

VIDEO-TAPE IN INTERVIEWING:
AN ANALYSIS OF RATINGS
AND ATTITUDES

by

ALEC JOHN LEE

B.Comm., University of British Columbia, 1969

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF BUSINESS ADMINISTRATION

in the Department

of

Commerce and Business Administration

WE ACCEPT THIS THESIS AS CONFORMING TO
THE REQUIRED STANDARD

The University of British Columbia

August, 1972

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study.

I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of COMMERCE & BUSINESS ADMINISTRATION

The University of British Columbia
Vancouver 8, Canada

Date September 1, 1972

ABSTRACT

This study was intended to examine three aspects of interviewing. The first was an exploration for any systematic differences between ratings given by interviewers in actual face-to-face interviews and ratings given by group (N=3) and individual observers of video-taped interviews. The second aspect was a comparison between group (N=3) and individual ratings in terms of mean variable ratings, reliability and halo error. The third section sought attitudes of interviewees and rating viewers of the utility of video-tape in employment interviews.

The interviewees for the first and third aspects of this study were thirty-four fourth-year Commerce students from the University of British Columbia. The video-tape viewers and interviewers were thirty members of the Bank of Montreal's managerial staff. For the second section the interviewees were three fourth-year Commerce students from the University of British Columbia and the video-tape viewers were one hundred and eighty-five Commerce undergraduate students also from the University of British Columbia. The viewers and interviewers were requested to evaluate interviewees along thirteen dimensions and to decide whether or not to call back the interviewees for further interviews.

Minimal differences were found between interviewer

and group observer ratings while individual observer ratings were found to be uniformly and significantly higher than both interviewer and group observer ratings. Reliability estimates were generally moderate to low with no significant differences existing between group and individual ratings. Halo error was possibly a contributing factor for group and individual ratings. The attitudes of interviewees toward the use of video-tape in interviewing were quite favorable. The viewers, on the other hand, displayed only a moderate enthusiasm.

In the discussion, attention was given to methods for revising the Bank of Montreal's interview ratings form and interviewing procedure.

.
.
.

Table of Contents

CHAPTER

1	INTRODUCTION	1
	Outline of the Paper	4
2	RESEARCH SUMMARY	6
	The Interview	6
	Future Research Directions	26
3	METHODOLOGY	41
	Differences in Ratings - Interviewers vs Viewers	41
	Groups vs Individuals	56
	Attitudes Toward the Use of Video-Tape in Interviews	63
4	RESULTS	66
	Interviewers vs Viewers	72
	Attitudes Toward Video-Tape In Inter- viewing	93
5	DISCUSSION	101
	Interviewers vs Viewers	101
	Group vs Individual Viewers	105
	Attitudes Toward Video-Tape	124
6	SUMMARY AND CONCLUSIONS	128
	BIBLIOGRAPHY	132
	APPENDIXES	
A	THE CAMPUS INTERVIEW	140
B	APPLICATION FOR EMPLOYMENT	146

C	VIDEOTAPED INTERVIEW QUESTIONNAIRE	148
D	VIDEO-TAPE VIEWERS QUESTIONNAIRE	154
E	SUMMARY SHEET	158

List of Tables

<u>Table</u>		<u>Page</u>
	<u>Chapter 2</u>	
1	The Reliability of Judgments of Individuals and of Groups	36
	<u>Chapter 3</u>	
1	Schedule of Interviewing and Viewing . . .	43
2	Interviewer Characteristics	45
3	Viewer Characteristics	47
4	Analyses of Variance	48
5	Description of Interviewees	59
6	Sampling Distribution for Data Acquisition	60
	<u>Chapter 4</u>	
1	Means, Standard Deviations, and Significant Differences for Interviewers', Groups', and Individuals' Ratings	67
2	Overall Comparison Among Samples	68
3	Convergent and Discriminant Validities of Interviewers-Groups, Interviewers-Individuals, and Individuals-Groups	69
4	Group, Individual and Interviewer Inter-viewer Intercorrelations Call Back - Reject Decision	71
5	Mean Traits Rating - Total Groups and Individuals	73
6	Mean Trait Ratings - Groups vs Individuals .	74
7	Overall Rating - Group and Individuals . . .	75
8	Call Back-Reject Decision - Groups vs Individuals	76
9	Rotated Factor Matrix - Individuals	77

<u>Table</u>		<u>Page</u>
10	Rotated Factor Matrix - Groups	78
11	Correlations Between Variables and Call Back-Reject Decision-Groups and Individuals	80
12	Canonical Correlations Between the Two Sets of Factor Scores	81
13	Intra-Rater Reliability - Groups and Individuals	82
14	Intra-Rater Reliability - Groups and Individuals - Pearson's r	84
15	Intra-Rater Reliability - Groups and Individuals - Percent Perfect Agreement .	85
16	Mean Variable Ratings: Test and Retest Groups and Individuals	86
17	Call Back-Reject Decision - Groups and Individuals - Intra-Rater Reliability . .	87
18	Call Back-Reject Decisions - Group Decision Shift	88
19	Call Back-Reject Decisions - Individuals Decision Shift	88
20	Inter-Rater Reliability - Group vs Individuals - Standard Deviation Scores .	89
21	Inter-Rater Reliability - Groups vs Individuals - Standard Deviation Scores .	90
22	Inter-Rater Reliability - Groups vs Individuals - Inter-Correlations	91
23	Extent and Cause of Interviewee Dis- traction	94
24	Extent of Difference in Interviewee's Be- haviour In a Video-Taped Interview Com- pared with a Face-to-Face Interview (N=34)	95
25	Enthusiasm for Video-Taped Initial Screen- ing Interviews in Selection of University Graduates for Employment (N=34)	97

<u>Table</u>		<u>Page</u>
26	How would Most Graduating Students React if Asked by a Company to Undergo a Video-Taped Interview Conducted by the Placement Office	97
27	Student (Interviewee) Rankings of the Three Most Serious Reservations or Objections re Video-Taped Screening Interviews	97
28	Degree of Realism in Portraying Interviewee Characteristics - Viewers (N=17)	99
29	Enthusiasm Toward the Possible General Use of Video-Taped Initial Screening Interviews for Selection of University Graduates - Viewers	100

ACKNOWLEDGEMENTS

It is a pleasant task to acknowledge the cooperation of colleagues and friends in the preparation of this manuscript. I wish to express my sincere appreciation to Dr. Larry F. Moore whose comments from afar strengthened the analyses and interpretations given here. Gratitude must also be given to Dr. Vance F. Mitchell, Dr. Gordon Walter, Dr. Ronald Taylor and Professor Merle Ace whose helpful and supportive observations cast light upon many a dark moment in my thoughts. Mr. George Strachan and Mr. James Watt from the Bank of Montreal deserve thanks for permitting me to study their organization. Their interest and encouragement clearly illustrated that the gap between the business world and the academic community may readily and profitably be bridged. Lastly, my wife Mary merits express recognition. She spurred me on to always go the next step. To her, my special thanks. To her, I dedicate this thesis.

CHAPTER 1

INTRODUCTION

Video-tape is being put to more and more uses in industry and industrial research. Applications range from industrial training (Stroh, 1969), information dissemination, surveillance, ego development (Kennedy, 1970) to improvements in self-acceptance (Walter, 1971) and industrial and psychological research (Wexley, Yukl, Kovacs and Sanders, 1972).

Moore and Craik (1972) saw the tremendous potential which video-tape may have for assisting in personnel selection. These researchers visualized the important contribution which instant playback, multiple viewing, transferability and other video-tape features could make in the complex and important process of interviewing. Indeed, many organizations have taken advantage of this potential and have instituted a system of interviewing which incorporates video-tape as a critical element. In its selection of volunteers for an overseas Canadian Goodwill Tour, the Canadian Public Service Commission recently utilized video-tape to record interviews in Vancouver for subsequent viewing by administrators in Ottawa. Ability Search, a Washington, D.C. firm that specializes in recruiting systems analysts and operations research personnel, use the video-taped

interview as a replacement for the more common resume. Through this organization candidates are interviewed and their video-tapes are sent to a number of participating companies for screening (Business Week, 1971).

In December, 1970, the personnel department of the British Columbia Regional Office of the Bank of Montreal expressed an interest in the possible use of video-tape as an aid to their personnel selection procedure and, more specifically, to interviewing. The Bank's original concern revolved around the possibility of video taping in one centre an interview with a candidate who wished employment with the Bank in another centre, and send to that centre the video-taped interview rather than the candidate himself. The principal question was not whether such a procedure would be feasible, as the foregoing examples indicate that it is, but whether or not a video-taped interview would introduce some bias which would not exist within a face-to-face interview. This issue formed the impetus for the present study. A research proposal was given to the Bank's Employment and Employee Relations Manager who permitted a study not only of this area but also of a number of other related topics. These other topics are briefly described below.

Interviewing research. Video-tape as well as audio-tape has been used by a number of researchers to explore aspects of the interview (Crissy, 1952; Kasl and Mahl, 1956; Wiens et al, 1966; Grant and Bray, 1969; Blakeney and

MacNaughton, 1971; Wexley, Yukl, Kovacs and Sanders, 1972). However, to date, no one has seriously questioned the use of video-tape or audio-tape for this type of research. McLuhan (1964) argued that the medium over which a message is communicated forms part of the message itself. If this is the case, then these researchers are not studying just the interview but the taped interview. With this possibility in mind, this paper addresses itself to the question of whether or not ratings made in a face-to-face interview differ from ratings made with a video-taped interview. The mode of interview presentation may be an influencing moderator affecting the outcomes of the above-stated research.

Groups vs Individuals. Panel interviews are often used for selection purposes (Taft, 1959; OSS, 1948). However, little research has taken place to directly examine differences in ratings as given by panels and as given by individual interviewers. Considerable work is reported which examines groups vs individual decisions (Maier, 1967; Lorge et al, 1958; Hall and William, 1970). These works, though, have taken place outside the context of the interview. This paper looks at groups vs individual differences within this context.

Interview Rating Forms. Many organizations evaluate interviewees on an interview rating form usually consisting of an accept-reject decision and a series of traits, each to be rated on a three-point, five-point or seven-point

rating scale. The Bank of Montreal is no exception. Very little work, though, has been devoted to analyzing these forms. This paper examines the Bank's rating form in terms of intra- and inter-rater reliability and halo error.

Attitudes Toward Video-Tape in Interviews. A recent paper by Moore and Craik (1972) looked at the attitudes which students and personnel administrators have toward the use of video taped interviews. None of the members of these samples were asked to formally assess the suitability of the candidates they viewed. Indeed, except for only looking at the taped interviews, the viewers were quite removed from the interview setting. This paper focuses upon the attitudes toward video-tape held by actual interviewees who were being video-taped and by viewers who formally rated the applicants' suitability.

Outline of the Paper

The areas explained above are grouped into three main subsections for analysis. The first area is aimed at determining the existence of any systematic differences between ratings given by interviewers in actual face-to-face interviews and ratings given by observers of video taped interviews. The second area is an examination of differences between group and individual raters in terms of mean variable ratings, overall ratings, call back-reject decisions, halo error and inter- and intra-rater reliability. The third area explores the attitudes of interviewees and

observing raters toward the use of video-tape in interviewing.

CHAPTER 2

RESEARCH SUMMARY

This review is segmented into two basic sections. The first is a summary of research on the employment interview. While this study does not directly investigate the findings reported in this section, it was felt to be worthwhile to provide such a summary in order to highlight and explain considerations which were made in designing this study's methodology and to offer a perspective within which this research took place. The second section is a review of some future directions which research in the employment interview may take. As is described, some of these directions are examined in this study.

SECTION 1: THE INTERVIEW

The employment interview has long been a topic of concern for both the personnel administrator and the researcher. Through the years, the utility of the interview has been sorely criticized and, as a job performance predictor, has received largely pessimistic reviews. Yet its widespread use still remains. Throughout the years at least 90-95% of organizations surveyed employed the personnel interview in their selection processes (Spriegel and James, 1958), and as Carlson et al (1971) point out, it is unlikely that this condition will change. The personnel interview has remained in existence not as a result of any major supporting empirical evidence

but through some process wherein it thrives today because it thrived yesterday.

Given this situation, then, the goal of research in the personnel interview should not be to discredit it so much as to improve it. This is the orientation followed in this paper.

The literature in personnel interviewing is dealt with chronologically, with heavier emphasis and more detail being given on the recent work. Comments on earlier studies are based upon the major research summary of Wagner (1949), Ulrich and Trumbo (1965), Mayfield (1964) and Wright (1969).

Interviewing Review

The preponderance of opinion and how-to-do-it manuals for interviewing became apparent in Wagner's (1949) summary. Of the 106 studies he reviewed only 25 offered quantitative information. Also, as Ulrich and Trumbo (1965) pointed out, within these 25 studies a number of methodological weaknesses could be readily located.

Ulrich and Trumbo (1965) drew largely similar conclusions as Wagner (1949). First, greater standardization in terms of information used as a basis for personnel selection was called for. By permitting variation in the type of information obtained and the form of interview structure employed, variation in decisional outcomes is inevitable. Structured interviews drawing comparable data should boost reliability and thereby serve to increase the upper boundaries of predictive validity.

Second, ancillary information (e.g. tests and credentials) should be employed more heavily in personnel decisions. This was closely linked to the first conclusion in that tests and the like are usually standardized and not so much suspect to individual interpretation as are data drawn solely from the interview. Finally, it was suggested that the scope of the interview be reduced. Rather than attempting to assess a large number of such complex psychological constructs as intelligence, leadership ability, aptitude, and so on in a half-hour interview, the interviewer should focus in upon only a few such areas. Ulrich and Trumbo suggested two important factors to which the interviewer could devote his interviewing time. These were personal relations (sociability) and motivation to work. These two traits received some empirical evidence supporting their presence in the interview (Woodworth et al, 1957; Rimland, 1960; Rundquist, 1947; Otis et al, 1962), while other traits have been generally found to be better predicted by alternate and likely more valid and reliable means.

Mayfield (1964), in another review, drew 15 conclusions from his research summary. These may be grouped into two basic categories, a) interviewer characteristics and b) methodological and structural properties of interviews.

Interviewer characteristics:

(1) Interviewers are generally consistent in their approaches to different interviewees.

(2) Interviewers frequently interpret or weight the same information generally differently.

(3) Interviewer predictions based on a combination of interview data and scores from tests of proven validity are usually no better (and sometimes worse) than predictions based on test scores alone. This raises questions as to the efficacy of Wagner's (1949) and Ulrich and Trumbo's (1965) suggestions that ancillary information should accompany interview data when making employment decisions.

(4) The attitudes of interviewers do affect the interpretations of what the interviewees say.

(5) In unstructured interviews, the interviewer generally talks more than the interviewee.

(6) Interviewers generally are influenced more by unfavourable than by favourable information.

(7) In an unstructured interview, interviewers tend to make their decisions fairly early.

Methodological and structural properties of interviews

(8) The interview can be reliably divided into various types of units. This permits a microanalytic approach to studying interview activities and dynamics.

(9) The intra-rater reliability of the interview appears to be relatively high. Test-retest time spans are of varying length, though, and no work has been done to isolate the effect which memory plays upon the retest outcomes.

(10) An unstructured interview with no prior information usually results in low inter-rater reliability for a general suitability rating.

(11) Material is not consistently covered in an unstructured interview.

(12) Inter-rater reliability is generally higher for a structured interview than for an unstructured interview. Evidence also seems to indicate that "structuring an interview increases inter-rater reliability when interviewers from the same company use the same form, but that two different structured forms may lead to completely different ratings when used with the same interviewee."

(13) Although reliabilities may be satisfactory in some types of interview situations, validities are generally quite low.

(14) Only the intelligence trait of an interviewee has been found to be judged satisfactorily. This appears to contradict Ulrich and Trumbo's (1965) conclusion that motivation to work and personal relations (sociability) show the greatest evidence of validity in interviewer decisions. Each reviewer has excluded the conclusion of the other. Mayfield made no mention of motivation to work and sociability and Ulrich and Trumbo did not consider the predictive validity of intelligence. A look at the bibliographies of each review reveals that neither included those research studies relevant to the other's conclusion.

(15) Answers given by interviewees are affected by the form in which the question is asked.

Mayfield also discussed some of the overwhelming methodological problems which arise when attempting to compare studies of interviews. There is frequently a variation among studies in the amount of structure imposed on rating forms and in the types of traits or behavioural characteristics on which subjects are asked to make judgments. Interview length is often quite variable as is the type of job for which applicants are being rated. Furthermore, the criteria of success in prediction are frequently different and often difficult to interpret. These criteria range from using job performance measures as dependent variables to validating traits by using other measures of the same traits. Crissy (1952) dealt with this issue at great length.

Throughout his review Mayfield assumed that the most frequently used type of employment interview is unstructured. However, he offered no interpretation of what "unstructured" means and gave no empirical evidence to support this assumption.

Two of the above conclusions drawn by Mayfield (i.e. interviewers are generally influenced by unfavourable rather than favourable information and interviewers tend to make decisions early in the interview) were derived from the series of studies presented by Webster (1964).

This most important book really set the stage for the recent developments in interviewing research. Basically, what

Webster and his associates did was to address themselves to problems of identifying the processes and conditions in an interview which produce employment decisions. Rather than looking at the validity or reliability of decisions they looked at how the decisions themselves were developed.

Webster's (1964) work summarized nine years of his research and that of his colleagues and former graduate students at McGill University (notably, B.M. Springbett, D. Sydiaha, C.W. Anderson, Areta Crowell and Patricia Rowe).

Webster's analysis drew seven major conclusions.

1. Interviewers develop a stereotype of a good candidate and seek to match men and stereotypes. This finding was initially reported by Sydiaha (1958) and later supported by Rowe (1963). Rowe's research also pointed out that stereotypes tend to be "good" rather than "bad."

2. A bias is established early in the interview and tends to be followed either by a favourable or by an unfavourable decision. This finding was initially reported by Springbett (1954) and later substantiated by Sydiaha (1961), Anderson (1960) and Lambert et al (1960). Furthermore, Anderson showed that interviewers tend to speak more when a decision to hire is made than when a decision to reject is made.

3. Interviewers are more influenced by unfavourable than by favourable information. This finding was first suggested

by Crissy and Regan (1951). Later evidence (Springbett, 1958; Bolster and Springbett, 1961, and Rowe, 1960) lent credence to this discovery as did the more recent works of Miller and Rowe (1967), Mayfield and Carlson (1966) and Blakeney and McNaughton (1971).

4. Interviewers seek information to support or refute hypotheses and when satisfied, they turn their attention elsewhere. Crowell (1961) basically suggested that interviewers change the emphasis they place on parts of information in order to confirm early impressions. Webster reported that the evidence for this finding was not as conclusive as for the other ones.

5. Empathy relationships are specific to individual interviewers. Sydiaha (1962) discussed the problems which may arise by treating the extent to which an empathic relationship can be obtained between interviewer and applicant as a basis for selection decisions. The most notable problem arises out of the evidence he reported showing that the empathically based decision may not be consistent from one individual to another. Largely, Sydiaha expressed caution against using common sense or intuition as a selection decision guide. He stated that with this approach "the decision making cues will be unspecified, unknown or specific to the interviewer".

6. Feeding information to a judge piece by piece affects the decision. Crowell (1961) reported evidence which suggests

that when judges¹ are given all information simultaneously, their decisions are different from and better than when information is given piece by piece. Her experiment was performed in a laboratory setting and, as Webster stated, "generalization ... to the employment interview must be made tentatively with a good deal of caution". No other research which aims at replicating this study could be located.

7. Experienced interviewers rank applicants in the same order although they differ in the proportion they will accept. This finding was reported by Rowe (1963) and represented the first major piece of research which recognized differences among interviewers and, hence, treated interviewers as an independent variable affecting selection decisions. Rowe found an ordered pattern of acceptance among judges. "Those who accept a small proportion of candidates accept individuals who are most frequently accepted by all judges; those who accept a large proportion are favourably disposed toward men accepted by more selective interviewers." More experienced interviewers were found to be more selective.

¹The term "judges" is employed rather than "interviewers" since the research involved persons who made judgments based on highly selective and controlled items of written or graphic information and not on information gathered in an actual personal interview. "Interviewers" per se were not used in the samples nor were interviews per se actually carried out.

This series of conclusions marked a change in the direction taken when researching the interview. The processes which led to decisional outcomes became the focus rather than the earlier approach of analyzing the outcomes themselves in terms of their validities and reliabilities. This is not to say that validity and reliability lost importance. They only were placed in a more causative perspective.

More current research on the employment interview has been performed notably by two research teams. These teams were respectively Eugene Mayfield and Robert Carlson of the Life Insurance Agency Management Association (LIAMA) and Milton Hakel and his associates at the University of Ohio and the University of Minnesota.

The LIAMA team undertook their program of research in an attempt to understand the mechanism of the interview and to improve the use of this selection device in the life insurance industry (Carlson et al, 1971). Their experimental designs were basically the same as those employed in the McGill studies; decisional outcomes were dependent variables affected by controlling and modifying processes and influences within the interview structure.

It is worth noting at this point that the LIAMA group and the McGill group generally used a paper and pencil approach when presenting the interviewees to the interviewers (raters). That is, the interviewees were not physically present but were described on paper. Mayfield and Carlson (1966) described

this approach as being quick, thereby permitting a wide scope of applicant information to be given, and experimentally convenient, in that it "allows control of the many outside variables which otherwise might affect the results". They further stated that "results obtained by this method could later be compared to those obtained when information is presented by other means... to determine if the mode of presentation has any effect". Video-tape may certainly be one of these "other means" with which to compare.

Carlson et al identified four main classes of influences which operate to affect or limit the interviewer's decisions. These were:

1. The physical and psychological properties of the interviewee;
2. The physical and psychological properties of the interviewer;
3. The situation/environment in which the interviewer works;
4. The task or type of judgment the interviewer must make.

A detailed summary of their research findings may be found in their recent paper (Carlson, Thayer, Mayfield and Peterson, 1971). Some of the more important results are as follows:

1. Using both photographs to represent physical appearance and personal history descriptions to depict interviewees' backgrounds, Carlson (1967) found that the rating of the photograph alone had a minimal effect on the mean rating of the overall applicant. Appearance accounted for less than 5%

of the variance in the mean rating of the applicant while the personal history accounted for about 40%. Furthermore, photographs were found to be most influential when they complemented the personal history description. However, the importance of appearance cannot be underestimated. In a brief summary of studies exploring the effects of appearance on interpersonal perception, Hakel et al (1970) stated that appearance alone can result in lasting and well-structured interpersonal impressions. It should be noted that in the studies Hakel et al describe actual physical presence formed the basis of appearance rather than the photograph technique employed by Carlson. This may explain some of the discrepancy in emphasis placed on the personal characteristic.

2. In an investigation of the extent to which interviewer experience affects decisional outcomes, Carlson (1967) found that there is little difference between the extent to which experienced interviewers agree with each other and the extent to which inexperienced interviewers agree with each other. He postulated some reasons for this occurrence as being that (1)"managers (interviewers) need not share the same or highly similar experiences which would be necessary to increase inter-rater agreement" and (2) there is usually little systematic feedback which would serve to increase inter-and intra-rater agreement. The only report of where experience does affect outcomes is described in (4) below. This work tended to limit the veracity of conclusions as to the effects of interviewer experience as reported by Rowe (1963).

3. A study by Carlson, Schwab and Henemann (1970) showed that inter-rater agreement was higher in structured interviews than in unstructured or semi-structured interviews hence showing that with more structure the likelihood of valid selection is greater. The researchers controlled for three conditions of structure (structured, unstructured and semi-structured) and within each condition gave 6 male interviewers 5 job applicants to rank-order. Inter-interviewer agreement in terms of these rankings was highest for the structured group and lowest for the unstructured group. An earlier study by Schwab and Henemann (1969) also supported these conclusions.

4. When interviewers are behind a stipulated quota Carlson (1967) found that they generally tend to evaluate applicants higher than if no quota existed. Furthermore, he reported that inter-rater agreement was higher when raters were either extremely ahead of or behind schedule, although the ratings in these conditions were somewhat impaired. One other interesting finding which emerged here was that more experienced interviewers were less susceptible to the pressure conditions than less experienced interviewers. The former were more consistent in ratings with and without a quota than were the latter.

5. Carlson et al (1971) reported a recent study wherein it was found that when interviewers did not take notes or follow an interview guide the accuracy of recall of applicant

characteristics was lower than if they had followed such procedures. Furthermore, when his accuracy was low the interviewer tended to evaluate the applicant quite favourably indicating the existence of a "halo strategy."

6. Mayfield and Carlson (1966) postulated the hypothesis that stereotypes, consisting of features specific to each interviewer and general to an associated group of interviewers, form a major basis for employment decisions. A later study (Carlson et al, 1971) supported this hypothesis and lent further credence to Sydiaha's (1958) work, reported earlier. Furthermore, Carlson et al (1971) found that when the stereotype was in operation, inter-rater agreement was higher than when it was not. This was due to the finding that approximately 70% of the factors considered relevant to making decisions were common to the raters in the study, while 30% were specific to each rater. This gave empirical support to Rowe's (1963) contention that stereotypes tend to be "good" rather than "bad". However, it must be recognized that this does not ensure higher validities based on job behaviour criteria (Mayfield, 1964).

7. On the basis of two research studies, Carlson (1968, 1970) argued that interviewers do not rate in terms of an absolute standard but rather with respect to relative comparisons. Applicants being interviewed by one interviewer were

evaluated according to one another. It was found that when an average applicant was being considered by an interviewer who had just evaluated three or four very unfavourable applicants, the average applicant was rated very favourably. This finding is not consistent with other work performed by Hakel et al (Hakel, Ohnesorge and Dunnette, 1970) who reported that while "contrast effects" exist, they account for very minor amounts of variance.

From these findings the LIAMA group began to propose ways of improving the selection interview. The two major applied implications were stated as follows:

"First, the selection interview should be made an integral part of an over-all selection procedure, and to accomplish this, new and additional materials are needed. The new materials should include a broad-gauge, comprehensive, structured interview guide; standardized evaluation and prediction forms that aid the interviewer in summarizing information from all steps in the selection process; and an evaluation system that provides feedback to the interviewer in language similar to the preemployment job behaviour predictions he must make. The second major applied implication is that an intensive training program for interviewers is necessary if interviewers are to initially learn enough in common to increase the probability of obtaining general validity from the selection interview." (Carlson, et al, 1971, p. 273)

A second main stream of research in employment interviewing was pursued by Milton Hakel and his associates at the University of Ohio and the University of Minnesota. These investigators have followed a rather different course from the LIAMA team by focusing upon the theoretical and empirically

founded notion of interpersonal perception. In a recent monograph, Checklists for Describing Job Applicants (Hakel and Dunnette, 1970), the McGill studies (Webster, 1964) were interpreted in terms of this framework and a model of Interpersonal Perception was developed on which to base future studies of employment interviewing. While their research summary is not detailed here, an exploration is made of a number of important findings from their research. These are as follows:

1) As described earlier, Carlson (1968, 1970) offered evidence to support the argument that interviewers do not rate in terms of an absolute standard but rather with respect to relative comparisons. For example, an average applicant will be evaluated favourably when he is preceded by unfavourable applicants or unfavourably when preceded by favourable applicants. Rowe (1967) offered evidence which showed a similar result and led her to conclude that "whether an individual is accepted or rejected for a job may well depend more on the characteristics of the previous applicants than on his own traits". (p. 173).

However, Hakel, Ohnesorge and Dunnette (1970) provided evidence to severely limit the heretofore postulated influence of "contrast effects." They concluded that, indeed, contrast effects are present. Yet they only account for 1.2% of the decision variance for a group of 97 employment interviewers and 1.9% of the decision variance for a group of 102 male

psychology students, amounts which, they stated, are for practical purposes "nearly trivial." Furthermore, a re-evaluation of Rowe's (1967) data shows that contrast effects account for only .7% of decision variance. Contrast effects are not as potent as previously believed.

This finding is more consistent with the notion of "stereotypes" and their influence on decisional outcomes. Mayfield and Carlson (1966), Carlson et al (1971) and Sydiaha (1958) drew the conclusion that interviewers base decisions in large part upon the "stereotype" of an ideal or at least acceptable candidate. Furthermore, the stereotype consisted partly of features specific to each interviewer and, in the main, of features common to all interviewers (Carlson et al, 1971). This would suggest that the standard of evaluation is more absolute than relative, an idea quite contrary to the notion of contrast effects which implies a relative standard.

2) Recognizing that stereotypes play an important role in employment decisions, Hakel, Hollmann and Dunnette (1970) explored the extent to which interviewers' stereotypes are accurate. They took three samples of raters; interviewers (N = 14), Certified Public Accountants who interview (N = 23), and students (N = 20) and compared how well they identified the interests of accountants. The test instrument which they used was a 57 - item forced-choice test constructed using data from the CPA scale of the Strong Vocational Interest Blank. Their findings were essentially two-fold. First,

among these three samples, rater accuracy was not significantly different and in each case was quite low. Stereotypes were used as a basis for decisions but, unfortunately, were sorely lacking in accuracy. Second, factor analysis of the data yielded two distinct clusters. The first cluster consisted largely of CPA's while the second was populated by mostly interviewers and students. The authors concluded that the characteristics of raters result in an impression of a somewhat unique stereotype. CPA's have a similar background and hence form a similar stereotype (although this background is moderated by other factors such as age). Likewise, interviewers and students have formed stereotypes based on their exposure to the accounting profession; an exposure which is largely similar for most lay people. This interpretation is consistent with the theory and research underlying inter-personal perception (Hakel and Dunnette, 1970).

3. In his review of research on the employment interview, Wright (1969) supported emphasis on macroanalytic studies which would show the existence of similarities or differences across interviews. Hakel and Schuh (1971) have performed the only recent piece of research which suits Wright's suggestion. In this study, the authors identified 22 attributes judged to be important, frequently considered and favourable by seven diverse occupational groups. Their study identified some important similarities among occupations in terms of interviewing. As the authors suggest, incorporation of these 22 items into interviewer training programs would enhance interviewer

transferability. Also, these items could be used as a base for patterned interview guides for general use as recommended by Carlson et al (1971).

4. Considerable work has been performed isolating the favourability of applicant information as an independent variable and examining its effects upon decisional outcomes (Bolster and Springbett, 1961; Miller and Rose, 1967; Mayfield and Carlson, 1966; Carlson, 1968; Webster, 1964). However, until recently (Hakel, Dobmeyer and Dunnette, 1970) no work had been reported which examined the actual content area to which favourableness or unfavourableness had been attributed. This is an important concern since, as Hakel et al state, "it is difficult to believe that all content categories are of equal importance in determining ratings of overall suitability". These authors varied the extent of favourability of three content dimensions, scholastic standing, business experience and interest and activities, and developed twelve descriptive combinations in résumé-form. Two samples (CPA interviewer, N = 22; students, N = 20) were given the task of evaluating these résumés as to overall suitability. The findings demonstrated that content moderates the evaluation of job applicants. From among the three content dimensions manipulated, scholastic standing played a major role in interviewers' decisions. This finding does not infer that scholastic standing is the most important interviewee attribute. Rather, it suggests that at a specific level, among the three dimensions Hakel et al studied, scholastic standing was most

important. More important, at a general level, this research demonstrated that content categories have differential influences on decisional outcomes.

While they have contributed much to an understanding of what goes on during employment interviews, Hakel and his associates have only begun to explain some of the causative factors affecting decisional outcomes. Future research based on their model and Checklists as described in their recent monograph will undoubtedly shed more light in this regard.

CONCLUSION

Recent research has taken the direction suggested at the start of this research summary; that research on the employment interview should not be aimed at discrediting it so much toward improving its predictive capability.

Prior to the recent series of investigation, summarized above, the interview was felt to have low reliability and even less validity. However, current findings suggest ways to reduce these negative characteristics. Furthermore, Ghiselli (1966) has demonstrated that a skilled interviewer can elicit information quite adequately and can capably use it to predict future performance. Similar findings are gradually being reported (Grant and Bray, 1969; Palacios, Newberry and Bootgin, 1966). The present study has attempted to incorporate many of these findings into its design wherever relevant, as is described in Chapter 3.

SECTION 2: FUTURE RESEARCH DIRECTIONS

What further directions can research of the employment interview take? Undoubtedly there are many, as future research will demonstrate. However, five such proposals are explored here.

I. Different Samples

Webster's studies were conducted primarily with Personnel Selection Officers in the Canadian Armed Forces acting as interviewers. Carlson et al's analyses were largely made upon insurance agents. Replication of their analyses with other samples would offer more information in terms of the generalizability of their findings.

II. Interpersonal Perception

Hakel and Dunnette (1970) proposed a model of interpersonal perception which could be employed as a basis for analysis of decision making in the employment interview. These authors stated that "learning about another person can be viewed as essentially a matter of gathering and processing information about that person in the context of other information he has about himself, others he has observed and stereotypes he has formed". Involved in this framework is a highly complex network of processes and conditions, drawing not only upon perception theory but also upon personality theory and expectancy theory. Its utility as an interpretive system of the dynamics of interviewing is clearly shown in the Hakel and Dunnette monograph (1970).

The use of such an approach adds an important dimension to the study of the interview. Not only does it provide a

theoretical basis but it also offers methodologies for the examination of various constructs (e.g. perceptual accuracy and impression formation) which affect the decision making processes in interviews. Hakel and Dunnette (1970) demonstrated its theoretical usefulness as they interpreted several of Webster's (1964) findings in the light of interpersonal perception theory and related studies in person perception. Furthermore, the research reported in the monograph employed methodologies used in studies of person perception.

This approach, then, offers utility both in a theoretical and practical sense. With such a theoretical framework, findings in interviewing research may be tied together thereby facilitating interpretation and identifying interrelationships. To date, such a framework was clearly missing. As Lewin (1945) stated, "there is nothing so practical as a good theory". Hakel and Dunnette proposed such a theory.

III. Evaluation of Rating Forms

Much work has been done investigating factors influencing decisional outcomes. However, very little has been devoted to analyzing the actual rating form used by interviewers. Hakel et al used only overall suitability as the dependent variable in their microanalytic studies. Rowe (1970) advocated the use of a rank order technique rather than the accept-reject or trait-rating approach. She claimed that this approach would enhance intrarater reliability. Other researchers (Schwab and Heneman, 1969; Carlson, Schwab and Heneman, 1970) also adopted this approach. Carlson (1967) utilized an accept-

reject decision and a rank order approach. He also stated that subjects were to evaluate a "predicted behavior rating scale" although he offered no description. Springbett (1954 cited in Webster, 1964) used a six-point descriptive-anchored acceptance-rejection scale. Crowell (1961 cited in Webster, 1964) asked interviewers to either accept or reject applicants as did Sydiaha (1961) and Rowe (1963).

Other studies have used trait-rating scales (e.g. Wagner, 1949) as dependent variables. Besides Wagner's summary of reliability and validity of interviews using these scales as dependent variables, Rowe (1970) addressed the problem of how rating-forms (be they ratings of traits and/or overall suitability, accept-reject decisions, or rank-orderings) affect inter- and intra-rater reliability. Using rankings, she reported intra-rater reliability scores as high as .812, with 17 of 20 coefficients significant at least at the .05 level. Wagner reported reliability coefficients as high as .96 for the intelligence trait, .87 for sociability and .77 for self-confidence. For overall ability, a maximum rating of .85 was reported (Scott, Bingham and Whipple, 1916). The method for achieving this coefficient was a correlation between sets of rank-ordering. While this supports Rowe's contention, it must be taken rather lightly due to methodological inadequacies, as described by Wagner.

It would seem that a simple accept-reject decision would be most suitable. After all, this reflects the primary function of the interview. However, it appears that most

interviewers like to have some record of the basis for acceptance or rejection of candidates and consequently include traits in their rating forms. Unfortunately, as Wagner (1949) pointed out the reliability and validity for most trait ratings is at best meagre. Also, very frequently there is little, if any, evidence to suggest that the traits examined have any bearing on potential success. Thirdly, as Wonderlic (1942) stated "few (interviewers) follow a well defined pattern and the interview generally amounts to a disorganized conversation resulting in a series of impressions based upon impulsive reactions". There usually is no systematic procedure in the interview for forming adequate trait impressions. Finally, there is no evidence to show the differential influence which trait ratings have upon overall suitability.

Perhaps the most constructive direction for exploring interview rating forms was taken by Maas (1965) using the procedure proposed by Smith and Kendall (1963). One problem with many trait rating forms is that there is often little agreement among raters as to the meaning of certain traits and as to the interpretation of the trait levels. What might be "good" to one rater might be "very good" to another. What might be rated as "1" by one rater might be rated "3" by another. Maas (1965) addressed this problem by constructing a "patterned scaled expectation interview". Employing Smith and Kendall's (1963) technique, Maas rigorously identified a series of traits which were deemed to be important for the

position of Orientation Counsellor. Around these traits he then developed written examples of on-the-job behaviours to illustrate three levels of each trait - a high degree of the trait, an average degree, and a low degree. Instead of the traditional rating adjectives (e.g. good, very good, satisfactory, etc.) for each trait level, then, Maas employed behavioural descriptions of trait levels. Interviewers were asked to rate each candidate on each trait "by making analogies from the candidate's responses, to behaviour that might be expected of the candidate, were he actually on the job."

(p.432). A total of 2,268 interviews were conducted to study inter-rater reliability using two different rating scales; traditional adjective rating scales and the scaled expectation technique. Using patterned interviews with both types of scales, Maas found significantly higher inter-rater reliability coefficients with the scaled expectation technique (.65 - .72) than with the adjective scales (.34 - .35). This was in agreement with the study reported by Smith and Kendall (1963) employing the same technique in a non-interview setting.

Maas' study was performed in an educational setting. The present study examines the rating form used by the Bank of Montreal, i.e. in an industrial setting. The rating form is studied in terms of inter- and intra-rater reliability as well as halo error.

IV. Modes of Applicant Presentation - Video-Tape

As described earlier, the LIAMA group and the McGill group generally employed a paper and pencil approach when presenting candidates. The interviewees were not physically present but were described in written form. Mayfield and Carlson (1966) describe this approach as being quick, thereby permitting a wide scope of applicant information to be given, and experimentally convenient, in that it permits control over extraneous variables. They further stated that "results obtained by this method could be compared later to those obtained when information is presented by other means...to determine if the mode of presentation has any effect". (p.43).

Certainly other modes of information preparation have been utilized. Kasl and Mahl (1956) used tape recordings of actual interviews as did Wiens, Molde, Holman and Matarazzo (1966). Findings from the latter study suggest that interview interaction measures can be reliably gathered from tape recorded interviews.

However, only one recent study (Blakeney and MacNaughton, 1971) has attempted to replicate any of these decision-making studies using a mode of presentation different from the paper and pencil technique. These authors used tape recordings to test the veracity of Bolster and Springbett's (1961) conclusion that there was a modified primacy effect operating in interviews.¹

¹ This "modified primacy effect" was based on Bruner's (1957) hypothesis that a "gating" phenomenon exists in interviews,

Their findings did not fully support Bolster and Springbett's and the question was raised as to whether or not the difference occurred as a result of using an alternate mode of applicant presentation. Indeed, interview research would be thrown in a turmoil if this question were answered affirmatively.

Certainly the issue must be raised as to whether or not the experimental convenience of the paper and pencil technique compensates for its lack of realism. The interview has utility mainly in terms of social interaction, a condition completely eliminated with paper and pencil. Possibly other means such as tape recordings or video tapes could be employed. Blakeney and MacNaughton (1971) have demonstrated that the content of audio-taped interviews can be capably manipulated. Furthermore, Grant and Bray (1969) demonstrated that audio tapes of relatively unstructured interviews can offer reliable and valid information. As a basis for establishing intra-rater reliability, taped

i.e., interviewers decrease the range of stimuli they perceive as the interview progresses. Bolster and Springbett (1961) looked at this phenomenon in terms of the effects of placement of unfavourable information at various stages throughout the interview. They argued and supported the hypothesis that if unfavourable information comes early in the interview a rejection is more likely than if it comes later. Blakeney and MacNaughton hypothesized that (a) if negative information comes in the first third of the interview the ratings of applicants will be the lowest, (b) if negative information comes during the second third of the interview ratings will be intermediate, (c) if it comes in the last third ratings will be highest. However, they found no significant difference between the rating when negative information was presented during the first third of the interview and rating when negative information was presented during the second third.

playbacks can serve a useful function. As Crissy (1952) described, a design for estimating intra-rater reliability could consist of making "soundscripts" of completed interviews. After a time period, the interviewer could re-appraise each interviewee on the basis of the playbacks. Using either audio-or video-tape recordings this method is feasible.

No work in these regards has been performed with video-tape specifically, although some investigations have been made. Moore and Craik (1971) explored personnel administrators' and students' (potential interviewees) attitudes towards the use of video-tape as an aid to employment interviewing. The most relevant aspect of their research here is the respondents perceptions of how realistically playback of video-tape interviews portrays a number of important interviewee characteristics. These characteristics include appearance, manner, voice, expression, force or drive, intelligence, interest, social sensitivity, experience, knowledge of field, nervousness, stress and judgment. No significant differences between the two samples were found. More than sixty-seven percent of the administrators checked "somewhat realistic" or "very realistic" on all characteristics except for social sensitivity, knowledge of field, stress and judgment. Over sixty-seven percent of the students checked "somewhat realistic" or "very realistic" for all characteristics except judgment. For the characteristics which less than sixty seven percent of the respondents in both samples checked as being realistic, an "unable to judge" response was very highly rated, indicating

that these characteristics were probably less prominent in the interviews shown. This research is important in that it demonstrated that characteristics of interviewees commonly held to be important by raters are adequately portrayed over video-tape.

However, a more important issue exists; namely whether or not the ratings given by viewers of video taped interviews are in any way different from ratings given by face-to-face interviewers. This has importance in both practice and research. If an organization chooses to use video-tape in its personnel selection process it is important to have knowledge of the differences between face-to-face and video-taped interviews in order to compensate for them.

Furthermore, if interviewing research is conducted using video-tape or films, as has been done or proposed (Wexley, Yukl, Kovacs and Sanders, 1972; Cline and Richards, 1961; Crissy, 1952), then it is necessary to recognize that perhaps the video-taped interview is different in some ways than the actual face-to-face interview. McLuhan (1964) argued that the medium over which a message is transmitted forms part of the message itself. If this is the case, then researchers using video-tape are not studying the interview per se but the videotaped interview.

This issue is examined here.

V. Group vs Individual Raters

Much work in social psychological research has been performed exploring differences between groups and individuals

in terms of decision making. However, very little research has focused upon these differences in terms of decision-making in the employment interview.

Most of the early comparative work was directed at exploring the validity and reliability of group versus individual decisional outcomes. Table 1 provides a summary of the reliabilities found in some of these studies.

Hall, Mouton and Blake (1963) provided a concise classification of research findings comparing groups and individuals in terms of decision-making outcomes. They segmented three historically accepted theoretical positions. The first is the notion of "pooled products". Here, an average of individual decisions is taken as being the more correct than any one person's decision. This statistical pooling was seen by other researchers as being too simplistic and not able to properly explain the group and individual differences. This led to the second position termed the "emergent product." Here, the stand taken was that discussion and, generally, interpersonal activity in a group "carried the group toward a correct rather than an incorrect decision." The important point here, then, is the positive effect of interaction among group participants largely of the form elucidated by Maier (1967) in his description of "group assets." The third position is called the "compromise product". The key notion here is "bargaining" leading to compromise in contrast to the "integration of the best ideas" of the participants as reflected

Table 1

The Reliability of Judgments of Individuals
and of Groups

Type of Judgment and Author	Individuals	Groups Size of Groups					
		5	10	20	40	50	60
Weights: Gordon, 1924	.41	.68	.79	.86		.94	
Weights: Stroop, 1932	.38	.68	.85	.92		.96	
Weights: Bruce, 1935	.50	.67	.83	.86	.87		.88
Numerosity of shot: Bruce, 1935	.82	.87	.94	.94	.84		
Personality Traits Smith, 1931	.37	.46	.49	.49		.51	
Esthetic Judg- ments Eysenck, 1939	.47	.77	.86	.94		.98	

Zajonc, 1966, p.100

Group coefficients were based on statistical pooling.

in the "emergent product" position. What is argued here is that as a result of largely political activities in a group, a group's decision will be more mediocre than the average of the individual decisions. The forces operating here resemble those headed by Maier's (1967) "group liabilities" classification.

Increased concern for group functioning and processes arose primarily as a result of the Human Relations movement and the consequent push to involve subordinates to take active roles in the decision making process. A wealth of research has been undertaken to explore the elements of effective group functioning. A summary of research findings in this regard is found in Hall and Williams (1970), Cartwright and Zander (1968) and Lorge et al (1958).

As stated earlier, there is little work reported which investigated differences between groups and individuals in terms of employment decisions. A number of studies have used groups (Schwab and Heneman, 1969; Carlson, Schwab and Heneman, 1970), yet only one offered reliability data (Howell and Vincent, 1970) and another examined group vs individual accuracy in judging personality (Cline and Richards, 1961). Howell and Vincent (1970) reported reliability coefficients of .89 and .92 for three member boards and .91 to .94 for boards of four members. In the exploration of aspects of interpersonal perception, Cline and Richards (1961) had a sample of 186 students view six filmed interviews and first rate the interviewers on personality factors first as individual raters

and then as groups of three. Comparisons were made between and among (a) the independent individual predictions, (b) the group consensus predictions, (c) the accuracy of an "artificial group" derived through a statistical combination of the independent predictions of these same three persons and (d) the accuracy of the "best judge" from each group. Their findings showed that least accurate ratings were obtained from individuals. They also found no significant differences between the other three procedures. Although, in terms of time, and procedural difficulties, the artificial group appeared most satisfactory.

Recently, Hollowman and Hendrick (1971) compared group consensus scores to averaged individual scores for decision accuracy when group size was varied. For groups of several sizes (3,6,9,12,15) group scores were more accurate than averages of individual scores in completing a complex decision making task requiring group interaction and discussion.

In summary, then, these studies suggested that groups offer more reliable and accurate decisions than individuals. This conclusion however must be only tentatively accepted. Campbell (1968) showed that with a complex realistic situation, the Change of Work Problem (Maier, Solem and Maier, 1957), "the quality of the group solution was inferior to the nominal group's composite score and was even inferior to the average individual solution" (p.209). Campbell stated that group participation and discussion tended to be inhibitory rather

than beneficial. He further concluded that comparisons between group and individual decisions hinged upon the type of problem used for analysis.

In terms of the employment decision type of problem no research has been reported which directly compares individuals and groups. Group or panel interviewing is widely used (OSS, 1948; Taft, 1959) yet its effectiveness compared to individual interviewing has not been examined.

This paper explores differences between group and individual decisions in the interview setting.

Summary

This research summary was intended to highlight the findings of recent investigations examining the employment interview and to focus upon a number of areas where future research could be directed and where the research reported in this paper is justified. Specifically, this paper intends to analyze three major questions related to the interview.

These are:

1. What differences if any exist between ratings given by viewers of video-taped interviews and by actual interviewers in face-to-face settings? This would provide information regarding considerations to make when video-tape is intended to be used for either employment decisions or research on the employment interview.

2. What differences exist between ratings of interviewees as given by group and individual raters? No research has been reported which looks at these differences in the

interview setting. The types of dimensions analyzed here would have a bearing on some of the psychometric properties of the Bank of Montreal's rating form.

3. What attitudes do interviewees and rating viewers have toward the use of video-tape in interviewing? This would provide information on what people who are actually involved in the interview process think about this mode of interview presentation.

CHAPTER 3

METHODOLOGY

As stated in Chapter 1, the main purposes of the study were three-fold. The first aim was to determine the existence of any systematic differences between ratings given by interviewers in actual face-to-face interviews and ratings given by observers of video-taped interviews. The second purpose was to examine differences between group and individual raters in terms of mean variable ratings, overall ratings, call back-reject decisions, halo error and inter- and intra-rater reliability. The third intention was to explore the attitudes of interviewees and observing raters toward the use of video-tape in interviewing. To obtain the necessary data, each of these three purposes required somewhat different methods as explained below.

I. Differences in Ratings - Interviewers vs Viewers

Procedure

In order to investigate the existence of any systematic differences between ratings given by interviewers in face-to-face interviews and ratings given by viewers of video-taped interviews, thirty-six interviews were conducted between six experienced interviewers from the Bank of Montreal and thirty-six fourth year Commerce students from the

University of British Columbia. The opportunities for which these interviews were held were positions in the Bank's management training program; a program designed to lead trainees to a branch managership after a period of three years. The six interviewers from the Bank each interviewed and rated six different students. These students were all seriously interested in the Bank and were largely representative of the labour market from which the Bank draws its management-trainees. Each of these interviews were video-taped on a one-half inch v-t system using a split-screen technique.¹ Due to technical difficulties in the video-taping, two of the interviews were omitted from the analysis, leaving a final sample size of thirty-four.

At the end of videotaping the 34 interviews, six groups of three persons and six individuals each viewed and rated approximately six different interviews. The viewings were arranged in such a way that no viewer saw the same interviewer on tape more than once, and no group and individual saw the same tape together more than once. This provided the maximum mix of responses and reduced bias due to any possible redundancies. A schedule of which group and which individual viewed which taped interview is shown on Table 1.

¹With this split-screen technique two cameras were used, one for the interviewer and one for the interviewee. Using a special effects generator, the images from both cameras were played on one screen. In our case, the interviewer occupied the left half of the screen and the interviewee the right half. MacDonald (1971) reports evidence supporting the utility of this technique when video-taping interviews.

Table 1

Schedule of Interviewing and Viewing

<u>Interview Number</u>	<u>Inter- viewer</u>	<u>Group</u>	<u>Individual</u>	<u>Interview Number</u>	<u>Inter- viewer</u>	<u>Group</u>	<u>Indiv- idual</u>
1	1	1	1	19	1	4	4
2	2	1	2	20	2	4	5
3	3	1	3	21	3	4	6
4	4	1	4	22	4	4	1
5	5	1	5	*23	5	4	2
6	6	1	6	24	6	4	3
7	1	2	6	25	1	5	3
8	2	2	1	26	2	5	4
9	3	3	2	27	3	5	5
*10	4	2	3	28	4	5	6
11	5	2	4	29	5	5	1
12	6	2	5	30	6	5	2
13	1	3	5	31	1	6	2
14	2	3	6	32	2	6	3
15	3	3	1	33	3	6	4
16	4	3	2	34	4	6	5
17	5	3	3	35	5	6	6
18	6	3	4	36	6	6	1

* Due to technical difficulties these two interviews were eliminated.

The groups were allowed to discuss as much as they wished. To avoid any influence which the group discussion may have upon the individual viewer, the groups and individuals viewed tapes separately.

The Samples

(a) The interviewees: All of the students who participated as interviewees were male undergraduates in their fourth year of Commerce at the University of British Columbia. 44.1% (N=15) were students of finance or accounting. 47.1% (N=16) were enrolled in Industrial Relations and Organizational Behaviour. The remaining 8.8% (N=3) were in other miscellaneous fields (Transportation, Marketing). Their average age was 27.2 years (median = 23.0, mode = 22.0, range = 21-49.) The average grade level achieved by these students was 70.0% (median = 70.0%, mode = 70.0%, range = 60% - 77%). The average number of interviews which these students had been to during the year prior to this study was 5.4 (range = 0-16). Furthermore, each interviewee had received fairly extensive prior exposure to video-tape in their classes at U.B.C. As a consequence, any bias due to nervousness or fear related to the video-taping was minimal. This bias, called the "reactivity effect" was discussed by Walter and Miles (1971).

(b) The interviewers: Details of interviewers' characteristics are shown in Table 2. As may be seen, each interviewer has had similar experience and training. This reduces

Table 2

Interviewer Characteristics

Interviewer Number	Sex	Age	Position	Interviewing Experience	Number of inter- views <u>this study</u>
1	Male	29	Employment Manager	Has been formally inter- viewing for 1½ years. Took a 2-week interview- ing course in Detroit	6
2	Male	30	Assistant Employment Manager	Had been formally inter- viewing for 1 year. Took course in Detroit	6
3	Female	24	Personnel Officer	Had been formally inter- viewing for 2 years. Took course in Toronto (1 week)	6
4	Male	34	Manpower Manager	Had been formally inter- viewing for 4 years. Took course in Detroit	5
5	Male	29	Assistant Manpower Manager	Had been formally inter- viewing for 6 months. Took course in Vancouver (1-week)	5
6	Male	29	Accountant	Had been formally inter- viewing for 1 year. Took course in Vancouver	6

any bias which may exist due to variations in interviewer experiences (Carlson, 1967a; Rowe, 1963). Also, since each interviewer has received principally the same in-house training (i.e. they have all been exposed to the Bank's directives and policy statements dealing with employee selection) their stereotypes of ideal candidates should be roughly equivalent thereby reducing bias due to stereotype variations (Rowe, 1963; Sydiaha, 1958; Mayfield and Carlson, 1966; Carlson et al, 1971). Furthermore, interviewers were permitted to take notes and, as described below, were trained to conduct similarly structured interviews with consequent similar interview guides. This permits accuracy of recall of applicant characteristics (Carlson et al, 1971) and reduces the likelihood of ratings being based on a halo strategy. Finally, each interviewer was given exposure to the video-tape prior to the actual interviewing. This was designed to reduce the reactivity effect due to the video-taping (Walter and Miles, 1971).

(c) The viewers: All of the viewers (individuals and groups) were managers with the Bank of Montreal and all had had some involvement in employment interviewing, if not with the initial screening interview, then at least for post initial employment interviews. The viewers' average age, average number of years with the Bank and average number of years they have been interviewing is as in Table 3.

Table 3

Viewer Characteristics

	Group	Individual	Combined
Age (years)	35	45	38
Number of years with Bank	15	25	18
Number of years they have interviewed	6	14	8

As with the interviewers, the viewers have had roughly equivalent experience and in-house training thereby reducing biases due to experience variations and stereotype variations. Also, for all viewers, note-taking was permitted.

(d) intra-sample biases: A bias may be introduced into the data if there were any significant differences among rating-variance within any of these three sub-samples. To test for the existence of these differences an analysis of variance was applied to the overall ratings given by each interviewer, group and individual, in each of the six successive interviews rated. The tests (see Table 4) showed no significant differences indicating that ratings within each of the three sub-samples were internally consistent.

The Measure

The measure employed here is the Bank of Montreal's standard rating form for evaluating University graduates for the management training program. The form consists of twelve

Table 4

Analyses of Variance

(a) Interviewers	of Sq.	D.f.	Est. of Var.	F
Total	6.97	33		
Between	.81	5	.162	
Within	6.16	28	.220	.736
(b) Groups				
Total	5.53	33		
Between	1.27	5	.254	
Within	4.26	28	.152	1.67
(c) Individuals				
Total	10.62	33		
Between	1.12	5	.224	
Within	9.50	28	.339	.661

With $df_1 = 28$ and $df_2 = 5$, the F-value should be greater than 2.56(.05), 3.75(.01) and 5.66(.001) to identify a significant difference.

Source: Blalock, H.M. Social Statistics. McGraw-Hill: New York, 1960.

trait and personality factors and an overall factor for which the rater is asked to evaluate the interviewer on a five-point scale. Also, a space is given for the rater to check his accept-reject decision. The measure is shown on Exhibit 1. As may be seen, each trait or personality factor is accompanied with a brief definition as is each rating term.

Numerous weaknesses associated with this measure were recognized. First, no reliability or validity data have been accumulated to demonstrate its utility. Second, no attempt has been made to determine whether the factors have any bearing on potential success. Third, no rationale as to why a number of the factors were included could be located. However, when reviewing the literature, a number of studies were located which provided empirical support for the inclusion of a number of the factors. First, though, it should be stressed that this measure was selected primarily because it is typical of the types of rating forms so commonly used by organizations for personnel selection. Also, it formed the basis of investigation for the third part of the study wherein an exploration of some psychometric properties was made.

As discussed in Chapter 2, a number of studies show evidence supporting the inclusion of a number of these factors. Ulrich and Trumbo (1965) review research which show that sociability and motivation to work may be ably measured in the interview. Mayfield (1964) states that intelligence

GRADUATE RECRUITING - CAMPUS INTERVIEW

EXHIBIT 1
BANK OF MONTREAL

Instructions overleaf

CONFIDENTIAL

1 DIVISION _____ DATE _____
NAME (Surname First) _____ RECRUITER _____
UNIVERSITY _____ DEGREE _____ MAJOR _____

2 **DO NOT** complete this section if a **COMPLETED** resume or information sheet is attached to this form.

ADDRESS (Residence while attending university) _____
CITY _____ POSTAL ZONE _____ PROVINCE _____ PHONE NUMBER _____
MARITAL STATUS: SINGLE ☐ MARRIED ☐ AGE _____ NO. OF DEPENDANTS _____

3 MOST RECENT ACADEMIC STANDING _____ NUMBER IN CLASS _____ AVERAGE MARK — ALL COURSES _____
AVERAGE MARKS: 1st Yr. _____ 2nd Yr. _____ 3rd Yr. _____ 4th Yr. _____ 1st Yr. Post Grad. _____ 2nd Yr. Post Grad. _____

4 EVALUATION

	ATTITUDE	APPEARANCE	INTEREST	INTELLIGENCE	LEADERSHIP	MATURITY	MOTIVATION	PERSUASIVENESS	SELF-CONFIDENCE	SELF-EXPRESSION	SOCIABILITY	POTENTIAL
EXCELLENT												
SUPERIOR												
AVERAGE												
MARGINAL												
UNSATISFACTORY												

FACULTY EVALUATION (If Available) ☐ EXCELLENT ☐ SUPERIOR ☐ AVERAGE ☐ MARGINAL ☐ UNSATISFACTORY
RECRUITER'S OVERALL EVALUATION ☐ ☐ ☐ ☐ ☐

SHOULD CANDIDATE BE CONSIDERED FURTHER?

☐ YES ☐ NO

COMPLETE SECTION 5
ONLY IF "YES"

5 AREAS OF INTEREST FOR DIVISIONAL INTERVIEW: GENERAL BANKING ☐ ADMINISTRATION ☐ CREDIT ☐
MARKETING ☐ PERSONNEL ☐ OTHER ☐
WAS SALARY DISCUSSED? ☐ YES ☐ NO RATE \$ _____ REACTION: FAVOURABLE ☐ UNFAVOURABLE (Explain in Comments below) ☐

IS CANDIDATE WILLING TO ACCEPT INITIAL PLACEMENT IN ANOTHER DIVISION? YES ☐ NO (Explain in Comments below) ☐

LOCATION PREFERENCE, IF ANY (Number first three choices)

☐ B.C. ☐ Alta. ☐ Man.-Sask. ☐ Ont. ☐ Que. ☐ Mtl. ☐ Atlantic

COMMENTS: _____

(If necessary, continue overleaf)

H.O. 1510-23914

SIGNED _____

COMMENTS (continued)

RATING FACTORS :

ATTITUDE	— outlook in general.
APPEARANCE	— physical appearance, neatness, posture, dress.
INTEREST	— indications of sincere interest in Bank of Montreal.
INTELLIGENCE	— mental ability, judgment, alertness, organization of thoughts.
LEADERSHIP	— degree of leadership experience, extracurricular positions held.
MATURITY	— social behaviour and emotional stability.
MOTIVATION	— initiative, drive enthusiasm, energy, desire to succeed, aggressiveness.
PERSUASIVENESS	— ability to influence others.
SELF-CONFIDENCE	— ease, self-assurance, interest in challenge.
SELF-EXPRESSION	— ability to express thoughts clearly, concisely, effectively.
SOCIABILITY	— ability to work and get along with others, warmth, response.
POTENTIAL	— likelihood of success in management of Bank of Montreal.

RATING TERMS :

EXCELLENT	— definitely stands out, exceptional, makes immediate and lasting impression.
SUPERIOR	— well above average, a significant asset.
AVERAGE	— normal for a person of similar age, education and experience.
MARGINAL	— does not meet minimum standard.
UNSATISFACTORY	— unsuitable.

GENERAL INSTRUCTIONS :

- A separate form is to be completed for each applicant interviewed during and **immediately** following the interview.
- All forms are to be returned to the Personnel Manager, at the conclusion of interviewing at each university or technical institute visited.

SECTION 1: Complete: the DIVISION responsible for graduate recruiting at the institution being visited; the DATE of the interview; the NAME of the candidate being interviewed (surname first followed by one First name and initials); the name of the RECRUITER; the name of the UNIVERSITY or technical institute being visited; the DEGREE or certificate title the individual hopes to obtain; the MAJOR area of course study.

SECTION 2: Complete this section only if the information is **NOT** shown on a resume or information sheet supplied by the student or the placement office. Such sheet is to be **STAPLED** to this form. When necessary complete: the ADDRESS of the student while he is attending university or the technical institute being visited; the CITY, POSTAL ZONE and PROVINCE of this address, and the PHONE NUMBER where the student resides; MARITAL STATUS; Student's AGE; and NO. OF DEPENDENTS (if applicable).

SECTION 3: Complete MOST RECENT ACADEMIC STANDING showing the position in the class of the applicant (upper, middle or lower third as a minimum); the total NUMBER IN THE CLASS; AVERAGE MARK IN ALL COURSES for the last set of examinations; AVERAGE MARKS for each year of university or technical institute studies completed.

SECTION 4: Complete the EVALUATION using the rating factors and rating terms as defined above. Where possible an overall FACULTY EVALUATION (usually available from Placement Officer) should be completed and in all cases the RECRUITER'S OVERALL EVALUATION **MUST** be completed.

IMPORTANT: A DECISION **MUST** BE MADE BY THE INTERVIEWER AS TO WHETHER THE CANDIDATE IS TO BE CONSIDERED FURTHER.

SECTION 5: is only to be completed where the candidate **IS** to be considered further. Indicate AREAS OF INTEREST FOR DIVISIONAL INTERVIEW to facilitate the selection of individuals the candidate is to see during the divisional visit. If SALARY is DISCUSSED indicate this and note the RATE. An UNFAVOURABLE reaction **MUST** be explained in COMMENTS section below. Determine whether the candidate is willing to commence employment in ANOTHER DIVISION and explain the reasons for a **NO** answer in COMMENTS section below. Indicate FIRST THREE CHOICES by numbering "1", "2", "3", as to Division in which to commence employment. All COMMENTS favourable or unfavourable are to be shown. Additional comments if necessary, include on the reverse side of this form. This form must be SIGNED by the recruiter.

may be judged satisfactorily. Wagner (1949) reviews studies which report that self-confidence, sociability, intelligence and overall ability may be reliably measured through the interview. Furthermore, Howell and Vincent (1970) demonstrated that self-expression, appearance, maturity, self-confidence, sociability and interest as well as a number of other factors may be independently assessed through an employment interview. Only three factors, attitude, leadership and persuasiveness, have no empirical support for their use. However, since they formed part of the Bank's rating form they are included here.

The next question which arises is how well does an interviewer discriminate between these dimensions and how much of his ratings is based on a "halo strategy". This will be discussed more fully later on in this chapter.

Of Statistical Concern

In order to compare ratings among interviewers, group raters and individual raters, the Pearson Product-Moment Correlation was originally used. However, as Brown, Lucero and Foss (1962) discuss, this statistic has a limitation in situations where a five-point classification scale is used as the basis for measurement. They suggest that the r 's value is lost if the measures are coarsely grouped or if distributions are limited or skewed. They recommend the use of the "Percent Perfect Agreement" (PPA) statistic as a better way of examining relatedness between samples. They state that "it is the measure most consistent with Technical

Recommendations (of the American Psychological Association, 1954) and is the only measure of reliability proposed by Goodman and Kruskal (1954) in their general discussion of measures of correlation for classificatory variables".

As a result, the principal measure of correlation employed in this study is the Percent Perfect Agreement. To offer a comparison, the Pearson r will also be given with each PPA in the presentation of the analyses. However, when tables are developed which exhibit intercorrelations among factors and sub-samples (e.g. the Campbell-Fiske design) only the Pearson r was derived due primarily to the convenience of computer analyses. For most major computations, though, the PPA was formulated.

The Interview

The format of interviews performed in this study was the same as the Bank of Montreal's initial screening interview. Essentially, the format was semi-structured (Carlson, Schwab and Heneman, 1970) with structure surrounding the overall interview plan and steps III, IV and VI (see below) within the plan. Beyond this plan, non-directive probing was in existence (unstructured) wherein the interviewers could ask or probe for any additional information they felt pertinent for their evaluations.

The overall plan is as follows, with the suggested breakdown of steps to be observed in an interview and with the suggested time for each step.

Step I	Review resume	1 minute
Step II	Establish rapport	2 minutes
Step III	Evaluate student	
	- education and work experience	- 3 min
	- personal history	- 2 min
	- potential	- <u>5 min</u> 10 minutes
Step IV	Provide information	7 minutes
Step V	Questions and answers	5 minutes
Step VI	Inform student of future consideration	1 minute
Step VII	Record results and evaluate	<u>4 minutes</u>
		30 minutes

A more detailed explanation of each step is given in Appendix 1.

The Bank's suggestion was that the interview itself (i.e. excluding Step VII - evaluation) should take 26 minutes. The average length of the interviews in this study was 26.9 minutes. The interviews conducted in this study were basically the same in format (as described above) although there was variation in overall length (range: 18-38 minutes) and in the amount of time devoted to any particular step. This, though, is to be expected.

Furthermore, no quota restrictions were given to the interviewers. They were simply asked to select from among their interviewees those who would be suitable for further consideration and those who would be rejected. This avoids any bias in ratings which may arise due to the presence of quotas (Carlson, 1967).

To place this initial screening interview in perspective, a description of the Bank of Montreal's hiring system is given below.

The Bank's Hiring System

As stated earlier in this chapter, the opportunities for which these interviews were held were positions in the Bank's three-year training program (called the Special Development Program) leading to a branch managership.

The hiring procedure at the Special Development Program level relies almost entirely upon the interview. Besides references and application form data the decision to hire or reject is made on the basis of information gathered through interviewing.

The applicant first completes the application form (see Appendix 2) and then proceeds to the initial screening interview, details about which were given earlier. At this stage the interviewer makes the decision to either reject the candidate or recommend him for further consideration. The decision to hire is not made here.

If the applicant is recommended he then goes down to the Bank's Personnel Department at the British Columbia Regional Office in Vancouver for the second interview. This interview generally lasts from 45 minutes to 1 hour and consists mainly of describing the Bank and its opportunities to the interviewee. Here the interviewer goes into considerably more depth than in the initial screening interview, talking about performance appraisal, working conditions,

organizational structure, other programs, types of courses the recruit would have to take, and so on.

If recommended for further consideration the applicant comes back for a third interview with either a line officer or a staff officer. Line officers include bank managers, accountants, loan officers, credit officers, etc. Staff officers include personnel managers, systems analysts, and so on. This interview usually lasts about one-half hour with the format being roughly equivalent to the initial screening interview. The principal difference between this third interview and the initial screening interview is that the former deemphasizes the "provide information" phase (step IV) and stresses the acquisition of more data regarding the interviewee. This third interview is primarily designed to introduce the applicant to the actual internal workings of the Bank.

If recommended again, the candidate comes back for a fourth interview with a line officer (if the third interview was with a staff officer) or with a staff officer (if the third interview was with a line officer). This interview is very informal and unstructured. Usually it includes a luncheon. At this stage, the applicant is generally considered to be accepted for employment and is consequently given a more relaxed reception. This interview may even be considered as part of the induction process rather than the selection process. However, in the Bank it is considered as part of the latter, since there is the possibility that the

applicant could be rejected at this stage.

Graphically, the Bank's selection process for recruits to the Special Development Program may be depicted as on Exhibit 2.

II. Groups vs Individuals - Psychometric Properties of the Measure

This section of the study was primarily designed to explore some of the more important psychometric properties of the principal measure (the Bank's rating form - Exhibit 1), namely inter- and intra-rater reliability and halo error. Furthermore, ratings of interviewees were obtained from both group raters and individual raters to examine the moderating effects which these alternative sources of judgments may have upon these psychometric properties. While the measure has some empirical support, albeit indirect (as described earlier), and while it is of the type most commonly employed in personnel selection, the possibility still exists for it to be psychometrically lacking. This section of the study was intended to provide more definite and meaningful evidence as to its utility and limitations.

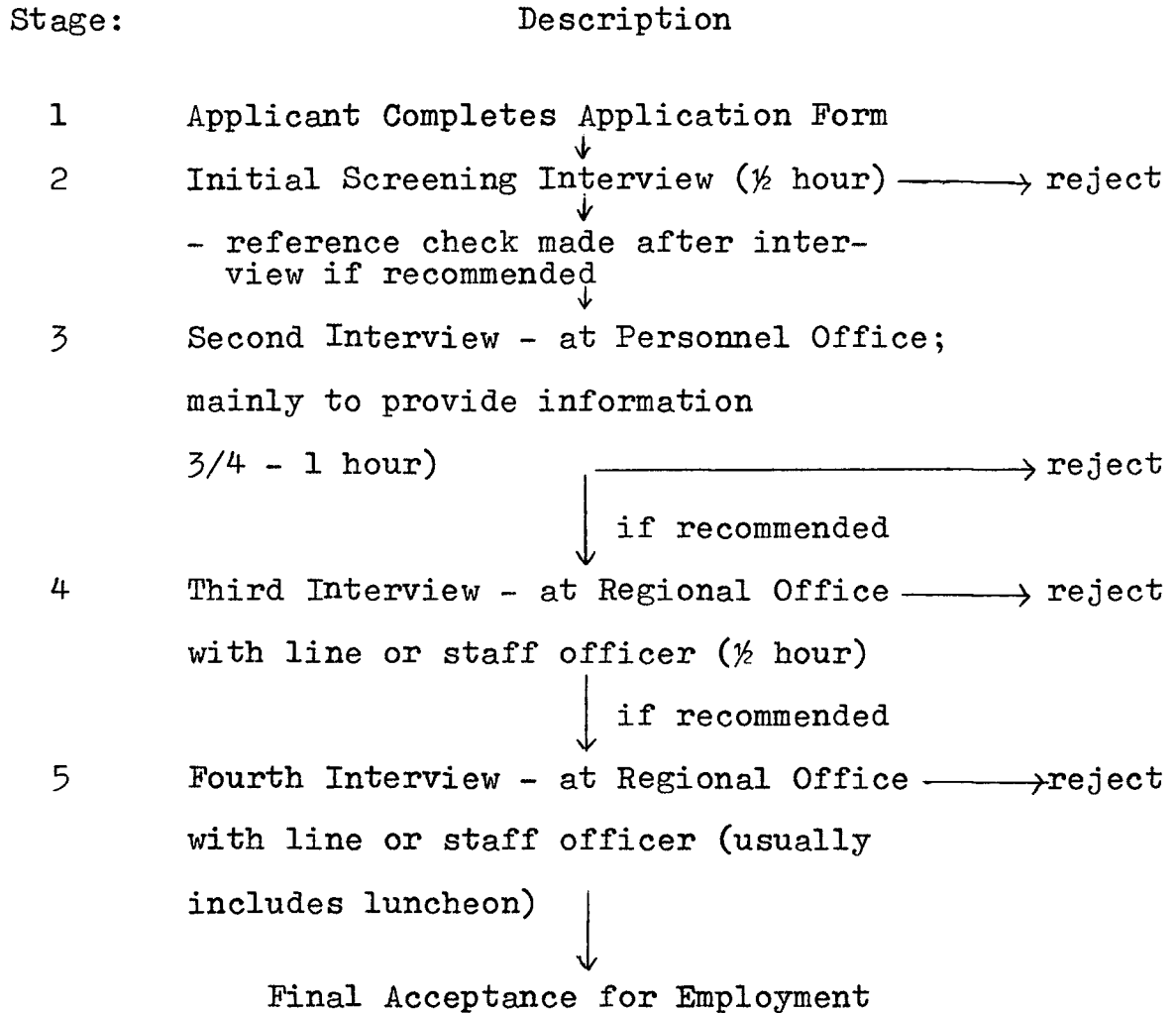
Finally, the group and individual ratings were compared to ascertain the presence and direction of mean rating differences.

Details of the Samples

To obtain the necessary data, 185 students from the Faculty of Commerce and Business Administration at the

Exhibit 2

The Bank of Montreal's Selection Process
for Recruits to the Special Development
Program



University of British Columbia were asked to provide rate - rerate data of interviewees from three video-taped interviews. Since the tapes last an average of roughly 27 minutes each, it was virtually impossible to expect the students to view three consecutive tapes twice. Consequently, the samples were broken into three approximately equal sub-samples, with each sub-sample viewing and re-viewing one tape. Furthermore, the three sub-samples were again segmented into group raters (of size 3) and individual raters (yielding 6 sub-samples).

By using video-taped interview, the group and individual raters were exposed to exactly the same thing thereby eliminating bias due to candidate inconsistency, a condition which Maas (1965) found contributed significantly to low inter-rater reliability.

The three tapes selected contained interviewees who had been rated by the interviewer and the Bank of Montreal viewers as being low, average and high in terms of suitability for employment. If each of the raters from the Bank evaluated the interviewee as unsatisfactory or marginal in their overall ratings then he was designated as being "low". If they all evaluated the interviewee as superior or excellent then he was designated as being "high". If they all evaluated the interviewee as average then he was designated "average". Each of the 34 interviewees were rated in this manner and from among them three were selected. For each of the three selected the interviewer was the same, thereby avoiding bias

due to interviewer-differences. A summary description of the three interviewees is shown in Table 5. It may be noted that there was a discrepancy between the "low" interviewee and the other two in terms of the number of interviews they had been to in the year prior to this study and their field of study at university. These differences were unavoidable and were recognized as a limitation of this analysis.

Table 5
Description of Interviewees

<u>Interviewee</u>	<u>Age</u>	<u>Grade Level</u> <u>%</u>	<u>Number of prior</u> <u>Interviews</u>	<u>Field of</u> <u>Study</u>
Low	23	65	12	Finance
Average	24	70	1	I/R - O/B
High	22	68	3	I/R - O/B

Three tapes were felt to provide a better representation of interviewees than only one tape. Moreover, if only one tape was used and the interviewee was either definitely unsatisfactory or definitely satisfactory then not only would the dispersion of ratings invariably be minimal but also the intra-rater agreement would be high thereby distorting the true picture of ratings given interviewees. Using three taped interviews reduced the probability of this distortion's occurrence. A table summarizing this sampling distribution is given below (Table 6).

Table 6

Sampling Distribution for Data Acquisition

Interviewee	Individual Viewers	Group Viewers
	N	N
Low	36	8
Average	34	10
High	<u>34</u>	<u>9</u>
Total	104	27

The time lapse between viewings (and associated ratings) was one week. Comparisons between group ratings and individual ratings were made for each interviewee.

To reduce the effects of memory upon rate-rerate decisions two procedures were followed. First, the measure for the rerate decisions was altered in format from the measure for the rate decisions. Second, where possible, additional video-taped interviews were shown to ss throughout the time lapse between viewings to confuse the importance of the interview on which these data were accumulated. This additional showing affected approximately one-third of the individual viewers and one-third of the group viewers.

Groups vs Individuals - Rating Differences

As is discussed in Chapter 5, risk may have an impact on group vs individual decisional outcomes. Likewise, the importance of negative information compounded by group processes may account for the differences between group and ratings. Indeed, the actual group processes themselves may result in differences. This aspect of group and individual

employment decisions is examined here. Due to inadequate control over the independent variables affecting decisions in this regard, no hypotheses were developed for empirical investigation. Rather, this section of the study was exploratory in nature with post-hoc discussion being given to offer alternative reasons for the findings.

The procedure involved comparing the mean variable ratings, the mean overall ratings and the call-back-reject decisions between groups and individuals. The data used here were the ratings given by the 185 students in their initial viewing of the three video-taped interviews.

Intra-rater Reliability

Intra-rater reliability is an estimate of the stability of ratings over time. On the basis of this estimate conclusions may be made as to the accuracy of the measure being employed. The intra-rater reliability estimates were obtained by computing the PPA and Pearson's r between the rate - rerate decisions for both groups and individuals.

Inter-rater Reliability

Inter-rater reliability provides information on the extent to which raters agree as to the amounts of any particular trait or personality variable which the interviewee possesses. This, then, offers an estimate of the extent to which a measure has internal stability.

Inter-rater reliability was calculated by two methods. First, the standard deviation scores for each variable evaluated by individual raters on their first viewing of the

interviews² were compared with the equivalent scores of group raters. If the standard deviation is lower for one sub-sample then inter-rater reliability is higher for that sub-sample and vice versa. This approach is the one used by Hollingworth (1922), Cottell (1910) and Norsworthy (1910) in their evaluations of inter-rater agreement in estimating traits of character (see Symonds, 1931, pages 112-113, for a discussion). More recently, Carlson and Mayfield (1967) and Carlson (1967, 1968) utilized this method to estimate inter-interviewer agreement scores. Second, an average inter-correlation score was computed by summing over correlations between raters for all rated variables and dividing by the number of correlations observed. This approach is of the type used by Carlson, Schwab and Heneman (1970).

Halo Error

Measures of the kind used by the Bank of Montreal and employed in this study are notorious for permitting halo error to distort trait ratings. Thorndike and Hagen (1969) defined halo error as being "the tendency of raters to base evaluations of a person being rated upon general favorable-ness toward that individual and not to differentiate degree of possession of specific traits". While procedures were followed to reduce the existence of halo error (see earlier

²Since this section deals with inter-rater agreement and is not concerned with stability over time only data acquired from the first viewing of the three taped interviews is used for analysis. The re-rate decisions are not included here.

discussion), it was unreasonable to expect that it would be non-existent. As a result, the amount of halo error present in the measure was determined by factoring the thirteen variables (including the overall rating) using a principal component factor analysis program with a varimax rotation procedure. Comparisons between group and individual factor matrices was made by employing a factor stability check which tests the hypothesis that the difference in factor spaces occupied by the group data and the individual data is not significantly different from zero. Canonical correlations between these sets of factors were computed and tested using a chi-square statistic and its probability. This procedure was contained in the factor analysis program mentioned above.

III. Attitudes Toward the Use of Video-Tape in Interviews

This section of the study was intended to go one step further than Moore and Craik (1972) in assessing attitudes toward video-tape use in interviewing. These authors focused upon answers to three basic questions:

- a. How realistically does playback of video taped interviews portray a number of important interviewee characteristics?
- b. In an overall sense, how effective is the video-taped screening interview?
- c. What is the general reaction to the suggestion that video taped interviews be used for initial screening of university graduates seeking employment?

The two samples whose attitudes Moore and Craik measured were a group of students and a group of professional

recruiters and personnel administrators. Both samples were asked to view a series of video-taped interviews and complete a questionnaire containing attitudinal items. None of the members of these samples were asked to formally assess the suitability of the candidates they viewed. In this sense, then, they were quite removed from the interview setting.

This study attempted to overcome the sampling condition in Moore and Craik's research by having actual interviewees and actual raters and interviewers provide their attitudes. This important dimension added considerable depth to the assessment of attitudes toward video-tape use in interviewing. Not only were uninvolved observers views obtained (Moore and Craik, 1972) but also were those of persons actually participating in the interviewing process.

Further to the questions asked by Moore and Craik, a number of other queries were raised. These are as follows:

1. How do viewers who are rating the interviewee's suitability feel as to the realism with which the video-tape portrays a number of important interviewee characteristics?
2. How do interviewees feel as to the extent to which their behaviours were either better or worse in the video-taped interviews than what they would have been had the interviews not been video-taped?
3. How distracting to the interviewees was the video-tape?

The samples and interviews were the same as those described earlier in this chapter. The measures used were modifications of the questionnaire used by Moore and Craik (1972) and are shown in Appendix 3. The interviewees were asked to complete their questionnaires immediately after

their interviews. The group and individual viewers were asked to fill out their forms on their own time after they had viewed all of their assigned interviews. This procedure for the viewers had limitations and resulted in a return of 17 of the 24 possible questionnaires (71% return). However, the number was deemed to be sufficiently large to offer an adequate representation of the sub-sample's attitudes.

CHAPTER 4

RESULTS

This chapter is segmented into three major sections, although the first two converge upon each other. This first part compares ratings given by interviewers in a face-to-face setting with ratings given by observers of video-taped interviews. The second section presents differences between group and individual raters in terms of mean trait ratings, overall ratings, call back-reject decisions, halo error and inter- and intra-rater reliability. The third part reports attitudes of interviewees and observing raters toward the use of video-tape in interviewing.

I. Interviewers vs Viewers

Table 1 presents means and standard deviations of the interviewer, group and individual ratings of the interview Ss. Mean trait ratings by interviewers following live interviews agreed closely with the mean ratings of groups following videotape playback and discussion. The mean group ratings were significantly higher than the mean interviewer ratings for intelligence and self-confidence. Individual viewers differed from the interviewers to a greater extent. The mean individual ratings were significantly higher than the mean interviewer ratings for six traits. Inspection of

TABLE 1

Means, Standard Deviations, and Significant Differences^a for Interviewers' (I), Groups' (G), and Individuals' (I') Ratings

Trait	Interviewers			Groups			Individuals		
	M	SD	I vs. G	M	SD	G vs. I'	M	SD	I' vs. I
Attitude	3.20	.69		3.35	.65		3.53	.61	2.05*
Appearance	2.94	.49		3.06	.42	1.95*	3.27	.45	2.85**
Interest	2.85	.61		3.00	.78		3.12	.59	
Intelligence	3.12	.33	2.87**	3.40	.50		3.47	.51	3.41**
Leadership	3.12	.54		3.15	.70		3.27	.57	
Maturity	3.20	.48		3.27	.62	2.29*	3.62	.65	2.97**
Motivation	3.15	.56		3.12	.81		3.27	.57	
Persuasiveness	3.00	.43		3.15	.66		3.15	.66	
Self-confidence	3.17	.52	2.02*	3.44	.56		3.41	.61	
Self-expression	3.17	.58		3.21	.59	1.93*	3.50	.66	2.15*
Sociability	3.17	.52		3.32	.53		3.44	.66	
Potential	2.97	.58		2.85	.86		3.09	.71	
Overall	3.00	.43		3.12	.69		3.27	.57	2.18*

^at test, (two-tailed)

*p < .05

**p < .01

the mean interviewer ratings show them to be uniformly lower than the mean individual ratings and lower than all of the corresponding mean group ratings except motivation and potential. Furthermore, with the exception of self-confidence, group ratings were consistently lower than individual ratings. This data is summarized in Table 2.

Table 2
Overall Comparison Among Samples

Comparison	Method	Direction
Mean ¹	Interviewers vs Individuals	Individuals > Interviewers**
	Interviewers vs Groups	Groups > Inter- viewers
	Individuals vs Groups	Individuals > Groups**
S.D. ²	Interviewers vs Individuals	Individuals > Interviewers**
	Interviewers vs Groups	Groups > Inter- viewers***
	Individuals vs Groups	Individuals > Groups

** p<.05

*** p<.01

1. Mann-Whitney Test used to determine significance of differences (Siegel, 1956).
2. Sign-Test used to determine significance of difference (Siegel, 1956).

Table 3 summarizes the analysis of convergent and discriminant validity following the Campbell-Fiske (1959) procedure. Convergent validity, indicated when two or more independent measures tend to agree in the measurement of a given variable, is shown by the correlations in columns one,

TABLE 3
Convergent and Discriminant Validities of Interviewers- Groups,
Interviewers-Individuals, and Individuals-Groups

	Heteromethod block								
	Interviewers- groups			Interviewers- individuals			Individuals- groups		
	Valid- ity coef.	Highest hetero- trait value	No. of hetero- trait values higher ^a	Valid- ity coef.	Highest hetero- trait value	No. of hetero- trait values higher ^a	Valid- ity coef.	Highest hetero- trait value	No. of hetero- trait values higher ^a
	1	2	3	4	5	6	7	8	9
Attitude	31*	56	7	38*	56	6	28*	59	7
Appearance	46*	39	0	21	41	9	24	39	3
Interest	51*	47	0	39*	41	1	13	48	16
Intelligence	25	48	12	02	34	20	05	42	20
Leadership	60*	40	0	09	54	12	28*	52	2
Maturity	32*	44	9	36*	57	4	11	43	17
Motivation	16	53	15	26	57	11	26	48	9
Persuasiveness	32*	53	11	00	40	19	30*	42	3
Self-confidence	45*	41	0	15	32	14	34*	42	7
Self-expression	51*	60	1	24	45	6	04	46	22
Sociability	44*	60	7	39*	45	2	01	64	24
Potential	42*	42	0	30*	57	10	37*	64	12
Overall	10	53	21	38*	56	5	31*	43	10
Median	42	48		26	45		26	46	

^aNumber of heteromethod off-diagonal coefficients in corresponding row and column higher than validity coefficient; maximum = 24

* $p < .05$

four and seven of Table 3. Minimal requirements for convergent validity were met for 10 of 13 traits in the interviewer-group rating comparison (or significantly different from zero). In the interviewer-individual and individual-group comparisons less than half the traits satisfied the criterion for convergent validity.

The first test for discriminant validity, requiring that the validity coefficients for a trait measured by two different methods should be higher than the correlations between that trait and other traits measured by different methods, is met by five traits in the interviewer-group comparison. This test, however, is not met by any trait in the interviewer-individual or individual-group comparisons.

The second test for discriminant validity not treated in Table 3, requires that measures of a given trait made with independent methods correlate higher than correlations between the given trait and other traits when measured by a common method. All heteromethod trait intercorrelations were exceeded by at least 50% of the intertrait correlations within methods except for appearance, interest and leadership, all within the interviewer-group heteromethod block. These were exceeded by 8%, 25% and 11% respectively. Furthermore, intercorrelations between traits within the three mono-method blocks were quite high (for interviewers, median $r = .40$; for groups, median $r = .53$; for individuals, median $r = .45$). Clearly, the second test for discriminant validity was not met.

An examination of the patterns of trait inter-correlations within and between rating methods for similarity

constitutes the third test for discriminant validity. No pattern similarity could be identified.

The extent of agreement among the three samples in terms of the decision to reject the candidate or call him back for a second interview is reported in Table 4.

Table 4
Group, Individual and Interviewer Intercorrelations¹
Call Back - Reject Decision

	Interviewer	Individual
Individual	.02	
Group	.38	.76**

¹ Kendall's Q (Blalock, 1960)

** $p < .01$

These findings indicate low agreement between individual viewers and interviewers and moderate agreement between group viewers and interviewers. A high correlation, however, exists between the two samples of viewers. Possibly video-tape may have had a bearing on this outcome.

In summary, then, the above results point out four major findings:

(1) decisions made by groups of managers after exposure to video-tape playbacks of candidates' interviews differed little compared to the decisions of interviewers having the typical level of training and experience;

(2) convergent and discriminant validity analysis generally revealed low convergent validity coefficients,

high intertrait correlations and inadequate fulfillment of the discriminant validity criteria for the individual-interviewer and the individual-group comparisons. However, in the main, the interviewer-group association approached fulfillment of all key criteria except the second and third tests for discriminant validity;

(3) individual ratings tended to be uniformly and significantly higher, statistically, than both group ratings and interviewer ratings;

(4) group-individual differences in terms of the call back-reject decision were not significantly different from zero ($r = .76$, $p .01$).

II. Group vs Individual Viewers

Further research was prepared with non-professional interviewer samples (i.e. students, as explained in chapter 3) to further test findings (3) and (4) above, as well as to explore the inter- and intra-rater reliability and existence of halo error. Data is reported for each of the three interviewees studied.

Mean Trait Rating Differences

Table 5 reports the mean trait ratings given by group raters and individual raters on the basis of the total individual N and the total group N (i.e. ss were collapsed along interviewers).

Table 5

Mean Traits Rating - Total Groups and Individuals

	Individuals N = 104	Groups N = 27
Attitude	3.22	3.07
Appearance	3.51	3.44
Interest	2.89	2.22
Intelligence	3.25	3.33*
Leadership	2.63	2.44
Maturity	3.17	2.96
Motivation	3.01	2.70
Persuasiveness	2.60	2.41
Self-Confidence	2.86	2.81
Self-Expression	3.27	3.33*
Sociability	3.37	3.44*
Potential	3.03	2.70

*These are ratings where group means were higher than individual means.

A Wilcoxon Matched-Pairs Sign-Ranks test (Siegel, 1956) was used to test the significance of the differences between group and individual ratings. Consistent with finding (3) above was the result that individuals tended to rate higher than groups ($p .02$). Only three traits had ratings in the inconsistent direction; intelligence, self-expression and sociability.

However, when the data is segmented in terms of interviewee a rather different pattern emerges. Table 6

reports the mean trait rating differences between group raters and individual raters for each of the three interviewees.

Table 6

Mean Trait Ratings - Groups vs Individuals

Interviewee	Low		Average		High	
	I	G	I	G	I	G
Traits	N=36	N=8	N=34	N=10	N=34	N=9
Attitude	2.81	2.75	3.38	3.10	3.50	3.33
Appearance	3.39	3.00	3.41	3.30	3.74	4.00*
Interest	2.58	1.88	2.91	2.00	3.20	2.78
Intelligence	2.94	2.75	3.35	3.60*	3.47	3.56*
Leadership	2.02	1.63	2.73	2.50	3.18	3.11
Maturity	2.81	2.50	3.26	3.10	3.47	3.22
Motivation	2.53	2.25	3.12	2.60	3.41	3.22
Persuasive- ness	2.25	1.75	2.50	2.30	3.05	3.11*
Self-Confid- ence	2.11	1.75	3.09	3.00	3.41	3.56*
Self-Expres- sion	2.97	3.00*	3.12	3.30*	3.73	3.67
Sociability	3.11	3.13	3.26	3.10	3.74	4.11*
Potential	2.36	2.25	3.08	2.40	3.68	3.44

*These are ratings where group means were higher than individual means.

Again, a Wilcoxon Matched-Pairs Sign-Ranks test (Siegel, 1956) was used to test the significance of the differences between group and individual ratings for each of the three interviewees. The low interviewee was consistently rated higher by individual raters than by group

ratars ($p < .01$) except for self-expression. The average interviewee was again rated higher by individual ratars than by group ratars ($p < .05$), except for intelligence and self-expression. The high interviewee was rated higher by individual ratars than by group ratars on only 7 of the 12 traits. The difference here was found to be non-significant.

Overall Rating

When the overall rating is separately considered a similar pattern is found as above. Table 7 shows the overall ratings given each interviewee by both group and individual ratars.

Table 7

Overall Rating - Group and Individuals

<u>Interviewee</u>	Group	N	Individuals	N	Signif.
Low	2.00	8	2.47	36	$p < .05$
Average	3.00	10	3.00	34	-
High	3.33	9	3.47	34	-
Total	2.97	27	2.81	104	-

The difference in the mean overall ratings given by groups and individuals for the low interviewee was statistically significant, with individuals rating more favourably than groups. For the average and high interviewees no significant difference was found between the group and individual mean overall ratings.

Call Back-Reject Decisions

A comparison between group and individual raters in terms of the call back-reject decision revealed no significant differences; a result consistent with finding (4) discussed earlier. Table 8 summarized this data.

Table 8

Call Back-Reject Decision - Groups vs Individuals

Interviewee	Groups		Individuals		Signif.
	Percent Call Back	N	Percent Call Back	N	
Low	12.5	8	19.4	36	-
Average	50	10	62	34	-
High	78	9	97	34	-
Total	48	27	59	104	-

Although no significant differences emerged, it may be noted that, consistent with earlier findings, a uniformly larger proportion of individual raters than group raters chose to call back interviewees.

Halo Error

To test for halo error a factor analysis using a principal component factor analysis program with a varimax rotation procedure was performed on all thirteen variables (including the overall rating) for both groups and individuals. The rotated factor matrix for individuals is shown in Table 9 and for groups is shown in Table 10.

Table 9
Rotated Factor Matrix - Individuals

Variable	Factor		h^2
	<u>1</u>	<u>2</u>	
Attitude	* .71	.17	.54
Appearance	-.02	* .82	.67
Interest	* .54	.29	.37
Intelligence	* .61	.03	.37
Leadership	* .74	.29	.63
Maturity	* .79	-.09	.63
Motivation	* .60	.36	.49
Persuasiveness	* .67	.32	.56
Self-Confidence	* .73	.29	.63
Self-Expression	* .55	.34	.42
Sociability	.37	* .67	.59
Potential	* .65	* .43	.61
Overall	* .71	* .50	.74

*loadings above .40

Two factors emerged from this analysis with individual raters, with the first factor containing all variables except appearance and sociability (the criterion for factor loadings was .40), although sociability loaded quite highly (.37). The total amount of variance accounted for by the two factors was 55.7% (Factor 1 = 47.52%; Factor 2 = 8.15%). As may be noted, appearance was the most outstanding variable in terms of its inconsistency with the loading pattern set by other variables.

For groups a somewhat similar result arises, as in Table 10.

Table 10

Rotated Factor Matrix - Groups

Variable	Factor		h^2
	<u>1</u>	<u>2</u>	
Attitude	.20	*-.62	.43
Appearance	.03	*-.83	.68
Interest	.25	*-.69	.54
Intelligence	*.61	-.23	.42
Leadership	*.69	-.39	.62
Maturity	*.88	.08	.78
Motivation	*.45	*-.55	.50
Persuasiveness	*.76	-.38	.72
Self-Confidence	*.75	-.33	.70
Self-Expression	*.42	*-.49	.42
Sociability	*.57	-.28	.40
Potential	*.43	*-.70	.67
Overall	*.79	-.36	.76

Again, 2 factors emerged from this analysis, with the first factor containing all but attitude, appearance and interest. The total amount of variance accounted for by the two factors was 58.5% (Factor 1 = 48.1%; Factor 2 = 10.4%). Again, appearance was the most outstanding variable in terms of its inconsistency with the loading pattern set by other variables.

From the communality columns of Tables 9 and 10 it may be seen that the variables were not too well accounted for by the two factors. The range of communality estimates was from .37 - .74 for individuals and from .42 - .78 for groups.

While the factor loadings for both the group and individual data are not immediately interpretable, one trend does seem to emerge. Factor 1 has moderate to high loadings on all variables with the notable exception of appearance, indicating that the raters were not discriminating amongst the remaining variables. Factor 2 for both samples has extremely high loadings on appearance, as well as on sociability, potential and overall (for individuals) and attitude, interest, motivation, self-expression and potential (for groups). It is interesting to note that appearance is the only directly observable variable included in the rating form. The remaining variables must all be inferred from the exchange of communications in the interviews. However, recognizing the low amount of variance accounted for by the two factors (55.7% for individuals; 58.5% for groups), this interpretation should not be taken as conclusive.

The possible existence of halo error was further explored by correlating each of the variables with the final call back-reject decision. Table 11 shows that for groups each variable with the exceptions of attitude, leadership, potential and overall correlated minimally or moderately ($p < .05$) with this decision. Potential and overall had the most outstanding correlations. For individuals, however, each variable had a highly significant ($p < .01$) correlation with the call back-reject decision. Again, though, potential and overall held, by far, the highest relationship.

Table 11

Correlations Between Variables and Call Back-Reject
Decision-Groups and Individuals

	Groups N=27	Individuals N=104
Attitude	.50**	.47**
Appearance	.35	.41**
Interest	.29	.36**
Intelligence	.41*	.40**
Leadership	.32	.59**
Maturity	.37*	.45**
Motivation	.38*	.53**
Persuasiveness	.54**	.55**
Self-Confidence	.36	.57**
Self-Expression	.41*	.47**
Sociability	.40*	.37**
Potential	.68**	.68**
Overall	.62**	.69**

* $p < .05$

** $p < .01$

It should be noted that "potential" and "overall" relate to the whole person rather than to any particular aspect of him. It is possible then that the call back-reject decisions were based on general impressions or total reactions without differentiating specific traits. While the

data is not clear-cut enough to make definitive conclusions, it appears that halo error could have been contributing to the ratings especially those given by groups.

Canonical correlations between the factor sets of groups and individuals demonstrated no significant differences. For factor 1 the correlation between groups and individuals was .99 with the chi probability being less than .00. For factor 2 the correlation between groups and individuals was .74 with the chi probability being less than .00. This evidence indicated that the factor spaces occupied by the group data and the individual data were not significantly different from each other. Table 12 summarizes this information.

Table 12

Canonical Correlations Between the Two Sets of Factor Scores

	Factor 1	Factor 2
Canonical r	.999	.738
Chi Square	980.03	101.26
d.f.	4	1
p	0.0	0.0

Intra-Rater Reliability

Table 13 reports the Pearson's r and the Percent Perfect Agreement statistic between the rate-rater decisions for each variable for both groups and individuals.

For 7 of the 13 variables the individual r is greater

than the group r . For 9 of the 13 variables the individual PPA is greater than the group PPA. However, using a Wilcoxon Matched-Pairs Signed-Ranks Test (Siegel, 1956) no significant differences were found between groups and individuals for either statistic.

Table 13

Intra-Rater Reliability - Groups and Individuals

	Groups (N=27)		Individuals (N=104)	
	r	PPA ¹	r	PPA ¹
Attitude	.30	50	.30**	49
Appearance	.45*	68	.63**	72
Interest	.34	50	.62**	61
Intelligence	.13	50	.47**	64
Leadership	.82**	82	.69**	66
Maturity	.21	54	.59**	72
Motivation	.59**	46	.48**	55
Persuasiveness	.68**	64	.64**	71
Self-Confidence	.65**	68	.67**	62
Self-Expression	.58**	68	.56**	55
Sociability	.51**	43	.50**	55
Potential	.69**	60	.73**	66
Overall	.54**	68	.59**	71
Median	.54	60	.59	64

¹Percent Perfect Agreement

* $p < .05$

** $p < .01$

More detailed estimates of intra-rater reliability are presented in Tables 14 and 15. Table 14 summarizes the Pearson's r statistics derived from both group and individual data for each of the three interviewees. Table 15 summarizes the Percent Perfect Agreement statistics for the same data. Attention may be given to the comparison between the underlined estimates on Table 14 and the corresponding underlined estimates on Table 15. These comparisons highlight the differences between the Pearson r method and the PPA method of calculating intra-rater reliability.

Using a Wilcoxon Matched-Pairs Signed-Ranks Test (Siegel, 1956) to test the significance of the differences between group and individual Pearson's r 's for each of the three interviewees (Table 14), a significant difference was found to exist for the High-interviewee ratings ($p < .01$). For the High interviewee, individual raters tended to have a significantly higher level of intra-rater reliability than groups.

A similar test of the PPA data (Table 15) yielded no significant differences between groups and individuals.

A comparison of the test-retest mean variable ratings given by both groups and individuals (see Table 16) indicated that the retest ratings tended to be significantly more critical than the initial ratings. Using a Wilcoxon Matched-Pairs Signed-Ranks Test the significance of the difference between test and retest ratings was found to be $p < .01$ for both groups and individuals.

Table 14

Intra-Rater Reliability - Groups and Individuals
Pearson's r

Interviewee	Low		Average		High	
	Group N=8	Indiv- idual N=36	Group N=10	Indiv- idual N=34	Group N=9	Indiv- idual N=34
Attitude	.76	.27	.15	.15	-.16	.25
Appearance	<u>.00</u>	.80	.51	.69	.40	.28
Interest	.71	.65	<u>.00</u>	.49	.35	.65
Intelligence	-.33	.50	.17	.30	<u>.00</u>	.29
Leadership	.49	.45	.90	.58	<u>-.13</u>	.52
Maturity	.64	.23	.53	.40	<u>-.58</u>	.78
Motivation	.71	.38	.30	.23	.24	.59
Persuasiveness	.87	.75	.50	.24	.58	.24
Self-Confid- ence	.58	.52	.48	.45	-.06	.43
Self-Expres- sion	.53	.68	.46	.22	.79	.45
Sociability	.32	.66	.11	.15	.07	.46
Potential	.75	.66	.34	.64	.62	.43
Overall	.73	.62	<u>.00</u>	.11	<u>.00</u>	.41
Median	.64	.62	.34	.30	.24	.43

Table 15

Intra-Rater Reliability - Groups and Individuals
Percent Perfect Agreement

Interviewee	Low		Average		High	
	Groups	Indiv- iduals	Groups	Indiv- iduals	Groups	Indiv- iduals
	N=8	N=36	N=10	N=34	N=9	N=34
Attitude	75.0	55.6	30.0	47.3	44.4	44.1
Appearance	<u>75.0</u>	83.3	70.0	82.3	55.5	50.0
Interest	62.5	75.0	<u>40.0</u>	58.8	44.4	50.0
Intelligence	50.0	75.0	50.0	58.8	<u>44.4</u>	55.9
Leadership	75.0	66.7	90.0	55.8	<u>77.7</u>	73.5
Maturity	62.5	64.7	60.0	70.6	33.3	79.4
Motivation	62.5	50.0	40.0	44.1	44.4	70.6
Persuasiveness	75.0	75.0	60.0	61.7	66.6	73.5
Self-Confidence	75.0	72.2	50.0	47.3	77.7	61.8
Self-Expression	62.5	61.1	60.0	47.3	88.8	55.9
Sociability	37.5	66.7	40.0	41.2	55.5	55.9
Potential	62.5	66.7	60.0	64.7	66.6	64.7
Overall	62.5	69.4	<u>70.0</u>	70.6	<u>66.6</u>	70.6
Median	62.5	66.7	60.0	58.5	55.5	61.8

Table 16

Mean Variable Ratings: Test and Retest
Groups and Individuals

	Groups		Individuals	
	Test	Retest	Test	Retest
Attitude	3.07	3.00	3.22	3.17
Appearance	3.44	3.40	3.51	3.44
Interest	2.22	2.15	2.89	2.82
Intelligence	3.33	2.89	3.25	3.13
Leadership	2.44	2.48	2.63	2.63
Maturity	2.96	2.89	3.17	3.05
Motivation	2.70	2.37	3.01	2.89
Persuasiveness	2.41	2.15	2.60	2.63
Self-Confidence	2.81	2.70	2.86	2.80
Self-Expression	3.33	3.04	3.27	3.07
Sociability	3.44	3.15	3.37	3.23
Potential	2.70	2.52	3.03	2.98
Overall	2.81	2.66	2.97	2.92

A test of intra-rater reliability of the ultimate decision to reject or call back the interviewee yielded the following data (Table 17).

Groups appeared to be more reliable than individuals for the low and average interviewee. For the high interviewee, the opposite occurred. It is interesting to note the high reliabilities in the Low Interviewee-Group cell and the High Interviewee-Individual cell. The remaining estimates

were at best mediocre. Again, though, the trend towards being more critical in the retest situation than in the initial test situation occurs. Bivariate matrixes for the total group decisions (Table 18) and for the total individual decisions (Table 19) demonstrate this trend.

Table 17

Call Back-Reject Decision - Groups and Individuals
Intra-Rater Reliability*

	Group	N	Individual	N
Interviewee				
Low	1.00	8	.64	36
Average	.50	10	.25	34
High	.23	9	.95 ¹	34
Total	.58	27	.65	104

*phi coefficient (Siegel, 1956)

¹this estimate is somewhat inaccurate due to the mechanics of computing the phi coefficient. The procedure requires occupancy of all four cells in a 2 x 2 matrix. This data for this estimate had two vacant cells limiting final computation. The cell structure was as follows:

		Time 2 (Retest)		
		Call Back	Reject	
Time 1	Call Back	33	0	33
(Test)	Reject	1	0	1
		34	0	34

As may be seen on Table 18, 18.5% of the groups shifted their decisions from "call back", at Time 1, to "reject" at Time 2, while only 3.7% changed in the opposite direction. Also, as on Table 19, 12.5% of the individuals shifted their decisions from "call back" at Time 1 to "reject" at Time 2, while only 3.8% changed in the opposite direction.

Table 18

Call Back-Reject Decisions - Group
Decision Shift

		Time 2 (retest)		
		Call Back	Reject	Total
Time 1 (Test)	Call Back	8 (29.6%)	5 (18.5%)	13 (48.2%)
	Reject	1 (3.7%)	13 (48.2%)	14 (51.8%)
	Total	9 (33.3%)	18 (66.7%)	27 (100%)

Table 19

Call Back-Reject Decisions - Individuals
Decision Shift

		Time 2 (retest)		
		Call Back	Reject	Total
Time 1 (Test)	Call Back	48 (46.2%)	13 (12.5%)	61 (58.6%)
	Reject	4 (3.8%)	39 (37.5%)	43 (41.4%)
	Total	52 (50.0%)	52 (50.0%)	104 (100.0%)

In summary then, except for a few isolated exceptions, the intra-rater reliability is not high. A trend, however,

in the direction of being more critical in Time 2 ratings as compared with Time 1 ratings appears to consistently occur.

Inter-Rater Reliability

The first method of comparing inter-rater reliability between groups and individuals was to examine differences between the standard deviations of each variable for both groups and individuals. Table 20 summarizes the data for these statistics.

As may be seen, individual raters had a wider dispersion of ratings for four variables, attitude, leadership, motivation and self-expression. For the remaining variables, the dispersion of scores for groups was higher than for individuals, indicating higher inter-rater reliability for individual raters.

Table 20

Inter-Rater Reliability - Group vs Individuals Standard Deviation Scores

	Groups (N=27)	Individuals (N=104)
Attitude	.62	.72
Appearance	.70	.57
Interest	.80	.68
Intelligence	.68	.57
Leadership	.75	.81
Maturity	.71	.67
Motivation	.78	.83
Persuasiveness	.93	.66
Self-Confidence	.92	.92
Self-Expression	.68	.78
Sociability	.80	.78
Potential	.87	.84
Overall	.79	.67

Using a Wilcoxon Matched-Pairs Signed-Rank Test (Siegel, 1956) to explore the significance of the difference between group and individual standard deviation scores, no significant difference was found.

A more detailed examination of the three interviewees yielded similar results (Table 21).

Table 21
Inter-Rater Reliability - Groups vs Individuals
Standard Deviation Scores

Interviewee	Low		Average		High	
Trait	Groups (N=8)	Indivs. (N=36)	Groups (N=10)	Indivs. (N=34)	Groups (N=9)	Indivs. (N=34)
Attitude	.71	.58	.57	.74	.50	.66
Appearance	.53	.60	.48	.56	.71	.51
Interest	.64	.69	.94	.62	.44	.59
Intelligence	.46	.41	.70	.60	.53	.56
Leadership	.52	.65	.53	.79	.33	.52
Maturity	.76	.62	.57	.51	.67	.71
Motivation	.70	.97	.70	.64	.67	.56
Persuasiveness	.70	.73	.95	.51	.60	.42
Self-Confidence	.46	.67	.67	.83	.53	.70
Self-Expression	.76	.77	.67	.69	.50	.67
Sociability	.83	.71	.57	.62	.60	.71
Potential	.89	.72	.70	.71	.53	.47
Overall	.76	.61	.47	.49	.50	.51

Using the Wilcoxon Matched-Pairs Signed-Ranks Test (Siegel, 1956), no significant differences were found between group raters and individual raters for each of the three interviewees.

From a relative standpoint, then, no conclusions may be drawn as to the comparative inter-rater reliability between groups and individuals.

The second method of estimating inter-rater reliability was computing average intercorrelation scores by summing over correlations between raters for all rated variables and dividing by the number of correlations observed. This method provided more absolute information on the inter-rater reliability of the samples. Table 22 summarizes the data for each of the three interviewees.

Table 22
Inter-Rater Reliability - Groups vs Individuals
Inter-Correlations

	Groups	Individuals
<u>Interviewee</u>		
Low	.47	.33
Average	.39	.19
High	.29	.21

Inter-rater reliability, using this method, is quite low, although groups provide uniformly higher estimates than individuals.

In summary, then, the above results point out six major findings:

(1) Individual ratings tended to be uniformly and significantly higher, statistically, than group ratings. This applies to all variables being rated including the

"overall" variable. However, when the data was segmented in terms of interviewee, a pattern emerged which indicated that the more unfavourable the interviewee, the higher were the individual ratings vs the group ratings.

(2) There were no significant differences between the call back-reject decisions of individual raters and of groups raters. A uniformly higher percentage of the individual raters, however, decided to call back each of the three interviewees.

(3) The group and individual raters did not appear to discriminate among the rated variables, except perhaps for appearance. Possible existence of halo error is exemplified by the high correlations of "potential" and "overall" with the decision to call back or reject. "Potential" and "overall" relate to the total person rather than any specific trait.

(4) Intra-rater reliability was not found to be high. Furthermore, except for the high interviewee Pearson product moment correlations, there was no significant difference between groups and individuals.

(5) When exploring intra-rater reliability, it was found that retest (Time 2) ratings were significantly more critical than the initial test ratings (Time 1). This finding emerged for both groups and individuals for the rated variables as well as for the call back-reject decision.

(6) Inter-rater reliability was found to be quite low for both groups and individuals. When the relative inter-rater reliability between group raters and individual

raters was examined by comparing standard deviation scores, no significant differences were located. When comparing rater-intercorrelations, groups were uniformly more reliable than individuals (although the coefficients were low).

III. Attitudes Toward Video-Tape in Interviewing

Interviewees

The questionnaire asked the interviewees to indicate the extent to which they found themselves distracted by various aspects of the video-taping procedure. Table 23 summarizes the responses given.

The presence of the cameraman and the noise of the video-tape equipment provided minimal distraction. The presence of the video-tape equipment proved distracting to 24% of the interviewees. This distraction lasted for less than half of the interview. The main source of distraction was the knowledge of being video-taped. Again, however, this proved distracting for less than half of the interview.

29% of the interviewees felt that the video-taped interview would be better than the face-to-face interview. 65% felt that it would be the same. Only 6% indicated that the face-to-face interview would be more effective.

The respondents were asked how much better or worse certain aspects of their behaviour were in the video-taped interview as compared with how they felt they would have been had the interview not been video-taped. Table 24 summarizes the responses.

Table 23

Extent and Cause of Interviewee Distraction

	Very Dis- traced (distracted throughout the whole interview)	Quite Dis- traced (distracted for at least half of the interview)	Somewhat Distracted (distracted for less than half of the interview)	Not Dis- traced at all
	%	%	%	%
a. presence of the video- tape equip- ment			24	76
b. presence of the camera-man			3	97
c. the know- ledge that you were being video- taped			44	56
d. the noise of the video- tape equip- ment		3	3	94

More than 80% of the interviewees felt their behaviour was about the same or better in the video-taped interview for all aspects but nervousness and verbal expression. 20% felt they were more nervous and 32% were less able to express themselves verbally in the video-taped interview.

76% (N=26) of the interviewees stated that they would be willing to undergo a video-taped interview which, at their request, could be sent to companies in which they were interested. 12% (N=4) said that they would not be willing to

Table 24

Extent of Difference in Interviewee's Behaviour In
a Video-Taped Interview Compared with a Face-to-Face
Interview (N=34)

	much better	slightly better	about the same	slightly worse	much worse
	%	%	%	%	%
Nervousness	3	9	68	20	
Honesty		6	85	9	
Judgment		9	88	3	
Voice intonation		23	62	15	
Ability to express myself verbally		12	56	32	
Manner		29	56	15	
Appearance		6	85	6	3
Force or drive		32	59	9	
Interest	3	44	53		
Social sensitivity		23	74	3	
Intelligence		6	85	9	
Overall behaviour		18	76	6	

undergo such a video-taped interview. The remaining 12% (N=4) were undecided.

Interviewees were asked to indicate on a seven-point scale their degree of enthusiasm toward the possible general use of video-taped initial screening for company selection of university graduates for employment.

As summarized in Table 2⁵, the response percentages

tended toward the enthusiastic side of the scale.

Table 25

Enthusiasm for Video-Taped Initial Screening
Interviews in Selection of University Graduates
for Employment (N=34)

		<u>%</u>
Very Enthusiastic	7	29
	6	26
	5	21
	4	12
	3	3
	2	6
Not Enthusiastic at all	1	3
Total		100%

The mean response was 5.4 on the seven-point scale. This may be compared with 3.6 for a student group of video-tape observers and 4.3 for an administrative group of observers (Moore and Craik, 1972).

The respondents were then asked how they felt most graduating students would react to the suggestion, by a company, that the student undergo a video-taped interview to be conducted by a University Placement Centre representative and forwarded to the company for examination. As shown in Table 26, slightly more interviewees felt that students would be favourable to such a suggestion than unfavourable.

From a list of possible objectionable factors concerning the use of video-tape screening interviews, the respondents were asked to rank the three most serious

objections. Table 27 summarizes these rankings.

Table 26

How Would Most Graduating Students React if Asked
by a Company to Undergo a Video-Taped Interview
Conducted by the Placement Office

	Very Un- favourable	Somewhat Unfavour- able	Somewhat Favourable	Very Favourable	Don't Know
%	9	35	29	21	6

Table 27

Student (Interviewee) Rankings of the Three Most
Serious Reservations or Objections re Video-Taped
Screening Interviews

Objection #	Rank	Possible Objection
1		- Many important personal character- istics cannot be conveyed
2		- This is just one more step toward the de-personalization of the employment relationship
3	2	- No assurance that v-t interview will be kept confidential
4		- The video-taped interview is unethical
5		- Most companies do not have the expensive video-playback equipment
6	3	- Once an interview is made there is no way of changing it
7	1	- The student may be forced to become an "actor"
8		- In front of a camera, most people do not act natural
9		- The technique will be too expensive

The first objection related to a concern that a person's real characteristics cannot be conveyed due to a forced role the interviewee must adopt. The second and third objections relate principally to moral and ethical matters.

Viewers

The viewers (both groups and individuals) were asked whether video-taped interviews were more or less effective than face-to-face interviews. 47% of these respondents felt video-taped interviews to be more effective. The remaining 35% felt that both were about the same.

The viewers were then asked to indicate on a four-point descriptive scale how realistically they felt the video-taped interview conveyed the actual amounts of several interviewee characteristics. The responses are shown in Table 28. More than 67% of the sample checked "somewhat realistic" or "very realistic" on all characteristics except appearance, knowledge of field, nervousness, and stress. For appearance and knowledge of field the "unable to judge" category was quite high (24% in both). Nervousness and stress, on the other hand, both had a high percentage of responses in the "somewhat unrealistic" category.

The viewers were then asked to indicate on a seven-point scale their degree of enthusiasm toward the possible general use of videotaped initial screening interviews for company selection of university graduates for employment.

Table 29 shows that these respondents tend to be only somewhat more enthusiastic than unenthusiastic. This is in marked contrast to the interviewee group (see Table 24) where significantly more enthusiasm was exhibited.

Table 28

Degree of Realism in Portraying Interviewee Characteristics - Viewers (N=17)

Characteristic	Degree of Realism				
	Very Un-realistic	Somewhat Unreal-istic	Somewhat Real-istic	Very Real-istic	Unable to Judge
	%	%	%	%	%
Appearance		12	59	6	23
Manner		6	59	23	12
Voice		6	23	53	18
Force or Drive		12	53	23	12
Intelligence		6	35	53	6
Potential		12	53	18	18
Interests		12	59	29	
Attitude		18	47	35	
Sociability		18	29	47	6
Self-Expression		6	18	71	6
Knowledge of Field		12	35	29	23
Self-Confidence	6	6	29	59	
Nervousness		35	29	35	
Motivation		12	41	41	6
Stress		35	18	29	18
Maturity		6	59	23	12
Judgment		12	53	18	18
Persuasiveness		6	53	35	6
Leadership		12	53	12	23

Table 29

Enthusiasm Toward the Possible General Use of
Video-Taped Initial Screening Interviews for
Selection of University Graduates - Viewers

Very En- thusiastic	7	6	5	4	3	2	1	Not en- thusiastic at all
% response	6	12	23	23	12	12	12	

In addition, the viewers were asked to elaborate on any strong personal objections with regard to the use of video-tape in screening interviews. The open ended written responses revealed three primary objections.

1. the viewer is unable to ask questions
2. the interview is not "live" or personal using video-tape
3. it can be expensive.

In summary, the attitudes toward the use of video-tape for the initial screening interview are widely divergent. On one hand, the interviewees exhibit an enthusiasm and state that, with video-tape, most of the various aspects of their behaviour were realistically portrayed. On the other hand, the viewers (i.e. Bank employees) display a hesitancy and, on the average, only a moderate enthusiasm.

CHAPTER 5

DISCUSSION

As in earlier chapters, this chapter is segmented into the three major areas; interviewers vs viewers, groups vs individuals, and attitudes toward the use of video-tape in interviewing.

I. Interviewers vs Viewers

The results relevant to this section showed four principle findings, two of which are explored in the next section. The two other findings, discussed here, are as follows:

(1) Group decisions made after exposure to video-tape playbacks of candidates' interviews differed little compared to the decisions of interviewers in face-to-face settings.

(2) Individual decisions made after exposure to video-tape playbacks of candidates' interviews were uniformly and significantly higher, statistically, than decisions of interviewers in face-to-face settings.

Each of these two findings is discussed separately.

Groups vs Interviewers

It has been argued by many researchers (Webster, 1964; Wright, 1969; Ulrich and Trumbo, 1965; Mayfield, 1964), that

biasing factors such as preconceived stereotypes, early decisions, the strong influence of negative information, and contrast effects play a significant role in the employment interview. While these biases exist, their effects on the employment decision remain unclear and largely uninvestigated. Two recent studies (Hakel, Ohnesorge and Dunnette, 1970; Carlson, 1970) have moved in this investigative direction by their examination of the influence of contrast effects. Rowe (1967) advanced the finding that employment decisions are made in the context of previous judgments. Hakel et al (1970), however, found that such contrast effects on rater's evaluations of employment resumes following previous high or low resume ratings accounted for very minor amounts of total decision variance. Carlson (1970) indicated that while the quality of the previous job applicant had a significant effect on the evaluation of test results, there was a negligible effect on the employment decision. These two studies brought under question the importance of a biasing factor, contrast effects, which had earlier been felt to provide a major influence on the employment decision. In like fashion, perhaps the finding that group-viewers' ratings differ little from interviewers' ratings suggests that some other biasing effects exert only trivial influence on the interview decision.

Considerable work has been performed exploring the comparative characteristics of groups and individuals (Holloman and Hendrick, 1971; Lorge, Fox, Davitz and Brenner, 1958;

Maier, 1967). One principle finding relates to the error-correcting propensity of social interaction in group decision-making. Erroneous assumptions and decision errors are more likely to be recognized and corrected in a group than by an individual. From this might be inferred that early decisions (Springbett, 1958) and inaccurate stereotypes (Hakel, Hollman and Dunnette, 1970) are less likely to occur with groups than with individuals. If this is the case, then there is a possibility that these two biases may affect the interview decision given by an individual rater in a relatively minor way. Since group-viewer ratings were found to differ little from individual-interviewer ratings perhaps the biases that are minimally present in groups may in fact be minimally present with individual interviewers.

Individuals vs Interviewers

The finding that individual viewers' decisions were uniformly and significantly higher, statistically, than the interviewers' decisions may suggest the possible existence of a "halo strategy" on the part of the individual viewers. Rowe (1963) reported that more experienced interviewers were found to be more selective and thereby more critical than less experienced interviewers. Indeed, in contrast with the interviewers, the viewers in the present study had no formal interviewer training. Also, they were farther removed from the interview setting than the interviewers, who personally screened candidates every day. As a result, these viewers

may be considered as being "less experienced" than the interviewers. The outcome of such a condition may well have been that the individual viewers were less selective and, as a consequence, rated more leniently than the interviewers. Furthermore, in line with discussion in the previous section, the error-correcting propensity of group activities may have served to mitigate this halo strategy, thereby resulting in the absence of any statistically significant differences between group viewers and interviewers.

Certainly evidence exists in the literature to support such an interpretation. Its specific veracity, however, could well serve as the basis for some future research.

There is, of course, the possibility of the six individual viewers having the tendency to rate high (the generosity error (Thorndike and Hagen, 1969) or the error of leniency (Kerlinger, 1964)) with the individual interviewers not holding such a tendency. The analyses of variance reported in Table 4 of Chapter 3 show that the F-probabilities are very small demonstrating no statistically significant differences and high internal consistency among raters in each sub-group. However, the probability of such a generosity error or error of leniency being committed by all members in one sub-group (individual viewers) and no members of the other sub-group (interviewers) is indeed quite small ($p = .014$).

One remaining question relates to whether or not video-tape itself provides a "media effect" resulting in

significant differences between interviewer ratings and viewer ratings. The close agreement between the interviewer and group ratings (two very diverse rating units) suggests that any such "media effect" is very weak or non-existent, indicating minimal effective difference between the interview and the video-taped interview.

II. Group vs Individual Viewers

The discussion of findings in this section is segmented into three areas. First of all is a comparison of group vs individual mean trait ratings, overall ratings and reject-call back decisions. Second, a discussion of groups vs individuals in terms of inter- and intra-rater reliability and halo error is presented. Third, a look is taken at the rating form used by the Bank of Montreal.

Groups vs Individuals - Ratings

In terms of trait ratings, overall ratings and reject-call back decisions, individual ratings tended to be statistically significantly higher than groups ratings for both samples studied. For the managerial sample the individuals' trait ratings were higher than the groups' trait ratings at the .05 level. The difference, however, in terms of reject-call back decisions was small ($r = .76$).

For the student sample, the individuals' trait ratings were higher than the groups' trait ratings at the .02 level. The individuals' overall rating for the low interviewee was higher than the groups' overall rating for the low interviewee at the .05 level, while for the average and high interviewees and for the aggregate of all three interviewees no significant differences were found (Table 7, Chapter 4). Also, no significant differences were found between groups and individuals in terms of the reject-call back decision, although for each interviewee and for the aggregate of interviewees a uniformly larger proportion of individual raters than groups raters chose to call back interviewees.

The general finding that individuals are more lenient in their ratings than groups may have significant importance to the personnel selection process. Many organizations employ panels of interviewers to screen candidates likely because panels are seen as being more reliable and valid than individuals. There is evidence to suggest that this is the case (Zajonc, 1966). However, as stated in Chapter 2, practically all of the group vs individual research from which this reliability and validity data was derived was conducted outside the interview setting. Within the setting, perhaps some other, as yet unexplored processes, unique to personnel selection,

operate to result in rating outcomes of the order found here.

That individuals are more lenient than groups appears to be fairly certain, since the finding emerged with two separate samples. The next issue of concern, then, is why this outcome came about. Two explanations are advanced below.

(a) One aspect of interviewing which consistently appears in the literature is that interviewers are more influenced by unfavourable than by favourable information about a candidate (Crissy and Regan, 1951; Springbett, 1958; Rowe, 1960; Bolster and Springbett, 1961; Mayfield and Carlson, 1966; Miller and Rowe, 1967; Blakeney and McNaughton, 1971). Webster (1964) described "an attitude of caution on the part of the interviewer who develops a high sensitivity to negative information with respect ... to its detection" (p. 90). The individual interviewer, then, focuses his attention quite heavily upon identifying negative information. In groups, this attention is further amplified with a consequent concomitant increase in the amounts of negative information perceived. Given this broader base of negative information on which to make a hiring decision, the groups become more critical and, hence, give lower ratings.

This explanation is further reinforced by the finding that, when the data was segmented in terms of interviewee, a pattern emerged which indicated that the more unfavourable the interviewee, the broader the gap between the individual and group ratings. Groups became disproportionately more critical of the unfavourable interviewee than did the in-

dividuals. Group members were readily able to collectively perceive this increase in the negative attributes of the candidate. On the other hand, the individuals, being constrained by having the perceptual capacities of only one person, were much less influenced by this increase.

(b) A second interpretation links this aspect of unfavourable information to the dynamics in process within groups. The likelihood is quite strong that an emphasis upon unfavourable information may serve as a norm of behaviour within a personnel decision-making group. Pressure likely exists in such a group to conform to this norm for fear of reprisal or for fear of being perceived as having unprofessionally low standards. The interviewer whose standards are too low is likely the one who meets with these reprisals since he is the one who has the highest probability of permitting entry of unfavourable personnel into the organization. As Webster (1964) states, "the interviewer is criticized because of misfits hired; praise for hiring good employees rarely occurs" (p. 90). Pressures to conform, based upon professional expectations, then, are quite strong.

At another level, the group setting evokes a personal need for social acceptance in each group member (Walter, 1972). Here, the concern shifts away from giving primary emphasis upon the hiring decision and towards satisfying social needs. To minimize the personal risk attached to social rejection and to satisfy needs for social acceptance, members are moved to conform to perceived group norms. If the perceived group

norm is to have high evaluative standards then the behavioural outcome is to be more critical in one's ratings. The group further serves to positively reinforce the conforming behaviours of members by offering social support when these behaviours are exhibited.

The process operating here is much like the one described by Brown (1965) as he interprets Stoner's (1961) findings of the "risky shift" in terms of value theory.

"Stoner's subjects were graduate students in the School of Industrial Management at M.I.T. and when members of the school first heard about the outcome of Stoner's experiment they argued that it could be explained by the fact that the field of industrial management sets a positive value on the ability to take risks. It is part of the role of an industrial management student to favor risky decisions, they held. The assertion may be true but it will not of itself account for Stoner's result. The subject is equally a student of industrial management when he answers the questions individually and when he agrees to a group decision following discussion. The effect is an increase in riskiness of the same subjects. One must argue, therefore, that the value of the role is more salient, more firmly engaged, when the management student is talking with peers. That seems reasonable enough. The student alone would be less concerned to manifest ideal role behaviour than would the student in the presence of other students. In the group each one has an audience to play to and that audience values riskiness." (p. 698)

In Brown's (1965) discussion, the audience values riskiness. In this study, the audience values having high evaluative standards or being critical by focusing upon unfavourable information. The effect is an increase in the amount of attention given to this unfavourable information.

Besides the two aspects of professional expectations and social need satisfaction exerting pressure to conform,

uniformity within the groups in this study also stems from similarities of members (they are all Bank managers) and similarities of environments in which they function; two further factors which Walter (1972) identifies as being instrumental in facilitating group uniformity.

With this explanatory hypothesis, then, not only is the group perceiving more negative information than an individual (as with the first explanation) but also each individual member in the group is perceiving more negative information than he would as an individual. The amount of negative information perceived as a result of group interaction is not simply concomitant (a function of the number of members in the group) but rather gestalt.

Risky-Shift

One interesting implication of the finding that individual raters are more lenient than group raters is its relationship to what would be predicted by the "risky-shift" model.

The "risky-shift" phenomenon refers to the situation wherein participants in a group assume a more risky stance compared with their initial individual position on a particular decision matter. Basically, groups are seen as being more risky than individuals.

This discovery was first reported by Stoner in 1961 in an unpublished master's thesis and later substantiated by Wallach, Kogan, Bem and others (Wallach, Kogan and Bem, 1962; Marquis, 1962; Wallach, Kogan and Bem, 1964; Bem, Wallach and

Kogan, 1965; Wallach and Kogan, 1965; Kogan and Wallach, 1967).

This notion of the "risky shift" could have significant effects on employment interviewing. If groups are used, one might expect their decisions to be more risky. Depending upon the interpretation of risk, in this setting, this could prove costly.

The key question, then, is how does risk relate to the employment interview. More specifically, what is the risky-alternative in employment decision-making? Perhaps this question should be looked at in terms of Type I versus Type II error. Type I error is reflected by the hiring of an unsuitable candidate while Type II error is the failure to hire a suitable candidate. What must first be determined is which of these errors involves more risk. Springbett's (see Webster, 1964) interpretation of the import of negative information on the hiring decision provides some ideas in this regard. As Webster (1964) states:

"Springbett ... impressed by the apparent predominance given to negative information, suggests an attitude or set on the part of the interviewer is created by the system of awards and punishments that marks the relationship between the employment and the production departments. He points out that two facts are clear: punishment is more certain than reward and only one type of error is punished. As to the first, the interviewer is criticized because of misfits hired; praise for hiring good employees rarely occurs. This situation produces an attitude of caution on the part of the interviewer who develops a high sensitivity to negative evidence with respect both to its detection and to the weight attached to it." (p. 90)

This interpretation suggests Type I error as being the more risky alternative.

Furthermore, it would seem that organizations would desire to minimize costs of training and selection, or at least offset these costs by ensuring that trained personnel will remain with the organization in a productive capacity. At the same time, the "costs" associated with a low interviewer performance rating or a lowering of esteem on the part of the interviewer's superior or peers towards him as the outcome of committing a Type I error may be higher than any personal costs suffered as a result of Type II error. Besides, the likelihood of others in the organization knowing that the interviewer released a suitable candidate is quite slim. Even if others did find out, the interviewer is always able to rationalize his actions by declaring that the candidate performed very poorly in the interview.

These interpretations would suggest that more risk is associated with committing a Type I error than with a Type II error. That is, the risky stance is taken by being more lenient in evaluative ratings and thereby increasing the probability of permitting entry to a larger proportion of unsuitable candidates. The "risky shift" model, then, would predict that groups would be more lenient in their ratings than individuals; a prediction completely inconsistent with the findings reported here.

If, however, the risky alternative is taken as being the Type II error then the findings are consistent with the model's prediction. One could argue that the costs of a Type II error are always or usually unknown, hence, there is

always a greater Type II risk operating in selection decisions. The framework within which risk is defined here, however, is from an organizational point of view. In the prior discussion, risk was approached from personal and interpersonal points of view. Quite possibly Type II error is the more risky alternative to the overall organization. For the individual, though, Type I error definitely appears more hazardous. This discussion brings under question the utility of the "risky-shift" as a predictive model. Should risk only be defined in terms of decisional outcomes, as Wallach, Kogan, Bem and others have done, or should it be approached in terms of outcomes of interpersonal relations and intragroup processes. The findings reported here suggest the "risky-shift" to be spurious as a predictor and largely dependent on the interpretation given to the risk-alternatives.

Groups vs Individuals - Reliability and Halo

(a) Intra-rater reliability

Generally, the intra-rater (test-retest) reliabilities found in this study were only meagre to moderate. The median trait correlations were found to be .54 for groups and .59 for individuals (see Table 13, Chapter 4). While these coefficients were not as high as is customarily required, they were both significantly different from zero ($p < .01$). The Percent Perfect Agreement statistics were only slightly better (median: 60% for groups; 64% for individuals). The "overall" variable alone showed similar results as measured by the

Pearson product moment correlation ($r = .54$ for groups; $r = .59$ for individuals). When measured by the PPA, however, the coefficients became more respectable (68% for groups; 71% for individuals). These, however, were still not as high as they should be.

These findings cast suspicion on the traits themselves; their definitional clarity, their degree of overlap or commonality, and their realism and relevance to the raters.

Noteworthy is the finding that the groups were not significantly different from the individuals, as may have been predicted. Zajonc (1966) reports data which suggests the advantages of groups over individuals in terms of reliability. Here, however, little difference was found to exist.

The incidence of meagre to moderate intra-rater reliabilities may be explained when viewing the decision-shift between test and retest ratings. For both groups and individuals, ratings were significantly more critical ($p < .01$) in the retest setting than in the test setting. As with explanations for findings reported earlier in this study, perhaps the impact of unfavourable information may have had a bearing on this outcome. In the test setting, viewers were largely influenced in their ratings by the unfavourable aspects about the candidate. In the retest setting, the same process operates, only this time the original perceptions of negative information (from the test setting) are further reinforced and also coupled with additional negative evidence. This whole process serves to accentuate the weighting given to the

unfavourable aspects of the candidate, and results in more critical retest ratings.

It is possible that the one-week time span may have been inadequate in terms of lessening the impact of memory on retest ratings. After such a time lapse, the viewers may well have remembered the negative aspects of the candidate. However, if memory was a contributing influence, one would expect the intra-rater reliability correlations to be higher. It appears that some other influence operated to result in the more critical retest ratings and the consequent lower reliabilities. Indeed, further work in this regard is justified to examine this explanation's veracity.

(b) Inter-rater reliability

As with the intra-rater reliabilities, inter-rater agreement was quite low as shown on Table 22, Chapter 4, although here groups' coefficients were uniformly higher than those of individuals. When comparing the dispersion scores (standard deviations) for both samples, though, no statistically significant differences were located.

These findings again indicate a lack of definitional clarity among the traits. Furthermore, they suggest a stereotype inconsistency in that the raters may have held divergent images of what is expected of an applicant for this type of position. This divergent image was not only held by the student sample but also by the managerial sample. The data in Column 4 on Table 3, Chapter 4 are indications of reliability in accord with the method suggested by Crissy (1952). In the

foregoing data, the individual viewers showed minimal agreement (median $r = .26$) with the individual interviewers. It is important to recall that both of these managerial subsamples have similar experience and are equally aware of the nature of the job for which the interviews were held.

(c) halo error

The incidence of halo error was suggested by the concurrently high loadings of most traits and the "overall" variable on one general factor, and by the high correlations between the global variables of "potential" and "overall" and the call back-reject decision.

In the factor analysis, the only trait which consistently gave very high loadings on a second factor for both group and individual raters was appearance. As discussed in Chapter 4, it is interesting to note that appearance is the only directly observable trait on which ratings must be given. Ratings for the remaining traits must all be inferred from the exchange of communications in the interview. It is quite possible, then, that if each of the traits were tied to directly observable behaviours, as Maas (1965) recommends, that separate factors for each trait may emerge, much like that reported by Howell and Vincent (1970).

Furthermore, contrary to what might be predicted, no differences were found to exist between the factor space occupied by group and individual raters. With the error-correcting propensity of groups (Holloman and Hendrick, 1971),

one might expect groups to discriminate among the traits more so than the individuals. The very high canonical correlations between group and individual factor spaces (see Table 12, Chapter 4) indicate that such an expectation was not met.

Restructuring the Bank's Interview Procedure

The foregoing findings regarding reliability and halo error indicate that possibly a restructuring of the Bank of Montreal's interview procedure is in order. This re-organization could focus upon three aspects of the Bank's personnel selection process; the rating form, the interview itself, and interviewer training.

(a) the rating form:

Maas (1965) proposed a procedure where ratings are made using scaled examples of on-the-job behaviour rather than using a traditional adjective rating scale (as was used here). As discussed in Chapter 2, Maas' findings showed quite markedly the improvement in reliability by using the former approach instead of the latter. The procedure he followed in establishing this technique was as below:

1. "the traits to be evaluated were established by a committee of interviewers who were familiar with the job to be performed."

The managers with the Bank may pursue this activity by first of all divorcing themselves from the current rating form and turning to the job itself. A guiding question they

could follow is: what characteristics should a trainee hold to successfully complete the duties assigned to his position? One additional question is, of course, - why? This procedure, of necessity, implies the existence of an adequate job description.

2. "examples of on-the-job behaviour were written to illustrate three levels of each trait - a high degree of the trait, an average degree ... and a low degree..."

Here, the approach to be taken is to address the question, what will the trainee be doing to demonstrate these three levels of each trait? For instance, if leadership was felt to be a critical trait, an example of behaviours which related to each level would be:

If an interpersonal conflict arises between himself and one of his subordinates the trainee will:

- | | |
|----------------|--|
| high degree | (a) identify and confront the problem immediately with the intention of achieving a resolution mutually satisfactory to himself and his subordinate; |
| average degree | (b) identify the problem and wait for an "appropriate" time for its resolution; |
| low degree | (c) ignore the problem all together or have the subordinate released or transferred with no explanation. |

This is a fabricated item and is not intended to be one actually used by the Bank. It does however alter trait

definitions away from being nebulously interpreted¹ toward being tied to specific behaviours. Perhaps for each trait, three or four such items could be employed.

3. "independent judges, not knowing which examples were written for which traits and levels, reallocated the examples back into traits and levels".
4. "only examples with complete agreement as to trait and level were retained".
5. "these examples were arranged on a continuous vertical graphic rating scale, ... putting each example at its proper scaled level for the trait".

One further procedure, not specifically identified by Maas (1965) is to construct questions around each item. The method recommended by Kahn and Cannell (1957) in their Chapters 7 and 8 would provide a useful guideline. Again, these questions should be agreed upon as being relevant and useful by all interviewers concerned. The outcome of the above overall practice would be a standardized interview guide of the type recommended by Carlson et al (1971). Once the guide is developed, reliability and, if possible, validity checks should be made, similar to those researched in this study.

(b) the interview:

The key aspect of a redesigned interview is standard-

¹The definition of leadership as shown on the current form used by the Bank is: "degree of leadership experience, extra-curricular positions held".

ization. The interviewers should all be asking essentially the same core questions and thereby receiving comparable responses from interviewees. Furthermore, as Maas (1965) recommends, each interviewer should rate each candidate on each trait by making analogies from the candidates responses to the standardized questions, to behaviour that might be expected of the candidate, were he actually on the job. This procedure, coupled with the intensive and rigorous preliminary activities described earlier should serve to increase the reliability and comparability of the interviews as well as the confidence which the interviewers place in their rating decisions.

(c) interviewer training:

Carlson et al (1971) stress the importance of an intensive training program for interviewers. They emphasize that such a program is critical "if interviewers are to initially learn enough in common to increase the probability of obtaining general validity from the selection interview".

One program which would be useful was attempted as an addendum to this study. While no empirical data was obtained to demonstrate its utility, the program appeared to be beneficial according to the reports obtained from the interviewers involved. The procedure was as follows:

In order to construct a questionnaire relating to the interviewers' interviewing abilities, the approach recommended by Robert F. Mager (1962, 1968) was utilized. Discussions with the Employment and Employee Relations Manager, the

Personnel Manager and a number of interviewers, all from the Bank, resulted in identifying necessary steps involved in performing an employment interview. Initially, the job was broken down into its various component steps. These were found to consist of:

1. Create an atmosphere of rapport;
2. Demonstrate an interest in the interviewee;
3. Gather information relating to the interviewee's suitability for the job;
4. Ask interviewee to describe his conception of the job;
5. Correct interviewee's misconceptions (if any);
6. Give additional information regarding the job; hours, pay, mobility, etc.;
7. Generally improve or at least maintain the interviewee's image of the Bank or whatever organization is concerned.

Around these steps, a questionnaire (see Appendix 4) was constructed which aimed at identifying the extent to which interviewers satisfied these requisite task components.

At the end of each interview, the interviewer rated himself on the questionnaire and the interviewee evaluated the interviewer. This was felt to be an important aspect of the program since, very rarely, do interviewers find out how well they communicated or generally "came across" to the interviewee; the person about whom the interviewer must make an evaluation, and the person for whom the interviewer must at least leave a favourable impression.

Furthermore, at the end of each showing of the video-

taped interviews, the viewers (both groups and individuals) evaluated the interviewer on the questionnaire.

Responses to the questionnaire indicated areas of interviewer strength and weakness and, more important, identified interviewer training needs.

For each interviewer, a summary was made of his ratings of interviewees and their comparison with the interviewees' self-ratings and with groups' and individuals' ratings of the interviewees. The form used here was the Bank's rating form (see Exhibit 1, Chapter 3). Also summarized were the interviewees', the groups; and the individuals' evaluations of the interviewer, as well as the interviewer's self-ratings. All this information was presented to each interviewer in a numeric and descriptive form, called a Summary Sheet (see Appendix 5 for an example). At the same time as this information was given back to the interviewer he was shown some of his video-taped interviews to highlight those areas of strength and weakness described in the Summary Sheet. With the video-taped interviews the interviewer served as his own example. He was encouraged to make notes and ask questions. He was also invited to stop or replay the tape at any time. Concurrent with the viewing session, the Employment and Employee Relations Manager of the Bank, other interviewers and this author offered suggestions for the interviewer to consider for self-improvement.

This describes the extent of the training program actually administered in the Bank. One final critical stage,

and one for which inadequate time and resources were available for its implementation, would be to video-tape more interviews with the same interviewers using the same procedure to determine whether or not there was an improvement in interviewing ability. Responses given to the questionnaire by interviewees and viewers would provide adequate measures of any changes.

One important feature of this program was that a highly structured and intensive feedback component was coupled with the video-tape playbacks of interviews. This meets with the suggestion given by Weber (1969) and Walter and Miles (1972). In a study of group decision-making Weber (1969) discovered that groups provided with definite instructions and direction to guide their viewing "experienced greater increases in personal agreement with group decisions, satisfaction with personal performance, perceived adequacy of group decision-making procedures and personal understanding of group decisions than groups which lacked instructions for viewing" or groups which received no feedback at all. Walter and Miles (1972) found that the amount of personal change experienced by participants was positively related to the degree of feedback structure imposed during playbacks. This evidence indicates that a trainer cannot rely solely on the trainee's insight to perceive sufficient information from video-tape viewings to develop personal improvement. Rather, it suggests that participants should be furnished with a "viewer's guide" for playback sessions to offer direction

regarding what important personal behaviours to observe and what each behaviour means. Such a guide was employed here.

As stated earlier, no empirical evidence was gathered to support or refute the utility of this training program for interviewers. It did, however, meet with the approval of all involved, and, by way of self-reporting, the interviewers did indicate that they learned a considerable amount about their behaviour in the interview setting.

III Attitudes Toward Video-Tape

In the study reported by Moore and Craik (1972), 66% of the student group and 53% of the administrator group felt that most graduating students would react unfavourably to the suggestion that they take part in a video-taped interview. The results of this study, however, showed that 76% of the interviewees (who were all graduating students) would be willing to undergo such a video-taped interview. Furthermore, 76% of the interviewees exhibited marked enthusiasm for the use of video-taped initial screening interviews in selecting university graduates.

The fears expressed by Moore and Craik's (1972) samples that many of a person's "real" characteristics or qualities cannot be transmitted effectively via video-tape may well be allayed by the data found here. The interviewees from this study stated that, except for nervousness and verbal self-expression, their behaviours were quite parallel to what

they would have been without the video-tape, and that in no way were they as threatened by the medium as might have been expected. The most important objection expressed by the interviewees was that the student may be forced to become an "actor". This, however, is a common objection held towards non-video-taped face-to-face interviews. The more important aspect of their objections is that they did not feel that many important personal characteristics cannot be conveyed over video-tape and that in front of a camera, most people do not act naturally. This is in marked contrast to the objections held by Moore and Craik's samples, and, indeed, lends support to the possible utility of video-tape in interviews. Furthermore, again in opposition to Moore and Craik's samples, there was a definite concern shown by the interviewees for ethical and control issues. The second and third objections given by this sample related to the confidential nature of the video-taped interview (objection 3) and its permanence (objection 6).

The distractive effect of the video-tape equipment and noise and the presence of the operator reported by the interviewees as being minimally present throughout the interview. The most significant cause of distraction was the knowledge that they were being video-taped and this, too, served to distract for only approximately one-fourth of the interviews. The distraction problem can in most instances be overcome by offering some non-threatening exposure before actual video-taping takes place. Specific techniques which may be employed

to minimize this problem's probability of occurrence are threefold. First of all, the participants could have all the video-tape equipment fully explained to them. This gives them a non-threatening introduction to the equipment and should facilitate their willingness to work with it. Second, a few non-interview related situations could be taped and played back. Johnston (1967) mentioned the technique of allowing trainees to do anything they want to do in front of the camera, from making face to reciting Shakespeare. When they see themselves they usually are initially shocked by seeing what they are really like. Very quickly most people learn to accept themselves and behave naturally. Third, the interviewees could be given the opportunity to do some video-taping themselves. Let them become the operators. Again, this should make them feel more relaxed and less threatened in the presence of the camera.

The foregoing evidence suggests that video-tape may well be accepted as part of the interviewing process. Interviewees exhibited little defensive behaviour regarding its use and further stated that they were minimally distracted and that their behaviours were not too different from what they would have been in the absence of video-tape. The major area of concern related to ethical matters which could easily be accommodated by a formal contractual arrangement between the interviewee and the interviewing organization regarding privacy and restricted use of the video-taped interview.

The viewers, on the other hand, showed considerably

less enthusiasm toward video-tape use. While, in general, the interviewee characteristics were described by the viewers as being realistically conveyed over the video-tape, they still showed a lack of acceptance for the medium. Perhaps this may be explained by looking at their major objections. Besides a concern for cost, the viewers seemed to be quite occupied with the fact that they were removed from the interview setting and could not ask questions or probes. This latter concern by the viewers exemplifies the statement by Webster (1964) that interviewers look for different things. This may serve to explain why the viewers, especially the individuals, differed from the interviewers in their judgments. Certainly, differing perceptions is a source of inter-interviewer disagreement and rating error.

Perhaps this problem could be rectified by the adoption of a more structured interview containing questions which all interviewing personnel have seemed to be important and which all such personnel would ask if they were in the role of interviewer. The procedure for establishing such an interview structure would be the same as was earlier described.

Overall, these findings indicate that video-tape can serve a purpose in employment interviewing, principally in the supplementary role as discussed by Moore and Craik (1972). Furthermore, if adequate measures are taken, the primary issues of concern stated by both interviewees and viewers can be readily alleviated.

CHAPTER 6

SUMMARY AND CONCLUSIONS; RECOMMENDATIONS FOR FUTURE STUDY

The major findings emerging in this study may be summarized as follows:

(1) Decisions made by groups of managers after exposure to video-tape playbacks of candidates' interviews differed little compared to the decisions of interviewers.

(2) Individual ratings tended to be uniformly and significantly higher, statistically, than interviewer ratings.

(3) In general, individual viewer ratings tended to be uniformly and, in most cases, significantly higher, statistically, than group ratings. This finding emerged for both the managerial sample and the student sample.

(4) Intra- and inter-rater reliability tended to be low to moderate. Furthermore there was little difference between group and individual reliabilities except for the rater intercorrelations (inter-rater reliability) where groups were uniformly more reliable than individuals (although the coefficients were low).

(5) The group and individual raters did not appear to discriminate among the rated variables, except perhaps for appearance. Possible existence of halo error was exemplified by the high correlations of "potential" and "overall" with

the decision to call back or reject. "Potential" and "overall" relate to the whole person rather than any specific trait.

(6) When exploring intra-rater reliability, it was found that retest (Time 2) ratings were significantly more critical ($p < .01$) than the initial test ratings (Time 1). This finding emerged for both groups and individuals for the rated variables as well as for the call back-reject decision.

(7) The interviewees exhibited an enthusiasm for the use of video-tape. Furthermore, they stated that several elements of the video-tape offered minimal distraction and that, with video-tape, most of the various aspects of their behaviour were realistically portrayed. On the other hand, the viewers displayed a hesitancy and, on the average, only a moderate enthusiasm.

These findings and the discussion surrounding them indicate a number of areas where further research would be in order. Three such areas, seen by this author as being of particular importance are as follows.

(1) One interpretation of the finding that groups were more critical than individuals in their ratings of candidates was that the dynamics of group interaction led members to conform to a group norm that emphasizes high evaluative standards. Further research could address the issue of whether or not such a norm exists in a personnel decision-making group and, if so, what type and amount of

influence does it have on the rating decision.

(2) Further to the second recommendation above, perhaps the "risky shift" could be examined from the point of view of personal risks based on the attempted satisfaction of group member's social needs through conformity to group norms, rather than from the more traditional point of view of decisional outcomes. When interpreting risk in terms of decisional outcomes, this study found the "risky shift" to be an inadequate predictor of group vs individual differences. When treated in terms of personal risks, however, the "risky-shift" became a viable alternative explanation.

(3) Rating viewers of the video-taped interviews expressed as an objection the fact that they were not a direct part of the interviewing process and, as a result, could not ask questions or probes to interviewees. Furthermore, the fact that they wished to ask questions other than or in addition to those asked by the interviewer suggests that they were focusing upon different candidate attributes. This condition may explain the low levels of agreement between individual viewers and interviewers. Would the restructuring of the interview process, as described in Chapter 5, have led to high levels of inter-rater agreement? Considerable evidence exists to show that higher inter-rater agreement is found with more structured interviews (Carlson, Schwab and Heneman, 1970; Schwab and Heneman, 1969). Would this be the case if the video-taping procedure, as utilized in this study, were employed? Also, would the viewer's attitudes toward video-tape become more favourable?

Furthermore, the reliability data found in this study suggest that the Bank of Montreal's interviewing procedure suffers from numerous sources of error, notably the different perceptual focii of the raters and the ambiguity and lack of definitional clarity surrounding the rating factors. As is consistently shown in the literature (Anderson, 1954; Maas, 1965; Carlson, Schwab and Heneman, 1970), restructuring this interviewing procedure along the lines described in Chapter 5 would in all likelihood result in more respectable reliabilities and more confidence in interviewers' ratings.

These are only three areas where this study indicates that further research may be carried out. The discussion in Chapters 2 and 5 provide suggestions for research in other directions. Only with such research can meaningful answers be found and can the "state of the art" of interviewing be improved.

Bibliography

- American Psychological Association. Technical Recommendations for Psychological Tests and Diagnostic Techniques. Psychological Bulletin Supplement. March, 1954, 51, Number 2, Part 2. (cited in Brown, Lucero and Foss, 1962).
- Anderson, C. W. The Relation Between Speaking Times and Decision in the Employment Interview. Journal of Applied Psychology, 1960, 44, 4, 267-268.
- Bem, D. J., Wallach, M. A., and Kogan, N. Group Decision Making Under Risk of Aversive Consequences. Journal of Personality and Social Psychology, 1965, 1, 453-460.
- Blakeney, R. N., and MacNaughton, J. F. Effects of Temporal Placement of Unfavorable Information on Decision Making During the Selection Interview. Journal of Applied Psychology, 1971, 55, 138-142.
- Blalock, H. M. Social Statistics. McGraw-Hill: New York, 1960.
- Bolster, B. I., and Springbett, B. M. The Reaction of Interviewers to Favorable and Unfavorable Information. Journal of Applied Psychology, 1961, 45, 2, 97-103.
- Brown, B. W., Lucero, R. J., and Foss, A. B. A Situation Where the Pearson Correlation Coefficient Leads to Erroneous Assessment of Reliability. Journal of Clinical Psychology, XVIII, 1962, 95-97.
- Brown, Roger. Social Psychology. New York: Free Press, 1965.
- Bruce, R. S. Group Judgments in the Fields of Lifted Weights and Visual Discrimination. Journal of Psychology, 1935, 1, 117-121. (cited in Zajonc, 1966).
- Bruner, J. S. On Perceptual Readiness. Psychological Review, 1957, 64, 123-152.
- Business Week, July 10, 1971, p. 36.
- Campbell, John P. Individual Versus Group Problem Solving in an Industrial Sample. Journal of Applied Psychology, 1968, 52, 4, 205-210.
- Campbell, David T., and Fiske, Donald W. Convergent and Discriminant Validation by the Multitrait Multimethod Matrix. Psychological Bulletin, 1959, 56, 81-105.

- Carlson, Robert E. Selection Interview Decisions: The Effect of Interviewer Experience, Relative Quota Situation And Applicant Sample on Interviewer Decisions. Personnel Psychology, 20, 1967a, 259-280.
- Carlson, Robert E. Selection Interview Decisions: The Relative Influence of Appearance and Factual Written Information on an Interviewer's Final Rating. Journal of Applied Psychology, 1967b, 51, 6, 461-468.
- Carlson, Robert E. Employment Decisions: Effect of Mode of Applicant Presentation on Some Outcome Measures. Personnel Psychology, 1968, 21, 193-207.
- Carlson, Robert E. Effect of Interview Information in Altering Valid Impressions. Journal of Applied Psychology, 1971, 55, 1, 66-72.
- Carlson, R. E. and Mayfield, E. C. Evaluating Interview and Employment Application Data. Personnel Psychology, 1967, 20, 441-460.
- Carlson, Robert E., Schwab, Donald P., and Heneman, III, Herbert, G. Agreement Among Styles of Selection Interviewing. Journal of Industrial Psychology, 1970, 5, 1, 8-17.
- Carlson, R. E., Thayer, P. W., Mayfield, E. C., and Peterson, D. A. Improvements in the Selection Interview. Personnel Journal, April, 1971, 268-275.
- Cartwright, D., and Zander, A. Group Dynamics, Research and Theory. New York: Harper-Row, 1968.
- Cattell, J. McK. American Men of Science (2nd Ed., 1910), Appendix (cited in Symonds, 1931).
- Cline, Victor B., and Richards, James M., A Comparison of Individuals Versus Groups in Judging Personality. Journal of Applied Psychology, 1961, 45, 3, 150-155.
- Crissy, W. J. E. The Employment Interview - Research Areas, Methods and Results. Personnel Psychology, 1952, 5, 73-85.
- Crissy, W. J., and Regan, J. Halo in the Employment Interview. Journal of Applied Psychology, 1951, 35, 338-341.
- Crowell, A. R. Decision Sequences in Perception. Unpublished Doctoral dissertation, McGill University, 1961, (cited in Webster, 1964).
- Eysenck, H. J. The Validity of Judgments as a Function of the Number of Judges. Journal of Experimental Psychology, 1939, 25, 650-654 (cited in Zajonc, 1966).

- Ghiselli, Edwin E. The Validity of a Personnel Interview. Personnel Psychology, 1966, 19, 389-394.
- Goodman, Leo A., and Kruskal, William H. Measures of Association for Cross-Classifications. Journal of the American Statistical Association, 1954, 49, 732-764. (cited in Brown, Lucero and Foss, 1962).
- Gordon, K. H. Group Judgments In the Field of Lifted Weights. Journal of Experimental Psychology, 1924, 3, 398-400 (cited in Zajonc, 1966).
- Grant, D. L., and Bray, D. W. Contributions of the Interview to Assessment of Management Potential. Journal of Applied Psychology, 1969, 53, 24-34.
- Hakel, Milton D., and Dunnette, Marvin D. Checklists For Describing Job Applicants. Industrial Relation Center, University of Minnesota, 1970.
- Hakel, M. D., Dobmeyer, T. W., and Dunnette, M. D. Relative Importance of Three Content Dimensions in Overall Suitability Ratings of Job Applicants' Resumes. Journal of Applied Psychology, 1970, 54, 1, 65-71.
- Hakel, M. D., Ohnesorge, J. P., and Dunnette, M. D. Interviewer Evaluations of Job Applicants' Resumes As a Function of the Qualifications of the Immediately Preceding Applicants. Journal of Applied Psychology, 1970, 54, 1, 27-30.
- Hakel, Milton D., Hollmann, Thomas D., and Dunnette, Marvin D. Accuracy of Interviewers, Certified Public Accountants, and Students in Identifying the Interests of Accountants. Journal of Applied Psychology, 54, 2, 1970, 115-119.
- Hakel, M. D., and Schuh, A. J. Job Applicant Attributes Judged Important Across Seven Diverse Occupations. Personnel Psychology, 1971, 24, 45-52.
- Hall, E. J., Mouton, J. S., and Blake, R. R. Group Problem Solving Effectiveness Under Conditions of Pooling and Interaction. Journal of Social Psychology, 1963, 59, 147-157.
- Hall, J., and Williams, M. S. Group Dynamics Training and Improved Decision Making. Journal of Applied Behavioural Science, 1970, 6, 1, 39-68.
- Hollingworth, H. L. Judging Human Character. Appleton: New York, 1922. (cited in Symonds, 1931).
- Holloman, C. R., and Hendrick, H. W. Problem Solving in Different Sized Groups. Personnel Psychology, 1971, 24, 489-500.

- Howell, M. A., and Vincent, J. W. Factor Analysis of Interview Data. Journal of Applied Psychology, 1970, 54, 4, 313, 315.
- Johnston, William R. Closed Circuit TV in Sales Training. Training and Development Journal. April, 1967, 21, 4, 18-25.
- Kahn, R. L., and Connell, C. F. The Dynamics of Interviewing. New York: Wiley, 1957.
- Kasl, S. V., and Mahl, G. F. A Simple Device for Obtaining Certain Verbal Activity Measures During Interviews. Journal of Abnormal and Social Psychology, 1956, 53, 288-390.
- Kennedy, Joanna B. Use of Audio-Visual Techniques in Training The Hard-Core. Training and Development Journal, February, 1970, 20=32.
- Kerlinger, Fred N. Foundation of Behavioral Research. New York: Holt, Rinehart and Winston, 1964.
- Kogan, N., and Wallach, M. A. Risky-Shift Phenomenon in Small Decision-Making Groups: A Test of the Information-Exchange Hypothesis. Journal of Experimental Social Psychology, 1967, 3, 75-84.
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., and Fillenbaum, S., Evaluational Reactions to Spoken Languages, Journal of Abnormal and Social Psychology, 1960, 60, 44-51.
- Lewin, Kurt, 1945 (given in a speech by E.E. Lawler, III, 1972).
- Lorge, I., Fox, D., Davitz, J., and Brenner, M. A Survey of Studies Contrasting the Quality of Group Performance and Individual Performance, 1920-1957. Psychological Bulletin, 1958, 55, 6, 337-372.
- MacDonald, J. Brian. The Use of Video Tape in the Personnel Interview. Unpublished manuscript. University of British Columbia, 1971.
- McLuhan, Marshall. Understanding Media. Toronto: McGraw-Hill, 1964.
- Maas, James B. Patterned Scaled Expectation Interview: Reliability Studies On a New Technique. Journal of Applied Psychology, 1965, 49, 6, 431-433.
- Mager, Robert F. Preparing Instructional Objectives. Belmont, Calif.: Fearon, 1962.

- Mager, Robert F. Developing Attitude Toward Learning. Belmont, Calif.: Fearon, 1968.
- Maier, N. R. F. Assets and Liabilities in Group Problem Solving: The Need for an Integrative Function. Psychological Review, 1967, 74, 4, 239-249.
- Maier, N. R. F., Solem, A. R., and Maier, A. Supervisory and Executive Development. New York: Wiley, 1957.
- Marquis, D. G. Individual Responsibility and Group Decisions Involving Risk. Industrial Management Review, 1962, 3, 8-23.
- Mayfield, Eugene C. The Selection Interview - A Re-evaluation of Published Research. Personnel Psychology, 1964, 17, 239-260.
- Mayfield, Eugene C., and Carlson, Robert E. Selection Interview Decisions: First Results From a Long-Term Research Project. Personnel Psychology, 1966, 19, 41-53.
- Miller, James W., and Rowe, Patricia M. Influence of Favorable and Unfavorable Information Upon Assessment Decisions. Journal of Applied Psychology, 1967, 51, 5, 432-435.
- Moore, L. F., and Craik, J. C. Video Tape and the Screening Interview. Personnel Journal, 1972, 51, 3, 173-178.
- Norsworthy, N., The Validity of Judgments of Character. Essays in Honor of William Jones, 1910 (cited in Symonds, 1931).
- Office of Strategic Services. Assessment of Men. New York: Rinehart, 1948.
- Otis, J. L., Campbell, J. H., and Prien, E. P. Assessment of Higher-Level Personnel: VII. The Nature of Assessment. Personnel Psychology, 1962, 15, 441-446.
- Palacios, M. H., Newberry, L. A., and Bodtzin, R. R. Predictive Validity of the Interview. Journal of Applied Psychology, 1966, 50, 1, 67-72.
- Rimland, B. A Follow-up Analysis of the New Composite System for Selecting NROTC Regular Students. USN Bureau of Naval Personnel Technical Bulletin, 1960, No. 60-8 (cited in Ulrich and Trumbo, 1965).
- Rowe, Patricia M. Individual Differences In Assessment Decisions. Unpublished Doctoral Dissertation, McGill University, 1960 (cited in Webster, 1964).

- Rowe, Patricia M. Individual Differences in Selection Decisions. Journal of Applied Psychology, 1963, 47, 304-307.
- Rowe, Patricia M. Order Effects In Assessment Decisions. Journal of Applied Psychology, 1967, 51, 2, 170-173.
- Rowe, P. M. Differences in Selection Decisions Among Personnel Interviewers. Canadian Personnel and Industrial Relations Journal, 1970, 19-23.
- Rundquist, E. A. Development of an Interview for Selection Purposes. In G. A. Kelly (Ed.), New Methods in Applied Psychology. College Park: University of Maryland, 1947, 85-95 (cited in Ulrich and Trumbo, 1965).
- Schwab, Donald P. and Heneman III, Herbert, G. Relationship Between Interview Structure and Interviewer Reliability In an Employment Situation. Journal of Applied Psychology, 1969, 53, 3, 214-217.
- Scott, W. D., Bingham, W. V., and Whipple, G. M. Scientific Selection of Salesmen. Salesmanship, 1916, 4, 106-108 (cited in Wagner, 1949).
- Siegel, Sidney. Non-Parametric Statistics. New York: McGraw-Hill, 1956.
- Smith, M. Group Judgments In the Field of Personality Traits, Journal of Experimental Psychology, 1931, 14, 562-565 (cited in Zajonc, 1966).
- Smith, Patricia C., and Kendall, L. M. Retranslation of Expectations: An Approach to the Construction of Unambiguous Anchors for Rating Scales. Journal of Applied Psychology, 1963, 47, 2, 149-155.
- Spriegel, W. R., and James, V. A. Trends in Recruitment and Selection Practices. Personnel, 1958, 35, 3, 42-48.
- Springbett, B. M. Series Effects in the Employment Interview. Unpublished Doctoral Dissertation, McGill University, 1954 (cited in Webster, 1964).
- Springbett, B. M. Factors Affecting the Final Decision in the Employment Interview. Canadian Journal of Psychology, 1958, 12, 13-22.
- Stoner, J. A. F. A Comparison of Individual and Group Decisions Involving Risk. Unpublished Master's Thesis, M.I.T. School of Industrial Management, 1961 (cited in Brown, 1965).

- Stroh, Thomas F. The Uses of Video Tape in Training and Development. New York: American Management Association, 1969, 11.
- Stroop, J. B. Is the Judgment of the Group Better than that of the Average Member of the Group? Journal of Experimental Psychology, 1932, 15, 550-560.
- Sydiaha, Daniel. The Relation Between Actuarial and Descriptive Methods in Personnel Appraisal. Unpublished Doctoral Dissertation, McGill University 1958 (cited in Webster, 1964).
- Sydiaha, Daniel. Bales' Interaction Process Analysis of Personnel Selection Interviews. Journal of Applied Psychology, 1961, 45, 6, 393-401.
- Symonds, P. M. Diagnosing Personality and Conduct. New York: Appleton-Century, 1931.
- Taft, R. Multiple Methods of Personality Assessment. Psychological Bulletin, 1959, 56, 333-352.
- Thorndike, R. L. and Hagen, E. Measurement and Evaluation in Psychology and Education. Wiley: New York, 1969 (3rd Edition).
- Ulrich, Lynn and Trumbo, Don. The Selection Interview Since 1949. Psychological Bulletin, 1965, 63, 2, 100-116.
- Wagner, R. F. The Employment Interview: A Critical Summary. Personnel Psychology, 1949, 2, 17-46.
- Wallach, M. A., Kogan, N. The Roles of Information, Discussion and Consensus in Group Risk Taking. Journal of Experimental Social Psychology, 1965, 1, 1-19.
- Wallach, M. A., Kogan, N., and Bem, D. J. Group Influence on Individual Risk Taking, Journal of Abnormal and Social Psychology, 1962, 65, 2, 75-86.
- Wallach, M. A., Kogan, N., and Bem, D. J. Diffusion of Responsibility and Level of Risk Taking in Groups. Journal of Abnormal and Social Psychology, 1964, 68, 3, 263-274.
- Walter, Gordon A. Augmentation of Task Group Performance Through Video Tape Feedback and Social Learning. Unpublished Doctoral Dissertation, Berkeley, Calif., 1971.
- Walter, G. and Miles, R. The Impact of Task Group Participational Plus Video Tape Feedback on Self Acceptance Scores, 1972 (submitted for publication).

- Weber, R. J. The Effects of Videotape Feedback on Interaction Behavior and Role Perceptions in Small Decision-Making Groups. Unpublished Doctoral Dissertation, Berkeley, Calif., 1969.
- Webster, E. C. Decision Making In the Employment Interview. Industrial Relations Centre, McGill University, Montreal, 1964.
- Wexley, K. N., Yukl, G. A., Kovacs, S. Z., and Sanders, R. E. Importance of Contrast Effects in Employment Interviews. Journal of Applied Psychology, 1972, 56, 1, 45-48.
- Wiens, N., Molde, D. A., Holman, D. C., and Matarazzo, J. D. Can Interview Interaction Measures be Taken from Tape Recordings? Journal of Psychology, 1966, 63, 249-260.
- Wonderlic, E. F. Improving Interview Technique. Personnel, 1942, 18, 232-238 (cited in Wagner, 1949).
- Woodworth, D. G., Barron, F., and MacKinnon, D. W. An Analysis of Life History Interviewers' Ratings for 100 Air Force Captains. USAF Personnel and Training Research Center Technical Note, 1957, No. 57-129 (cited in Ulrich and Trumbo, 1965).
- Wright, Orman R. Summary of Research on the Selection Interview Since 1964. Personnel Psychology, 1969, 22, 391-413.
- Zajonc, R. B. Social Psychology: An Experimental Approach. Wadsworth: Belmont, Cal., 1966.

APPENDIX A

THE CAMPUS INTERVIEW

APPENDIX A

The Campus Interview

Bank of Montreal

This is primarily a screening interview in which you must decide if the student should be invited to the divisional office for a second "look".

- o Allow sufficient time to review your day's schedule of interviews before the first interview starts.
- o Keep on schedule. If you don't students may either skip the interview or miss classes.

The following is a suggested breakdown of steps to be observed in an interview with times suggested for each step:

Step I	Review resume	1 min.
Step II	Establish rapport	2 min.
Step III	Evaluate student -	
	o Education and work experience	3 min.
	o Personal history	2 min.
	o Potential	<u>5 min.</u>
		10 min.
Step IV	Provide information	7 min.
Step V	Questions and answers	5 min.
Step VI	Inform student of future consideration	1 min.
Step VII	Record results and action	<u>4 min.</u>
		30 min.

Each of these steps are explained in detail below:

(I) Review Resume

In almost all cases you will be provided with a resume on each student. Study this before you meet him. It will usually outline vital information on the student along with some information on his background, education and work experience. Use it to assist you in planning your interview and also to avoid asking for information he has already supplied.

(II) Establish Rapport

There are many methods used to gain the confidence and acceptance of a student, each varying according to the personality of the interviewer and the student. The student may be nervous or tense. Put him at ease -

- (a) Rise from your chair, give him a friendly greeting, a firm handshake and introduce yourself by your first and last name.
- (b) Use some, not too much, small talk.
- (c) Offer him a cigarette or permit him to smoke if he so desires and generally provide an informal but businesslike atmosphere conducive to effective communication.
- (d) Switch to the meat of the interview quietly by introducing a broad series of questions on a topic you think the candidate will talk freely about.
- (e) Let the student do most of the talking. In so doing he will reveal his personality and the features he considers important about himself and his future.
- (f) Your job is to listen critically and with understanding, interspersing brief comments only to guide the interview into the areas that will help you to make a proper appraisal of the applicant as a person.

(III) Evaluate Student (see also Section 6)

This is the prime purpose of the interview - to obtain an accurate evaluation of the student while he is trying to favourably impress you.

(a) Education and Work Experience

The student should be encouraged to discuss his educational background and work experience. In addition to providing an indication of the student's previous training this provides him with subjects he knows well and can discuss easily.

Statements rather than questions will usually produce more information. If the student is permitted to discuss his background and experiences, a greater insight into his personality will be obtained.

This period of the interview will provide the interviewer with the opportunity to evaluate the student's state-

ments, note any inconsistencies and shortcomings, observe his manner, and consider his experience.

(b) Personal History

The interviewer should, if possible, consider the influence of the student's home background from the standpoint of early advantages or disadvantages effecting his development during the formative years. Only if good rapport has been established, should questions be asked pertaining to the student's home environment. His answers may uncover basic reasons for pursuing a higher education, his motivation, attitude towards himself, and his ambitions. If the information is not provided readily by the student or rapport is not extremely well-established, it would be best for the recruiter (particularly if he is inexperienced) not to investigate this subject in depth. Instead, questions should concentrate on the student's educational background and work experience.

The extent of the student's extra-curricular activities and participation in sports should be explored, as should any hobbies he enjoys. This may provide indication of leadership ability and sociability.

While the student discusses his personal history you may be able to assess if his goals are compatible with his qualifications.

(c) Potential

The student should be questioned as to his career objectives and how he feels these can be fulfilled in the Bank. It should be determined if he is casually interested in banking or if he has been considering it as a career for some time. Generally, this may be determined by exploring his understanding of banking. Enthusiasm is also an indicator of potential, as those with a positive attitude are less likely to be discouraged by the difficulties encountered through demanding job assignments and responsibilities.

(IV) Provide Information

In this part of the interview you should turn from evaluating the student to describing the Bank and its opportunities. Here you will do most of the talking. The student should be informed of the Bank of Montreal, our development programme and the salaries and benefits offered by the Bank. Be honest and do not oversell.

(a) The Bank of Montreal

In addition to the "quick facts" and organization charts included in the Recruiter's Guide, you should stress

our new management philosophy, i.e. to be "the most successful bank in Canada - which to us means the most profitable." Students should be informed of our promotion from within policy, and the fact that we are prepared to give heavy responsibility to those with demonstrated ability.

(b) The Development Programme

As a recruiter you must be thoroughly familiar with the Special Development Programme for graduates. The nature and extent of assignments and objectives of this programme should be covered as outlined in Section 9 of the Recruiter's Guide. Following this, the initial employment location should be discussed. If he is not willing to relocate, reasons should be explored and noted on the Campus Interview form.

(c) Salaries and Benefits

The student should be informed of the basic starting salary for his level of education, and that merit increases are provided at regular intervals based on performance. The Bank's benefits should be covered, including our policy regarding initial moves, transfers and the Tuition Refund Plan.

(V) Questions and Answers

You should answer all the student's questions if possible at the time of the interview. Be honest and frank with all students particularly with those who are to be considered further. If a question is asked which you cannot answer at the time, obtain the answer later from the Divisional Office if necessary, and forward a reply to the student. In the interests of good relations, this procedure is to be followed even if the student does not warrant further consideration.

As a general rule, the question and answer period should be dealt with as briefly as possible without loss of clarity. The few minutes available will not be sufficient to satisfy the student and, therefore, no time should be wasted.

(VI) Inform Student of Future Consideration

Close the interview by informing the student that time is running out and give him a chance to ask one or two last questions.

All students should be informed that they will receive a letter, mailed within ten day of the interview informing them of the outcome of the interview.

If the student is to remain a candidate tell him that his resume will be considered thoroughly by management

prior to a possible visit to the Divisional Office.

No firm offer of employment should be made at the time of the campus interview. Thank the student for his interest in the Bank and tell him how pleased you are to have had the chance to talk with him.

(VII) Record Results and Action

Immediately following the interview you must complete the Graduate Recruiting - Campus Interview form. Section 1, 2 (where necessary), 3 and 4 are to be completed on all students interviewed. As a recruiter, you must decide whether the student is to be considered further or not. If the student is to be considered further, section 5 is to be completed. Any comments you believe would assist the division in selecting students for a "second look" should be included in the "comments" section.

In making your decision watch for the following:-

- o Any bias that may have resulted from an incident in the interview that either very much disturbed or pleased you.
- o Tendency to select candidates of a certain image pattern.
- o Evaluations resulting from "halo" rating.
- o Ability of the candidate to fit into the Bank and still maintain his individuality.

APPENDIX B

APPLICATION FOR EMPLOYMENT



Bank of Montreal

Application for Employment

Please print all particulars clearly

Divisional Office Use Only

APPENDIX B -147-

Division _____ Entry Date _____

Branch _____

Salary _____ Allowance _____

Show Mr., Miss, Mrs. Last Name Given names as on birth certificate—Underline or indicate name by which you are known.

Address Number and Street City or Town Province Telephone

Check boxes applicable

Single	<input type="checkbox"/>	Languages	1 _____	2 _____	3 _____
Married	<input type="checkbox"/>	Speak Fluently or	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Widowed	<input type="checkbox"/>	Good working knowledge	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Separated	<input type="checkbox"/>	Write well or	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Divorced	<input type="checkbox"/>	Write fairly well	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Married female applicants—give maiden name _____

Person to be contacted in case of emergency Relationship Telephone

Birth date _____ Day Month Year	Height _____	If married, does spouse work? _____
Age _____ Sex _____	Weight _____	Name of spouse's employer _____
Social Insurance No. _____	No. of brothers _____ Sisters _____	No. of dependent children _____

Do you, have any physical defects or diseases? Yes ☐ No ☐ If yes, explain _____

Have you ever been guilty of a criminal offence? Yes ☐ No ☐ If yes, explain _____

Have you applied to us before? Yes ☐ No ☐ If yes, date _____ Location _____

Have you worked for us before? Yes ☐ No ☐ If yes, where? _____ From _____ 19____ to _____ 19____

Have you ever been discharged or requested to resign from any position you have held? Yes ☐ No ☐ If yes, explain _____

If your application is accepted, when could you commence work? _____ Will you accept transfers periodically? Yes ☐ No ☐

Do you have any debts? Yes ☐ No ☐ If yes, give particulars of amounts, creditors, etc. _____

Additional sources of income? Yes ☐ No ☐ If yes, explain _____

List present or past affiliations in High School, University, professional or community activities (excluding religious, national or racial groups), mentioning offices held, if any.

What hobbies or recreational activities do you enjoy? _____

List three persons other than relatives or previous supervisors from whom we may request references.

Name	Full Address	Occupation	Years known

Circle highest educational level achieved	High School				University				Graduate School				
	10	11	12	13	1	2	3	4	1	2	3	4	
	Name				Address				From		To		Diploma or Degree and Major Course
Month									Year	Month	Year		
High School_____													
Business School_____													
University_____													
Graduate School_____													
Other_____													

If diploma or degree not obtained state reason_____

Average marks obtained during each of last 3 years of education _____ Final_____

List any scholastic honours including scholarships and awards_____

What were your two best subjects? First_____Second_____

Do you plan to continue formal education? Yes ☐ No ☐ If yes, when, where and what type? _____

Indicate any special skills you have acquired through job experience or specific training.

Typing ☐ Filing ☐ Posting Machine ☐
Shorthand ☐ Bookkeeping ☐ Other ☐ Specify_____
Dictating machines ☐ Cash Experience ☐

List particulars of last three employers from whom we may request references. If no previous permanent employment give information concerning part-time and summer work. If employed at present may we contact your employer? Yes ☐ No ☐

Dates	1 Name of Company 2 Address of Company and Telephone No. 3 Name of Supervisor and Title	Salary		Job Title or type of work performed	Reason for Leaving
		Starting	Final		
From Month Year	1	\$	\$		
To Month Year	2	per	per		
	3				
From Month Year	1	\$	\$		
To Month Year	2	per	per		
	3				
From Month Year	1	\$	\$		
To Month Year	2	per	per		
	3				

Which previous position provided the greatest satisfaction?_____

Why are you interested in banking?_____

Additional information which you feel may be of assistance in assessing your application _____

I certify that all information contained herein is correct to the best of my knowledge.

APPENDIX C

VIDEOTAPED INTERVIEW QUESTIONNAIRE

APPENDIX C

Videotaped Interview Questionnaire

1. How many employment interviews have you been to this year?

2. To what extent did you find yourself distracted by each of the following: (please check the appropriate response)

	Very Dis- traced (distracted throughout the whole interview)	Quite Dis- traced (distracted for at least half of the interview)	Somewhat Distracted (distracted for less than half of the interview)	Not Distracted at all
	(1)	(2)	(3)	(4)
a. presence of the video- tape equipment	_____	_____	_____	_____
b. presence of the camera- man	_____	_____	_____	_____
c. the know- ledge that you were being video-taped	_____	_____	_____	_____
d. the noise of the video- tape equipment	_____	_____	_____	_____

3. In this question we are not concerned with such factors as the interviewer, questions asked, room etc., but rather with the effects of being video-taped in an interview as compared with not being video-taped. Please keep this in mind as you respond.

For each of the following behavioural characteristics please check the extent to which you felt your behaviour was better or worse in this video-taped interview as compared to how you feel it would have been had the interview not been video-taped.

In this video-taped interview, my behaviour along each characteristic was:

	<u>much better</u>	<u>slightly better</u>	<u>about the same</u>	<u>slightly worse</u>	<u>much worse</u>
	(1)	(2)	(3)	(4)	(5)
nervousness:	_____	_____	_____	_____	_____
honesty:	_____	_____	_____	_____	_____
judgment:	_____	_____	_____	_____	_____
voice into- nation:	_____	_____	_____	_____	_____
ability to express myself verbally:	_____	_____	_____	_____	_____
manner:	_____	_____	_____	_____	_____
appearance:	_____	_____	_____	_____	_____
force or drive:	_____	_____	_____	_____	_____
interest:	_____	_____	_____	_____	_____
social sensitivity:	_____	_____	_____	_____	_____
intelligence	_____	_____	_____	_____	_____
overall behaviour:	_____	_____	_____	_____	_____

4. After having experienced a video-taped interview, how do you feel the video-taped interview compares to non-video-taped interviews? (check one)

Video-taped interviews are:

- a. less effective than non-video-taped interviews _____ (1)
 b. more effective than non-video-taped interviews _____ (2)
 c. about the same as non-video-taped interviews _____ (3)

5. Would you be willing to undergo a video-taped interview which, at your request, could be sent to companies you are interested in?

yes _____ (1)
 no _____ (2)
 undecided _____ (3)

6. How enthusiastic are you toward the possible general use of video-taped initial screening interviews for company selection of university graduates for employment?
 (please check the appropriate response)

very enthusiastic _____ not enthusiastic at all _____
 (7) (6) (5) (4) (3) (2) (1)

7. How do you feel that most graduating students would react to the suggestion, by a company, that the student undergo a videotaped interview conducted by a University Placement Center representative and forwarded to the company for examination?

very unfavourable _____ (1)
 somewhat unfavourable _____ (2)
 somewhat favourable _____ (3)
 very favourable _____ (4)
 don't know _____ (5)

8. In which of the following areas do you have reservations or objections regarding videotaped screening interviews?
 (Please rank the three most serious objections you have, then check any others that apply)

- 1 _____ Many important personal characteristics cannot be conveyed through the videotaped interview.
 2 _____ This is just one more step toward the de-personalization of the employment relationship.
 3 _____ There is no assurance that a videotaped interview will be kept confidential -- too easy to duplicate and use for unauthorized purposes.
 4 _____ The videotaped interview is unethical.
 5 _____ Most companies do not have the expensive video playback equipment necessary.
 6 _____ Once an interview is made, there is no way of changing it.
 7 _____ The student may be forced to become an 'actor' if wants to get a job.
 8 _____ In front of a camera, most people do not act natural
 9 _____ The technique will be too expensive.
 10 _____ Other (specifiy) _____

VIDEOTAPE INTERVIEW EVALUATION

The videotape technique may or may not be useful in employment interviewing. The following questionnaire is designed to assess some of the aspects on which the videotape interview may be evaluated.

Please answer the questions honestly and candidly. There are no right or wrong answers. You are asked not to discuss this evaluation until all those involved have completed their questionnaires. We are interested in how you feel. Thank you.

-
1. In your experience, and after having seen a number of videotaped interviews, how do you feel the videotaped interview compares to the face-to-face interview technique (i.e. without video-tape)? (check one)

Generally, videotaped interviews are:

- less effective than face-to-face interviews ____ (1)
- more effective than face-to-face interviews ____ (2)
- about the same as face-to-face interviews ____ (3)

2. In the whole, how realistically do you feel the videotaped interviews conveyed the actual amount of the following personal characteristics possessed by the interviewees? (Please check the appropriate responses).

	very un- realistic	somewhat unrealistic	somewhat realistic	very realistic	unable to judge
	(1)	(2)	(3)	(4)	(5)
appearance	_____	_____	_____	_____	_____
manner	_____	_____	_____	_____	_____
voice	_____	_____	_____	_____	_____
force or drive	_____	_____	_____	_____	_____
intelligence	_____	_____	_____	_____	_____
potential	_____	_____	_____	_____	_____
interest	_____	_____	_____	_____	_____
attitude	_____	_____	_____	_____	_____
sociability	_____	_____	_____	_____	_____
self-expression	_____	_____	_____	_____	_____
knowledge of field	_____	_____	_____	_____	_____
self-confidence	_____	_____	_____	_____	_____
nervousness	_____	_____	_____	_____	_____
motivation	_____	_____	_____	_____	_____
stress	_____	_____	_____	_____	_____
maturity	_____	_____	_____	_____	_____
judgment	_____	_____	_____	_____	_____
persuasiveness	_____	_____	_____	_____	_____
leadership	_____	_____	_____	_____	_____

3. How enthusiastic are you toward the possible general use of video-taped initial screening interviews for company selection of university graduates for employment? (please check)

very enthusiastic

not enthusiastic
at all

(7) (6) (5) (4) (3) (2) (1)

APPENDIX D

VIDEO-TAPE VIEWERS QUESTIONNAIRE

Video-tape Viewer's Evaluation of the Interviewer

In this series of questions we are interested in your perceptions of how the interviewer behaved in this interview. For each question please check the appropriate response.

1. To what extent do you think the interviewer was interested in the interviewee?

The interviewer was:

very interested in the interviewee	_____	(1)
quite interested in the interviewee	_____	(2)
somewhat interested in the interviewee	_____	(3)
not interested in the interviewee at all	_____	(4)

2. In terms of the content of this interview, how knowledgeable do you think the interviewer was of the banking field (i.e. banking training programs, possible available positions, etc.)

The interviewer was:

very knowledgeable	_____	(1)
quite knowledgeable	_____	(2)
somewhat knowledgeable	_____	(3)
not knowledgeable at all	_____	(4)

3. To what extent do you think the interviewer made the interviewee feel at ease?

The interviewer made the interviewee feel:

very much at ease	_____	(1)
quite at ease	_____	(2)
slightly at ease	_____	(3)
not at ease at all	_____	(4)

4. How well do you think the interviewer communicated information to the interviewee?

The interviewer:

communicated information very well	_____	(1)
communicated information quite well	_____	(2)
communicated information reasonably well	_____	(3)
did not communicate information well at all	_____	(4)

5. To what extent do you feel the interviewer spoke too much or too little in this interview?

The interviewer:

spoke much too much	_____	(1)
spoke more than enough	_____	(2)
spoke about the right amount	_____	(3)
did not speak quite enough	_____	(4)
spoke much too little	_____	(5)

6. Please check the extent to which each of the following paired-adjectives describe the questions which the interviewer asked the interviewee.

On the whole, the interviewer's questions were:

	ex- treme- ly _____	very _____	some- what _____	some- what _____	very _____	ex- treme- ly _____	
relevant	_____	_____	_____	_____	_____	_____	irrelevant
	(1)	(2)	(3)	(4)	(5)	(6)	
difficult to answer	_____	_____	_____	_____	_____	_____	easy to answer
	(1)	(2)	(3)	(4)	(5)	(6)	
vague	_____	_____	_____	_____	_____	_____	clear
	(1)	(2)	(3)	(4)	(5)	(6)	

7. To what extent do you feel that the interviewer gathered sufficient or insufficient information about the applicant? (please check)

The interviewer:

gathered enough information about the applicant	_____	(1)
could have gathered somewhat more information about the applicant	_____	(2)
could have gathered a lot more information about the applicant	_____	(3)

8. Please list any additional items of information you would like to have obtained about the applicant.

9. As far as you can tell, how favourable or unfavourable do you think the interviewee's image of the Bank of Montreal was before he came to the interview?

The interviewee's image of the bank was:

very favourable	_____	(1)
quite favourable	_____	(2)
neutral	_____	(3)
quite unfavourable	_____	(4)
very unfavourable	_____	(5)

10. To what extent do you think the interviewer made the interviewee's image of the Bank more favourable or unfavourable?

The interviewer made the interviewee's image of the Bank:

much more favourable	_____	(1)
somewhat more favourable	_____	(2)
remain about the same	_____	(3)
somewhat more unfavourable	_____	(4)
much more unfavourable	_____	(5)

11. In terms of this interview, how good an interviewer do you think he (she) is?

a very good interviewer	_____	(1)
quite a good interviewer	_____	(2)
an average interviewer	_____	(3)
quite a bad interviewer	_____	(4)
a very bad interviewer	_____	(5)

APPENDIX E

SUMMARY SHEET

APPENDIX E

-159-

INTERVIEWER - Jim Jones

FACTOR	SELF	INTERVIEWEE	GROUP	INDIVIDUAL
Attitude	3.25	3.50	3.42	3.65
Appearance	3.09	3.65	3.00	3.50 #
Interest	2.83	3.50	3.17	3.33 #
Intelligence	3.33	3.33	3.59	3.83
Leadership	3.59	3.50	3.42	3.70
Maturity	3.33	3.65	3.42	3.83
Motivation	3.17	3.50	3.33	3.50
Persuasiveness	3.00	3.50	3.17	3.17
Self-Confidence	3.17	3.50	3.42	3.83 #
Self-Expression	3.17	3.00	3.17	3.74
Sociability	3.42	3.33	3.33	3.83
Potential	2.83	3.50	3.09	3.50 #
Overall	2.17	2.33	2.25	2.41 #

2nd Interview

(Halliburton)	1	2	2	1
(McNish)	2	2	2	2
(Saunders)	2	1	2	2
(Gordon)	1	1	1	1
(Royston)	1	1	1	1
(Dupuis)	2	1	2	1

1 = yes

2 = no

KEY FOR RATING FACTORS:

Excellent 5
 Superior 4
 Average 3
 Marginal 2
 Unsatisfactory 1

OVERALL RATING

Excellent 4
 Superior 3
 Average 2
 Marginal 1

When looking at these ratings it would be worthwhile to keep in mind a few points:

- (1) this exercise is for training and informational purposes only and is in no way to be used as a formal performance evaluation;
- (2) from a preliminary review of the data it appears that groups tend to be more critical than individual raters;
- (3) the group and individual ratings were made after viewing video-taped interview, not after actually engaging in an interview.

Evaluation of the Interviewees

Generally, your ratings are much more critical than those of the viewers and of the interviewees. At this point it is not possible to determine whether this fact is a result of the video-tape or whether it is something which is peculiar to yourself.

You particularly are more critical of the interviewee in terms of appearance, interest, self-confidence and potential (see asterisks on the Summary Sheet). The appearance factor may be quite readily explained by the video-tape. The viewers did not obtain as good a view of the interviewee as you did while you were in his presence.

In like fashion, interest and self-confidence are two factors which may be influenced by the interviewee's facial characteristics and gestures which the video-tape viewers would not perceive as well as yourself.

Conditions such as the above may have affected your general impression of the applicant in a manner quite different from that of the video-tape viewers. This could consequently have a significant bearing upon your overall rating and upon your rating of the interviewee's potential (both of which are again quite lower than those of the viewer's ratings).

Evaluation of the Interviewer

For each of these factors your ratings were generally quite high. The data points out that your knowledge of the Bank's programmes, the extent to which you put the applicant at ease, your interest and the relevance and clarity of your questions were very positive. Overall, you spoke about the right amount; an amount (36%) which comes within the acceptable range given your interview structure. Also, you improved the applicants' average image of the Bank; an image which was on the favourable side of neutral.

Your ability to communicate information met with some disagreement, particularly from the part of the group. The group's rating, though, was still on the positive side.

Overall, your ability as an interviewer was consistently rated as being quite good especially from the point of view of the interviewees, whose impressions may be considered as being most important in an interview.

General

There was very little variation amongst scores for each of the factors considered in your analysis. You were consistently rated quite highly.

I hope that this data proves to be of some value to you Jim, and I thank you very much for participating in the study.

EVALUATION OF THE EXPERIMENT

5114

JONES

1. How interested were you in the applicant?

	High				Low
	1	2	3		4
self	1.83				
interviewee	1.83				
group	2.15				
individual	1.66				

2. How knowledgeable were you?

	High				Low
	1	2	3		4
self	1.17				
interviewee	1.32				
group	1.82				
individual	1.50				

3. To what extent did you put the applicant at ease?

	High				Low
	1	2	3		4
self	1.66				
interviewee	1.66				
group	1.66				
individual	1.53				

4. How well did you communicate information?

	High 1	2	3	Low 4
self	2.33			
interviewee	1.33			
group	2.17			
individual	1.50			

5. To what extent did you speak too much or too little?

	Much too much 1	2	the right amount 3	4	much too little 5
self	3.17				
interviewee	3.00				
group	3.00				
individual	3.00				

5a What percentage of the time did you speak

	0%	25%	50%	75%	100%
self	31%				
interviewee	43%				
group	30%				
individual	29%				
actual	36%				

6. Your questions were:

	relevant			irrelevant		
	extremely 1	very 2	some- what 3	some- what 4	very 5	extremely 6
self		2.17				
interviewee		2.00				
group		2.00				
individual		1.83				

	vague				clear	
	1	2	3	4	5	6
self				4.50		
interviewee				4.66		
group					5.33	
individual					5.00	

7. The interviewees' image of the Bank was:

	very favourable		neutral		very unfavourable	
	1	2	3	4	5	
self			2.50			
interviewee			2.66			
group			3.00			
individual			2.66			

8. You made the interviewees' image of the Bank

	much more favourable 1	2	about the same 3	4	much more unfavourable 5
self			2.33		
interviewee		1.83			
group			2.50		
individual			2.33		

9. How good an interviewer are you (in terms of these interviews)?

	very good inter- viewer 1	2	an average inter- viewer 3	4	very bad inter- viewer 5
self			2.66		
interviewee		1.50			
group			2.00		
individual		1.83			