather than to the presence of feminine ones. The results were analyzed to determine the specific components of masculinity related to fear of success. High self-confidence, decisiveness, analyticalness, and independence were related to low fear of success levels. It appears that it is the lack of such characteristics rather than the fear of being unfeminine that causes fear of success. Sadd, Miller and Zeitz (1979) also found a strong negative correlation between masculinity and fear of success; femininity and fear of success were unrelated.

Bem and Lenney (1976) examined the avoidance of stereotypically gender inappropriate activities. It was hypothesized that sex-typed individuals engage in high levels of gender-stereotypic behaviour because they are motivated to maintain behaviours consistent with cultural definitions of gender-appropriateness. It was questioned whether gender-inappropriate activity was motivationally problematic for sex-typed individuals, or whether they would be willing to be thus engaged if the situation was structured to encourage it.

* Although there is controversy in the literature as to the precise definition of the terms "expressive" and "instrumental", within the context of this paper, "expressive" will be used interchangeably with "feminine sex-typed" traits or characteristics, and "instrumental" with "masculine sex-typed" traits or characteristics.

Seventy-two undergraduate students (24 sex-typed, 24 androgynous, and 24 cross-sex-typed) were asked to choose from listed paired activities. Students were asked to select one activity from each pair for which they would be paid to model for photographing. (To reduce self-consciousness due to performance anxiety, it was emphasized that there was no interest in how well they could perform the activities, and they were told the photos would be used only on another campus.) Fifteen pairs required choices between activities with opposite gender connotations; the gender inappropriate activity always paid more. In the remaining pairs, both activities were either masculine, feminine or neutral, and one of each pair was arbitrarily assigned more money.

Sex-typed participants were significantly more likely to choose behaviours consistent with their own sex than were androgynous or cross-sex-typed participants. (Individuals in the latter two categories did not differ from one another in this regard.) This tendency to select gender appropriate activities, and to reject inappropriate ones was maintained by sex-typed subjects *even when* such choices cost them money. "Moreover, actually engaging in cross-sex behaviour caused sex-typed subjects to report greater psychological discomfort and more negative feelings about themselves" (Bem & Lenney, p. 48).

-- Hypotheses and Rationale --

Hypotheses

1) IT IS THE GENERAL HYPOTHESIS THAT IT IS THE SEX-TYPING OF THE SUPERVISORS, AND NOT THEIR BIOLOGICAL SEX PER SE, THAT INFLUENCES THE EXTENT TO WHICH THEY DELAY AND DISTORT. NO SIGNIFICANT DIFFERENCES ARE EXPECTED ACROSS SEX FOR SUPERVISORS OF SIMILAR SEX-TYPES. DIFFERENCES ARE EXPECTED WITHIN SEXES ACROSS SEX-TYPES.

2) IN COMPARISON WITH INSTRUMENTAL SUPERVISORS, IT IS EXPECTED THAT EXPRESSIVE SUPERVISORS WILL POSITIVELY DISTORT MORE ON THEIR EVALUATION OF POOR PERFORMANCE.

3) IN COMPARISON WITH INSTRUMENTAL SUPERVISORS, IT IS EXPECTED THAT EXPRESSIVE SUPERVISORS WILL DELAY MORE ON THEIR FEEDBACK CONCERNING POOR PERFORMANCE.

Rationale

By theoretical definition, sex-typed females attribute to themselves those expressive characteristics that are traditionally assigned to females: affection, warmth, tenderness, sympathy, compassion and sensitivity (Bem, 1981). Conversely, sex-typed males ascribe to those instrumental characteristics traditionally attributed to males: assertiveness, forcefulness, dominance and willingness to take a stand (Bem, 1981). Expressive traits are conventionally associated with strong interpersonal skills and with being carefully attuned to the needs and feelings of others. However, possessing only these expressive traits, without any of the traditionally masculine, instrumental ones, would make it difficult for a supervisor to provide prompt and accurate negative feedback*. Such an expressive supervisor would be quite concerned with the interpersonal consequences of providing such feedback -- How would the negative results be received? Would the subordinate attempt to retaliate against the supervisor? How would such actions reflect on the supervisor? How would the feedback affect both the self-esteem of the recipient and the subsequent supervisory relationship?

Theoretically, expressive supervisors would feel more uncomfortable giving negative feedback than would instrumental supervisors, as the content activates dysfunctional cognitive schemata which are inconsistent with the nurturant and caring behaviours stereotypically attributed to females. As a result, these female supervisors might delay and positively distort the PER results to soften the blow to the subordinate. Such behaviour would also reduce their own feelings of discomfort and low-self-esteem arising from behaving in a stereotypically discordant fashion.

* In this thesis "negative feedback" refers to feedback concerning performance set at the 25th percentile.

At the other end of the spectrum, instrumental traits are congruent with being goal directed, and with being able to accomplish the task required. Instrumental supervisors would be less likely to delay and distort than would expressive supervisors, as, by definition they have the necessary independent, instrumental skills. The feedback concerning poor performance would activate cognitive schemata more consistent with their focus on goal setting and task accomplishment. Possible subordinate surprise, anger, and retaliation would not have the same inhibitory impact on provision of prompt and accurate feedback. There would be less focus on how feedback would be received, and more on accomplishing the appraisal task.

Androgynous supervisors, male and female, by definition possess both instrumental and expressive characteristics. They would be able to draw on the traditionally instrumental ones to deliver timely and accurate results, while using the traditional expressive ones to deliver the negative feedback in an empathic and concerned manner.

Much research has investigated the differential reactions of observers to the assertive behaviours of males and females. Specifically, it has been hypothesized that, as a result of traditional sex-role stereotyping, females' assertions would elicit more negative reactions than would those of males (e.g., Kelly & Worell, 1977; MacDonald, 1982). The results however have been mixed (e.g., Bean, 1997; Kelly, Kern, Patterson & Kean, 1980; Schroeder, Rakos & Moe, 1983). Kern et al. (1985) suggested that only a certain subset of observers react differentially to the assertive behaviours of males and females. They expected that female assertiveness would elicit more negative reactions than would similar assertive behaviour in males from those observers with more traditional sex-role stereotypes (i.e., those with conservative attitudes towards the role of women). "Specifically, individuals who sex-type females into a traditional female role would evidence biases against females' assertions, males' non-assertions and possibly females' empathic assertions" (p. 65).

They administered the Attitudes Towards Women Scale (Spence and Helmreich, 1972) to 1,846 undergraduate students. From this original group, 209 students with conservative attitudes towards the roles of women and 253 students with liberal views about the roles of women were selected. These participants viewed videotapes of males and females modelling assertive, empathic-assertive, and submissive behaviours in response to an unreasonable request.

As predicted, conservative observers devalued all the assertions of the female models, and females portraying assertive roles were rated significantly lower on such characteristics as social skill, likability and attractiveness. Of interest, females portraying assertive roles were rated significantly more competent, but less likeable than were the non-assertive models.

Such findings have implications for this study in that expressive supervisors would anticipate such adverse reactions and would delay and distort so as to avoid or prevent it. In contrast to their instrumental or androgynous counterparts, the expressive female supervisors in this study might be more likely to feel that providing accurate and prompt negative feedback face-to-face would be incompatible with standards of stereotypic feminine behaviour. They would see themselves, and see others as seeing them, as violating the standards of appropriate feminine behaviour. Anxiety would result from having to choose between behaviours they feel would result in their being viewed either as a competent supervisor, or as a likable person (Kern et al., 1985). Padgett (1988) found that the more managers wanted to be liked by their subordinates, the more likely they were to feel uncomfortable giving negative feedback for fear the employees would not like them. Additionally, as previously mentioned, the research of Bond and Anderson (1987), and of Deaux (1974) suggests the reluctance to deliver bad news is related to maintaining a good public image. Thus, in the current study, it is anticipated that female expressive supervisors would feel anxious when confronted with the requirement to give accurate face-to-face negative feedback, and they would employ strategies to reduce such distress. Delay and positive distortion of performance feedback constitute such anxiety reducing strategies. Male instrumental supervisors would not have to employ such tactics as there would be no such discrepancy between stereotypic social expectation and their behaviours. Androgynous supervisors would be able to avoid this conflict by using both their instrumental and their expressive skills.

Supervisor Appraisal and Effectance Motivation

According to Bandura's motivational model (1977, 1982), human behaviour can be explained and changes predicted in terms of motivational self-efficacy. The theory further

suggests that this expectancy can be subdivided into **self-efficacy expectancy** - the belief that one is or is not capable of learning and performing an action or group of actions, and **outcome expectancy** - the belief that a specific behaviour will or will not lead to a specified outcome. **Outcome value** (Teasdale, 1978) is a further addition to this theory, and suggests that the value an individual places on a specified outcome will also be an influential motivational component.

Research has clearly established that change in self-efficacy and changes in behaviour are highly correlated, and that self-efficacy is a substantive predictor of behaviour (for reviews see Bandura, 1986; Maddux & Stanley, 1986). The research on outcome expectancy is not as conclusive, nor as well defined (e.g., Maddux & Rogers, 1983; Maddux, Sherer & Rogers, 1982). Although the theoretical difference between self-efficacy expectancy and outcome expectancy seem logical, this clear differential is often lost in research, particularly in terms of accurately defining the outcome expectancy measure (e.g., Davis & Yates, 1982; Manning & Wright, 1983).

Little direct research has been done on outcome value (Teasdale, 1978), although its importance in the self-efficacy model is frequently implicitly inferred (Maddux, Norton & Stoltenberg, 1986). Most research in self-efficacy expectancy assumes that targeted behavioural changes are valued by the participants - a depressed person wishes to become less depressed, a smoker wishes to quit smoking. However, individual differences in outcome value may increase predictive accuracy.

Maddux, Norton and Stoltenberg (1986) manipulated self-efficacy expectancy, outcome expectancy, and outcome value to assess their relative impact on behavioural intentions. They predicted all three variables would be independent and approximately equivalent predictors of intentions. The subjects ($N=88$) were "Introduction to Psychology" students who were given experimental credit for their participation.

This experiment used a 2X2X2 factorial design, with high and low levels for each of the independent variables: self-efficacy expectancy, outcome expectancy and outcome value. These variables were manipulated by written communications presented as part of a series of brochures on interpersonal effectiveness training. The brochures describe the "broken-record" technique, the repetition of an assertive response without variation or qualification in the face of continuing

opposition (p.785). Arguments and "research" were cited to support the position taken in each brochure: high self-efficacy essay versus low self-efficacy expectancy essay, high outcome expectancy essay versus low outcome expectancy essay, or high outcome value essay versus low outcome value essay. Questionnaires were administered to determine the effectiveness of each manipulation, and four items were included to measure behavioural intentions (i.e., that subjects intended to use this technique). The high self efficacy essay argued that the technique is easy to learn and perform, whereas the low self-efficacy essay took the opposite position. The high outcome-expectancy essay stated that the broken record technique is highly effective, whereas the low outcome-expectancy essay suggested it is relatively ineffective. In the high outcome value condition it was argued use of the technique would substantially increase the chances of success in problematic interpersonal encounters; the low outcome value proposed that consistent use would probably lead to increased interpersonal friction.

Results suggested that self-efficacy expectancy, outcome-expectancy, and outcome value may be considered three independent and non-redundant variables. Surprisingly, outcome expectancy had a main effect on behavioural intentions, whereas self-efficacy expectancy did not. Maddux et al. (1986) caution that this study assessed behavioural intentions, as opposed to actual behaviours, and thus generalizability may be limited. Nonetheless, they propose that self-efficacy expectancy may not always be the best predictor of behavioural intentions, and that outcome expectancy should also be considered independently of self-efficacy.

In a theoretical paper, Leary and Atherton (1986) examined self-efficacy theory within the context of self-presentation theory, specifically social inhibition and anxiety with self-efficacy theory. They proposed two new categories of efficacy expectation based on the work of Schlenker and Leary (1982) that social anxiety is the result of people's concerns as to how they are perceived and evaluated by others. **Self-presentational efficacy** expectancy was proposed as the subjective probability of behaving in a manner intended to convey a particular impression. **Self-presentational outcome expectancy** was proposed as the probability of making a particular impression given the performance of a particular behaviour. If both expectancies are high, social anxiety should be low. If either or both of the expectancies are low, social anxiety should occur (given that the individual has some motivation to make a particular impression). "The

higher one's general self-efficacy and self-esteem, the more likely one is to feel capable of behaving in ways that will make the desired impression, and the less socially anxious one is likely to feel" (p. 259). Such a model of understanding could contribute to this present study, as supervisors would be expected to behave in a manner congruent with their self-presentational concerns, in some cases adhering to stereotypic behavioural expectations.

Other earlier studies had also examined the correlation between self-efficacy expectancy and social anxiety. Jennings (1985) used the General and Social Self-Efficacy Scales (Sherer, Maddux, Mercandante, Prentice-Dunn, Jacobs, & Rogers, 1982) to examine the correlations between social anxiety and general self-efficacy ($r = -.41$), and between social anxiety and social self-efficacy ($r = -.63$). These significant relations remained even after other variables that related to social anxiety were removed. Morris (1985) also found that social anxiety correlated negatively with the strength of the self-efficacy response within a wide range of social situations ($r = -.35$ to $-.65$). It appears that one's social anxiety in a given situation could have a profound effect on one's self-efficacy expectancy and behaviours.

In examining the relation between outcome expectancy and social anxiety Leary and Atherton (1986) felt that an individual might feel capable of performing a certain action, however they might still feel anxious because they believe the action will not achieve the desired effect. Some individuals are effective in most situations, but they may fail to perform with a certain population or under a certain set of conditions. Others may be predispositionally socially anxious, characterized generally by low social outcome expectancy. The combination of self-efficacy expectancy and outcome expectancy may have very different effects upon the resulting reaction.

Using data gathered from graduate students in education, Arch (1992) examined sex differences in **affective control efficacy** -- the belief that one can cope with one's own affective responses. She hypothesized that affective control efficacy would be a more important determinant for women than for men of willingness to engage in a possibly intimidating scenario. Both men and women reported feeling anxious under the situations described in the scenario, however the one variable that showed a significant sex difference was that of affective control efficacy. Thus, how anxious the females expected to feel in a given situation was not related

to their confidence of being able to perform the necessary task, but rather how well they expected to be able to handle those feelings of anxiety. Arch suggests that an important factor in women's deciding whether or not to participate in a given situation may be whether they believe they have the ability to handle their emotional responses (p. 8). If there is little assurance that they will be able to control emotions sufficiently to effectively deal with a situation, then there is a compelling reason not to participate. Thus, within the context of the current research, female supervisors who anticipate feeling emotionally overloaded or overwhelmed by behaving in a certain manner, will refrain from such actions even when they believe they have the abilities or skills required.

The self-efficacy model has implications for the career development and progression of both men and women, and it may become increasingly useful in understanding sex differences in career-related behaviours (Stikel and Bonet, 1991). Vasil (1993) examined sex differences in research productivity and self-efficacy beliefs among a sample of university academic staff. It was found that male academics were significantly more confident in their ability to perform research activities than were their female counterparts. These self-efficacy beliefs were positively correlated with productivity; for example, male academic members published significantly more articles in a three year period in refereed journals than females (4.9 vs. 2.3 respectively). Results also showed that a greater proportion of females than males did not hold doctoral degrees (44% vs. 20%). Vasil hypothesized that completion of the doctorate would enhance self-efficacy expectations with respect to research activities for these women. Other researchers (e.g., Landino and Owen, 1988; Schoen and Winocur, 1988) found similar results with respect to self-efficacy and sex differences in the performance of male and female academics.

Researchers have examined differences between degree of self-efficacy of females and males in female- and male- dominated occupations. Using university students, Betz and Hackett (1981) found that males reported equivalent self-efficacy for both female- and male-dominated jobs, whereas females reported significantly higher levels for the female-dominated than for the male-dominated positions. Matsui, Ikeda, and Ohnishi (1989) replicated these results with Japanese university students. Matsui (1994) expanded this research, examining the relation

between instrumentality and expressiveness, and self-efficacy in female- and male-dominated fields. Although results paralleled those previously mentioned, it was found that sex did not make a substantial contribution to predicting levels of self-efficacy for the two types of occupations. Rather, whereas self-efficacy for female-dominated occupations was significantly associated with both instrumentality and expressiveness, self-efficacy for the male-dominated occupations was associated with instrumentality only (p. 177). Although generalizability to this present study may be limited due to cultural differences, the suggestion remains that expressive supervisors may experience lower levels of self-efficacy expectancy than do their androgynous or instrumental counterparts.

Long (1989) examined women's sex-role orientation, self-efficacy, coping strategies, and stress in male- and female-dominated occupations. She found that, compared with low-masculine women, high-masculine women reported higher self-efficacy and greater problem-focused (as opposed to emotion-focused) coping strategies. Low-masculine women were significantly less self-efficacious regardless of occupation. As well, there was some support for the theory that differences in coping and efficacy can be a function of both feminine sex-role orientation and sex-typed occupation. Specifically, she found that for high-feminine women, self-efficacy and coping were not differentiated by occupational role. However, for low-feminine women, there was an effect; low-feminine women in nontraditional roles had greater self-efficacy and greater problem-focused coping than similar women in traditional jobs.

Singer (1989) examined 220 (114 male, 106 female) business and political science university students in their final year, to study gender differences in leadership aspiration within the framework of instrumentality-valence and self-efficacy. The instrumentality-valence model (e.g., Mitchell & Beach, 1976, cited in Singer, 1989) assumes that occupational choice is based upon both the perceived instrumentality of an occupation in providing certain outcomes (e.g., pay, social contacts, achievement fulfilment), and the perceived desirability (ie., valence) of such outcomes. Three self-efficacy expectancies were measured: 1) effectiveness -- "If you were in a leadership position, how effective would you be as a leader?", 2) ability-match -- "How well does your own ability fit requirements for leadership position?", and 3) ease of success -- "How easy would it be for you to succeed in a leadership position?".

Male students in this study had stronger leadership aspirations than their female counterparts. There were significant gender differences on all three measures of self-efficacy; males consistently expected themselves to be more effective as leaders, to have better leadership abilities, and to succeed more readily as leaders. Interestingly, males felt they were more likely to succeed if competence was the key requirement for effective leadership. Females, on the other hand, felt their success in leadership positions was contingent upon dispositional (e.g., intelligence) and situational (e.g., favourable organizational characteristics) factors. Such beliefs could affect the behaviours of supervisors required to provide performance evaluation feedback, with males generally feeling more prepared and capable of dealing with difficult or unpleasant situations (e.g., providing feedback to a poorly performing employee).

Napier and Latham (1986) examined the outcome expectancies of people who conduct performance appraisals. They hypothesized that appraisers who are high in self-efficacy may experience feelings of futility, because they felt that the primary consequence of giving employees positive or negative feedback was "nothing" -- a function of outcome expectancy. They further found that supervisors often feared the negative consequences of providing negative feedback, particularly aversive consequences for the appraiser. If the consequences of one's actions are perceived as neutral, then they should have little impact on subsequent actions (as behaviour was neither reinforced or punished); however, if there are negative outcomes to providing accurate feedback of poor performance, then valid appraisals are less likely to occur. Napier and Latham caution against a rigid interpretation of these results, as they are primarily descriptive and heuristic in nature, and they were based on small samples.

Leary and Atherton (1986) proposed that anxiety and defensiveness may be coeffects of self-efficacy expectancy. "Aversive experiences... create expectation of injury that can activate both fear and defensive conduct" (Bandura, 1977, p. 61). For those worried about how they might be perceived, passive interaction may minimize social risks (Leary, Knight & Johnson, 1986). These two concepts may be related in another way, as people often use anxiety as a cue for self-efficacy expectancy. One can end up with a circular scenario, low self-efficacy expectancy increases social anxiety, and anxiety decreases self-efficacy expectations by signalling to the individual that s/he is not in control and by cuing for low performance output (Leary &

Atherton, 1986).

Maddux, Norton and Leary (1988) tested the utility of integrating self-efficacy theory with anxiety, as understood in self-presentation theory. They also examined the individual predictive value of self-efficacy, outcome expectancy, and outcome value. Results supported the usefulness of distinguishing between **presentational self-efficacy expectancy** (the expectancy of executing a particular self-presentational behaviour) and **presentational outcome expectancy** (the expectancy that a behaviour will achieve ones' interpersonal goal). Self-efficacy emerged as the most powerful predictor of situation specific social anxiety. However, social anxiety was not associated with outcome value for the social goals used. The researchers suggested that social anxiety may be influenced less by the importance of a goal, than by a person's doubt in achieving a desired goal, and the value of the impression in attaining that goal.

Transposing to this context, all three motivational variables may play a role in determining whether or not supervisors of differing sex-types will provide prompt and accurate feedback to subordinates who perform poorly. Thus, if a female sex-typed supervisor does not feel that she has the ability to perform this task, it is unlikely much effort or many risks will be undertaken. Even if such supervisors do feel capable of providing prompt negative feedback but feel that such communication will generate negative interpersonal outcomes (e.g., employee grievances), they will be less likely to make the effort and to endure the outcome related anxiety.

Padgett (1988) found that the behaviour of managers in (negative) feedback sessions was influenced by the anticipated reactions from the employee. When managers expected a negative employee reaction and felt it would be difficult to motivate the employee in the future, they were less likely to provide accurate evaluation feedback. Thus, self-efficacy expectancy and outcome expectancy, separately and in combination, could play significant roles in determining how supervisors of varying sex-typing will provide negative feedback of varying accuracy and promptness.

-- Hypotheses and Rationale --

Self-Efficacy Expectancy

4) IT IS EXPECTED THAT IN COMPARISON TO EXPRESSIVE SUPERVISORS, INSTRUMENTAL AND ANDROGYNOUS (MALE AND FEMALE) SUPERVISORS WILL REPORT HIGHER LEVELS OF SELF-EFFICACY EXPECTANCY WITH RESPECT TO BOTH PROVIDING PROMPT AND ACCURATE FEEDBACK OF NEGATIVE EVALUATION RESULTS.

The provision of prompt and accurate negative feedback requires, at least in part, that a supervisor be assertive, independent, and willing to take a stand. In this study, by the operational definition (Bem, 1981), the instrumental individuals possess such characteristics -- sex-typed males, cross-sex-typed females, and both male and female androgynous supervisors . Expressive supervisors (sex-typed female supervisors and cross-sex-typed males) do not report such skills, and they therefore would be unlikely to feel as able or as ready to respond promptly and accurately in a negative feedback situation -- *self-efficacy expectancy*. If, within the context of their own gender identity, they feel unable to perform the actions necessary to achieve this end, they are more likely to use a distortion/delay strategy to make their supervisory behaviour congruent with the gender traits they possess.

Outcome Expectancy

5) IT IS EXPECTED THAT IN COMPARISON TO EXPRESSIVE SUPERVISORS, INSTRUMENTAL AND ANDROGYNOUS (FEMALE AND MALE) SUPERVISORS WILL REPORT HIGHER LEVELS OF OUTCOME EXPECTANCY WITH RESPECT TO PROVIDING BOTH PROMPT AND ACCURATE FEEDBACK OF NEGATIVE EVALUATION RESULTS.

It is expected there will logically be some discrepancy between supervisors of differing sex-types as to the *outcome expectancy* of giving prompt and accurate negative feedback. The consequences of prompt and accurate negative feedback will more likely be anticipated by expressive supervisors to lead to poor supervisory relationships with the subordinates involved, and, further, such individuals might feel unqualified or unwilling to deal with a subordinate's outbursts or anger. Thus, such supervisors would be less likely to give prompt and accurate feedback, believing that even if they do, it will not achieve the desired end (i.e, improved employee performance), or that the resulting cost: benefit ratio is too high.

In contrast, instrumental and androgynous supervisors would more likely feel that giving prompt and accurate feedback is necessary both for their position and for the future of the employee. An effective way to get the job done and to aid the employee in improving future performance is to provide timely and reliable feedback. Thus, regardless of the subordinate's reaction, they would proceed with communicating the negative results to the employee. In addition, by definition, androgynous supervisors (male and female) also possess the skills necessary to deal with such possible unpleasant interpersonal reactions, thereby they would report high expectancies of desirable outcomes.

Outcome Value

6) IT IS EXPECTED THAT IN COMPARISON TO EXPRESSIVE SUPERVISORS, INSTRUMENTAL AND ANDROGYNOUS (FEMALE AND MALE) SUPERVISORS WILL REPORT HIGHER LEVELS OF OUTCOME VALUE WITH RESPECT TO PROVIDING BOTH PROMPT AND ACCURATE FEEDBACK OF NEGATIVE EVALUATION RESULTS.

It is expected that expressive supervisors will differ from instrumental or androgynous supervisors with regard to their evaluation of the value and consequences of giving subordinates prompt, frank, and accurate information regarding the level of their work performance -- *outcome value*. Although they may agree that such feedback is necessary in the long term for improving employee effectiveness and productivity, expressive supervisors are likely to place

greater value on the short-term disruptive interpersonal consequences in the workplace. They are more likely to delay and/or distort negative feedback so as to avoid the anticipated detrimental effects of disrupted and/or negative interactions on office effectiveness and productivity.

Instrumental and androgynous male and female supervisors, in contrast, are more likely to focus on the positive and long-term results of providing timely and accurate feedback. First, it is part of being an effective supervisor; getting the job done, and done well, is highly valued. Second, the long-term cumulative benefits of maximizing workforce efficiency, so as to maintain the organization in a growth mode that will benefit supervisors and subordinate alike, is also highly valued.

PILOT RESEARCH

Pilot Project #1 -- Classification Agreement Between the Bem Sex Role Inventory-Short Form and the Personal Attributes Questionnaire

The Bem Sex Role Inventory (BSRI; Bem, 1974) and the Personal Attributes Questionnaire (PAQ; Spence, Helmreich & Stapp, 1975) are both extensively used in the role and gender identity research. A recent PYSCHINFO search revealed 653 citations (English and French publications only) for the BSRI and 221 for the PAQ. In order to help determine which questionnaire would be used in the final study, the following analyses were undertaken to examine the degree of agreement between sex-type classifications on the Bem Sex Role Inventory-Short Form (BSRI-S; Bem 1981) and on the PAQ.

Method

Volunteer subjects were recruited from undergraduate Social Psychology and Industrial Psychology classes ($N=103$; after deletions due to missing data and outliers, $N_{\text{final}}=95$). Students were asked to fill out a battery of tests including the BSRI-S (Bem 1981) and the PAQ (Spence, et al., 1975).

Subjects were matched to age, years work experience, and months supervisory experience. A correlation between the BSRI-S sex-type classification and the PAQ classification was calculated using the contingency coefficient.

RESULTS

Matching on Age, Work Experience and Supervisory Experience

Results of t -tests showed that there were no significant differences ($p > 0.10$) between males and females on age (males $M=23.1$, $SD=5.7$; females $M=23.0$, $SD=6.6$), years work experience (males $M=6.9$, $SD=5.8$; females $M=6.8$, $SD=5.0$), and months supervisory experience (males $M=11.9$, $SD=20.6$; females $M=18.1$, $SD=36.4$).

Classification Agreement: BSRI-S and PAQ

Correlation between the classification of subjects on the BSRI-S and on the PAQ also produced non-significant results (contingency coefficient = .299, $p > 0.10$). Crosstab results showed a marked discrepancy between classification using the BSRI-S and those made using the PAQ (see Table 1). Of the participants classified according to both sex-type scales, the PAQ classified 4 (4.2%) as feminine, 4 (4.2%) as masculine and 87 (91.6%) as androgynous; there were no undifferentiated individuals. The BSRI-S classified 19 (20.0%) as feminine sex-typed, 34 (35.8%) as androgynous, 23 (24.2%) as masculine sex-typed, and 19 (20.0%) as undifferentiated.

TABLE 1

CORRELATION AND COMPARISON OF CLASSIFICATIONS USING THE BSRI WITH THOSE USING THE PAQ

		PAQ Classifications		
		Feminine	Androgynous	Masculine
<u>BSRI</u> Classifications	Feminine Sex-Typed	2	17	19 20.0%
	Androgynous		31	3 34 35.8%
	Masculine Sex-Typed		22	1 23 24.2%
	Undifferentiated	2	17	19 20.0%
		4 4.2%	87 91.6%	4 4.2%
				95 100%

Contingency Coefficient = 0.299, $p > 0.10$

DISCUSSION

The low correlation between the classification of subjects on the BSRI-S and on the PAQ is consistent with previous research (e.g., Cunningham & Antill, 1980; Day & Korabik, 1991; Gaa & Liberman, 1981). The PAQ also showed little variability in classification, classifying 91.6% of the subjects as androgynous. This suggests that, at least for this group, the BSRI-S seemed the more effective measure.

Pilot Project #2 -- Factor Analytic Examination of the
Bem Sex Role Inventory-Short Form

The Bem Sex Role Inventory (BSRI; Bem, 1974) has been extensively examined and criticized. One area of particular interest has been the validity of the BSRI masculine and feminine scales. Many researchers (e.g., Bledsoe, 1983; Gaudreau, 1977; Pedhauzer & Tetenbaum, 1979; Waters et al., 1977) have examined the BSRI using factor analysis to examine the construct validity of these scales; items on each scale should correlate higher with their own scale than with any other.

Gaudreau (1977) performed factor analysis on the intercorrelations of 64 variables: the 60 items on the BSRI, the gender of the subject, a masculine score, a feminine score, and an androgynous score. Four main factors were found: Factor I - called the masculinity factor, Factor 2 - called the femininity score, Factor 3, and Factor 4 - called the maturity factor. The masculinity factor was comprised of 17/20 masculine BSRI items (masculine, self-reliant and athletic were missing), two feminine BSRI items and the masculine score. The feminine factor was made up of 13/20 of the BSRI items (missing were feminine, child-like, does not use harsh language, flatterable, gullible, shy and soft spoken), the femininity score, and six adjectives from the filler scale. Four items loaded onto Factor 3: feminine, masculine, athletic and the actual gender of the subject. The maturity factor was made up of a few items from each scale: reliable, self-sufficient, self-reliant, and sincere loaded positively, flatterable, gullible, child-like, and inefficient loaded negatively. Most traits which loaded on this last factor did not load significantly on either the masculine or feminine factors.

Gaudreau (1977) maintained that these results suggest that as most of the items do load on common factors, they support two dimensions for masculinity and femininity rather than bipolar theory. However, he suggested changes on the characteristics for the masculine and feminine scales would reduce the multidimensionality of the test. It was suggested that the three masculine items missing from Factor I be eliminated, along with self-sufficiency and analytical

(item loadings of .31 and .32 respectively). The feminine items missing from Factor 2 were to be eliminated and the three filler adjectives with the highest loadings (helpful, friendly and tactful) were to be added to this scale.

Waters et al. (1977) factor analyzed the 40 masculine and feminine items on the BSRI ($N=252$; half males, half female). Their results were consistent with those of the Gaudreau (1977) study. They too found four factors: one with loading pertaining to the biological sex of the subject, one with predominantly feminine traits, and two factors with predominantly masculine traits. Waters et al. agreed that the results support the idea of masculinity and femininity as independent constructs. They too felt that certain items should be eliminated to increase the interpretability and homogeneity of the scales. The studies agreed on which items should be eliminated, with the exceptions of competent (Waters et al., 1977), and self-sufficient and self-reliant (Gaudreau, 1977).

In 1986 Waters and Popovitch replicated the 1977 study to determine the reliability of the factors over time. The same four factors emerged; those two corresponding to masculine traits were virtually identical as was the factor relating to the sex of the respondent. Slight variations were found in the expressive/affective factor, with 9/14 of the previous items loading with values of at least .40. The authors felt that "the stability of the four factor solutions indicate that these factors are not anomalies or item specific and should be taken into account when using the scale" (p. 1326).

Moreland, Gulanick, Montague and Harren (1978) conducted two studies on the BSRI. Study I evaluated a factor sample comparative to the Gaudreau (1977) study. Study II examined the internal consistency of a factor analyzed derived scale compared to the original BSRI scale. The results of Study I were almost identical to the two studies cited above, both in the number of factors and in the items which loaded on these factors.

In Study II ($N=600$; 314 males, 286 females) Moreland et al. calculated the correlations between the response to each item and the original BSRI scales, and with the scales found in their first study. Of the feminine BSRI items (.76), 14/20 correlated more with the feminine scales than with others, 3/20 correlated more with social desirability, and 2/20 correlated negatively with masculinity. Of the masculine items (.86) 19/20 most highly correlated with the

masculine scale, and one negatively correlated with the feminine scale. The social desirability items were the least internally consistent (.68), with only 11/20 items being more correlated with this scale than with others. The correlations with the factors found in Study I were much higher: for the emotional/expressive items (.89) and for the instrumental items (.86). In each case 16/16 items correlated most highly with their own scale.

Moreland et al. (1978) suggested that these results concur with the previous studies that masculinity and femininity are two separate continua. They also suggest that, as these results substantiate previous factor analytic studies (e.g., Gaudreau, 1977), the non-loading characteristics should be eliminated to increase the homogeneity of the scales.

Gruber and Powers (1982) dispute the suggestion that items can be deleted from the BSRI without loss. Firstly, all of the studies cited included items not included on the BSRI itself. All included the biological gender of the subject, and Gaudreau (1977) also included masculinity, femininity and androgyny scores. Gruber and Powers argue that, as a result, the resulting patterns are not a clean indication of the relations among the items on the inventory. The introduction of variables not found on the BSRI itself might alter the factor pattern as items not on the BSRI might rate highly with those that are. They further argue that although the elimination of certain characteristics would reduce multidimensionality, "concerns about the multidimensionality of the items are of little practical use if the power of the instrument is appreciably reduced" (p.285).

They administered the BSRI to 1117 subjects (224 males, 893 females), who were classified into four categories using the median split technique. Gruber and Powers used discriminant step-wise analysis to quantify the contribution that the items in question made to the sex-type classification. They compared the original sex-type with the discriminant function's predicted sex-type. There was 88.99% agreement between the classifications made by the BSRI and those made by the discriminant function. When the items suggested by Waters et al. (1977) were eliminated, the agreement dropped to 80.75 %, with those suggested by Moreland et al. (1978) to 76.01%, and with those by Gaudreau (1977) to 75.92%.

Gruber and Powers (1982) went on to examine the level of significance of each variable in classification. With the exception of the trait "feminine", all of the suggested eliminees were

significantly discriminant in the "all items" analysis. These results suggest that there could be significant loss in the discriminatory abilities of the BSRI if any of the items are deleted.

Bem (1979) acknowledges the multidimensionality of the scales, but she feels that the elimination of the items solely to guarantee unidimensionality would defeat the purpose of the test.

"The culture has arbitrarily clustered together heterogeneous collections of attributes into the two categories prescribed as more desirable for one sex or the other... If the culture groups a hodge podge of attributes...then that (is) what sex-types individual will take as the standard for their behaviour. The purpose of the BSRI is to discriminate between those individuals for whom this hodge podge does form a unitary cluster and those for whom it does not" (p.1049).

Far from eliminating the BSRI from use in sex-typing research, many of the concerns raised resulted in Bem's development of the short form of the BSRI (BSRI-S). The internal consistency and homogeneity of this test, comprised of half of the original items, have been increased by the elimination of certain traits from the scales. For example, masculine, feminine and athletic, terms which loaded with the biological gender of the subject, have been eliminated. The results of the BSRI-S correlate strongly with those of the original BSRI (Bem, 1981).

This study examines the factor analytic structure of the BSRI-S using both exploratory and confirmatory procedures.

Method

Volunteer subjects were recruited from undergraduate Psychology and Industrial Psychology classes ($N=280$). Data were collected over two consecutive years, 177 the first year and 103 subjects the second year. All participants (after deletions due to missing data and outliers, N final=257; 50 males, 207 females) were asked to fill out a battery of tests including the BSRI-S.

In order to determine the clearly interpretable solution, a series of exploratory factor analyses (EFA) were performed on the 30 items of the BSRI-S (see Table 2). Four solutions are consistently found in the research on the BSRI (long form), so factor solutions were generated for five through two factor models. Oblique rotation was chosen as the factor correlation matrix showed numerous moderate correlations between the variables.

Confirmatory Factor Analysis (using LISREL) was used to test the hypothesized two-factor structure of the Feminine and Masculine subscales of the BSRI-S. The 10 filler items of the BSRI-S were dropped for the CFA procedures as theoretically they are not intended to load on a common factor (supported by the EFA results), and their inclusion would have reduced the degree of fit between the hypothesized and generated models.

RESULTS AND DISCUSSION

The assumptions of multivariate normality, outliers, homogeneity of variance-covariance matrices, linearity and singularity or multicollinearity, and missing data were evaluated and the necessary adjustments made.

Exploratory Factor Analyses

The factor pattern matrix was examined, and loadings $\geq .30$ were considered and interpreted (Tabachnik & Fidell, 1989). The factor models were evaluated and compared using the chi-square statistic, the ratio of chi-squared to degrees of freedom, the percentage of variance explained, the number of iterations required, and the percentage of residual $\geq .05$. The eight-factor solution generated by the computer as well as the five, four, and two-factor models were rejected in favour of the three-factor solution.

In comparison to previous research, a clear solution was obtained for a three-factor solution (see Table 3.) All of the feminine items loaded on the Feminine Factor, and nine of the ten masculine items loaded on the Masculine Factor. However, one of the masculine items

TABLE 2

ITEMS ON THE BEM SEX ROLE INVENTORY - SHORT FORM

- | | |
|--|-------------------------------------|
| ● 1. Defends own belief | ◆ 16. Understanding |
| 2. Moody | 17. Secretive |
| ● 3. Independent | ◆ 18. Compassionate |
| 4. Conscientious | ◆ 19. Eager to soothe hurt feelings |
| ◆ 5. Affectionate | 20. Conceited |
| ● 6. Assertive | ● 21. Dominant |
| ● 7. Strong Personality | ◆ 22. Warm |
| ● 8. Forceful | ● 23. Willing to take a stand |
| 9. Reliable | ◆ 24. Tender |
| ◆ 10. Sympathetic | ● 25. Aggressive |
| 11. Jealous | 26. Adaptable |
| ● 12. Has leadership abilities | ◆ 27. Loves children |
| ◆ 13. Sensitive to the needs of others | 28. Tactful |
| 14. Truthful | ◆ 29. Gentle |
| ● 15. Willing to take risks | 30. Conventional |

- ◆ = Items used for calculating Feminine Raw Score
● = Items used for calculating Masculine Raw Score

TABLE 3

ITEMS LOADING FROM THE PATTERN MATRIX FOR THE THREE-FACTOR SOLUTION

Item	Factor*		
	1.	2.	3.
18. Compassionate	.78		
24. Tender	.76		
13. Sensitive to Needs of Others	.74		
10. Sympathetic	.73		
22. Warm	.72		
29. Gentle	.72		
5. Affectionate	.70		
19. Eager to Soothe Hurt Feelings	.62		
16. Understanding	.60		
27. Loves Children	.35		
28. Tactful			
30. Conventional			
25. Aggressive		.74	
21. Dominant		.70	
6. Assertive		.69	
8. Forceful		.66	
7. Strong Personality		.65	
12. Has Leadership abilities		.64	
23. Willing to Take a Stand		.56	
15. Willing to Take Risks		.55	
1. Defends own Beliefs		.35	
9. Reliable			-.56
4. Conscientious			-.51
14. Truthful			-.50
11. Jealous			.46
3. Independent			-.42
20. Conceited		.30	.41
2. Moody			
17. Secretive			
26. Adaptable			

NOTE: Only loading $\geq .30$ were interpreted and reported

- * Factor 1 - Feminine Factor
- Factor 2 - Masculine Factor
- Factor 3 - Composed primarily of items from filler questions

(Item 3, "independent") did not load on the Masculine Factor, loading instead on Factor Three, a factor comprised of six of the filler items. This suggests that Item 3 does not contribute to the discriminability of the Masculine Factor, and that perhaps another trait should be considered as replacement.

Six of the filler items grouped together on a third factor, Factor 3. Bem (1974) originally intended the 10 filler items to be used as a measure of Social Desirability. However, this was subsequently shown to be an unreliable measure, and the items were retained for filler and distraction purposes only. It is therefore not surprising that six of the filler items should load significantly on the same factor. One of the filler items (Item 20, "conceited") cross-loaded, loading significantly on both the Masculine Factor and Factor Three. It should also be noted that "conceited" loaded more strongly on Factor Three (-.42) than on the Masculine Factor, there barely reaching the minimum loading requirement (.30).

Confirmatory Factor Analysis

Results of the CFA suggest that the hypothesized factorial model of the feminine and masculine subscales is minimally adequate. Based on the standardized solution, all items except one loaded on the proposed factor with values $\geq .30$. The exception was Item 3 ("independent"), the same item which failed to load significantly on the Masculine Factor in the EFA (see Figure 1).

Indices used to assess goodness-of-fit suggest that this model could still be improved (see Table 4). For the GFI, AGFI, CFI, and TLI, values $> .90$ are considered indicative of a good fit. Thus, the values in the case, 0.84, 0.81, 0.77, and 0.85 respectively, suggest a less than ideal fit between the hypothesized model and the observed data. The RMR represents the average discrepancy between items in the sample and in the hypothesized model; values $< .05$ are considered acceptable. The RMR in this study (.08) further suggests a problem of fit. The T -values for the items were all > 2.00 , suggesting that all are important to the model. Item 3 ("independent") had the lowest t -value (3.43).

FIGURE 1

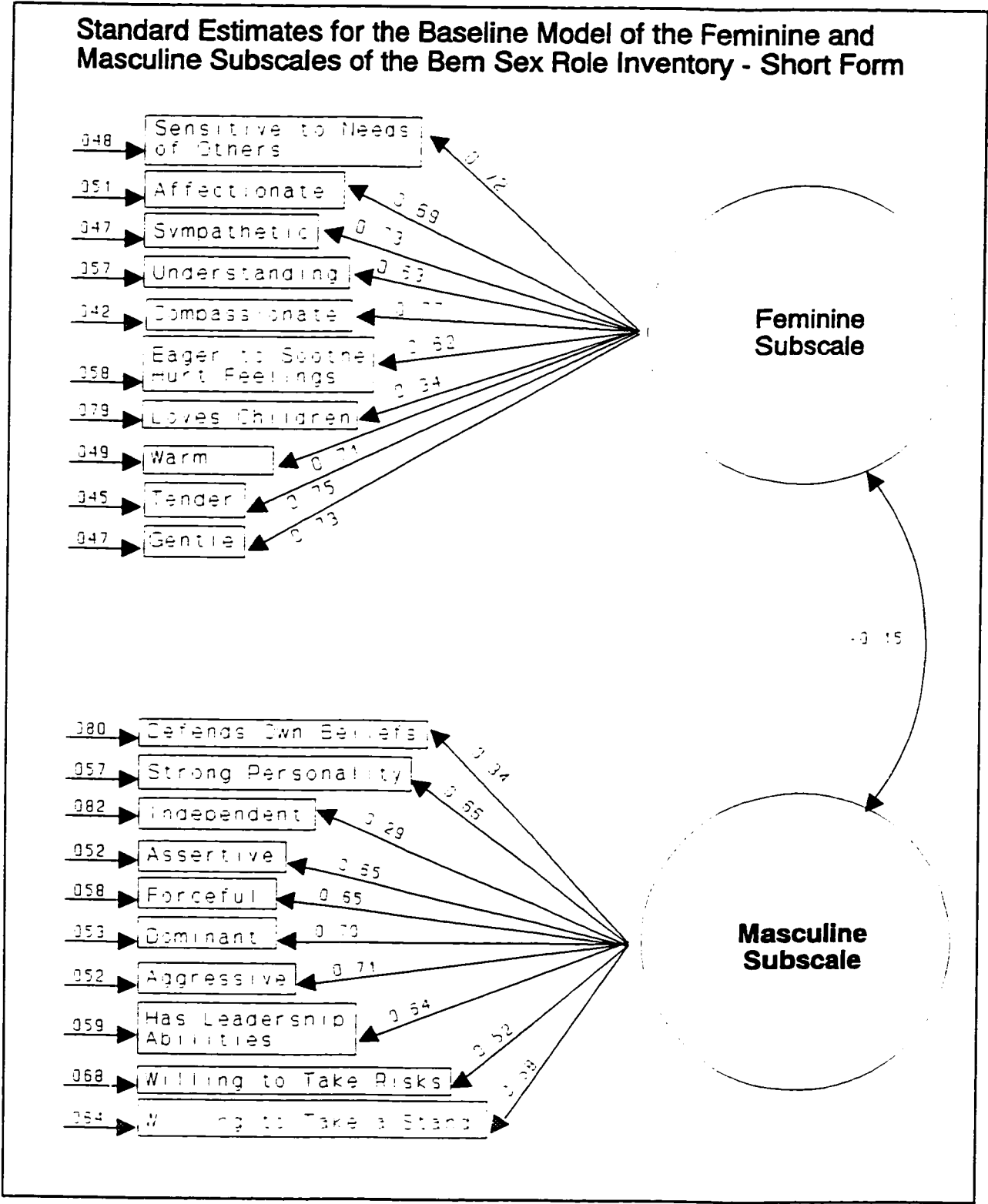


TABLE 4

SUMMARY OF INDICES USED TO ASSESS GOODNESS-OF-FIT

χ^2	435.8 (df = 169, $p < .001$)
GFI ^a	0.84
AGFI ^b	0.81
CFI ^c	0.77
TLI ^d	0.85
RMR ^e	0.08

^a Goodness-of-Fit Index

^b Adjusted Goodness-of-Fit Index

^c Comparative Fit Index

^d Tucker-Lewis Index

^e Root Mean Square Residual

The BSRI-S appears to be a improvement over the original long version of this measure, at least in terms of the factor structure. However, although the hypothesized model appears to adequately reflect that of the observed data, CFA results suggest that this continues to be a less than ideal model. The BSRI-S will be used in this study to determine sex-typing, but further research is required to clarify or to find alternative constructs under-lying gender.

Pilot Project #3 -- The Impact of Sex and Sex-typing on
Motivational Expectancies, and on the Delay and Distortion
of Negative Performance Appraisal Feedback

This pilot study was conducted to test the planned procedure to follow-up the research of Benedict and Levine (1988), who found that when faced with the requirement to deliver negative feedback to subordinates, female supervisors delay and distort more than do male supervisors.

Hypotheses

Within the context of the current study, the overriding assumption is that it is not one's sex that influences how a person reacts, but rather one's gender or sex-typing, the characteristics or traits one possesses and can thus draw upon to cope with a given situation -- in this instance, one's gender or sex-type which contain expressive or instrumental skills and resources.

- i) IN COMPARISON WITH EXPRESSIVE SUPERVISORS, IT IS EXPECTED THAT INSTRUMENTAL SUPERVISORS WILL BE LESS LIKELY TO DELAY AND DISTORT THE PROVISION OF NEGATIVE FEEDBACK.
- ii) IN COMPARISON WITH EXPRESSIVE SUPERVISORS, IT IS EXPECTED THAT INSTRUMENTAL SUPERVISORS WILL DEMONSTRATE HIGHER SELF-EFFICACY MOTIVATION REGARDING THE DELIVERY OF NEGATIVE FEEDBACK.

Method

Subjects and Recruitment

Subjects were recruited from university MBA and MPA courses with the requirement that they have at least six months supervisory experience. All participants were asked to complete the short form of the Bem Sex Role Inventory (BSRI-S; Bem, 1981), and to provide demographic data -- sex, age, years work experience, and months supervisory experience. It was explained that we were attempting to establish a profile on individuals pursuing post-graduate studies in business. All students were further asked to volunteer an hour of their time to help with an additional research project, establishing the validity and reliability of a computer science aptitude test. It was explained that this would involve the administration and scoring of the test and the completion of post-test questionnaires.

One hundred and thirty-eight students completed the BSRI-S, 61 males and 77 females, mean age=36.5 years (SD=6.7). Of that number, twenty-two volunteers participated in the follow-up study, 13 males and 9 females, mean age=32.2 years.

Measures

Independent Variables --

Sex: The subject's sex was included to compare with the results of Benedict and Levine (1988).

Sex-Type: Sex-type was assessed using the BSRI-S (Bem, 1981). See Appendix 1.

Dependent Variables --

Supervisor distortion of subordinate performance.

The evaluative judgement of the volunteer supervisors was assessed using three scales on the Supervisory Performance Evaluation Questionnaire (see Appendix 3), derived from the work of Benedict and Levine (1988):

- 1) "Please check the point on the scale which best describes the performance of the individual whose work you supervised."
- 2) "Please rate the overall quality of the work performed by the individual whom you supervised on the following scale."
- 3) "Based on your experience with the work performed by the individual whom you supervised, please circle your response to the following statement: This individual is capable of performing computer programming tasks."

As the actual performance of the confederate was constantly at the 25th percentile, the distortion score was calculated as the difference between the rating given and the 25th percentile.

Supervisor delay: Ending performance trials.

This was the first of two delay measures. The volunteer supervisors were told that they could stop testing when they felt they had sufficient information to make an accurate assessment of the confederate volunteer's abilities (minimum 1 trial, maximum 10 trials). The confederate subordinate's performance remained at the 25th percentile over repeated trials. It was felt that the more trials administered, the greater the hesitation on the part of the volunteer supervisor to accept these results as the confederate volunteer's maximum level of performance.

Supervisor delay: Scheduling the feedback interview.

This was the second of the delay measures. Delay was assessed using the number of days between the testing session and the date scheduled for face-to-face feedback. Holidays, weekends, and other reasonable delays (e.g., being out of town) were subtracted from the total number.

Appraisal Effectance Motivation.

The volunteer supervisors' motivation to provide prompt and accurate feedback was assessed using the Appraisal Effectance Motivation Questionnaire (see Appendix 3), derived from the work of Bandura (1977, 1982), and Teasdale (1978). This questionnaire has eight items; three measure self-efficacy expectancy, two measure outcome expectancy, and three measure outcome value. Each item is answered by indicating one's position on a 14-point scale, ranging from strongly disagree (1) to strongly agree (14).

Control Factors

Subordinate sex and ethnicity.

In an effort to eliminate possible confounding factors due to the volunteers' sex, or to ethnic interactions, all programmer candidates were male Caucasians.

Supervisor-subordinate interpersonal bonding

In the work place, the supervisor and supervisee know one another, and, in most cases, must continue to work together after the evaluation feedback. Such conditions do not exist in the experimental situation. In order for a bond to form between the supervisor and the confederate, a five to ten minute opportunity was scheduled to allow them to get to know each

other briefly. The supervisor-supervisee relationship was evaluated post-test to determine if they had been acquainted prior to the experiment, or if the volunteer supervisor believed the confederate had tried his best and was interested and invested in the test results.

Procedure

Students who had volunteered to participate in the "Computer Aptitude Study" were contacted by telephone and given an appointment. Upon his/her arrival, a member of the research team introduced the volunteer supervisor (hereafter referred to as the supervisor) to the confederate volunteer (hereafter referred to as the confederate) as the programmer candidate. The supervisors were told that the programmer candidates were students from the Computer Science Department who had volunteered to participate in this study. The procedure was then explained (see Appendix 2 for the "Supervisory Control Worksheet, the "Computer Science Aptitude Test Norms", the "Computer Science Aptitude Test", and the "Computer Science Aptitude Test Sheets A & B"). Particular attention was paid to explaining the "normative data", expressed in percentiles, as the confederates would consistently be performing at the 25th percentile, and it was important that the supervisors realize this represented a substandard level of performance. It was explained to the confederate, for the benefit of the supervisor, that each time test trial lasted three minutes, and that he was being marked on speed and accuracy (i.e., his score was the total correct). Finally, it was emphasized that the researcher was there only to oversee the procedure and answer questions, and that the decisions would be left up to the supervisor. As well, they were asked to complete the Supervisor Qualifications Sheet (see Appendix 2).

Following this introduction, the team member left the room for 5-10 minutes (on the pretext of making an important phone call), and she requested that they not start prior to her return. During this period the confederate attempted to establish rapport with the supervisor, emphasizing that he really wanted to do well and that, although this was an experimental measure, he hoped to use the results in the future (e.g., Co-op applications).

Upon her return, the team member apologized for the delay, and she indicated that they

could proceed with the testing. When the volunteer supervisor felt that s/he had sufficient information, s/he called a halt to the testing, the confederate volunteer was thanked, and he was told that he would be contacted later with the results.

After the confederate's departure, the supervisor was asked to set a date to provide face-to-face feedback to the confederate volunteer. The supervisors were deliberately informed of this requirement prior to completing the post-test questionnaires to increase the generalizability to the work force. The supervisors were then asked to complete the Supervisory Performance Evaluation sheet, the Appraisal Effectance Motivation Questionnaire, and the Post-Evaluation Questionnaire (see Appendix 3). The procedure was completed by re-introducing the volunteer confederate as a team member, and debriefing the supervisor as to the true purpose of the study.

Results

Subjects

A total of 22 students volunteered to participate in this study (13 females, 9 males). Male and females subject were matched for months supervisory experience (males - $\bar{M}=30.4$, $SD=23.4$; females - $\bar{M}=41.1$, $SD=29.0$), but there were significant differences in age, $t(20)=-2.46$, $p=.03$ males - $\bar{M}=29.4$, $SD=6.1$; females - $\bar{M}=36.3$, $SD=7.1$). Subjects were equally divided in terms of training in the use of performance evaluations, 11 with training, 11 without.

BSRI-S Norms

Norms for masculine raw score (MRS) and feminine raw score (FRS) of the BSRI-S were calculated on the basis of 138 questionnaires completed by MBA and MPA students. The data for the MRS and FRS scores were normally distributed, and thus the means and medians for these scores were essentially the same: MRS -- $\bar{M}=4.9$, $SD=0.76$, $Mdn=5.0$; FRS -- $\bar{M}=5.6$, $SD=.78$, $Mdn=5.6$. As Bem (1981) has presented the median in other normative samples, it will be used in this study. Using these norms, subjects for the evaluation study were classified

according to their sex-type: males -- 1 expressive, 3 instrumental, 3 androgynous, and 1 undifferentiated; females -- 3 expressive, 3 instrumental, 5 androgynous, and 2 undifferentiated.

Control Factors

Post-test questionnaires revealed that >96% of the volunteer supervisors indicated no personal knowledge of the confederates. As the confederates denied any prior personal interactions with the supervisors, it is possible that the few positive responses resulted from the information exchanged in the bonding session. (The post-test questionnaire did not specify *prior* personal knowledge.) Nearly 92% of the supervisors felt the confederates had tried their best, and >96% felt the confederates were interested in the test results.

Delay and Distortion Variables

Contrary to the work of Benedict and Levine (1988), MANOVA results revealed non-significant differences between the sexes on the two delay variables and the three distortion variables. Follow-up exploratory *t*-tests corroborated these results (see Table 5).

MANOVA results also showed non-significant differences between sex-types on the delay and distortion variables. *T*-tests, supported these results. As expected, *t*-tests comparing instrumental (both masculine sex-typed and androgynous) male and female supervisors revealed no significant differences. Nonetheless, a considerable difference was seen on the "overall performance" score, with instrumental males unexpectedly distorting more than instrumental females (see Table 6.) Contrary to expectations, there were no significant differences between expressive female and instrumental female supervisors (see Table 7). Despite this lack of significance, there were substantial mean differences between these two groups of female supervisors on all three distortion variables. A larger sample size, with increased power, might have yielded significance.

TABLE 5

**PILOT STUDY - DELAY AND DISTORTION VALUES FOR MALE AND FEMALE
VOLUNTEER SUPERVISORS**

VARIABLES	MEAN - MALES (<u>n</u> = 9)	MEAN - FEMALES (<u>n</u> = 13)	t - VALUE	p - VALUE
<u>Delay #1</u>				
# of Tests Administered	4.6 (<u>SD</u> = 1.2)	4.0 (<u>SD</u> = 1.2)	1.2	2.5
<u>Delay #2</u>				
# Days to Feedback Session	1.9 (<u>SD</u> = 1.9)	0.8 (<u>SD</u> = 10)	1.53	.16
<u>Distortion #1</u>				
Overall Performance	23.9 (<u>SD</u> = 27.4)	10.7 (<u>SD</u> = 8.3)	1.7	.11
<u>Distortion #2</u>				
Quality of Performance	23.3 (<u>SD</u> = 28.3)	31.2 (<u>SD</u> = 21.5)	-.73	.48
<u>Distortion #3</u>				
Ability to Perform Computer Programming Tasks	60.0 (<u>SD</u> = 13.9)	47.4 (<u>SD</u> = 15.3)	1.9	.07

TABLE 6

**PILOT STUDY - DELAY AND DISTORTION VALUE FOR INSTRUMENTAL*
MALE AND INSTRUMENTAL* FEMALE VOLUNTEER
SUPERVISORS**

VARIABLES	MEAN - MALES (<u>n</u> = 6)	MEAN - FEMALES (<u>n</u> = 8)	t - VALUE	p - VALUE
<u>Delay #1</u>				
# of Tests Administered	4.4 (SD = 1.3)	3.75 (SD = 1.2)	.93	.38
<u>Delay #2</u>				
# Days to Feedback Session	1.8 (SD = 1.6)	.75 (SD = 1.0)	1.4	.18
<u>Distortion #1</u>				
Overall Performance	31.6 (SD = 32.2)	9.0 (SD = 7.8)	1.5	.19
<u>Distortion #2</u>				
Quality of Performance	17.6 (SD = 26.0)	27.7 (SD = 24.6)	-.71	.5
<u>Distortion #3</u>				
Ability to Perform Computer Programming Tasks	59.2 (SD = 11.8)	50.0 (SD = 14.8)	1.2	.27

* Masculine sex-typed and androgynous combined

TABLE 7

PILOT STUDY - DELAY AND DISTORTION VALUES FOR INSTRUMENTAL VS EXPRESSIVE FEMALE VOLUNTEER SUPERVISORS

VARIABLES	MEAN - INSTRUMENTAL (<u>n</u> = 3)	MEAN - EXPRESSIVE (<u>n</u> = 3)	t - VALUE	p - VALUE
<u>Delay #1</u>				
# of Tests Administered	4.0 (<u>SD</u> = 2.0)	3.7 (<u>SD</u> = .58)	-.28	.8
<u>Delay #2</u>				
# Days to Feedback Session	.67 (<u>SD</u> = 0.6)	.67 (<u>SD</u> = 1.2)	0	1.0
<u>Distortion #1</u>				
Overall Performance	3.66 (<u>SD</u> = 2.8)	7.7 (<u>SD</u> = 6.2)	1.0	.36
<u>Distortion #2</u>				
Quality of Performance	25.4 (<u>SD</u> = 33.8)	38.8 (<u>SD</u> = 20.9)	.58	.6
<u>Distortion #3</u>				
Ability to Perform Computer Programming Tasks	38.7 (<u>SD</u> = 19)	41.3 (<u>SD</u> = 15.1)	.19	.86

Appraisal Effectance Motivation

A 2 X 4 factorial MANOVA revealed non-significant results for both sex and gender on the three motivation variables. Subsequent exploratory t -tests confirmed these results, revealing non-significant differences between male and female supervisors, and between expressive vs. instrumental female supervisors (see Tables 8 & 9). However, there was a marginally significant effect for outcome value, $t(4)=2.5$, $p=.07$, with expressive female supervisors placing greater value on the value of prompt and accurate feedback than their instrumental counterparts. Comparisons of instrumental (both masculine sex-typed and androgynous) male and female supervisors revealed a significant difference on outcome expectancy, $t(8.76)=3.21$, $p=.01$, with instrumental male supervisors having higher outcome expectancy for the provision of prompt and accurate feedback than their female counterparts (see Table 10).

TABLE 8

PILOT STUDY - MOTIVATION VALUES FOR MALE AND FEMALE VOLUNTEER SUPERVISORS

VARIABLES	MEAN - MALES ($n = 9$)	MEAN - FEMALES ($n = 13$)	t - VALUE	p - VALUE
Self-Efficacy Expectancy	5.9 ($SD = 2.2$)	4.4 ($SD = 3.0$)	1.3	.23
Outcome Expectancy	10.9 ($SD = 2.9$)	8.1 ($SD = 3.4$)	1.7	.10
Outcome Value	9.6 ($SD = 4.0$)	11.0 ($SD = 1.7$)	-.83	.43

TABLE 9

**PILOT STUDY - MOTIVATION VALUES FOR INSTRUMENTAL VS EXPRESSIVE
SEX-TYPED FEMALE VOLUNTEER SUPERVISORS**

VARIABLES	MEAN - INSTRUMENTAL (<u>n</u> = 3)	MEAN - EXPRESSIVE (<u>n</u> = 3)	t - VALUE	p - VALUE
Self-Efficacy Expectancy	5.0 (SD = 3.2)	2.3 (SD = .7)	-1.42	.23
Outcome Expectancy	10.8 (SD = 1.4)	8.8 (SD = 2.8)	-1.11	.33
Outcome Value	11.4 (SD = .2)	12.0 (SD = .3)	2.5	.07

TABLE 10

**PILOT STUDY - MOTIVATION VALUES FOR INSTRUMENTAL* MALE AND
INSTRUMENTAL* FEMALE VOLUNTEER SUPERVISORS**

VARIABLES	MEAN - MALES (<u>n</u> = 6)	MEAN - FEMALES (<u>n</u> = 8)	t - VALUE	p - VALUE
Self-Efficacy Expectancy	5.7 (SD = 2.3)	5.2 (SD = 3.4)	.3	.8
Outcome Expectancy	12.4 (SD = 1.1)	7.7 (SD = 3.9)	3.2	.01
Outcome Value	12.1 (SD = 1.4)	10.5 (SD = 1.8)	1.7	.11

* Masculine sex-typed and androgynous combined

Training

As the volunteers were divided equally between those who had and had not received some training in performance appraisals, it was decided to conduct some preliminary analyses to determine if this factor might impact on differences on the dependent variables (delay, distortion, and effectance motivation). T-tests comparing trained and untrained volunteers revealed only one significant difference, the number of days to the scheduled feedback session, $t(11.75)=2.89$, $p=.01$. Thus, those volunteers who had received no training in performance appraisal scheduled feedback sessions sooner than did their trained counterparts (see Table 11.)

TABLE 11

PILOT STUDY - DELAY, DISTORTION AND EFFECTANCE MOTIVATION VALUES FOR TRAINED VS UNTRAINED VOLUNTEER SUPERVISORS

VARIABLES	MEAN - TRAINED (<i>n</i> = 11)	MEAN - UNTRAINED (<i>n</i> = 11)	t - VALUE	p - VALUE
<u>Delay #1</u>				
# of Tests Administered	4.3 (SD = 1.0)	4.1 (SD = 1.3)	.55	.6
<u>Delay #2</u>				
# Days to Feedback Session	1.9 (SD = 1.7)	.4 (SD = .5)	2.89	.01
<u>Distortion #1</u>				
Overall Performance	22.6 (SD = 19.4)	21.8 (SD = 14.2)	.11	.91
<u>Distortion #2</u>				
Quality of Performance	35.5 (SD = 26)	42.0 (SD = 18.8)	-.68	.51
<u>Distortion #5</u>				
Ability to Perform Computer Programming Tasks	14.1 (SD = 4.0)	11.6 (SD = 3.7)	1.51	.15
Self-Efficacy Expectancy	5.1 (SD = 2.7)	4.7 (SD = 2.8)	.31	.76
Outcome Expectancy	10.0 (SD = 3.6)	8.3 (SD = 3.6)	1.12	.28
Outcome Value	9.9 (SD = 3.5)	11.1 (SD = 1.4)	-1.04	.32

Discussion

This study had several goals: to evaluate the experimental procedure as outlined by Benedict and Levine (1988), to begin to examine the effects of gender socialization or sex-typing on the delay and distortion variables, and to begin to examine how effectance motivation factors may impact on the delivery of prompt and accurate negative feedback.

This pilot project uncovered no major procedural obstacles. It was decided to improve the Effectance Motivation Questionnaire, listing questions about prompt feedback separately from those concerning accurate feedback, to avoid possible confounding. As well, it was determined that this questionnaire should be the last questionnaire given to the volunteers, as this one raised most suspicion as to the "real" reason or nature of the experiment. Two additional post-test questionnaires were to be added. One was to evaluate factors proposed by Benedict and Levine (1988) to explain sex differences found on delay and distortion variables (e.g., confidence in ratings, concern with impact of feedback). The second was to examine possible sex differences in performance attributions (i.e., ability, effort, test difficulty, and luck) made by supervisors to explain their subordinate's poor performance.

In many instances, empirical results did not support either previous research or tendered hypotheses, with some results opposite to expectations. However, as previously mentioned, all results must be considered very tentative due to the small sample size (and resulting low power). Interestingly, results failed to replicate the findings of Benedict and Levine (1988) that female supervisors were more likely to distort and delay than their males equivalents. That most differences (although non-significant) found were in the opposite direction with males delaying and distorting more than females, leads one to question the validity of biological sex being the crucial dividing factor between those who do, and those who do not, provide prompt and accurate feedback to their poorly performing subordinates.

Multivariate analyses revealed no significant differences due to sex-type with respect to delay, distortion, or effectance motivation variables. Exploratory procedures revealed a considerable, although non-significant, differences between instrumental male and instrumental female supervisors on their distortion of the overall performance of the volunteer. This result

is somewhat confusing, as by operational definition, both groups should be equivalent. It is possible that this finding is an artifact of the small sample size. The lack of differences between instrumental and expressive supervisors was somewhat surprising as the groups should be, by operational definition, quite different. However, here again, the cell sizes were very small (the cell sizes were too small for even exploratory analyses on their male equivalents). But, the distortion variables at least showed differences in the expected direction, with expressive females distorting more than instrumental females.

Motivational expectancies were evaluated as possible mediating factors in the delivery of prompt and accurate negative feedback. Few notable differences were found. There were no significant differences between expressive and instrumental female supervisors on these variables. Results indicate some suggestion that expressive female supervisors feel less capable than their instrumental equivalents of providing prompt and accurate negative results. Even if they were capable of behaving in such a manner, they are less sure that such actions would have the desired outcome. Thus, although the trend is suggestive that expressive female supervisors value the theoretical consequences of delivering prompt and accurate feedback to poorly performing subordinates, they may feel less capable and less confident of the outcome than instrumental females. Again, a more definitive answer would require additional research with a larger sample.

The lack of significant differences between those who have received training in performance evaluations and those who have not is somewhat surprising as one would have hoped that training would make individuals more effective. As the subjects were recruited from MBA and MPA classes, it is possible that, in classes, they had had exposure to issues in appraisal that would reduce the differences between those with and without formal training. Of interest, and somewhat counter-intuitive, supervisors without training delayed less than their trained peers. Perhaps they were not as aware of the potential consequences and problems associated with delivering feedback to someone who has performed well below expectations. As well, although the difference was statistically significant, one must question the practical

significance of only 1.5 days difference in scheduled appointments. Here again, a larger sample (preferably taken from a work place situation) would be needed to adequately assess the effectiveness of performance appraisal training.

Overall, the results of this pilot study are far from conclusive. However, the inconclusive and somewhat contradictory results should be interpreted with caution, particularly due to the small sample size. Future research using the method refined in this study will use supervisors from the Federal Public Service. The use of this subject pool will increase the applicability and generalizability of results: do individuals who work as supervisors behave in the same manner as business students when required to provide feedback in a poor performance situation?

ON SITE RESEARCH

The Effect of Sex and Sex-Typing on Motivation, Delay, and Distortion in Negative Feedback Situations

Method

Subjects and Recruitment

Participants were recruited over a two year period from Federal Service employees enrolled in management training courses; all subjects were employed within the same department. All participants were asked to complete the short form of the Bern Sex Role Inventory (BSRI-S; Bem, 1981), and to provide demographic data -- sex, age, years work experience, and months supervisory experience. It was explained that we were attempting to establish a profile on managers within the Federal Service. All were further asked to volunteer an hour of their time to help with an additional research project, establishing the validity and reliability of a computer science aptitude test. Their participation would involve the administration and scoring of the test and the completion of post-test questionnaires, however, they were not informed that they would be asked to schedule an interview for "subordinate" face-to-face feedback. It was explained that their department was very interested in the results of this research and hoped to make use of them in the future, and that the course facilitator had agreed to allow them to absent themselves for a hour from class if they wished to participate. Those who agreed to take part were asked to fill out a scheduling sheet with three possible times they could be available. To minimize inconvenience, time slots were offered both before and after class, and, based upon pre-recruitment discussion with the course coordinator, times particularly crucial to the course were not offered.

One hundred and thirty-seven individuals completed the BSRI-S, 61 males and 76 females, mean age=36.5 years (SD=6.7). Of that number, eighty-one volunteers participated in the follow-up study, 40 males and 41 females.

Measures

Independent Variables

Sex.

The supervisors' biological sex was included to compare with the results of Benedict and Levine (1988).

Sex-type.

Sex-type was assessed using the BSRI-S (Bem, 1981), yielding feminine sex-typed/expressive, masculine sex-typed/instrumental, androgynous, and undifferentiated sex-types. See Appendix 1.

Dependent Variables

Supervisor distortion of subordinate performance.

The evaluative judgement of the volunteer supervisors was assessed using four scales on the Supervisory Performance Evaluation Questionnaire (see Appendix 4), derived from the work of Benedict and Levine (1988):

- 1) "Please check the point on the scale which best describes the performance of the individual whose work you supervised." Scored using three verbal anchor points, "Excellent Performance", "Average Performance", and "Extremely Poor Performance".
- 2) "Please rate the overall quality of the work performed by the individual whom you supervised on the following scale." Scored using three verbal anchor points, "Excellent Performance", "Average Performance", and "Extremely Poor Performance".
- 3) "Based on your experience with the work performed by the individual whom you supervised, please circle your response to the following statement: This individual is capable of performing computer programming tasks." Scored on a 25 point scale, using

verbal anchor points. "Strongly Agree", "Agree", "Uncertain", "Disagree", and "Strongly Disagree".

4) "Please rate the overall quantity of the work performance of the individual you supervised on the following scale." Scored using three verbal anchor points, "Excellent Performance", "Average Performance, and "Extremely Poor Performance".

As the actual performance of the confederate was constantly at the 25th percentile, the distortion score was calculated as the difference between the rating given and the 25th percentile.

Item four was added subsequent to the pilot study. Based on comments, verbal and written, made by participants, it became clear that some supervisors were attempting to differentiate between the quality and the quantity of the work performed by the subordinate. (Based on the norms presented, both should have been the same.)

Supervisor delay: Ending performance trials.

This was the first of two delay measures. The volunteer supervisors were told that they could stop testing when they felt that had sufficient information to make an accurate assessment of the confederate volunteer's abilities (minimum 1 trial, maximum 10 trials). (See Appendix 2, Supervisory Control Worksheet.) The confederate subordinate's performance remained unchanged at the 25th percentile over repeated trials. It was felt that the more trials administered, the greater the hesitation on the part of the volunteer supervisor to accept these results as the confederate volunteer's maximum level of performance. They allowed more trials in the hope that the confederate would eventually show improvement.

Supervisor delay: Scheduling the feedback interview.

This was the second of the delay measures. Delay was assessed using the number of days between the testing session and the date scheduled for face-to-face feedback. Holidays, weekends, and other reasonable delays (e.g., being out of town) were subtracted from the total number.

Appraisal Effectance Motivation.

The volunteer supervisors' motivation to provide prompt and accurate feedback was assessed using the Appraisal Effectance Motivation Questionnaire (see Appendix 4), derived from the work of Bandura (1977, 1982), and Teasdale (1978). This questionnaire has seventeen items; six measure self-efficacy expectancy -- three for prompt feedback and three for accurate feedback, four measure outcome expectancy -- two for prompt feedback and two for accurate feedback, and six measure outcome value -- three for prompt feedback and three for accurate feedback. Thus, there are six measures of appraisal effectance motivation, three relating to prompt feedback and three relating to accurate feedback. Each item is answered by indicating one's position on a 14-point scale, ranging from strongly disagree (1) to strongly agree (14). This questionnaire was always administered last, as, during the pilot study, it was learned this is the measure that arouses the most suspicions as to the true nature of the research. Placing it last reduces the impact of such suspicions on the completion of the other questionnaires.

Supervisor Attribution of Confederate Subordinate Performance.

The causal attributions supervisors make about a subordinate's performance logically would have a impact on how that performance is viewed and evaluated, and feedback given. Four measures of attribution, two internal and two external, were included: confederate ability, confederate effort, task difficulty, and luck (see Appendix 4). The volunteer supervisors were

asked to what extent they felt each of these factors had influenced the performance of the person the supervised. Each factor was rated upon a seven point Likert-type scale, ranging from 1 to 7.

Supervisor Confidence and Concern re: Accuracy and Impact of Feedback.

Benedict and Levine (1988) proposed several possible factors to account for the differences between the delay and distortion scores of male and female subjects in their study: females might be less confident in their rating abilities than were males, and females might be more concerned with the effect poor ratings would have on subordinates than were males. Two questions were included to assess these possibilities, each one with a seven point Likert-type response scale (see Appendix 4). Additionally, in keeping with the previously cited research which suggests that people have greater difficulty delivering bad news when it is face-to-face, one question was included to ascertain possible supervisor sex differences in levels of anxiety related to such a task (see Appendix 4, Item 14, Appraisal Effectance Behaviours).

Control Factors

Subordinate sex and ethnicity.

In an effort to eliminate possible confounding factors due to *subordinate* sex, or to ethnic interactions, all programmer candidates were male Caucasians. The number of confederates was kept to a minimum in an effort to reduce the effects of individual confederate characteristics; five confederates were trained to perform at the 25th percentile, for two years of data collection.

Supervisor-subordinate interpersonal bonding.

In the work place, the supervisor and subordinate know one another quite well, and, in most cases, must continue to work together after the evaluation feedback. Such conditions do not exist in the experimental situation. In order for a bond to form between the supervisor and the confederate, a five to ten minute pre-experiment interlude was designed, allowing an opportunity for them to talk and exchange information.

Although the experimental situation was set up as to allow for the establishment of some rapport between the volunteer supervisor and the confederate, it was hoped that no supervisors would be personally acquainted with the confederate on an a priori basis. A prior friendship could have influenced the results in many ways (e.g., exacerbating delay and distortion scores, invalidating the study because with prior personal knowledge they would know the confederate was not a computer science student). A question was included asking whether or not the supervisor was personally acquainted with the confederate.

In the work place the results of performance evaluations can have a significant impact on the subordinate, and part of the bonding time was to convince the supervisor volunteer that the confederate felt the results were significant to him (e.g., "I'm hoping to be able to use these results in later job searches"). Two questions were included to assess whether or not the supervisor volunteer was satisfied about the confederate motivation to do well: the first, "Do you feel the volunteer tried his/her best on the work materials?", and the second, "Do you feel this volunteer is interested in or concerned about his/her performance results for this project?". Both questions were answered on a yes/no basis (see Appendix 4, Post Evaluation Questionnaire).

Procedure

Supervisors who had volunteered to participate in the "Computer Aptitude Study" were notified by letter of their appointment time and date (all of which occurred during the span of their course). Upon his/her arrival, a member of the research team introduced herself to the

volunteer supervisor (hereafter referred to as the supervisor), re-explained the pretext for the supervisor's participation, and outlined what s/he would be required to do (see Appendix 2 for the "Supervisory Control Worksheet, the "Computer Science Aptitude Test Norms", the "Computer Science Aptitude Test", and the "Computer Science Aptitude Test Sheets A & B"). Particular attention was paid to explaining the "normative data", expressed in percentiles, as the confederates would consistently be performing at the 25th percentile, and it was important that the supervisors realized this represented a substandard level of performance. The supervisor was told that the programmer candidate with whom s/he would be working was a student from the Computer Science Department at the University of Ottawa who had volunteered to participate in this study. When the confederate volunteer (hereafter referred to as the confederate) arrived, introductions were made. It was explained to the confederate, for the benefit of the supervisor, that each time test trial lasted three minutes, and that he was being marked on speed and accuracy (i.e., his score was the total correct). Finally, it was emphasised that the researcher was there only to oversee the procedure and answer questions, and that the decisions would be left up to the supervisor. As well, they were asked to complete the Supervisor Qualifications Sheet (see Appendix 2).

Following this introduction, the research team member then left the room for 5-10 minutes (on the pretext of making an important phone call), and she requested that they not start prior to her return. During this period the confederate attempted to establish rapport with the supervisor, emphasizing that he really wanted to do well and that he hoped to use the results in the future (e.g., job, or co-op applications).

Upon her return, the team member apologized for the delay, and she indicated that they could proceed with the testing. When the supervisor felt s/he had sufficient information, s/he called a halt to the testing, the confederate was thanked, and he was told that he would be contacted later with an appointment to discuss his results.

After the confederate's departure, the supervisor was asked to set a date to provide face-to-face feedback to the confederate volunteer. The supervisors were deliberately informed of this requirement prior to completing the post-test questionnaires to increase the generalizability to the work force. The supervisors were then asked to complete the Supervisory Performance Evaluation Sheet, the Appraisal Effectance Motivation Questionnaire, and the Post-Evaluation Questionnaire (see Appendix 4). The procedure was completed by re-introducing the confederate as a research team member, and debriefing the supervisor as to the purpose of the study.

RESULTS

Matching on Age, Work Experience, and Supervisory Experience

Eighty-one supervisors were recruited, 40 males and 41 females. Results of t -tests showed no significant differences ($p > .1$) between male and female supervisors on age (males $M=35.8$, $SD=6.6$; females $M=36.9$, $SD=7.9$), years work experience (males $M=14.7$, $SD=7.0$; females $M=13.5$, $SD=7.0$), or months supervisory experience (males $M=31.2$, $SD=42.5$; females $M=31.3$, $SD=42.0$).

Sample Size

Power calculations for 0.80 and 0.70, $\alpha = 0.05$, and both large and medium effect size suggest that the current sample size is sufficient for detecting group differences (Cohen, 1988). See Appendix 5.

BSRI-S Norms

Norms for masculine raw score (MRS) and feminine raw score (FRS) of the BSRI-S were calculated on the basis of 137 questionnaires completed by Federal Service managers (61 males, 76 females; see Table 12 for demographic data). The data for the MRS and FRS scores were normally distributed, and thus the means and medians for these scores were essentially the same: MRS -- $M=4.9$, $SD=0.76$, $Mdn=5.0$; FRS -- $M=5.6$, $SD=0.78$, $Mdn=5.6$. As Bem (1981) uses the median in her normative samples, we will use medians as well (see Table 13).

Using these norms, male and female supervisors were classified according to their sex-type. Of the 40 males, 10 were feminine sex-typed (or cross-sex-typed), 6 were androgynous, 13 were masculine sex-typed, and 11 were undifferentiated. Of the 41 females, 11 were feminine sex-typed, 17 were androgynous, 6 were masculine sex-typed (or cross-sex-typed), and 7 were undifferentiated. (See Figure 2).

TABLE 12

DEMOGRAPHIC DATA FOR BSRI-S MANAGER NORMATIVE SAMPLE (N = 137)

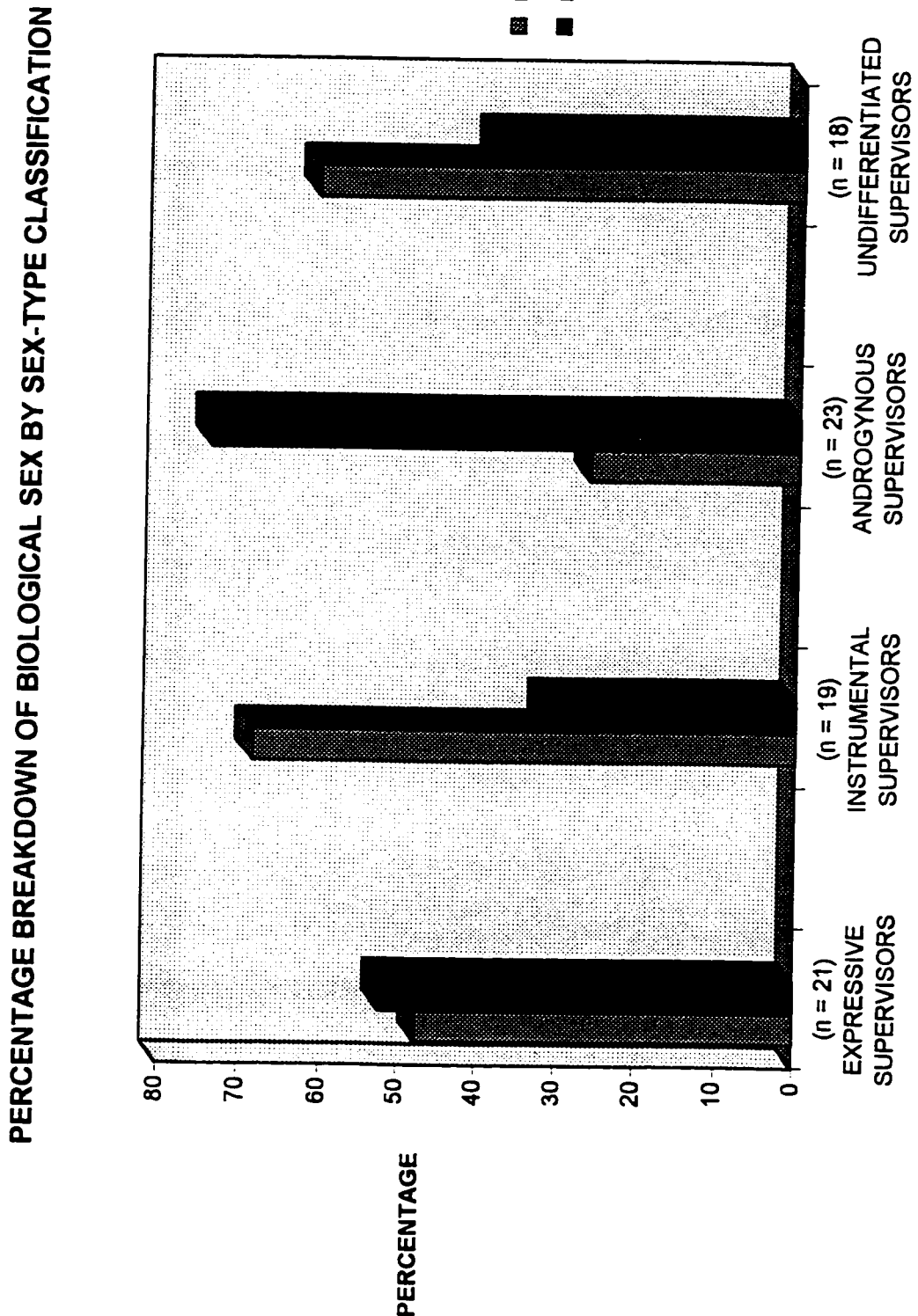
	MALES (<u>n</u> = 61; 44.5%)	FEMALE (<u>n</u> = 76; 55.5%)
Age	35.8 (<u>SD</u> = 5.7)	37.1 (<u>SD</u> = 7.3)
Years Work Experience	14.4 (<u>SD</u> = 6.4)	14.5 (<u>SD</u> = 7.3)
Months Supervisory Experience	32.1 (<u>SD</u> = 39.8)	34.8 (<u>SD</u> = 46.4)

TABLE 13

NORMS FOR BSRI AND BSRI-S

	FRS	MRS
<u>BSRI</u> (Bem, 1981)	4.90	4.95
<u>BRSI-S</u> (Bem, 1981)	5.50	4.80
<u>BRSI-S</u> (On-Site Study, Federal Service Supervisors)	5.60	5.50

FIGURE 2



Control Factors

Three supervisors (3.7%) indicated they were personally acquainted with their confederate subordinate. Confederates were questioned about these responses, and in each case they indicated they had had no prior personal knowledge of the supervisor. Perhaps these three supervisors responded affirmatively to this question as a result of the time spent talking with the confederate during the bonding component.

The majority of the subjects (91.5%) felt that the confederates had tried their best on the test, and 96.3% felt that the confederates were interested in the test results.

Delay and Distortion Measures

A 2 X 4 (sex by sex-type) factorial MANOVA showed non-significant main effects for biological sex of supervisor and for gender-role socialization (sex-type), and non-significant results for the sex by gender interaction on the four individual measures of feedback distortion and the two measures of feedback delay (see Table 14). Consistent with this, subsequent exploratory *t*-tests revealed non-significant differences between the female and male supervisors on these variables (see Table 15).

Results were further analyzed using broader classifications -- instrumental supervisors and androgynous (males and females), those with higher levels of instrumental characteristics; and expressive supervisors, those with higher levels of expressive characteristics. A 2 X 2 factorial MANOVA again indicated no significant main or interaction effect differences for either sex or the expanded gender classifications (see Table 16). Here again, exploratory *t*-tests revealed no significant differences between the expanded instrumental and expressive gender groupings on both delay variables and three distortion variables. However, a significant difference was noted (see Table 17) on the rating of the *quality* of the confederates' performance, $t(62)=2.91$, $p=.005$, with instrumental supervisors distorting the quality of performances less ($M=27$, $SD=26$) than their expressive counterparts ($M=46.4$, $SD=22.2$).

TABLE 14

MEAN DELAY AND DISTORTION VALUES FOR SUPERVISOR SAMPLE
(2 X 4 FACTORIAL MANOVA)

	<u>Delay #1</u> # of Test Trials Administered	<u>Delay #2</u> # Days to Feedback Session	<u>Distortion #1</u> Overall Performance	<u>Distortion #2</u> Quality of Performance	<u>Distortion #3</u> Quantity of Performance	<u>Distortion #4</u> Ability to Perform Computer Programming Tasks
<u>MALE</u>						
Expressive/Feminine Sex- Type ($n = 10$)	3.7 (<u>SD</u> = 1.4)	2.7 (<u>SD</u> = 1.9)	24.8 (<u>SD</u> = 27.4)	50.2 (<u>SD</u> = 24.3)	26.3 (<u>SD</u> = 26.1)	36.1 (<u>SD</u> = 18.1)
Instrumental/Masculine Sex-Typed ($n = 13$)	4.5 (<u>SD</u> = 1.2)	2.3 (<u>SD</u> = 1.9)	14.5 (<u>SD</u> = 21)	24.6 (<u>SD</u> = 22.2)	13.3 (<u>SD</u> = 12.4)	24.4 (<u>SD</u> = 17.9)
Androgynous ($n = 6$)	5.3 (<u>SD</u> = 3.2)	1.3 (<u>SD</u> = 0.8)	5.2 (<u>SD</u> = 18.0)	25.5 (<u>SD</u> = 37.3)	13.8 (<u>SD</u> = 33)	26.8 (<u>SD</u> = 10.1)
Undifferentiated ($n = 11$)	4.0 (<u>SD</u> = 1.5)	1.6 (<u>SD</u> = 1.0)	25.3 (<u>SD</u> = 21.5)	44.2 (<u>SD</u> = 25.1)	18.8 (<u>SD</u> = 10.3)	30.4 (<u>SD</u> = 19.7)
<u>FEMALE</u>						
Expressive/Feminine Sex- Typed ($n = 11$)	4.8 (<u>SD</u> = 1.6)	2.5 (<u>SD</u> = 2.5)	19.5 (<u>SD</u> = 23.1)	42.9 (<u>SD</u> = 20.7)	19.0 (<u>SD</u> = 19.9)	14.7 (<u>SD</u> = 12.5)
Instrumental/Masculine Sex-Typed ($n = 6$)	5.0 (<u>SD</u> = 1.1)	3.5 (<u>SD</u> = 2.6)	20.5 (<u>SD</u> = 21.8)	33.7 (<u>SD</u> = 28.3)	17.0 (<u>SD</u> = 7.6)	27.2 (<u>SD</u> = 12.7)
Androgynous ($n = 17$)	5.1 (<u>SD</u> = 1.7)	2.2 (<u>SD</u> = 2.2)	12.8 (<u>SD</u> = 10.6)	28.4 (<u>SD</u> = 26)	18.2 (<u>SD</u> = 21.8)	27.1 (<u>SD</u> = 15.6)
Undifferentiated ($n = 7$)	3.4 (<u>SD</u> = 1.0)	2.0 (<u>SD</u> = 1.4)	20.0 (<u>SD</u> = 13.8)	26.0 (<u>SD</u> = 17.7)	22.3 (<u>SD</u> = 13.5)	29.1 (<u>SD</u> = 16.4)

TABLE 15

DELAY AND DISTORTION VALUES FOR MALE VERSUS FEMALE VOLUNTEER SUPERVISORS

VARIABLES	MEAN - MALES (<u>n</u> = 40)	MEAN - FEMALE (<u>n</u> = 41)	t - VALUE	p - VALUE
<u>Delay #1 -</u>				
# of Tests Trials Administered	4.2 (SD = 1.8)	4.6 (SD = 1.6)	-1.14	.26
<u>Delay #2 -</u>				
# Days to Feedback Session	2.1 (SD = 1.6)	2.5 (SD = 2.3)	-1.03	.30
<u>Distortion #1 -</u>				
Overall Performance	18.8 (SD = 22.9)	16.7 (SD = 16.6)	0.46	.64
<u>Distortion #2 -</u>				
Quality of Performance	36.5 (SD = 27.6)	32.0 (SD = 23.9)	.79	.43
<u>Distortion #3 -</u>				
Quantity of Performance	18.2 (SD = 20)	18.6 (SD = 17.9)	-.11	.92
<u>Distortion #4 -</u>				
Ability to Perform Computer Programming Tasks	29.3 (SD = 17.6)	24.1 (SD = 15.2)	1.42	.16

TABLE 16

**MEAN DELAY AND DISTORTION VALUES FOR INSTRUMENTAL AND EXPRESSIVE GROUPS
(2 X 2 FACTORIAL MANOVA)**

	<u>Delay #1</u> # of Test Trials Administered	<u>Delay #2</u> # Days to Feedback Session	<u>Distortion #1</u> Overall Performance	<u>Distortion #2</u> Quality of Performance	<u>Distortion #3</u> Quantity of Performance	<u>Distortion #4</u> Ability to Perform Computer Programming Tasks
<u>MALE</u>						
Instrumental Grouping ^a (<u>n</u> = 19)	4.7 (<u>SD</u> = 2.0)	2.1 (<u>SD</u> = 1.7)	11.5 (<u>SD</u> = 20.1)	24.9 (<u>SD</u> = 26.8)	13.5 (<u>SD</u> = 20.1)	25.2 (<u>SD</u> = 15.6)
Expressive Grouping ^b (<u>n</u> = 10)	3.7 (<u>SD</u> = 1.4)	2.7 (<u>SD</u> = 1.9)	24.8 (<u>SD</u> = 27.4)	50.2 (<u>SD</u> = 24.3)	26.3 (<u>SD</u> = 26.1)	36.1 (<u>SD</u> = 18.1)
<u>FEMALE</u>						
Instrumental Grouping ^a (<u>n</u> = 23)	5.0 (<u>SD</u> = 1.6)	2.5 (<u>SD</u> = 2.3)	14.8 (<u>SD</u> = 14.2)	29.8 (<u>SD</u> = 26.1)	17.9 (<u>SD</u> = 18.9)	27.1 (<u>SD</u> = 14.6)
Expressive Grouping ^b (<u>n</u> = 11)	4.8 (<u>SD</u> = 1.6)	2.5 (<u>SD</u> = 2.5)	19.5 (<u>SD</u> = 23.1)	42.9 (<u>SD</u> = 20.7)	19.0 (<u>SD</u> = 19.9)	14.7 (<u>SD</u> = 12.5)

^a = sex-typed males, cross-sex-typed females, and androgynous males and females

^b = sex-typed females and cross-sex-typed males

TABLE 17

MEAN DELAY AND DISTORTION VALUES FOR INSTRUMENTAL^a VERSUS EXPRESSIVE^b SUPERVISORS

VARIABLES	MEAN - INSTRUMENTAL ^a GROUPING (<u>n</u> = 42)	MEAN - EXPRESSIVE ^b GROUPING (<u>n</u> = 21)	t - VALUE	p - VALUE
<u>Delay #1 -</u>				
# of Tests Trials Administered	4.9 (SD = 1.7)	4.2 (SD = 1.6)	-1.33	0.2
<u>Delay #2 -</u>				
# Days to Feedback Session	2.41 (SD = 2.1)	2.6 (SD = 2.6)	-0.35	0.73
<u>Distortion #1 -</u>				
Overall Performance	13.1 (SD = 16.8)	22.0 (SD = 24.7)	1.70	0.1
<u>Distortion #2 -</u>				
Quality of Performance	27.0 (SD = 26.1)	46.4 (SD = 22.2)	2.91	0.005
<u>Distortion #3 -</u>				
Quantity of Performance	15.6 (SD = 19.2)	22.5 (SD = 22.7)	1.26	0.2
<u>Distortion #4 -</u>				
Ability to Perform Computer Programming Tasks	26.2 (SD = 14.9)	24.9 (SD = 18.5)	-0.31	0.8

^a sex-typed males, cross-sex-typed females, and androgynous males and females

^b sex-typed females and cross-sex-typed males

To increase the power of measurement, the four distortion measures were combined to form one global distortion variable (reliability, $\alpha = 0.75$). The two delay measures were left separate, as combining resulted in low internal consistency ($\alpha = .15$). A 2 X 4 factorial MANOVA revealed a significant main effect for stereotypic gender-role socialization. No biological sex differences were found on these combined distortion scores (see Table 18). Interpretation of Pillai's criterion indicated significant differences between gender groups, $F(9,222)=2.28$, $p=.02$, with an overall effect size of 8.5%. Subsequent exploratory t -tests (see Table 19, and Figures 3 & 4) revealed significant differences between feminine sex-typed supervisors ($M=90.9$, $SD=56.6$) and both androgynous ($M= 55.5$, $SD=57.2$) and masculine sex-typed supervisors ($M=56.2$, $SD=48.0$) on the composite distortion score. The only significant difference found on the number of test trials allowed was between androgynous ($M=5.1$, $SD=2.1$) and undifferentiated supervisors ($M=3.7$, $SD=1.3$).

The delay and distortion variables were also analyzed by ANOVA, as this was done by Benedict and Levine (1988). A main effect of gender on the rating of the quality of the confederates performance was found, $F(4,74)=2.34$, $p=.041$. All other analyses revealed non-significant main and interaction effects for both sex and gender on the remaining distortion variables and the two delay variables.

Responses were also analyzed in terms of training experience, that is, was there a significant difference in the delay and distortion variables between those supervisors who had received formal training on the handling of performance appraisals and those who had none. Results revealed no significant effects for training on these variables. Subsequent exploratory t -tests revealed no significant differences between those supervisors who had received performance appraisal training and those who had not, on all delay and distortion variables except one. A marginally significant effect was found for feedback delay, $t(80)=1.87$, $p=.065$.

TABLE 18

MEAN VALUES FOR COMPOSITE DISTORTION SCORES AND FOR DELAY SCORES
(2 X 4 FACTORIAL MANOVA)

	Composite Distortion Score*	Delay #1 # of Test Trials Administered	Delay #2 # Days to Feedback Session
<u>MALE</u>			
Expressive/Feminine Sex-Type ($\underline{n}$ = 10)	101.3 ($\underline{SD}$ = 69.3)	3.7 ($\underline{SD}$ = 1.4)	2.7 ($\underline{SD}$ = 1.9)
Instrumental/Masculine Sex- Typed ($\underline{n}$ = 13)	52.3 ($\underline{SD}$ = 48.2)	4.5 ($\underline{SD}$ = 1.2)	2.3 ($\underline{SD}$ = 1.9)
Androgynous ($\underline{n}$ = 6)	44.5 ($\underline{SD}$ = 83.8)	5.3 ($\underline{SD}$ = 3.2)	1.3 ($\underline{SD}$ = 0.8)
Undifferentiated ($\underline{n}$ = 11)	88.7 ($\underline{SD}$ = 41)	4.0 ($\underline{SD}$ = 1.5)	1.6 ($\underline{SD}$ = 1.6)
<u>FEMALE</u>			
Expressive/Feminine Sex-Typed ($\underline{n}$ = 11)	81.5 ($\underline{SD}$ = 43.1)	4.8 ($\underline{SD}$ = 1.6)	2.5 ($\underline{SD}$ = 2.5)
Instrumental/Masculine Sex- Typed ($\underline{n}$ = 6)	64.1 ($\underline{SD}$ = 50.8)	4.9 ($\underline{SD}$ = 1.1)	4.0 ($\underline{SD}$ = 2.7)
Androgynous ($\underline{n}$ = 17)	59.4 ($\underline{SD}$ = 47.4)	5.1 ($\underline{SD}$ = 1.7)	2.2 ($\underline{SD}$ = 2.2)
Undifferentiated ($\underline{n}$ = 7)	68.3 ($\underline{SD}$ = 40.8)	3.4 ($\underline{SD}$ = 1.0)	2.0 ($\underline{SD}$ = 1.4)

* Overall performance, quality of performance, quantity of performance, and ability to perform computer programming.

TABLE 19

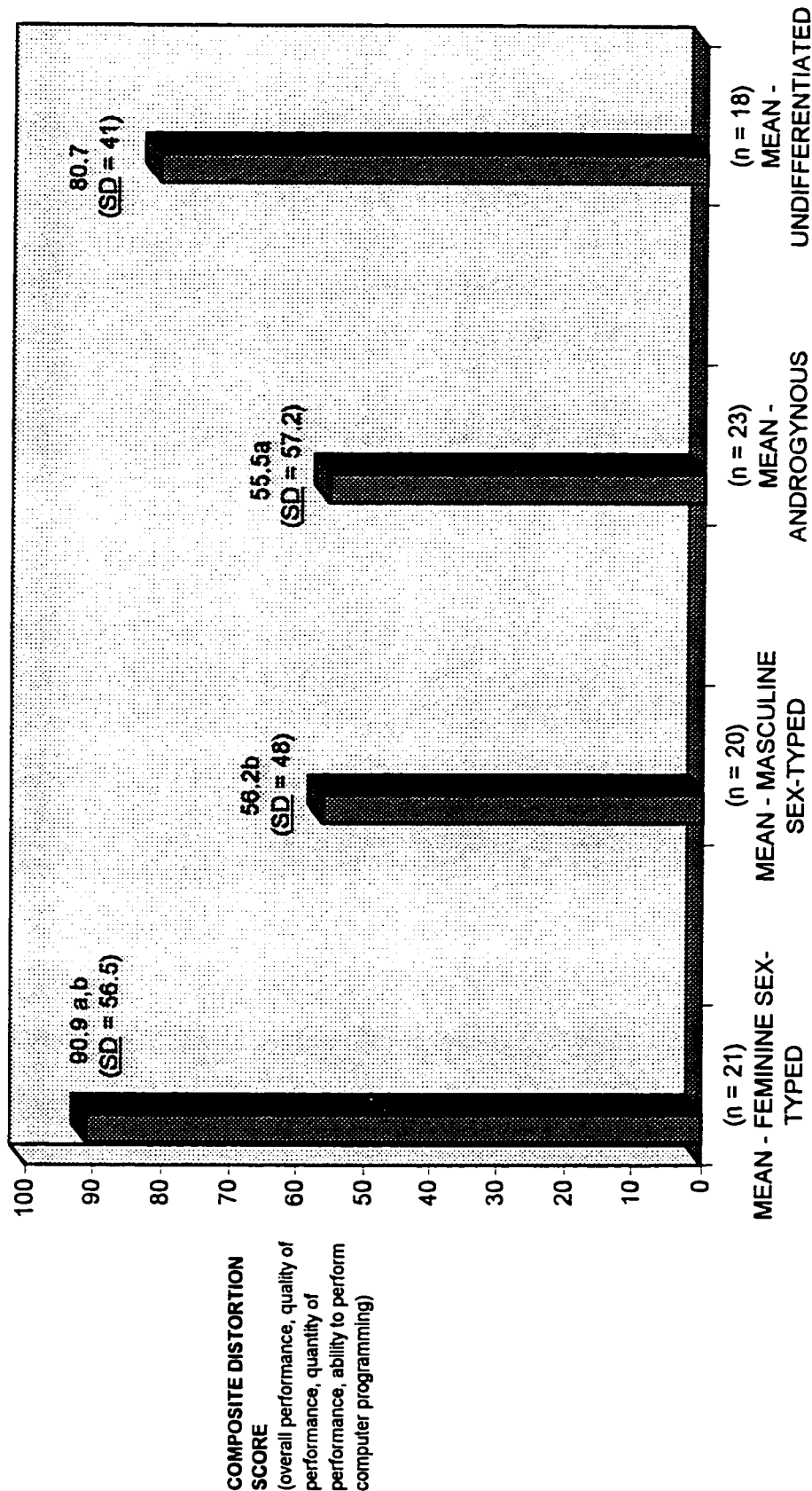
COMPOSITE DISTORTION VALUES (MEANS) AND # OF TEST TRIALS (DELAY) FOR GENDER GROUPINGS OF
VOLUNTEER SUPERVISORS

VARIABLES	MEAN - EXPRESSIVE (<u>n</u> = 21)	MEAN - INSTRUMENTAL (<u>n</u> = 19)	MEAN - ANDROGYNOUS (<u>n</u> = 23)	MEAN - UNDIFFERENTIATED (<u>n</u> = 18)
COMPOSITE DISTORTION SCORE (overall performance, quality of performance, quantity of performance, ability to perform computer programming)	90.9 ^{a, b} (<u>SD</u> = 56.5)	56.2 ^b (<u>SD</u> = 48)	55.5 ^a (<u>SD</u> = 57.2)	80.7 (<u>SD</u> = 41)
# OF TEST TRIALS ADMINISTERED	4.3 (<u>SD</u> = 1.6)	4.6 ^d (<u>SD</u> = 1.1)	5.1 ^c (<u>SD</u> = 2.1)	3.7 ^{c, d} (<u>SD</u> = 1.3)

^{a, b, d} Significant at $p < .05$

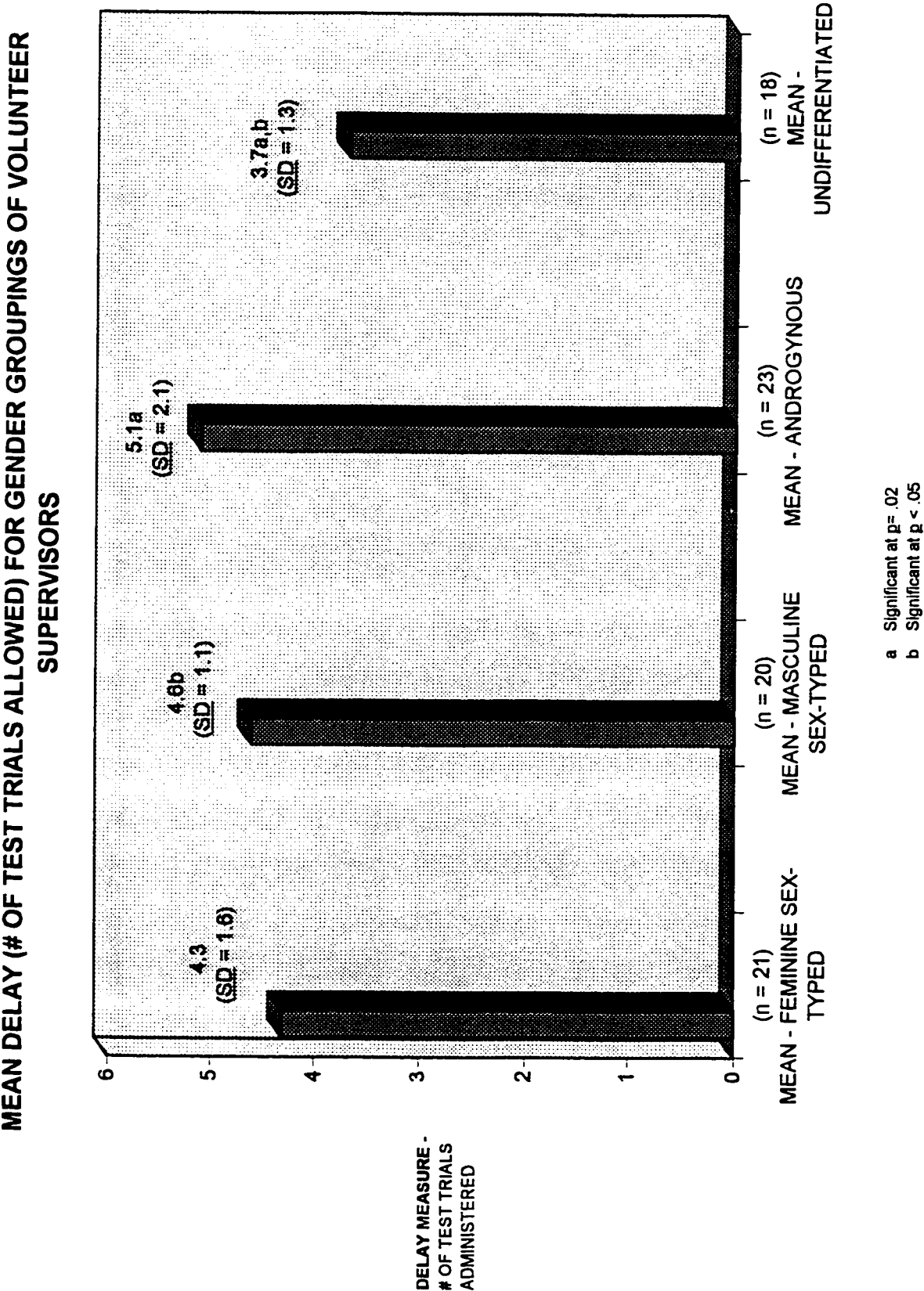
^c Significant at $p = .02$

FIGURE 3
COMPOSITE DISTORTION VALUES (MEANS) FOR GENDER GROUPINGS OF VOLUNTEER SUPERVISORS



a, b Significant at $p < .05$

FIGURE 4



Appraisal Effectance Motivation Measures

A 2 X 4 (sex by sex-type) factorial MANOVA showed non-significant main effects for biological sex and for gender, and non-significant results for the sex-by gender interaction on the six component measures of motivation (see Table 20): 1&2) self-efficacy expectancy for accuracy and promptness, 3&4) outcome expectancy for accuracy and promptness, and 5&6) outcome value for accuracy and promptness. Correlational analyses showed high correlations between the "prompt" and "accurate" values of each of the motivation variables (self-efficacy expectancy -- .76, outcome expectancy -- .75, and outcome value -- .76). Thus it was decided to combine the "prompt" and "accurate" variables, resulting in three motivation scores. These combined variables also produced increased internal consistency indicators in two of three instances: self-efficacy - $\alpha = .81$, outcome expectancy - $\alpha = .69$, and outcome value - $\alpha = .88$.

MANOVA results (2 X 4 factorial) using these combined motivation scores showed non-significant main effects for biological sex, and non-significant results for the sex by gender interaction on the three combined measures of motivation. There was a marginally significant effect due to stereotypic gender-roles, $F(9,222)=1.71$, $p=.09$. Subsequent exploratory t -tests revealed non-significant differences between the female and male managers on these variables (see Table 21).

Responses were further analyzed in terms of training experience, that is, was there a significant difference in the motivation variables between those supervisors who had received formal training in the handling of performance appraisals and those who had not. Analyses revealed a significant main effect. Interpretation of Pillai's criterion indicated significant differences between training groups, $F(3,78)=2.76$, $p=.048$. An examination of the structure coefficients indicated that self-efficacy expectancy was the most important variable in discriminating between training groups, distantly followed by outcome value. The SPSS MANOVA command was used to adjust for unequal sample sizes (Tabachnick and Fidell, 1989). Exploratory t -tests revealed significant differences between those supervisors who had received performance appraisal training and those who had not, such that trained supervisors reported higher self-efficacy expectancy (see Table 22).

TABLE 20

**MEAN MOTIVATION VALUES -- SELF-EFFICACY EXPECTANCY, OUTCOME EXPECTANCY,
AND OUTCOME VALUE -- FOR SUPERVISOR SAMPLE
(2 X 4 FACTORIAL MANOVA)**

	Self-Efficacy Expectancy - Accuracy	Outcome Expectancy- Accuracy	Outcome Value - Accuracy	Self-Efficacy Expectancy - Prompt	Outcome Expectancy - Prompt	Outcome Value - Prompt
<u>MALE</u>						
Expressive/Feminine Sex- Type ($n = 10$)	64.9 ($\underline{SD} = 21.4$)	45.6 ($\underline{SD} = 22.3$)	65.5 ($\underline{SD} = 25.7$)	60.9 ($\underline{SD} = 23.1$)	47.7 ($\underline{SD} = 17.5$)	64.0 ($\underline{SD} = 28.2$)
Instrumental/Masculine Sex-Typed ($n = 13$)	73.6 ($\underline{SD} = 17.1$)	61.7 ($\underline{SD} = 16.2$)	74.8 ($\underline{SD} = 15.3$)	70.5 ($\underline{SD} = 16.3$)	57.3 ($\underline{SD} = 16.8$)	72.7 ($\underline{SD} = 17.5$)
Androgynous ($n = 6$)	66.4 ($\underline{SD} = 16.4$)	65.3 ($\underline{SD} = 4.7$)	72.0 ($\underline{SD} = 24.0$)	62.2 ($\underline{SD} = 15.6$)	59.7 ($\underline{SD} = 12.8$)	76.3 ($\underline{SD} = 18.6$)
<u>FEMALE</u>						
Expressive/Feminine Sex- Typed ($n = 11$)	62.4 ($\underline{SD} = 16.6$)	63.9 ($\underline{SD} = 18.8$)	72.7 ($\underline{SD} = 18.2$)	57.9 ($\underline{SD} = 13.0$)	57.6 ($\underline{SD} = 21.2$)	70.4 ($\underline{SD} = 21.6$)
Instrumental/Masculine Sex-Typed ($n = 6$)	80.0 ($\underline{SD} = 11.9$)	53.3 ($\underline{SD} = 14.7$)	82.9 ($\underline{SD} = 10.4$)	79.0 ($\underline{SD} = 8.1$)	57.6 ($\underline{SD} = 13.3$)	77.2 ($\underline{SD} = 15.9$)
Androgynous ($n = 17$)	68.0 ($\underline{SD} = 14.8$)	63.1 ($\underline{SD} = 11.6$)	79.1 ($\underline{SD} = 12.1$)	64.9 ($\underline{SD} = 19.4$)	61.8 ($\underline{SD} = 12.9$)	78.7 ($\underline{SD} = 14.3$)

TABLE 21

MEAN APPRAISAL MOTIVATION SCORES FOR MALE VERSUS FEMALE SUPERVISORS

VARIABLES	MEAN - MALES (<u>n</u> = 40)	MEAN - FEMALE (<u>n</u> = 41)	t - VALUE	p - VALUE
Self-Efficacy Expectancy	67.8 (SD = 16.9)	66.2 (SD = 14.4)	.31	.66
Outcome Expectancy	54.2 (SD = 16.2)	58.8 (SD = 14.2)	.43	.18
Outcome Value	70.8 (SD = 19.9)	75.2 (SD = 15.4)	.10	.26

TABLE 22

MEAN APPRAISAL MOTIVATION SCORES FOR SUPERVISORS WITH AND WITHOUT PREVIOUS PERFORMANCE APPRAISAL TRAINING

VARIABLES	MEAN - WITH TRAINING (<u>n</u> = 30)	MEAN - WITHOUT TRAINING (<u>n</u> = 51)	t - VALUE	p - VALUE
Self-Efficacy Expectancy	73.1 (SD = 13.6)	63.4 (SD = 15.7)	2.83	.006
Outcome Expectancy	57.5 (SD = 15.3)	56.0 (SD = 15.5)	.41	.68
Outcome Value	75.7 (SD = 15.5)	71.6 (SD = 18.9)	1.01	.32

Supervisor Attribution Factors

A 2 X 4 (sex by sex-type) factorial MANOVA revealed a non-significant main effect for gender, and a non-significant interaction for sex and gender (see Table 23). A marginally significant effect was found for sex, $F(4,71)=2.24$, $p=.07$. Exploratory t -tests revealed significant differences between male and female supervisors on performance attributions of the ability of the confederate, and of the difficulty of the test (see Table 24). Thus, there was a tendency for male versus female supervisors to attribute confederate's poor results to ability and to the difficulty of the test.

TABLE 23

MEAN PERFORMANCE ATTRIBUTION VALUES FOR SUPERVISORY SAMPLE
(2 X 4 FACTORIAL MANOVA)

	Ability of Confederate	Difficulty of Test	Effort	Luck
<u>MALE</u>				
Expressive/Feminine Sex- Type ($\underline{n}$ = 10)	4.4 ($\underline{SD}$ = 1.3)	3.1 ($\underline{SD}$ = 1.7)	4.8 ($\underline{SD}$ = 1.1)	1.2 ($\underline{SD}$ = 0.4)
Instrumental/Masculine Sex- Typed ($\underline{n}$ = 13)	4.6 ($\underline{SD}$ = 1.0)	3.4 ($\underline{SD}$ = 1.9)	4.2 ($\underline{SD}$ = 2.0)	1.2 ($\underline{SD}$ = 0.3)
Androgynous ($\underline{n}$ = 6)	3.8 ($\underline{SD}$ = 1.2)	4.1 ($\underline{SD}$ = 1.7)	4.5 ($\underline{SD}$ = 1.5)	1.5 ($\underline{SD}$ = 1.2)
Undifferentiated ($\underline{n}$ = 11)	4.5 ($\underline{SD}$ = 1.3)	4.5 ($\underline{SD}$ = 1.6)	4.5 ($\underline{SD}$ = 1.6)	1.4 ($\underline{SD}$ = 0.5)
<u>FEMALE</u>				
Expressive/Feminine Sex- Typed ($\underline{n}$ = 11)	3.8 ($\underline{SD}$ = 1.1)	2.7 ($\underline{SD}$ = 1.3)	5.1 ($\underline{SD}$ = 1.1)	1.1 ($\underline{SD}$ = 0.3)
Instrumental/Masculine Sex- Typed ($\underline{n}$ = 6)	4.1 ($\underline{SD}$ = 1.9)	3.9 ($\underline{SD}$ = 2.1)	4.1 ($\underline{SD}$ = 1.9)	1.3 ($\underline{SD}$ = 0.5)
Androgynous ($\underline{n}$ = 17)	3.7 ($\underline{SD}$ = 1.3)	2.9 ($\underline{SD}$ = 1.4)	4.8 ($\underline{SD}$ = 1.9)	1.4 ($\underline{SD}$ = 0.7)
Undifferentiated ($\underline{n}$ = 7)	3.3 ($\underline{SD}$ = 0.8)	3.0 ($\underline{SD}$ = 1.5)	4.3 ($\underline{SD}$ = 1.1)	1.6 ($\underline{SD}$ = 0.8)

TABLE 24

MEAN PERFORMANCE ATTRIBUTION VARIABLES FOR MALE VERSUS FEMALE SUPERVISORS

VARIABLES	MEAN - MALES (<u>n</u> = 40)	MEAN - FEMALE (<u>n</u> = 41)	t - VALUE	p - VALUE
Ability of Confederate	4.4 (<u>SD</u> = 1.2)	3.8 (<u>SD</u> = 1.2)	2.44	.02
Difficulty of the Test	3.8 (<u>SD</u> = 1.8)	3.0 (<u>SD</u> = 1.5)	2.00	.05
Effort Exerted	4.5 (<u>SD</u> = 1.6)	4.7 (<u>SD</u> = 1.6)	-.60	.55
Luck	1.3 (<u>SD</u> = 0.6)	1.3 (<u>SD</u> = 0.6)	-.44	.66

Supervisor Confidence and Concern
Re: Accuracy and Impact of Feedback

The supervisors' confidence in ratings, concern with impact of feedback, and anticipated level of comfort with face-to-face feedback were analyzed using separate ANOVA's as there was no theoretical reason to group them together, and all inter-correlations were low (0.03 - 0.28). ANOVA results revealed a significant main effect for gender on concern with impact of feedback of results on the confederate, $F(3,72)=3.2$, $p=.028$. Additional ANOVA analyses, revealed a significant main effect for biological sex on level of anticipated comfort with face-to-face

feedback, $F(1,74)=4.56$, $p=.036$. There were no significant differences between groups according to sex or to gender on the supervisors' confidence in their ratings. Subsequent exploratory t -tests for gender revealed significant differences only for concern about the "impact of feedback" between expressive and undifferentiated, between androgynous and instrumental supervisors, and between androgynous and undifferentiated supervisors (see Table 25 and Figure 5). Additional exploratory t -tests for biological sex revealed significant differences between males and female supervisors on anxiety due to face-to-face feedback (see Table 26), with male supervisors reporting such anxiety 51.8% of the time, versus 67.5% for their female counterparts.

TABLE 25

MEAN SUPERVISOR VARIABLE SCORES (CONFIDENCE, CONCERN AND ANXIETY) FOR GENDER GROUPINGS
OF VOLUNTEER SUPERVISORS

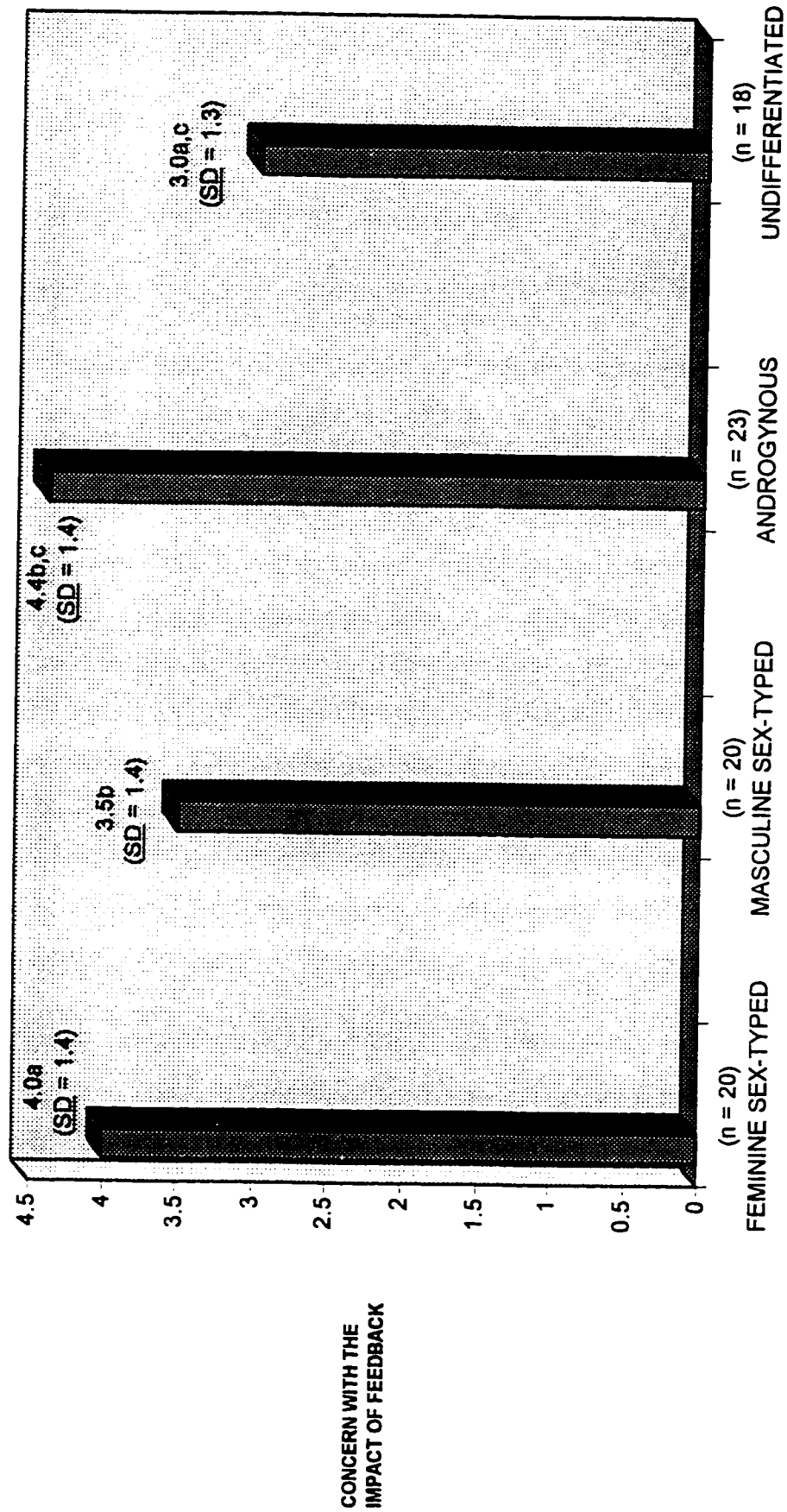
VARIABLES	EXPRESSIVE (n = 20)	INSTRUMENTAL (n = 19)	ANDROGYNOUS (n = 23)	UNDIFFERENTIATED (n = 18)
Confidence in Accuracy of Ratings	5.1 (<u>SD</u> = 1.9)	5.0 (<u>SD</u> = 1.4)	5.1 (<u>SD</u> = 1.9)	4.6 (<u>SD</u> = 1.7)
Concern With Impact of Feedback	4.0 ^a (<u>SD</u> = 1.4)	3.5 ^b (<u>SD</u> = 1.4)	4.4 ^{b,c} (<u>SD</u> = 1.4)	3.0 ^{a,c} (<u>SD</u> = 1.3)
Anxiety Caused Due Face-to-Face Feedback	60.8 (<u>SD</u> = 29.8)	55.2 (<u>SD</u> = 29.2)	64.6 (<u>SD</u> = 30.5)	57.8 (<u>SD</u> = 32.6)

^a Significant at $p < .03$

^b Significant at $p < .05$

^c Significant at $p = .003$

FIGURE 5
MEAN SUPERVISOR CONCERN SCORES FOR GENDER GROUPINGS OF VOLUNTEER SUPERVISORS



a Significant at $p < .03$
b Significant at $p < .05$
c Significant at $p = .003$

TABLE 26

**MEAN SUPERVISOR VARIABLES (CONFIDENCE, CONCERN AND ANXIETY) FOR
MALE VERSUS FEMALE VOLUNTEER SUPERVISORS**

VARIABLES	MEAN - MALES (<u>n</u> = 40)	MEAN - FEMALE (<u>n</u> = 41)	t - VALUE	p - VALUE
Confidence in Accuracy of Ratings	5.1 (SD = 1.6)	4.8 (SD = 1.8)	.92	.36
Concern With Impact of Feedback	3.5 (SD = 1.4)	4.0 (SD = 1.5)	-1.63	.11
Anxiety Caused Due to Face-to-Face Feedback	51.8 (SD = 30.5)	67.5 (SD = 28.0)	-2.44	.02

DISCUSSION

Distortion and Delay

Few of the results supported the hypotheses that there would be differences in delay and distortion scores due either to biological sex or gender-role orientation. More specifically, based on theory, we had expected that those volunteer supervisors with instrumental or masculine traits (i.e., masculine sex-typed, feminine cross-sex-typed, and androgynous) would be *less* likely to delay and distort than would those with primarily expressive or feminine traits (i.e., feminine sex-typed, and masculine cross-sex-typed). Not only were these hypotheses largely not supported, but this study also failed to replicate the results of Benedict and Levine (1988) to the effect that females delay and distort more than do males. This is not to say that distortion of ratings was not found. However, in terms of such behaviours, the volunteer supervisors behaved far more as one group of supervisors than as two, either in terms of sex or gender.

When asked to rate the overall performance of the confederate subordinates, the supervisors inflated their evaluation scores from the 25th percentile (actual performance) to, on average, the 43rd percentile (i.e., from poor performance to slightly below average performance). Their evaluation of the overall quantity of the work was similarly positively distorted to the 43rd percentile, and the overall quality ratings were stretched to the 59th percentile. In terms of capability to perform computer programming tasks, the confederate volunteers were given an average rating at the 52nd percentile. Thus all performance evaluations were positively distorted much closer to or within the average range. This occurred despite the performance norms having been explained prior to testing. Moreover, the supervisors had correctly indicated on the results sheet that actual test scores they saw corresponded to the 25th percentile. Some supervisors commented they felt the quality of the work produced was high, as the volunteer subordinates had made few or no errors on those items they did complete. However, it had been stressed to supervisors ahead of time that success on this test required *both* accuracy and speed. This would be equivalent to a professor saying that the quality of a student's work is good, for, although s/he failed the test having only answered 25/100 questions in the examination time allowed, no mistakes had been made.

Further analyses using the collapsed gender groups (grouping all supervisors with instrumental characteristics into one group, and all supervisors with expressive traits into another) revealed no significant differences between the groups on the delay and most distortion variables. However, exploratory analyses suggested a difference between these expanded instrumental and expressive supervisory categories in terms of their evaluations of the overall *quality* of their subordinate's work. Again all supervisors distorted, but this time to different degrees. Instrumental supervisors inflated these ratings to the 52nd percentile, whereas expressive supervisors distorted, on average, to the 71st percentile!

Researchers (e.g., Tesser & Rosen, 1975) have extensively investigated the general reluctance of individuals to deliver bad news (particularly in a face-to-face situation), and the possible reasons behind such behaviour, including concern with self, with the recipient, and with societal norms. By definition (Bem, 1981), those individuals who endorse predominantly feminine sex-typed characteristics, place considerable emphasis and energy on their interpersonal skills, and on being attuned and supportive to the needs of others. As such, the expressive supervisors may have been more concerned with the potential interpersonal impact of negative feedback on their confederate, and thus they distorted the evaluation of work quality considerably more than did their instrumental counterparts. Thus, the presence of instrumental traits may facilitate the delivery of accurate negative feedback, at least by reducing more of the distortion concerning the quality that they saw. These supervisors, notwithstanding their quantitative jobs, did not have experience with computer science aptitude tests. Their task was to evaluate performance on a task that was new to them. It is possible most supervisors distorted in that they may have given subordinates the benefit of any doubts they had. However, the likelihood that this actually happened is contradicted by their average certainty of the accuracy of their ratings.

Analyses (see Tables 15 & 17) revealed a significant difference on the quality of performance variable, again attributable to the gender orientation, not to the biological sex of the supervisors (as had been the case of Benedict and Levine, 1988). Although this finding does provide some support for the hypotheses in this study, it should be interpreted with caution: multivariate results revealed no significant differences, the cell sizes are unequal, and differences

were found on only one variable even after collapsing the gender groupings.

When the four distortion questions (see Table 19 and Figure 3) were combined to form one umbrella distortion factor, again the only differences found were due to gender (stereotypic gender-role socialization), and not to biological sex. This suggests that although, there were no gender differences on the individual distortion questions, when these are grouped together to give a better measure of overall positive distortion, a general tendency for differences becomes more apparent. Those supervisors with more expressive/feminine traits distorted considerably more (average score = 90.9) than did their counterparts with more instrumental/masculine characteristics (average score = 56.2). This lends some support to the hypothesis that it is gender (the characteristics and traits that one possesses) and not biological sex, per se, that influences the extent to which distortion occurs when supervisors are required to evaluate and to provide face-to-face feedback to a poorly performing subordinate. The expressive supervisors showed more positively distorted performance evaluations than their instrumental co-supervisors.

It is instructive to ask what factors or processes would logically and theoretically stand behind these motivationally induced distortions. Possibly the evaluation of poor subordinate performance, when face-to-face feedback is required, activates dysfunctional cognitive supervisor schema which plays an inhibitory role, antecedent to positive distortions of performance judgements. These positively distorted evaluations reduce inhibitory anxiety linked to anticipated face-to-face feedback. It is further possible that the activated dysfunctional cognitive schema pertain to the concern which the supervisor expected his/her confederate to have, particularly given that he had presented himself as a likeable individual who wanted to use these results (e.g., job interviews).

By definition, expressive supervisors endorse traits (e.g., compassionate, sensitive to the needs of others, eager to soothe hurt feelings) that could very well predispose them to behave in a way to protect their subordinate from a potentially negative reaction to a poor PER by distorting the results, even in a subtle fashion. They might rationalize that minor distortions are not significant, whereas the results of this study clearly indicate otherwise. When added together, even minor positive distortions can result in a substantial distortion of overall performance.

Self-presentational concerns may also have had an impact on the evaluations made by female expressive supervisors (approximately half of the expressive supervisors in this study). Other researchers (e.g., Bond and Anderson, 1978; Deau, 1974) have found that the reluctance to transmit bad news is related to concerns about public display. Kern et al. (1985) found that traditional observers were likely to devalue assertive behaviours on the part of females, and to rate such women as lower in social skill, likeability, and attractiveness. Thus, compared to male expressive supervisors or to instrumental supervisors, the female expressive supervisors in this study may have also been preoccupied with self-presentational concerns (particularly as all confederate subordinates were male). They might have made the assumption that they would be judged negatively for behaving in such an assertive manner, and thus tailor their judgements of subordinate performance and feedback strategy accordingly.

With respect to delay in the provision of feedback, in particular the scheduling of an appointment to deliver results, we had expected expressive supervisors to delay more than their instrumental counterparts. This was expected from a theoretical standpoint such that delay would provide a distancing mechanism, whereby expressive supervisors could separate themselves from the delivering of the bad news (e.g., Tesser, Rosen & Batchelor, 1972; Tesser, Rosen & Tesser, 1971). This position was not supported, and the results may be due, at least in part, to the research design used. First, in this scenario, delivering negative feedback would have had less impact on the supervisor-subordinate relationship than would be the case in a real organization, because supervisors realized they would not have to interact with the confederate subordinate again after the face-to-face feedback. Thus, the concern that expressive supervisors would have about the effect of such face-to-face feedback on the ongoing interactions would be mitigated to a considerable extent. Even though a bonding period had been incorporated to allow confederate subordinates to present themselves as pleasant and invested in the procedure (e.g., job search), such interpersonal effect would have little lasting impact as supervisors would realize there is no on-going relationship.

Second, when asked to schedule a face-to-face feedback appointment with the confederate, the supervisors were forced to choose a date after the testing procedure. That is, they were not offered the option of providing feedback immediately after testing. It was

logistically impossible for some, whereas others initially requested to do so. Perhaps had the "zero delay" condition been allowed, there would have been greater variability in these results. (These same limitations applied to the Benedict and Lerrine (1988) study.)

Appraisal Effectance Motivation Measures

Contrary to the proposed hypotheses, no substantial differences were found between male and female supervisors, nor between any of the stereotypic gender groups on any of the motivation variables (self-efficacy expectancy, outcome expectancy, and outcome value). Supervisors indicated that, on average, they expected to be able to provide prompt and accurate feedback more than sixty percent of the time to subordinates whom are performing poorly, and that they value the outcome of such behaviours more than seventy percent of the time. Interestingly, only slightly more than half the time did supervisors expect that the provision of prompt and accurate feedback would have the desired outcome expectancy results (e.g., allowing the employee to make timely changes in his/her behaviour). Thus, although they expected to be able to deliver prompt and accurate feedback under difficult circumstances, and said they value the potential outcome of such actions, they believed that their behaviours would have a significant positive impact on the responses or future behaviours of their employees only slightly over half the time.

Thus, in resumé, the average outcome value of giving prompt and accurate negative feedback to poorly performing subordinates seems to be positive only about half the time. These supervisors appear split on the utility of such an approach. These results are consistent with the findings of Napier and Latham (1986) who found that even supervisors who have high levels of self-efficacy re: their ability to conduct performance appraisals, still experience futility, as they feel the primary consequences of such actions are neutral or negative. It may also be that the impact of long-term versus short-term consequences of prompt and accurate performance feedback are unclear. In the experience of many supervisors, the negative *short term* consequences (e.g., challenge, anger, retaliation, further deterioration of performance) may be more important, more salient, or carry greater weight than the potential beneficial *long term*

consequences (e.g., improved performance, lateral transfer, more suitable placement) through improved fit between personal resources and job demands.

Supervisors in the public service who have given prompt and accurate feedback to employees and then failed to see positive changes in performance, may feel there is little point in such action, regardless of their ability to do so, as the employee is not likely to alter his/her performance. It is possible that with the down-sizing in the public sector, the positive long-term consequences of prompt and accurate face-to-face feedback for poorly performing subordinates will outweigh the negative, more short-term effects. One might speculate that these particular results would have been different had the supervisors come from companies within the private sector. In today's economy, employment in private companies may be more contingent upon performance than it is within the public service which, until recently, was a place of long-standing employment stability.

The Impact of Training on Delay and Distortion in Performance Evaluation Feedback

Do supervisors who have had training in the use and administration of personnel evaluations in their organization complete the measures in this study any differently than those who have not? In terms of the delay and distortion variables, analyses revealed no significant differences. There was, however, a trend for those who had received such training to schedule feedback sessions somewhat sooner (1.8 days) compared to those who had not (2.6 days), perhaps realizing that delay of such feedback is only postponing the inevitable and not helpful for the individual being evaluated. Although these results approach statistical significance ($p = .07$), it is questionable as to whether or not 0.8 days constitutes a practical difference.

The Impact of Training on Appraisal Effectance Motivation

Training appeared to have an effect on supervisors' self-efficacy expectancy to provide prompt and accurate feedback to poorly performing employees. Thus from Table 22 we can see

that training appears to increase supervisors' self-confidence in their ability to learn the skills required, and to provide feedback to employees who are performing substantially below expectations. But, training did not appear to impact upon outcome expectancy -- the belief that providing such feedback would have the desired outcome (e.g., improved employee performance), nor upon outcome value. This being the case, supervisors may conclude that it is not in their best interests to provide prompt and accurate feedback to poorly performing subordinates, even though they feel they could do it.

This result may be a characteristic of the population from which the sample was drawn, and may reflect the frustration and lassitude which reportedly currently permeates much of the Federal Public Service. Such supposition was supported anecdotally in many of the post-debriefing discussions we had with supervisors during which they described their frustrations within the Federal Public Service system in general, and the evaluation and appraisal system in particular. Yet putting this into a larger perspective, employees in many organizations in addition to the Canadian Public Service have reported little confidence in and great frustration within their company's performance management systems (Bretz & Milkovitch, 1990).

That supervisors with prior training in doing performance appraisals reported greater self-efficacy expectancy, but not greater outcome expectancy or outcome value, could further pertain to other personal behaviours or characteristics of the system in which the performance ratings are embedded. Supervisors frequently do not document occasions of poor subordinate performance (e.g., Buzzotta, 1989; Jenks, 1991), nor do they give informal feedback sessions as difficulties arise (e.g., Clopton, 1990; Hunter, 1988). Thus, when such negative feedback is given in a formal annual appraisal, subordinates are often surprised and/or angry, and they may react accordingly (e.g., launching grievance proceedings). Some research has indicated that people tend to evaluate their own performance as above average; Meyer (1980) found that 75% of engineers sampled felt their performance was in the top 25%! As well, supervisors are often required by management to justify performance rated as either "below average" or "above average", and they may consider the effort and energy required to outweigh the potential benefits or value of such actions. To spare themselves these energy and time-consuming consequences, supervisors, knowingly or unknowingly, may rate poor performance as "about average" or "low

average". As a result supervisors may feel they understand and use effective techniques in making appropriate performance ratings. However, when the costs of giving prompt and accurate feedback are felt to outweigh the benefits, supervisors would likely report higher levels of self-efficacy expectancy, with lower levels of outcome expectancy (as was the case in the supervisors of this study).

Supervisor Attribution Factors

The results in Table 24 provide some support for the idea that male versus female supervisors make differential attributions (i.e., ability, effort, task difficulty, or luck) as to the reasons for the poor performance of their subordinates. Statistically significant differences were found on both "ability of the confederate" and "difficulty of the test". Let us examine the relative endorsement of internal versus external attributions made by the supervisors in general. Conventional attribution theory (e.g., Fiske & Taylor, 1991) indicates that when males succeed it is usually attributed internally to their ability and effort, while their failures are attributed externally to variables such as luck (or lack thereof) and task difficulty. Conversely, when females succeed it is often more likely to be externally attributed to luck or ease of the task, whereas their failures tend to be explained via internal attributions such as lack of ability or lack of effort.

In this instance, when supervisors were asked to rate the factors which might have influenced the performance of the subordinate (see Appendix 4), both male and female supervisors consistently provided higher ratings for ability and effort (internal attributions) than they did for the external factors, test difficulty and luck. (The average internal attribution score was = 4.12; that of the external attribution factors = 2.35). On average the supervisors distorted their evaluations of the overall confederate performance to the 43rd percentile. Thus, they would have been trying to explain a near average performance -- consistent with internal attributions, and not to explain a failure (more likely to result in external attributions).

Although these results appear contradictory at first (afterall the confederates performed poorly!), they may not be when considered within the context of the "fundamental attribution

error" (e.g., Baron and Byrne, 1991), in which the observers over-emphasize dispositional or internal factors, and actors over-emphasize situational or external factors. Transposed to the scenario in this study, and consistent with the findings, supervisors attributed what they rated as a borderline performance more to internal (ability and effort) versus external factors. There were no significant differences between male versus female supervisors on either the extent of the internal attribution (males, $\bar{M}$ =4.5; females, $\bar{M}$ = 4.3), or on the external attributions (males, $\bar{M}$ =2.6; females, $\bar{M}$ =2.2). There was no support for biological sex differences in the fundamental attribution error on the part of supervisors regarding the confederates' performances.

Supervisor Confidence and Concern Re: Accuracy and Impact of Feedback

Two possible supervisor characteristics proposed by Benedict and Levine (1988) to account for the biological sex differences found in their study, were investigated in this research. The first was the level of confidence supervisors reported having in the accuracy of their ratings of the confederates' performance. No significant differences between male and female supervisors were found on this variable. These means reflected a confidence rating in the vicinity of "5" (corresponding to "very much") on the seven-point scale. These results then fail to support Benedict and Levine's proposal that when compared to male supervisors, female supervisors have less confidence in their rating of quantitative tasks. At the same time, however, this is disconcerting given that most performance evaluations were considerably distorted (18 points on average), and therefore they were *not* accurate reflections of the work performed. Thus, despite the (false) confidence of the supervisors in this study of their rating abilities, the performance evaluation ratings were nonetheless skewed in a positive direction, on average to a considerable extent.

Supervisor "concern with impact of feedback" on the confederate was also evaluated. There were no biological sex differences; however, consistent with our rationale, significant differences were found between stereotypic gender-role types (see Table 25 & Figure 5).

Undifferentiated individuals reported lower levels of concern, followed by instrumental, with androgynous and expressive supervisors reporting the higher levels of concern. It appears that the possession of stereotypic expressive traits increased levels of concern. Non-gender schematic (i.e., undifferentiated) supervisors showed the lowest levels of concern. Interestingly, the reported concern with the impact of feedback on the confederates did not translate into significant gender differences on most of the delay and individual distortion items. However, this concern does fit with the results of the global distortion measure seen earlier (see Table 19).

Anxiety due to the face-to-face feedback requirement was also evaluated. Female supervisors reported significantly higher levels of anxiety than their male counterparts. However, once again, this higher anxiety did not translate into longer delays before scheduling a face-to-face feedback appointment. Thus, the female supervisors in this study may have been able to overcome the anxiety, possibly by drawing on another set of beliefs (e.g., "The sooner the feedback is received the better", "I want to get this over with as soon as possible").

What could account for the substantial difference between the findings of Benedict and Levine (1988) and those of this study? Although the subjects in the Benedict and Levine study reportedly had more supervisory experience (males -- $\bar{M}=3.5$ years, females -- $\bar{M}=3.2$ years) than those in this study (both males and females $\bar{M}=2.6$ years), perhaps the most striking difference is in terms of the sample population at the time of testing. Benedict and Levine used university students; this study recruited supervisors from the Federal Public Service. Other researchers (e.g., McCarrey et al., 1989) have failed to find significant differences between male and female supervisors in the work place, in comparison with studies conducted with university students. It appears that there may be a transition that occurs as one progresses from the role of student to the role of manager. Thus, those who are working as supervisors have developed the skills necessary to meet the demands of that role. Regardless of biological sex or prior gender-related socialization experiences, these participants were first and foremost *supervisors* operating in a public sector organization. Moreover, in addition, there is likely a self-selection procedure, such that those who do not possess or who can not develop those skills necessary to be adequate supervisors in such an organization either drop out of such positions, or are not promoted.

Perhaps of equal interest to the lack of powerful overall differences between either sexes or gender groupings, was the striking distortion on the part of all the supervisors, from the 25th percentile to just below the 50th -- almost 20% on a 100 point scale! In very practical terms, if such performance management behaviour is consistent with their on-the-job appraisal behaviour, subordinates of such supervisors would not be receiving accurate feedback, and, moreover, they would not be aware of their very definite short-comings. Even those with training in the area of performance appraisal did not differ significantly on most variables, although training did increase appraisal self-efficacy expectancy motivation. Nonetheless, although clearly feeling capable of providing prompt and accurate feedback, and valuing the outcome of such behaviours, they failed to do so. Why not? The reluctance to deliver bad news and the desire to avoid the anticipated potential interpersonal ramifications of delivering such feedback to subordinates, may be so strong that even those with some training are unable or unwilling to avoid distorting the results in favour of the recipient. Haccoun (1996; Jeanrie & Haccoun, 1996) notes there is abundant evidence that performance often remains unchanged following training, on account of no transfer of training, questionable validity, unstable criteria of training outcome, and dismal or untested content validity.

This study focused on an expansion of the work of Benedict and Levine (1988) by increasing the number of variables linked empirically and logically to differential delay and distortion scores found previously between male and female subjects required to provide negative feedback. Due to limitations in time and resources, certain variables that might have provided a more complete assessment were not studied. First, a "no feedback requirement" was not included. Although the research is fairly conclusive that requiring face-to-face feedback increases avoidant behaviour (e.g., Ilgen & Knowlton, 1980), inclusion of such an element would have helped confirm that it is the feedback requirement per se that results in or exacerbates the distortion. Future investigation needs to compare feedback modality: 1) evaluation with the confederate receiving no feedback, 2) with the confederate's results being mailed, 3) with face-to-face feedback, and 4) requiring the supervisors to actually provide feedback. It would have been interesting to assess the correlation between the evaluation ratings they placed on paper and what was actually said. As well, content analysis of the feedback

would allow for study of how supervisors actually provide such feedback.

Finally, only male confederates were included in this study. Research including both male and females would allow investigation of the impact of biological sex of a subordinate on delay and distortion of performance feedback. Indeed, does negative distortion of incredibly high performance take place? Is this performance more attributed to luck or to ease of task when male versus female subordinates are studied?

This study focused on the social cognition and motivation of male and female supervisors when expected to provide face-to-face feedback after observing and scoring a poor (25th percentile) subordinate performance on a quantitative task requiring both speed and accuracy. Our point of departure was the work of Benedict and Levine (1988) who found, using business students, that all distorted and that females did so significantly more than males. Using essentially the same task and similar demand characteristics with actual supervisors in a quantitative Canadian public sector organization, we found overwhelming evidence of *similarity* in the evaluation of poor work behaviours across sex of supervisors. We found that all supervisors (irrespective of biological sex or stereotypic gender-related skill type) positively distorted their evaluation of poor subordinate performance.

From our investigation of supervisor beliefs and motivation when they knew they would have to actually provide prompt and accurate face-to-face feedback in poor performance situations, it seems possible to draw several conclusions. First, most supervisors felt they could learn and/or use the strategies required to provide such feedback. Second, although the value and utility (e.g., in terms of long-term subordinate career development) were positive in principle, nonetheless supervisors' expectancy that they would actually obtain such positive outcomes was negative. It is possible, even likely, that supervisors did not analytically distinguish positive long-term versus negative, disruptive short-term outcomes of prompt and accurate face-to-face feedback. The potentially more disruptive short-term interpersonal consequences (e.g., surprise, anger and retaliation) may have been more salient than the potentially more beneficial long-term gains.

In addition, we speculate supervisors may not separate process from outcome in giving

face-to-face feedback in poor performance situations. They likely avoid carrying out a difficult feedback task because they have limited or no experience in providing information that is unpleasant and threatening, in a manner that is supportive and provides alternative possibilities.

Despite some biological sex and gender outcomes which did support differences along these dimensions, the majority provided a picture of essential parallel across sex and gender both in the actual distortion and delay results, and in related beliefs, attributions and motivational expectancies.

It is our conviction that the participants in this organization were supervisors first, and males and females second. We speculate that a self-selection and promotional process has resulted in a group of people who have the necessary interests and skills to enact (most) required supervisory role behaviours. Those females and males who, by socialization, were poorly equipped to meet the requirements of such supervisory jobs have likely chosen, or been selected for, non-supervisory work, with the result that they are not present in our group of participants. The notion that male and female supervisors have more in common than not, is supported in the organizational literature (e.g., Davidson & Cooper, 1987; Dobbins & Platz, 1986; Rice, Instove & Adams, 1984; Ritchie & Moses, 1983; Simpson, McCarrey, & Edwards, 1987)) where participants in the research were in organizational subordinate and supervisory roles, and not that of students .

Contributions

In general terms, the contribution of this study lies in its divergent findings in comparison to those of Benedict and Levine (1988). Limits to what they found using a particular methodology are suggested by our study using the same method but on a different group of participants. We suggest the sex differences they found may be due to the student status of their participants, albeit with previous supervisory experience. A difference was that at the time of testing, their subjects were students in university taking business courses.

A second contribution is the indication of the magnitude of motivational distortion in performance evaluation judgements. The key to this seems to be, as noted by Napier & Latham

(1986), that even in the 1990's many supervisors, at least in this public service organization, are of the conviction that accurate face-to-face reporting of poor subordinate performance is not worth the expected negative consequences.

Conclusions

1. This study did not support the overall findings of Benedict and Levine (1988) that female supervisors delay and distort more than their male counterparts when required to provide face-to-face feedback to a poorly performing subordinate. Our findings support the work of other researchers (e.g., Davidson & Cooper, 1987; Dobbins & Platz, 1986; McCarrey et al., 1989; Rice, Instove & Adams, 1984; Ritchie & Moses, 1983; Simpson, McCarrey, & Edwards, 1987) which found few overall differences between supervisors on the basis of biological sex. Thus, although there may be sex differences when subjects are business students, these did not appear in our group of actual supervisors from an actual public sector organization.

2. Some support was found for the hypothesis that stereotypic gender-role orientation, as opposed to biological sex, could account for some previously measured sex differences in performance evaluations. Although all supervisors were found to distort results, expressive individuals (those endorsing feminine stereotypic characteristics) were more likely to positively distort the quality of their subordinate's poor performance than did their more instrumental counterparts.

As well, expressive supervisors produced higher composite distortion scores. This suggests that although there may be no significant differences when each individual item of feedback distortion is examined, when they are summed and considered together, there is a considerable cumulative effect in the results of the expressive supervisors as opposed to their instrumental counterparts.

3. In general, notable distortion was found to exist. Even those supervisors reporting high levels of self-efficacy expectancy to deliver prompt and accurate feedback still distorted their

evaluations. Present organizational training appeared to have a limited impact on the tendency to distort performance judgements. It also did not increase outcome expectancy nor outcome value. Thus regardless of their sex, gender or training background, supervisors still illustrated a strong tendency to distort negative feedback results. Further research is needed in this area.

REFERENCES

- Arch, E.C. (1992). Sex differences of self-efficacy on willingness to participate in a performance situation. Psychological Reports, 70, 3-9.
- Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. New York: Prentice-Hall.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37, 122-147.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. Psychological Review, 84, 191-215.
- Baron, R.A., & Byrne, D. (1991). Social Psychology: Understanding Human Interaction. MA: Allyn and Bacon.
- Bean, L.D. (1977). Interpersonal perception among college students as a function of sex and level of assertiveness. Dissertation Abstracts International, 38, 1292A.
- Bem, S.L. (1981.) Bem Sex Role Inventory: Professional manual. CA:Consulting Psychologists Press Inc.
- Bem, S.L. (1979). Theory and measurement of androgyny: A Reply to the Pedhauzer-Tetenbaum and Lockley-Colten critiques. Journal of Personality and Social Psychology, 37, 1047-1054.
- Bem, S.L. (1974). The measurement of psychological androgyny. Journal of Consulting and Clinical Psychology, 42, 155-162.
- Bem, S.L., & Lenney, E. (1976). Sex-typing and the avoidance of cross-sex behaviour. Journal of Personality and Social Psychology, 33, 48-54.
- Benedict, M.E., & Levine, E.L. (1988). Delay and distortion: Tacit influences on performance appraisal effectiveness. Journal of Applied Psychology, 73, 507-514.
- Betz, N.E., & Hackett, G. (1981). The relationship of career-related self-efficacy expectations to perceived career options in college women and men. Journal of Counselling Psychology, 28, 399-410.
- Bledsoe, J.C. (1983). Factorial validity of the Bem Sex Role Inventory. Perceptual and Motor Skills, 56, 55-58.

- Bond Jr., C.F., & Anderson, E.L. (1987). The reluctance to transmit bad news: Private discomfort or public display? Journal of Experimental Social Psychology, 23, 176-187.
- Bretz, B., & Milkovitch, G. (1990). The current state of performance appraisal research and practice: Identifying the gap. Centre for Advanced Human Resource Studies: Cornell University.
- Buzzotta, V.R. (1989). Improve your performance appraisals. Security Management, 33, 94-100.
- Cano, L., Soloman, S., & Homes, D.S. (1984). Fear of success: The influence of sex, sex-role identity and components of masculinity. Sex Roles, 10, 341-346.
- Clopton, J.D. (1990). The performance appraisal: T.G.I.O. (Thank God it's over). Sales and Marketing Management, 142, 124.
- Cohen, J. (1988). Statistical power for the behavioral sciences (2nd ed.). NJ: Lawrence Erlbawn Assoc.
- Cohen, S.L., & Jaffee, C.L. (1982). Managing human performance for productivity. Training and Development Journal, 36, 94-100.
- Conlee, M.C., & Tesser, A. (1973). The effects of recipient desire to hear news transmission. Sociometry, 36, 588-599.
- Cunningham, J.D., & Anthill, J.K. (1980). A comparison among five masculinity-femininity-androgyny instruments and two methods of scoring androgyny. Australian Psychologist, 15, 437-448.
- Davidson, M.J., & Cooper, C.L. (1987). Female managers in Britain: A comparative study. Human Resource Management, 26(2), 217-242.
- Davis, F.W., & Yates, B.T. (1982). Self-efficacy expectancies versus outcome expectancies as determinants of performance deficits and depressive affect. Cognitive Therapy and Research, 6, 23-35.
- Deaux, K. (1974). Anonymous altruism: Extending the lost letter technique. Journal of Social Psychology, 92, 61-66.
- Dobbins, G.H., Cardy, R.L. & Truxillo, J. (1988). The effects of purpose of appraisal and individual differences in stereotypes of women on sex differences in performance ratings: A laboratory and field study. Journal of Applied Psychology, 73, 551-558.

- Erikson, E.H. (1959). Identity and the life cycle. Psychological Issues, No.1. N.Y.: International Universities Press.
- Fiske, S.T., & Taylor, S.E. (1991). Social Cognition (3rd ed.). NY: McGraw-Hill
- Gaa, J.P., & Liberman, D. (1981). Categorization agreement of the Personality Attributes Questionnaire and the Bem Sex Role Inventory. Journal of Clinical Psychology, 37, 593-601.
- Gaudreau, P. (1977). Factor analysis of the Bem Sex Role Inventory. Journal of Consulting and Clinical Psychology, 45, 299-302.
- Gruber, K.J., & Powers, W.A. (1982). Factor and discriminant analysis of the Bem Sex Role Inventory. Journal of Personality Assessment, 46, 284-291.
- Haccoun, R.R. (1996, August). Transfer of training: The results of four field experiments. Poster session presented at the XXVI International Congress of Psychology.
- Hunter, W.L. (1988). Relieving the pain of performance appraisals. Management World, 17, 7-9.
- Ilgen, D.R., & Knowlton Jr., W.A. (1980). Performance attribution effects. Organizational Behaviour and Human Performance, 25, 441-456.
- Jeanrie, Ch., & Haccoun, R.R. (1996, August). Transfer of training: An extended view. Poster session presented at the XXVI International Congress of Psychology.
- Jenks, J. (1991). Do your performance appraisals boost productivity? Management Review, 80, 45-47.
- Jennings, S.R. (1985). The relationship of personality factors to social anxiety and disaffiliation. Unpublished doctoral dissertation, University of Texas, Austin.
- Kelly, J.A., Kern, J.M., Kirkley, B.G., Patterson, J.N., & Keane, T.M. (1980). Reactions to assertive versus unassertive behaviour: Differential effects for males and females and implications for assertive training. Behavior Therapy, 11, 670-682.
- Kelly, J.A., & Worell, J. (1977). New formulations of sex roles and androgyny: A critical review. Journal of Consulting and Clinical Psychology, 45, 1101-1115.
- Kern, J.M., Cavell, T.A., & Beck, B. (1985). Predicting differential reactions to males' versus females' assertions, empathic-assertions, and nonassertions. Behaviour Therapy, 16, 63-75.

- Landino, R.A., & Owen, S.V. (1988). Self-efficacy in university faculty. Journal of Vocational Behavior, 33, 1-14.
- Landy, F.J., & Farr, J. (1980). Performance rating. Psychological Bulletin, 87, 72-107.
- Leary, M., & Atherton, S.C. (1986). Self-efficacy, social anxiety, and inhibition in interpersonal encounters. Journal of Social and Clinical Psychology, 4, 256-267.
- Leary, M.R., Knight, P.D., & Johnson, K.A. (1987). Social anxiety and dyadic conversation: A verbal response analysis. Journal of Social and Clinical Psychology, 5, 34-50.
- London, M., & Poplawski, J.R. (1976). Effects of information of stereotype development in performance appraisal and interview context. Journal of Applied Psychology, 61, 199-205.
- Long, B.C. (1989). Sex-role orientation, coping strategies, and self-efficacy of women in traditional and nontraditional occupations. Psychology of Women Quarterly, 13, 307-324.
- MacDonald, M.L. (1982). Assertion training for women. In J.P. Curran & P.M. Monti (Eds.) Social skills training: A practical handbook for assessment and treatment (pp. 253-279). New York: Guildford.
- Maddux, J.E., Norton, L.W., & Leary, M.R. (1988). Cognitive components of social anxiety: An investigation of the integration of self-presentation theory and self-efficacy theory. Journal of Social and Clinical Psychology, 6, 180-190.
- Maddux, J.E., Norton, L.W., & Stoltenberg, C.D. (1986). Self-efficacy expectancy, outcome expectancy, and outcome value: Relative effects on behavioural intentions. Journal of Personality and Social Psychology, 51, 783-789.
- Maddux, J.E., & Rogers, R.W. (1983). Protection, motivation and self-efficacy: A revised theory of fear appeals and attitude change. Journal of Experimental Social Psychology, 19, 469-479.
- Maddux, J.E., Sherer, M., & Rogers, R.W. (1982). Self-efficacy expectancy and outcome expectancy: Their relationship and their effects on behavioral intentions. Cognitive Therapy and Research, 6, 207-211.
- Maddux, J.E., & Stanley, M.A. (1986). Self-efficacy theory in contemporary psychology: An overview. Journal of Social and Clinical Psychology [Special Issue], 4, 249-255.

- Manning, M.M., & Wright, T.L. (1983). Self-efficacy expectations, outcome expectancies and the persistence of pain control in childbirth. Journal of Personality and Social Psychology, 45, 421-431.
- Matsui, T. (1994). Mechanisms underlying sex differences in career self-efficacy expectations of university students. Journal of Vocational Behavior, 45, 177-184.
- Matsui, T., Ikeda, H., & Ohnishi, R. (1989). Relations of sex-typed socialization to career self-efficacy expectations of college students. Journal of Vocational Behavior, 35, 1-16.
- McCarrey, M., Ahmed, S., Gasse, Y., Conrad, G., Seguin, S., Major, M., & Mercier, P. (1989). The subjective culture of public sector women and men managers: A common instrumental/expressive value orientation or, two different worlds? Canadian Journal of Administrative Sciences, 6, 54-62.
- Meyer, H.H. (1980). Self-appraisal of job performance. Personnel Psychology, 33, 291-295.
- Mitchell, T.R., & Beach, L.R. (1976). A review of occupational preferences and choice research using expectancy theory and decision theory. Journal of Occupational Psychology, 49, 231-248.
- Moreland, J.R., Gulanick, N., Montague, E.K., & Harren, V.A. (1978). Some psychometric properties of the BSRI. Applied Psychological Measurement, 2, 249-256.
- Morris, K.J. (1985). Discriminating depression and social anxiety: Self-efficacy analysis. Unpublished master's thesis. Texas Tech, University.
- Napier, N.K., & Latham, G.P. (1986). Outcome expectancies of people who conduct performance appraisals. Personnel Psychology, 39, 827-837.
- O'Connor, K., Mann, D.W., & Bardwick, J.J. (1978). Androgyny and self-esteem in the upper middle class: A replication of Spence. Journal of Consulting and Clinical Psychology, 46, 1168-1169.
- Orlofsky, J.L. (1977). Sex-role orientation, identity formation, and self-esteem in college men and women. Sex-Roles, 3, 561-575.
- Padgett, M.Y. (1988). Performance appraisal in context: Motivational influences on performance ratings (Doctoral dissertation, Michigan State University, 1988). Dissertation Abstracts International, 49A/09, 2728A (DA 8824882).
- Pedhauzer, E.J., & Tetenbaum, T.J. (1979). Bem Sex Role Inventory: A theoretical and methodological critique. Journal of Personality and Social Psychology, 37, 996-1016.

- Public Service Commission of Canada. (1990). Annual Report 1989 (Cat. No. Sc 1-1989, ISBN 0-662-S7197-5, PSC 3X2 D906). Ottawa, ON: Supply and Services Canada.
- Rice, R.W., Bender, L.R., & Vitters, A.G. (1980). Leader sex, followers attitudes towards women, and leadership effectiveness: A laboratory experiment. Organizational Behavior and Human Performance, 25, 46-78.
- Ritchie, R.J., & Moses, J.L. (1983). Assessment center correlates of women's advancement into middle management: A 7-year longitudinal analysis. Journal of Applied Psychology, 68, 227-231.
- Sadd, S., Lenauer, M., Shaver, P., & Dunivant, N. (1978). Objective measurement of fear of success and fear of failure: A factor analytic approach. Journal of Consulting and Clinical Psychology, 46, 405-416.
- Sadd, S., Miller, F.D., & Zeitz, B. (1979). Sex roles and achievement conflicts. Personality and Social Psychology Bulletin, 5, 352-355.
- Schlenker, S.R., & Leary, M.R. (1982). Social anxiety and self-presentation: A conceptualization and model. Psychological Bulletin, 92, 641-669.
- Schoen, L.G., & Winocur, S. (1988). An investigation of the self-efficacy of male and female academics. Journal of Vocational Behaviour, 32, 307-320.
- Schroeder, J.E., Rakos, R.F., & Moe, J. (1983). The social perception of assertive behavior as a function of response class and gender. Behavior Therapy, 16, 63-75.
- Sherer, M., Maddux, J.E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R.W. (1982). The Self-Efficacy Scale: Construct and validation. Psychological Reports, 51, 663-671.
- Simpson, S., McCarrey, M., & Edwards, H.P. (1987). Relationship of supervisors' sex-role stereotypes to performance evaluation of male and female subordinates in non-traditional jobs. Canadian Journal of Administration Sciences, 4, 15-30.
- Singer, M. (1989). Gender differences in leadership aspirations. New Zealand Journal of Psychology, 18, 25-35.
- Spence, J.T., & Helmreich, R. (1980). Masculine instrumentality and feminine expressiveness: Their relationship with sex role attitudes and behaviors. Psychology of Women Quarterly, 5, 147-163.

- Spence, J.T., & Helmreich, R. (1972). The Attitudes Towards Women Scale: An objective instrument to measure attitudes toward the right and roles of women in contemporary society. Journal Supplement Abstract Service, 2, 66.
- Spence, J.T., Helmreich, R., & Stapp, J. (1975a). Ratings of self and peers on sex role attributes and their relation to self conceptions of masculinity and femininity. Journal of Personality and Social Psychology, 32, 29-39
- Spence, J.T., Helmreich, R., & Stapp, J. (1975b). The Personal Attributes Questionnaire: A measure of sex role stereotypes and masculinity and femininity. [Microfilm]. Princeton, NJ: ETS Collection , Educational Testing Services.
- Stickel, S., & Bonet, R.M. (1991). Gender differences in career self-efficacy: Combining a career with home and family. Journal of College Student Development, 32, 297-301.
- Tabachnick, B.G., & Fidell, L.S. (1989). Using Multivariate Statistics (2nd ed.). New York: Haper & Row.
- Tesser, A., & Rosen, S. (1975). The reluctance to transmit bad news. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (pp. 193-232). New York: Academic Press.
- Tesser, A., Rosen, S., & Batchelor, T. (1972). On the reluctance to communicate bad news: A role play extension. Journal of Personality, 40, 88-103.
- Tesser, A., Rosen, S., & Tesser, M. (1971). On the reluctance to communicate bad undesirable messages: A field study. Psychological Reports, 29, 651-654.
- Vasil, L. (1993). Gender differences in the academic career in New Zealand universities. New Zealand Journal of Educational Studies, 28, 143-153.
- Waterman, A.S., & Witbourne, S.K. (1982). Androgyny and psychosocial development among college students and adults. Journal of Personality, 50, 121-133.
- Waters, C.E., Waters, L.K., & Pincus, S. (1977). Factor analysis of masculine and feminine sex-typed items for the Bem Sex Role Inventory. Psychological Reports, 40, 567-5700.
- Waters, L.K., & Popovich, P.M. (1986). Factor analysis of sex-typed items from the Bem Sex-Role Inventory: A multiple replication across time. Psychological Reports, 59, 1323-1326.

APPENDIX 1.

Bem Sex Role Inventory - Short Form (BSRI-S)

INSTRUCTIONS

On the following page, you will be shown a number of personality characteristics. We would like to use those characteristics in order to describe yourself. That is, we would like you to indicate on a scale from 1 to 7, how true of you these various characteristics are. Responses are to be written on the computer answer sheet provided by filling in the appropriate numbered space corresponding to each characteristic. Please do not leave any characteristics unmarked.

EXAMPLE: Sly

Blacken in 1 if it is NEVER OR ALMOST NEVER TRUE that you are sly.

Blacken in 2 if it is USUALLY NOT TRUE that you are sly.

Blacken in 3 if it is SOMETIMES BUT INFREQUENTLY TRUE that you are sly.

Blacken in 4 if it is OCCASIONALLY TRUE that you are sly.

Blacken in 5 if it is OFTEN TRUE that you are sly.

Blacken in 6 if it is USUALLY TRUE that you are sly.

Blacken in 7 if it is ALWAYS OR ALMOST ALWAYS TRUE that you are sly.

Thus, if you feel that it is sometimes but infrequently true that you are "sly", never or almost never true that you are "malicious", and often true that you are "carefree", then you would answer as follows:

QUESTION #:

ANSWER SHEET:[illegible]

DESCRIBE YOURSELF

1	2	3	4	5	6	7
Never or almost true	Usually not true	Sometimes but infrequently true	Occa- sionally true	Often true	Usually true	Always or almost always true
1. Defends own beliefs		11. Jealous		21. Dominant		
2. Moody		12. Has Leadership Abilities		22. Warm		
3. Independent		13. Sensitive to the Needs of Others		23. Willing to Take A Stand		
4. Conscientious		14. Truthful		24. Tender		
5. Affectionate		15. Willing to Take Risks		25. Aggressive		
6. Assertive		16. Understanding		26. Adaptable		
7. Strong Personality		17. Secretive		27. Loves Children		
8. Forceful		18. Compassionate		28. Tactful		
9. Reliable		19. Eager to Soothe Hurt Feelings		29. Gentle		
10. Sympathetic		20. Conceited		30. Conventional		

APPENDIX 2.

Supervisory Control Worksheet
Computer Science Aptitude Test - Norms
Computer Science Aptitude Test - Test Booklet
Computer Science Aptitude Test - Answer Sheets 1 & 2
Supervisor Qualifications Sheet - Pilot Study
Supervisor Qualifications Sheet - On Site Study

SUPERVISORY CONTROL WORKSHEET

DATE: _____

SUPERVISOR CODE: _____

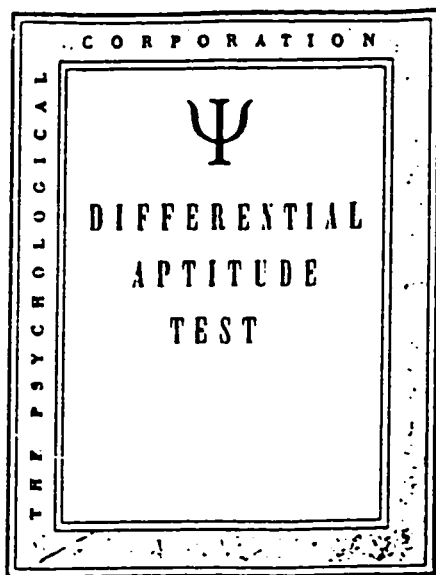
VOLUNTEER CODE: _____

	<u>RAW SCORE</u>	<u>PERCENTILE RANK</u>
Work Period 1:	_____	_____
Work Period 2:	_____	_____
Work Period 3:	_____	_____
Work Period 4:	_____	_____
Work Period 5:	_____	_____
Work Period 6:	_____	_____
Work Period 7:	_____	_____
Work Period 8:	_____	_____
Work Period 9:	_____	_____
Work Period 10:	_____	_____

COMPUTER SCIENCE APTITUDE TEST NORMS
FORMS 1 THROUGH 10 PERCENTILE NORMS

AUTUMN, 1989

<u>PERCENTILE</u>	<u>RAW SCORE</u>	<u>PERCENTILE</u>
99	80-100	99
75	52-60	75
50	49-50	50
25	40-45	25
1	0-25	1



BOOKLET 1

COMPUTER SCIENCE APTITUDE TEST

Do not open this booklet until you are told to do so. Wait for the examiner's instructions. Then:

On the SEPARATE ANSWER SHEET, print your name and fill in the other requested information in the proper spaces.

Then wait for further instructions.

DO NOT MAKE ANY MARKS IN THIS BOOKLET

Do not make
any marks in
this booklet

COMPUTER SCIENCE APTITUDE TEST

Mark your answers
on the separate
Answer Sheet

DIRECTIONS

Find the Answer Sheet for Part I of COMPUTER APTITUDE TEST . Print your name and fill in the other requested information in the proper spaces.

This is a test to see how quickly and accurately you can compare letter and number combinations. On the following pages are groups of these combinations; each Test Item contains five. These same combinations appear after the number for each Test Item on the separate Answer Sheet, but they are in a different order. You will notice that in each Test Item one of the five is underlined. You are to look at the one combination which is underlined, find the same one after that item number on the separate Answer Sheet, and fill in the space under it.

These examples are correctly done. Note that the combination on the Answer Sheet must be exactly the same as the one in the Test Item.

TEST ITEMS

V.	<u>AB</u>	AC	AD	AE	AF
W.	aA	aB	BA	Ba	<u>Bb</u>
X.	A7	7A	B7	<u>7B</u>	AB
Y.	Aa	Ba	<u>bA</u>	BA	bB
Z.	3A	3B	<u>33</u>	B3	BB

SAMPLES OF ANSWER SHEETS

	AC	AE	AF	AB	AD
V.	:	:	:		:
	BA	Ba	Bb	aA	aB
W.	:	:		:	:
	7B	B7	AB	7A	A7
X.		:	:	:	:
	Aa	bA	bB	Ba	BA
Y.	:		:	:	:
	BB	3B	B3	3A	33
Z.	:	:	:	:	

	AC	AE	AF	AB	AD
V	○	○	○	●	○
	BA	Ba	Bb	aA	aB
W	○	○	○	○	○
	7B	B7	AB	7A	A7
X	○	○	○	○	○
	Aa	bA	bB	Ba	BA
Y	○	○	○	○	○
	BB	3B	B3	3A	33
Z	○	○	○	○	○

If you finish the items in Part I before time is called, check your work. Do not turn to Part II until you are told to do so. Work as fast as you can.

You will have 3 minutes for each part of this test. Work as rapidly and as accurately as you can. If you are not sure of an answer, mark the choice which is your best guess.

DO NOT TURN OVER THE BOOKLET UNTIL YOU ARE TOLD TO DO SO.

DO NOT TURN THE PAGE UNTIL YOU ARE TOLD TO DO SO.

There are two Answer Sheets for the next test. Wait for further instructions from the examiner.

PART I

Appendix 2.
Page 2 - 7

1. <u>nv</u> nx xn vx xv	21. ar ra <u>ro</u> or oa	41. 7c 9b <u>9c</u> 9e 7b	61. HN HZ ZH ZN <u>NH</u>	81. 35 53 h3 <u>3h</u>
2. bl dl ld <u>lb</u> bd	22. lc lo ol <u>oc</u> co	42. 7c <u>2b</u> 7b 2d 7d	62. RR BR RB <u>BB</u> RP	82. bl <u>dl</u> ld lb
3. ar au ur ra <u>ru</u>	23. ls l3 3l 3s <u>sl</u>	43. <u>n3</u> Sn 3s ns 3n	63. CU <u>UU</u> UC US CC	83. lk <u>lk</u> kf lf
4. <u>wu</u> vu vw wv uw	24. ma cm ca <u>mc</u> am	44. 20 <u>25</u> 02 05 52	64. PR PB RB <u>RP</u> BP	84. 69 6d 9d c6
5. wm um mu wu <u>mw</u>	25. xv vx <u>vw</u> wx wv	45. ec ac <u>ca</u> ce ae	65. CK KJ JC KC <u>JK</u>	85. XX VX <u>VZ</u> ZV
6. <u>79</u> 76 67 69 97	26. <u>ud</u> un nd nu du	46. 2h h4 42 <u>4h</u> 24	66. Tl lT <u>ll</u> Tt TT	86. <u>j8</u> a8 8a 8j
7. ra <u>na</u> nr rn ar	27. fk <u>lk</u> kf lf kl	47. <u>av</u> va vo ao ov	67. SX sX sx <u>Xs</u> XS	87. 79 76 <u>67</u> 69
8. za mz <u>zm</u> az ma	28. pq <u>qg</u> gp gq qp	48. fa <u>fr</u> ra rf ar	68. LT Tt <u>lT</u> Tl tt	88. nr <u>ne</u> en rn
9. AV VN NV <u>NA</u> VA	29. 2u 2q <u>qu</u> q2 u2	49. <u>ma</u> cm ca mc am	69. Zz NZ zZ zn <u>ZN</u>	89. 4X 4V <u>Vx</u> V4
10. OQ <u>CQ</u> QC QO OC	30. 4l 44 <u>l4</u> ll 40	50. rc cr <u>co</u> oc or	70. GQ Qg <u>qq</u> qg QG	90. vn vz zv nv <u>z</u>
11. <u>CU</u> UU UC US CC	31. nr ne en rn <u>re</u>	51. <u>ch</u> ho hc oc oh	71. 4c <u>la</u> lc 4d 2d	91. B3 R8 <u>8B</u> RB E
12. 4H 4N NH N4 <u>HN</u>	32. bb <u>dd</u> ld db bd	52. se <u>rs</u> re es er	72. <u>S8</u> C3 S3 C8 C5	92. <u>OQ</u> CQ QC QO C
13. Rr RP <u>pR</u> PP rr	33. RB <u>RD</u> DR BR BD	53. <u>ar</u> au ur ra ru	73. A9 <u>7b</u> 79 9b b7	93. OD OB <u>BD</u> DO B
14. Aa A8 8a <u>8A</u> aA	34. MW MV VW <u>VM</u> WM	54. pq qg <u>gp</u> gq qp	74. l8 8l <u>7l</u> 78 17	94. ZY ZX XY <u>YZ</u> Y
15. LT Tl <u>lT</u> Tl tt	35. OD OB BD <u>DO</u> BO	55. am na nm <u>mn</u> an	75. b4 4d db d4 <u>bd</u>	95. OU OC <u>UC</u> UO C
16. Av <u>Vv</u> av VV AA	36. PR <u>PB</u> RB RP BP	56. gj <u>jg</u> pg jp gp	76. <u>u6</u> u4 4u 6u 46	96. Cc Oc OO cO <u>c</u>
17. 4d 3c <u>4a</u> 4c 3a	37. Dd Db <u>dB</u> bB DD	57. <u>tp</u> et ep pe pt	77. 3x 7x <u>73</u> 37 x7	97. Aa <u>A8</u> 8a 8A a
18. X7 V9 V5 X9 <u>V7</u>	38. EE Ef eF Fe <u>FF</u>	58. ra na nr rn <u>ar</u>	78. ls l3 3l 3s <u>sl</u>	98. Ze Zz <u>ZE</u> zE e
19. <u>A9</u> 7b 79 9b b7	39. Ze Zz <u>ZE</u> zE eZ	59. bb dd ld <u>db</u> bd	79. <u>en</u> dn de ed nd	99. BP Pb bP <u>p2</u> bE
20. <u>20</u> 5 02 05 52	40. <u>Zz</u> NZ zZ zn ZN	60. l8 <u>8l</u> la 8a a8	80. ni fi <u>fn</u> in nf	100. <u>Cz</u> CZ Zc zC c

STOP. YOU MAY CHECK YOUR WORK ON THIS TEST. DO NOT TURN TO ANY OTHER TEST.

PART II

Appendix 2.
Page 2 - 8

1. YZ VY <u>VX</u> XY ZY	21. Rr RP <u>pR</u> PP rr	41. wu vu vw wv <u>uw</u>	61. <u>18</u> 81 71 78 17	81. <u>se</u> rs re es
2. b9 c6 69 96 <u>6c</u>	22. LT IT IL <u>TL</u> TI	42. <u>er</u> ri ir ie re	62. Vv Ww Wv <u>wV</u> vv	82. <u>4X</u> 4V VX V4
3. ou <u>oa</u> ua uo ao	23. MW <u>MV</u> VW VM WM	43. <u>31</u> 23 32 13 21	63. Mm <u>MN</u> NN nn mM	83. zn zz nz nn
4. lc lo ol oc <u>co</u>	24. Uu Wu uW <u>WW</u> uU	44. 2u <u>2q</u> qu q2 u2	64. b9 c6 69 96 <u>6c</u>	84. LT IT <u>IL</u> TL
5. X7 <u>V9</u> V5 X9 V7	25. 3x xc c3 cx <u>3c</u>	45. xv <u>vx</u> vw wx wv	65. <u>4c</u> 1a 1c 4d 2d	85. 41 44 14 <u>11</u>
6. <u>Sc</u> 8c 8s cS c8	26. AV VN NV <u>NA</u> VA	46. ae et <u>ea</u> ta te	66. 2h h4 <u>42</u> 4h 24	86. us ue se su
7. ob bl ol tb <u>bo</u>	27. YX XX Yy Xy <u>xX</u>	47. <u>VI</u> SI SV VS IV	67. <u>YZ</u> VY VX XY ZX	87. PR <u>PB</u> RB RP
8. 5e 3d <u>4d</u> 2e 2d	28. EL <u>FL</u> FE LF LE	48. th <u>he</u> et eh ht	68. n3 Sn 3s <u>ns</u> 3n	88. <u>Rr</u> RP pR PP
9. rc dc <u>dr</u> rd cr	29. MN NM VN MV <u>NV</u>	49. za mz zm <u>az</u> ma	69. wo ro <u>rw</u> ow wr	89. SX <u>sX</u> sx Xs
10. ws sw <u>st</u> tw ts	30. EE Ef eF <u>Fe</u> FF	50. sx sa ax xs <u>xa</u>	70. ar <u>ra</u> ro or oa	90. ra <u>na</u> nr rn
11. wm um mu wu <u>mw</u>	31. S8 <u>C8</u> 8C 8S S5	51. Av <u>Vv</u> aV VV AA	71. ni fi fn <u>in</u> nf	91. OU OC UC UO C
12. pp qq <u>pq</u> pg qp	32. h6 h8 <u>86</u> 8h 6h	52. Mw wW <u>WM</u> MM mW	72. wu vu vw wv <u>uw</u>	92. RB RD DR <u>BR</u> B
13. nv nx xn vx <u>xv</u>	33. <u>4d</u> 3c 4a 4c 3a	53. 4H <u>4N</u> NH N4 HN	73. th he <u>et</u> eh ht	93. XX XO OO <u>OX</u> O
14. nu un <u>um</u> mn mu	34. Z4 Z1 14 <u>1Z</u> 4Z	54. <u>Dd</u> Db dB bB DD	74. am na <u>nm</u> mn an	94. HN <u>HZ</u> ZH ZN N
15. zn zz <u>nz</u> nn mn	35. <u>Qo</u> Qq OQ oq QQ	55. S8 83 S3 38 <u>3S</u>	75. 3x <u>7x</u> 73 37 x7	95. Av <u>Vv</u> aV VV A
16. <u>pg</u> gy py yp yg	36. xc ex <u>ec</u> ce xe	56. XO OO OX OV <u>XX</u>	76. <u>j8</u> a8 8a 8j ja	96. <u>OQ</u> CQ QC QO O
17. <u>59</u> 9Y 5Y Y9 95	37. <u>ar</u> ra ro or oa	57. S8 C8 <u>8C</u> 8S S5	77. 59 <u>9Y</u> 5Y Y9 95	97. Ze Zz <u>ZE</u> zE eZ
18. nu <u>on</u> ou un uo	38. 8c 8a 7a 6c <u>7c</u>	58. X7 <u>V9</u> V5 X9 V7	78. fk lk <u>kf</u> lf kl	98. <u>GQ</u> Qg qq qg Q
19. ud un nd <u>nu</u> du	39. us ue se <u>su</u> eu	59. L7 <u>L1</u> 17 1L 7L	79. <u>ma</u> cm ca mc am	99. Mm MN NN <u>nn</u> m
20. 41 <u>44</u> 14 11 40	40. <u>wo</u> ro rw ow wr	60. RB RD <u>DR</u> BR BD	80. nv nx xn vx <u>xv</u>	100. Qo Qq <u>OQ</u> oq QQ

STOP. YOU MAY CHECK YOUR WORK ON THIS TEST. DO NOT TURN TO ANY OTHER TEST.

COMPUTER SCIENCE APTITUDE TEST

ANSWER SHEET

A

FORM

76 46 6u u4 u6 4u
77 3x 37 7x 73 x7
78 13 61 31 3s 1s
79 ed nd de en dn
80 in fi ni fo nf

81 53 3h b3 sh 3s
82 bd dl lb bl ld
83 lf fk kl kf lk
84 9d dh d9 69 6d
85 XX VZ VX XV ZV

86 ja 8j a8 j8 8e
87 69 76 67 79 97
88 nr rn re en ne
89 4X 4V V4 Vx X4
90 zv zn nv vz vn

91 8R R8 B8 R8 8B
92 QO OQ CQ QC OC
93 DO OD BO BD OB
94 ZX YZ XY YX
95 OC OU CO UC UO

Appendix 2.
Page 2 - 9

96 OO Oe ee eo Ce Cc
97 Aa Ab 8a 8A aA
98 Ze zE eZ zE Ze
99 bp bP lP DP Pb pp

DO NOT WRITE IN THIS SPACE

RAW SCORE

PERCENTILE

NOMMS USED

PART I

26 nu ud du nd un
27 lf lk kl fk kf
28 qp qf pf pq pf
29 qu qu u2 q2 2q
30 40 14 Tt 41 41

BE SURE YOUR MARKS ARE HEAVY AND BLACK.

31 ne en rn nr re
32 bd bb bd dd db
33 BR BD RB RD DR
34 MV VM MW VW WM
35 BO BD DO OD OB

36 PB RB RP RP PR
37 bB Bd Db dB DD
38 EE Ee ef ff Ff
39 eZ Ze zE Zz ZE
40 zn zZ ZN Zz NZ

ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

41 7c 9a 9b 9c 7b
42 7b 2b 2d 7d 7c
43 ns n3 Sn 3s 3n
44 52 05 25 02 20
45 ca ce ee ec ac

46 h4 42 4h 24 2h
47 vu av av ov va
48 ar rf rf fa ra
49 cm ca am mc ma

51 uc ho oh ch bc
52 rs es er re se
53 ur ru au ra ur
54 qR pr qp gp gq
55 um am mn nn na

56 sp jp pg jg gj
57 et pe ep tp pt
58 rn ru nr na
59 hb bd ld dd db
60 1a 81 18 8n a8

61 IZ ZN NH ZH HN
62 RB RP RR RR RB
63 CU UC US UU CC
64 BP RB PB PR RP
65 KJ JC JC KC JK

66 Tt Tt Tt Tt Tt
67 SX XS sX sX sX
68 tt Tt Tt Tt Tt
69 zZ NZ zn Zz ZN
70 QC qn GQ Qg qg

71 4d 4c 1a 2d 1c
72 C5 C3 S8 S3 C8
73 79 b7 9b 7b A9
74 71 18 01 17 78

1 30 xv nx vx nv
2 bl ld lb dl bd
3 ra ur ar ru au
4 vw vu vv vw uv
5 uu um mu mw um

6 76 69 57 79 67
7 ur nu nr rn rn
8 az za zm mz ma
9 NV VN AV VA NA
10 QO QC OC OQ CQ

11 CU US CC UU UC
12 N4 NH 4N HN 4H
13 RP RP rr rR Rr
14 uA Aa 8A 8a 8A
15 TT Tt Tt LT lt

16 Av av AA VV Vv
17 Ja 4c 6a 3c 4d
18 X9 V9 X7 V7 V5
19 b7 A9 79 7b 9b
20 25 52 20 05 02

21 or ro oa ar
22 lc oc lo co ol
23 13 31 3s 1s
24 ma mc cm am ca

DATE

CITY

GRADE

MUSE

(SCHOOL)

DIFFERENTIAL APTITUDE TESTS
COMPUTER SCIENCE APTITUDE TEST

ANSWER SHEET

FORM A

PART II

PART I

DO NOT WRITE IN THIS SPACE

RAW SCORE	PERCENTILE
NORMS USED	

1	YZ	VX	VY	XY	ZY	26	NV	NA	VN	VA	AV	51	AV	VV	Vv	AA	av	76	8j	ja	aB	jB	8a
2	b9	69	6c	96	c6	27	XY	XX	xx	YX	Yy	52	WM	wW	NIM	Mw	mW	77	95	9Y	59	Y9	5Y
3	ua	uo	uo	ou	ou	28	EL	LE	LF	FL	FE	53	4H	NH	N4	HN	4N	78	lk	kf	fk	lf	kl
4	oc	ol	le	lo	co	29	MN	MV	VN	NM	NV	54	hb	Dd	dB	Db	DD	79	ca	am	mc	cm	ma
5	V7	X9	V9	V5	X7	30	ef	Ef	Fe	EE	FF	55	S3	S8	83	38	3S	80	xv	vx	nx	xn	nv

BE SURE YOUR MARKS ARE HEAVY AND BLACK.

6	8s	cS	8c	Sc	c8	31	C8	S5	BS	S8	8C	56	OX	XX	OO	XO	OV	81	re	er	es	se	rs
7	tb	ot	ob	bt	bo	32	h6	h8	8h	86	6h	57	C8	RS	S8	S5	8C	82	4V	V4	VX	X4	4X
8	3d	4d	2e	2d	5e	33	4d	4c	3a	3c	4a	58	V5	V7	X9	V9	X7	83	mn	zz	nz	zn	nn
9	dc	rc	rd	dr	cr	34	42	14	21	12	24	59	1.1	1L	1.7	7L	17	84	LT	IL	TL	IT	TI
10	sw	ts	tw	ws	st	35	QQ	oq	Qq	OQ	Qo	60	BD	RD	DR	BR	RB	85	14	40	41	44	11

11	nw	wu	mu	wm	um	36	xc	ce	ec	ex	xe	61	78	81	71	18	17	86	eu	us	se	ue	su
12	qq	qp	pq	pg	pp	37	oa	ra	ar	or	ro	62	vv	Vv	wv	wV	Ww	87	BP	RP	PR	PB	RB
13	xv	vx	nx	nv	xn	38	6c	8a	7a	7c	8c	63	Mm	nn	NN	mM	MN	88	Rr	rr	PP	pR	RP
14	mn	un	nu	mu	um	39	us	su	eu	se	ue	64	69	c6	96	6c	b9	89	XS	Xs	sx	SX	sX
15	nn	zz	zn	mn	nz	40	wo	ow	ro	rw	wr	65	1a	4d	1c	4c	2d	90	ra	na	rn	nr	ar

ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

16	pg	py	yg	gy	yp	41	vu	vw	uw	wv	wu	66	h4	24	2h	42	4h	91	OU	OC	UC	CO	UO
17	5Y	Y9	95	9Y	59	42	er	re	ir	ri	ie	67	ZX	Y2	XY	VX	VY	92	DR	BD	RB	BR	RD
18	un	ou	on	nu	uo	43	13	23	32	21	31	68	n3	3s	3n	ns	Sn	93	OV	XX	OO	OX	XO
19	nd	ud	nu	du	un	44	q2	qu	u2	2u	2q	69	wo	wr	rw	ow	ro	94	H2	ZN	HN	NH	ZH
20	41	40	11	14	44	45	vw	wv	vx	xv	wx	70	ur	ru	ro	or	oa	95	Vv	VV	av	AA	Av

21	rr	pR	PP	Rr	RP	46	et	te	ue	ta	ea	71	fn	in	nf	ni	fi	96	CQ	QO	OC	QC	OQ
22	TL	IT	TI	IL	IT	47	SV	IV	VS	VI	SI	72	wv	vw	vu	wu	uw	97	Ze	eZ	ZE	Zz	zE
23	MV	VM	VW	MW	WM	48	et	he	ht	eh	th	73	ht	et	he	th	eh	98	Qg	GQ	qq	QG	qg
24	uU	uW	Uu	WW	Wu	49	zu	mz	ma	az	zm	74	nm	na	un	um	mn	99	nn	MN	Mm	NN	mM
25	43	xc	34	cx	3c	50	ux	sx	sa	xa	xs	75	3x	37	7x	73	x7	100	uq	OQ	Qu	Qq	QQ

Appendix 2.
Page 2 - 10

PILOT STUDY

SUPERVISOR QUALIFICATIONS SHEET

SUPERVISOR CODE: _____

BIRTHDATE: _____ Sex (Please Circle): M F
Month Day Year

Have you ever supervised the work of others
in a non-volunteer, paid position? (that is,
assigned work, evaluated subordinates). YES ____ NO ____

Please give a description of the most recently held such supervisory position:

What is the longest period of time you have worked in such a supervisory position?

_____ Months _____ Years

Have you ever received training in how to use and conduct performance appraisals?

YES _____ NO _____

How many days are you regularly on campus during a normal week? _____

Please indicate which of the following days you are normally on campus:

SUN MON TUES WED THUR FRI SAT

ON SITE STUDY

SUPERVISOR QUALIFICATIONS SHEET

SUPERVISOR CODE: _____

BIRTHDATE: _____ Sex (Please Circle): M F
Month Day Year

Have you ever supervised the work of others
in a non-volunteer, paid position? (that is, YES ____ NO ____
assigned work, evaluated subordinates).

Please give a description of the most recently held such supervisory position:

What is the longest period of time you have worked in such a supervisory position?

_____ Months _____ Years

Have you ever received training in how to use and conduct performance appraisals?

YES ____ NO ____

In what training course are you currently enrolled?

Are you from out-of-town? YES ____ NO ____

When do you leave Ottawa? _____

APPENDIX 3.

Pilot Study Questionnaires

**Supervisor Performance Evaluation Sheet
Appraisal Effectance Motivation Questionnaire
Post-Evaluation Questionnaire**

SUPERVISOR PERFORMANCE EVALUATION

DATE: _____

SUPERVISOR CODE: _____

PROGRAMMER CANDIDATE CODE: _____

PROGRAMMER CANDIDATE SEX: M ____ F ____

PLEASE CHECK THAT POINT ON THE SCALE WHICH BEST DESCRIBES *THE PERFORMANCE OF THE INDIVIDUAL* WHOSE WORK YOU SUPERVISED:

Excellent Performance	Average Performance	Extremely Poor Performance
--------------------------	------------------------	-------------------------------

PLEASE RATE THE OVERALL *QUALITY OF THE WORK PERFORMED* BY THE INDIVIDUAL WHOM YOU SUPERVISED ON THE FOLLOWING SCALE:

0 100

Excellent Performance	Average Performance	Extremely Poor Performance
--------------------------	------------------------	-------------------------------

BASED ON YOUR EXPERIENCE WITH THE WORK PERFORMED BY THE INDIVIDUAL WHOM YOU SUPERVISED, PLEASE CIRCLE YOUR RESPONSE TO THE FOLLOWING STATEMENT:

This individual is capable of performing computer programming tasks.

<u>1 2 3 4 5</u>	<u>6 7 8 9 10</u>	<u>11 12 13 14 15</u>	<u>16 17 18 19 20</u>	<u>21 22 23 24 25</u>
Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree

BASED ON YOUR OBSERVATIONS OF THE INDIVIDUAL YOU RATED, HOW MUCH DO YOU RECOMMEND THAT THIS PERSON BE COMPENSATED FOR PARTICIPATING? (CIRCLE ONE)

NOTHING	\$5.00	\$10.00	\$15.00	\$20.00
---------	--------	---------	---------	---------

APPRAISAL EFFECTANCE MOTIVATION

As you know, performance appraisal involves many particular activities in a whole host of situations from organization to organization. In some situations one activity is called for whereas in another context a different course of action appears appropriate. You are to indicate the extent of your agreement on the following vertical scales. Because of the varying situational parameters there are no right or wrong answers. Please indicate your opinion on these issues. Proceed fairly rapidly through the questions but at the same time please do not be careless as we do want your true impressions.

1. In appraisal situations, I believe I can give negative feedback in a prompt and accurate fashion

1	- Strongly disagree	- I can do it less than 5% of the time
2		
3	- Disagree somewhat	- I can do it 15% of the time
4		
5	- Disagree slightly	- I can do it 30% of the time
6		
7	- 50/50 agree/disagree	- I can do it 50% of the time
8		
9	- Agree slightly	- I can do it 70% of the time
10		
11	- Agree somewhat	- I can do it 80% of the time
12		
13		
14	- Strongly agree	- I can do it more than 95% of the time

2. The next time I have to appraise a subordinate who's demonstrated poor performance, I believe it would be difficult for me to give prompt and accurate negative feedback.

1	- Strongly disagree	- I can do it more than 95% of the time
2		
3	- Disagree somewhat	- I can do it 80% of the time
4		
5	- Disagree slightly	- I can do it 70% of the time
6		
7	- 50/50 agree/disagree	- I can do it 50% of the time
8		
9	- Agree slightly	- I can do it 30% of the time
10		
11	- Agree somewhat	- I can do it 15% of the time
12		
13		
14	- Strongly agree	- I can do it less than 5% of the time

3. Giving prompt and accurate negative feedback during a performance appraisal would be an easy thing for me to learn.

1	-	Strongly disagree	-	It would take me over 3 years to learn this
2				
3	-	Disagree somewhat	-	It would take me over a year to learn this
4				
5	-	Disagree slightly	-	It would take me 6 months to learn this
6				
7	-	50/50 agree/disagree	-	It would take me two months to learn this
8				
9	-	Agree slightly	-	It would take me 3 weeks to learn this
10				
11	-	Agree somewhat	-	It would take me a week to learn this
12				
13				
14	-	Strongly agree	-	It would take me less than a day to learn this

4. For those supervisors who can do it, the provision of prompt and accurate negative feedback in appraisals allows subordinates ample time to improve their performance.

1	- Strongly disagree	- That's true less than 5% of the time
2		
3	- Disagree somewhat	- That's true only 15% of the time
4		
5	- Disagree slightly	- That's true only 30% of the time
6		
7	- 50/50 agree/disagree	- That's true 50% of the time
8		
9	- Agree slightly	- That's true 70% of the time
10		
11	- Agree somewhat	- That's true 80% of the time
12		
13		
14	- Strongly agree	- That's true over 95% of the time

5. If I were able to do it, the provision of prompt and accurate negative feedback to subordinates in performance appraisal would make it harder for them to claim ignorance of areas for improvement.

1	- Strongly disagree	- That's true less than 5% of the time
2		
3	- Disagree somewhat	- That's true only 15% of the time
4		
5	- Disagree slightly	- That's true only 30% of the time
6		
7	- 50/50 agree/disagree	- That's true 50% of the time
8		
9	- Agree slightly	- That's true 70% of the time
10		
11	- Agree somewhat	- That's true 80% of the time
12		
13		
14	- Strongly agree	- That's true over 95% of the time

6. I place a lot of value on appraisals that let subordinates know exactly where their performance needs to be improved.

1	- Strongly disagree -	That's true only 5% of the time
2		
3	- Disagree somewhat -	That's true only 15% of the time
4		
5	- Disagree slightly -	That's true only 30% of the time
6		
7	- 50/50 agree/disagree -	That's true 50% of the time
8		
9	- Agree slightly -	That's true 70% of the time
10		
11	- Agree somewhat -	That's true 80% of the time
12		
13		
14	- Strongly agree -	That's true over 95% of the time

7. Supervisors who have appraisal skills which let them promptly and accurately inform subordinates which subordinate performance areas need improvement, have more opportunities for promotions themselves.

1	- Strongly disagree	- That's true only 5% of the time
2		
3	- Disagree somewhat	- That's true only 15% of the time
4		
5	- Disagree slightly	- That's true only 30% of the time
6		
7	- 50/50 agree/disagree	- That's true 50% of the time
8		
9	- Agree slightly	- That's true 70% of the time
10		
11	- Agree somewhat	- That's true 80% of the time
12		
13		
14	- Strongly agree	- That's true over 95% of the time

8. Carrying out appraisals which let subordinates know specifically which performance areas need improvement is very important to me.

1	- Strongly disagree -	That's true only 5% of the time
2		
3	- Disagree somewhat -	That's true only 15% of the time
4		
5	- Disagree slightly -	That's true only 30% of the time
6		
7	- 50/50 agree/disagree -	That's true 50% of the time
8		
9	- Agree slightly -	That's true 70% of the time
10		
11	- Agree somewhat -	That's true 80% of the time
12		
13		
14	- Strongly agree -	That's true over 95% of the time

POST EVALUATION QUESTIONNAIRE

SUPERVISORY CODE: _____

Are you personally acquainted with the programmer candidate you supervised?

YES ____ NO ____

Did you experience any difficulty in the use of the appraisal forms?

YES ____ NO ____

Do you feel that the programmer candidate tried his/her best on the work materials?

YES ____ NO ____

Do you feel that this programmer candidate is interested in or concerned about his/her performance results for this project?

YES ____ NO ____

What made you decide to evaluate (ie stop testing) the programmer candidate when you did?

Why did you schedule the appointment with the programmer candidate when you did?

How do you feel that this programmer candidate will react when receiving your performance appraisal information?

How do you feel personally about presenting this information to the programmer candidate?

APPENDIX 4.

On Site Study Questionnaires

Supervisor Performance Evaluation Sheet

Volunteer Performance Factors

Supervisor Evaluation Factors

Post-Evaluation Questionnaire

Appraisal Effectance Motivation Questionnaire

SUPERVISOR PERFORMANCE EVALUATION

DATE: _____

SUPERVISOR CODE: _____

PROGRAMMER CANDIDATE CODE: _____

PLEASE CHECK THAT POINT ON THE SCALE WHICH BEST DESCRIBES *THE PERFORMANCE OF THE INDIVIDUAL* WHOSE WORK YOU SUPERVISED:

Excellent
Performance

Average
Performance

Extremely
Poor Performance

PLEASE CHECK THAT POINT ON THE SCALE WHICH BEST DESCRIBES THE *OVERALL QUALITY* OF THE WORK PERFORMANCE OF THE INDIVIDUAL YOU SUPERVISED:

Excellent
Performance

Average
Performance

Extremely
Poor Performance

BASED ON YOUR EXPERIENCE WITH THE WORK PERFORMED BY THE INDIVIDUAL WHOM YOU SUPERVISED, PLEASE CIRCLE YOUR RESPONSE TO THE FOLLOWING STATEMENT:

This individual is capable of performing computer programming tasks.

<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>	<u>11</u>	<u>12</u>	<u>13</u>	<u>14</u>	<u>15</u>	<u>16</u>	<u>17</u>	<u>18</u>	<u>19</u>	<u>20</u>	<u>21</u>	<u>22</u>	<u>23</u>	<u>24</u>	<u>25</u>
Strongly Agree					Agree					Uncertain					Disagree					Strongly Disagree				

PLEASE RATE THE ***OVERALL QUANTITY*** OF THE WORK PERFORMANCE OF THE INDIVIDUAL YOU SUPERVISED ON THE FOLLOWING SCALE:

Excellent Performance	Average Performance	Extremely Poor Performance
--------------------------	------------------------	-------------------------------

BASED ON YOUR OBSERVATIONS OF THE INDIVIDUAL YOU RATED, HOW MUCH DO YOU RECOMMEND THAT THIS PERSON BE COMPENSATED FOR PARTICIPATING? (CIRCLE ONE)

NOTHING \$5.00 \$10.00 \$15.00 \$20.00

**SUPERVISOR ATTRIBUTION OF CONFEDERATE
SUBORDINATE PERFORMANCE**

You are asked to indicate **the extent** to which you feel the performance of your volunteer was **due to the following factors**. Please circle the number which best corresponds to your response for each question.

1. To what extent was the volunteer's performance due to his/her ability?

1	2	3	4	5	6	7
not at all		somewhat		very much		totally

2. To what extent was the volunteer's performance due to the difficulty of the task?

1	2	3	4	5	6	7
not at all		somewhat		very much		totally

3. To what extent was the volunteer's performance due to his/her effort?

1	2	3	4	5	6	7
not at all		somewhat		very much		totally

4. To what extent was the volunteer's performance due to s/he being lucky?

1	2	3	4	5	6	7
not at all		somewhat		very much		totally

**SUPERVISOR CONFIDENCE AND CONCERN
RE: ACCURACY AND IMPACT OF FEEDBACK**

Please circle the number which best corresponds to your response to the following statements.

1. I am confident in the accuracy of my ratings of this volunteer's performance.

1	2	3	4	5	6	7
not at all		somewhat		very much		totally

2. I feel my ratings of this volunteer's performance will have a significant effect on him/her.

1	2	3	4	5	6	7
not at all		somewhat		very much		totally

3. In general, I feel more confident rating the performance of a female versus that of a male subordinate.

1	2	3	4	5	6	7
not at all		somewhat		very much		totally

POST EVALUATION QUESTIONNAIRE

SUPERVISORY CODE: _____

Are you personally acquainted with the programmer candidate you supervised?

YES ____ NO ____

Did you experience any difficulty in the use of the appraisal forms?

YES ____ NO ____

Do you feel that the programmer candidate tried his/her best on the work materials?

YES ____ NO ____

Do you feel that this programmer candidate is interested in or concerned about his/her performance results for this project?

YES ____ NO ____

What made you decide to evaluate (ie stop testing) the programmer candidate when you did?

Why did you schedule the appointment with the programmer candidate when you did?

How do you feel that this programmer candidate will react when receiving your performance appraisal information?

How do you feel personally about presenting this information to the programmer candidate?

APPRAISAL EFFECTANCE BEHAVIORS

As you know, performance appraisal involves many particular activities in a whole host of situations from organization to organization. In some situations, one activity is called for whereas in another context, a different course of action appears appropriate. you are to indicate the extent of your agreement by circling the appropriate number on the following horizontal scales. Because of the varying situational parameters, there are no right or wrong answers. Please indicate your opinion on these issues. Proceed fairly rapidly through the questions, but at the same time, please do not be careless as we do want your true impressions.

SELF-EFFICACY EXPECTANCY ☆

1. In appraisal situations concerning poor performance (30th ile), I **believe I can** give negative feedback in an **accurate*** fashion ____ % of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30% il that their output is at the 30% ile).

SELF-EFFICACY EXPECTANCY ☆

2. Giving **prompt** negative feedback concerning poor output during a performance appraisal would be an **easy thing for me to learn** for ____ % of such cases.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

OUTCOME EXPECTANCY ☆

3. **For those supervisors who can do it**, the provision of **prompt** negative feedback during performance appraisals provides subordinates the information necessary to make appropriate (training/transfer) career decisions ____ % of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

SELF-EFFICACY EXPECTANCY ☆

4. The next time I have to appraise a subordinate who's demonstrated **poor** performance (at the 30% ile), I believe **it would be difficult for me** to give **accurate*** negative feedback.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30th % ile that their output is at the 30th % ile).

☆ Question item measures this component of Appraisal Effectance Motivation measure. Labels not on questionnaires administered.

SELF-EFFICACY EXPECTANCY ☆

5. In appraisal concerning poor performance, I believe I can give negative feedback in a **prompt** fashion, ____% of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

OUTCOME EXPECTANCY ☆

6. Assuming I could do it, the provision of **accurate*** negative feedback to poor performing subordinates would make it harder for them to claim ignorance of areas necessary for information ____% of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30th % ile that their output is at the 30th % ile).

SELF-EFFICACY EXPECTANCY ☆

7. The next time I have to appraise a subordinate who's demonstrated poor performance (at the 30% ile), I believe it would be difficult for me to give **prompt** negative feedback ____% of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

OUTCOME VALUE ☆

8. ____% of the time, I place a lot of value on **accurate*** negative feedback in appraisals to poor performing subordinates so that they know exactly where their output must be improved.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30th % ile that their output is at the 30% ile).

☆ Question item measures this component of Appraisal Effectance Motivation measure. Labels not on questionnaires administered.

SELF-EFFICACY EXPECTANCY ☆

9. Giving **accurate*** negative feedback concerning poor output during a performance appraisal would be an **easy thing for me to learn** for ____% of such cases.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30th % ile that their output is at the 30% ile).

OUTCOME EXPECTANCY ☆

10. **Assuming I could do it**, the provision of **prompt** negative feedback to poor performing subordinates would make it harder for them to claim ignorance of areas necessary for improvement ____% of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

OUTCOME VALUE ☆

11. **In my organization**, supervisors who have appraisal skills which let them **promptly** inform poor performing subordinates of poor output domains that need improvement, **have more opportunities** for promotions themselves ____% of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

OUTCOME EXPECTANCY ☆

12. **For those supervisors who can do it**, the provision of **accurate*** feedback during performance appraisals provides subordinates the information necessary to make appropriate (training/transfer) career decisions ____% of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30th % ile that their output is at the 30% ile).

☆ Question item measures this component of Appraisal Effectance Motivation measure. Labels not on questionnaires administered.

OUTCOME VALUE ☆

13. Carrying out appraisals which **accurately*** let poor performing subordinates know specifically which performance areas need improvement is **very important to me** ____ % of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30th % ile that their output is at the 30% ile).

ANXIETY ☆

14. I usually get somewhat apprehensive when I have to give face-to-face performance feedback concerning poor performance (25% ile) on the part of one of my subordinates.

Do not agree at all	Agree slightly	Agree somewhat	Agree moderately	Agree completely
------------------------	-------------------	-------------------	---------------------	---------------------

OUTCOME VALUE ☆

15. ____ % of the time, I place a lot of value on **prompt** negative feedback in appraisals to poor performing subordinates so that they know exactly where their output must be improved.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

OUTCOME VALUE ☆

16. **In my organization**, supervisors who have appraisals skills which let them **accurately*** inform poor performing subordinates of poor output domains that need improvement **have more opportunities** for promotions themselves ____ % of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

* (i.e., tell a subordinate at the 30th % ile that their output is at the 30% ile).

☆ Question item measures this component of Appraisal Effectance Motivation measure. Labels not on questionnaires administered.

OUTCOME VALUE ✧

17. Carrying out appraisals which **promptly** let poor performing subordinates know specifically which performance areas need improvement is **very important to me** ____% of the time.

10%	20%	30%	40%	50%	60%	70%	80%	90%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

✧ Question item measures this component of Appraisal Effectance Motivation measure. Labels not on questionnaires administered.

APPENDIX 5.

Power Calculations

6 DV's

1 IV -- 2 levels

Large Effect Size -- .35 (Cohen, 1988)

Power -- .80

Enter values in this column

trial lambda =	15	implied v =	35.86
effect size =	0.35		
u =	6		

lower tabled v =	20	interpolated lambda =	16.14
upper tabled v =	60		
lower tabled lambda =	18.4		
upper tabled lambda =	15		

iterated v =	39.13
--------------	-------

ky =	6
kx =	1
max(kc,ka + kb) =	0

N =	46
-----	----

6 DV's

1 IV -- 2 levels

Medium Effect Size -- .15 (Cohen, 1988)

Power -- .80

Enter values in this column

trial lambda =	14.3	implied v =	88.33
effect size =	0.15		
u =	6		

lower tabled v =	60	interpolated lambda =	14.55
upper tabled v =	120		
lower tabled lambda =	15		
upper tabled lambda =	14.3		

iterated v =	90.01
--------------	-------

ky =	6
kx =	1
max(kc,ka + kb) =	0

N =	97
-----	----

6 DV's

1 IV -- 2 levels

Medium Effect Size -- 0.15 (Cohen, 1988)

Power -- .70

Enter values in this column

trial lambda =	11.6	implied v =	70.33
effect size =	0.15		
u =	6		

lower tabled v =	60	interpolated lambda =	12.02
upper tabled v =	120		
lower tabled lambda =	12.2		
upper tabled lambda =	11.6		

iterated v =	73.16
--------------	-------

ky =	6
kx =	1
max(kc,ka + kb) =	0

N =	80
-----	----

6 DV's

IV -- 4 levels

Large Effect Size -- .35 (Cohen, 1988)

Power -- .80

Enter values in this column

trial lambda =	15	implied v =	23.86
effect size =	0.35		
u =	18		

lower tabled v =	20	interpolated lambda =	17.58
upper tabled v =	60		
lower tabled lambda =	18.4		
upper tabled lambda =	15		

iterated v =	31.22
--------------	-------

ky =	6
kx =	3
max(kc,ka + kb) =	0

N =	20
-----	----

6 DV's

Appendix 5.

1 IV -- 4 levels

Page 5 - 6

Medium Effect Size -- .15 (Cohen, 1988)

Power -- .80

Enter values in this column

trial lambda =	22.5	implied v =	131.00
effect size =	0.15		
u =	18		

lower tabled v =	60	interpolated lambda =	22.29
upper tabled v =	120		
lower tabled lambda =	25		
upper tabled lambda =	22.5		

iterated v = 129.60

ky =	6
kx =	3
max(kc, ka + kb) =	0

N =	55
-----	----

Enter values in this column

trial lambda =	20.1	implied v =	115.00
effect size =	0.15		
u =	18		

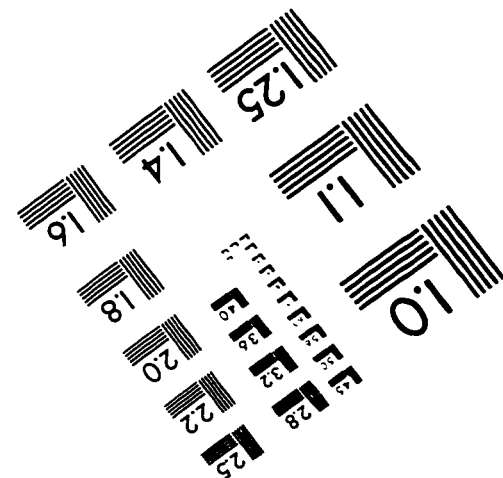
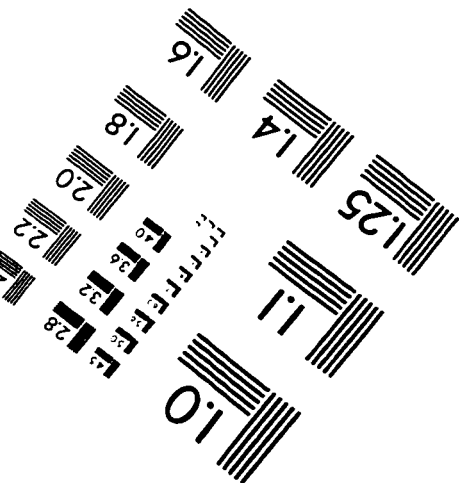
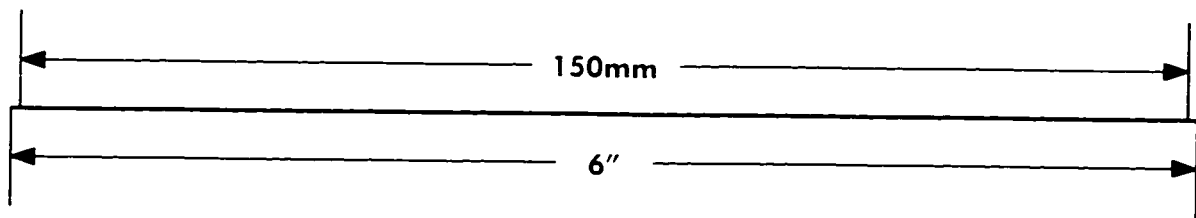
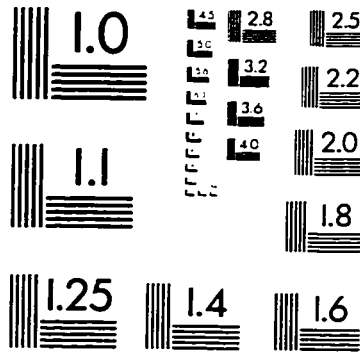
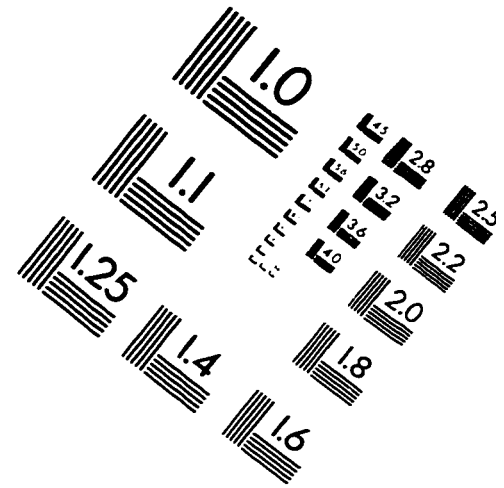
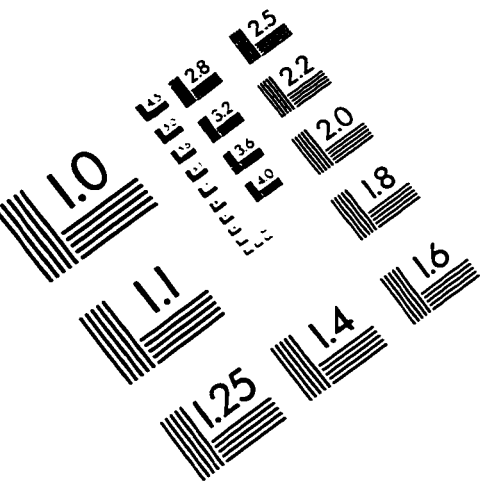
lower tabled v =	120	interpolated lambda =	22.60
upper tabled v =	∞		
lower tabled lambda =	22.5		
upper tabled lambda =	20.1		

iterated v = 131.70

ky =	6
kx =	3
max(kc, ka + kb) =	0

N =	55
-----	----

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved