1996

# The validity of situational interview ratings as a function of channel of communication

Seema Thakur

# THE VALIDITY OF SITUATIONAL INTERVIEW RATINGS AS A FUNCTION OF CHANNEL OF COMMUNICATION

---

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

---

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Psychology: Industrial/Organizational

---

by

Seema Thakur

June 1996

# THE VALIDITY OF SITUATIONAL INTERVIEW RATINGS AS A FUNCTION OF CHANNEL OF COMMUNICATION

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

by

Seema Thakur

June 1996

Approved by:

[redacted] 6/7/96

Matt L. Riggs, Ph.D., Chair, Psychology Date

[redacted]

Kenneth S. Shultz, Ph.D., Psychology

[redacted]

Thung-Rung Lin, Ph.D.,
Personnel Commission
Los Angeles Unified School District

ABSTRACT

This study investigated whether making nonverbal cues accessible to raters enhanced or eroded situational interview validity. Also investigated in this study, was the impact of behaviorally anchored rating scales on interview validity.

Eighty subjects used a situational interview to evaluate job candidates in three channels of communication: video, audio, and transcription. Seven of the questions included in the interview contained behaviorally anchored rating scales, and four did not. Interview ratings were correlated with job performance ratings to assess validity.

Fishers' r to z transformations were computed to compare bivariate correlations between job performance and interview ratings made in the three communication modes. Results revealed the correlation between interview ratings made in the video mode and job performance to be significantly higher than either the correlation between ratings made in the audio mode and job performance ($z=2.26^*, p=.0119$) or between ratings made in the transcription mode and job performance ($z=3.02^*, p=.0013$). No differences in validity were found between transcription and audio ratings. A hierarchical regression analysis showed that ratings made with behavioral anchors could explain variance in job performance beyond the variance in job performance explained by

ratings made without anchors (F=39.04*,p=.000). These results suggest that both nonverbal cues and behaviorally anchored rating scales contribute to the validity of the situational interview.

ACKNOWLEDGMENTS

This thesis is dedicated to my mom and dad. Thank you for teaching me to never give up. Most of all, thank you for being such loving and encouraging parents.

Thank you T.R. Lin for your support.

Thank you Matt Riggs for greeting me with a smile -- every single time. Thank you for coaching me through this project.

Thank you Sharon Cady for your soothing, calming personality -- just when I needed it most.

TABLE OF CONTENTS

LIST OF TABLES

# INTRODUCTION

The interview is by far the most commonly used selection procedure. It is used by virtually every company in the United States (Lin, Dobbins & Farh; Latham, Saari, Pursell, & Campion, 1980). Despite its widespread popularity, the employment interview often lacks reliability and validity (Latham, et. al., 1980). One reason is because interviewers rarely ask the same questions of different applicants. In addition, raters often disagree over whether interviewee responses are desirable or appropriate (Arvey & Campion, 1982). Furthermore, when the same questions are asked, they are often unrelated to job success. This results in low validity.

## The Situational Interview

There is hope for the interview. For the past decade, research has been investigating different kinds of selection interviews; such as interviews with structured formats. These have proven to be more predictive of job performance (Motowildo, Carter, Dunnette, Trippins, Werner, Burnett, Vaughan, 1992). One such interview is the situational interview (Latham, 1989). The situational interview uses a systematic job analysis known as the critical incident technique to develop interview content. The incidents are transformed into interview questions where job applicants are asked to indicate how they would behave in given situations. Two or more interviewers independently rate each

answer on a five-point likert-type scale. Behavioral statements created by job experts are used as benchmarks or illustrations of the different points on the likert scale, which facilitates objective scoring.

Results from research on the situational interview have been promising. Latham, Saari and colleagues (1980, 1984) conducted a series of concurrent and predictive validation studies of the situational interview that demonstrated validities ranging from .14 to .46. Weekley and Gier (1987) obtained a predictive validity coefficient of .47 when they administered the situational interview to select for a sales position. A study by Lin, Dobbins and Farh (1992) that investigated whether age and race similarity between interviewer and interviewee influenced interview outcomes, revealed a stronger same race effect for the conventional interview than for the situational interview. Their results suggest that the tight structure of the situational interview may minimize same race bias.

The situational interview also seems to have adequate face validity (Weekly & Gier, 1987). Managers using the interview displayed an overwhelmingly positive response to the situational interview format. Specifically, the ease of administration, the ease of interpretation, and the job relatedness of the questions were noted as benefits of the situational interview.

The format of the situational interview is tightly

structured, so that interviewers are limited to questions about applicants' intentions in certain job-related situations. This format implicitly assumes that the content of applicant's answers to the questions is what contributes to the validity of the interviewer judgements. That is, the strict design of the situational interview expects that visual and other nonverbal cues (such as mannerisms, physical attractiveness, dress and grooming) are to be excluded as sources of information when completing ratings on the job-related dimensions being assessed through the interview. However, it has been shown that visual and other nonverbal cues can have strong influences on the favorability of interviewer judgements (Baron, 1983; Wexely, Fugita & Malone, 1975). Thus, despite the structured format of the situational interview, which stipulates that interview dimensions are to be rated only according to content of interviewee answers to questions, visual cues may still be influencing rater judgements.

Understanding the effects of interviewee nonverbal behavior on interview ratings is important because it may lead to improvements in validity of the situational interview. For example, if certain nonverbal behaviors are found to affect interview outcomes, but do not predict future job performance, then a source of invalidity is isolated. Future interviewers can then be trained to be less influenced by interviewee nonverbal behavior. On the other hand,

it may be the case that visual cues significantly overlap with the knowledge, skills, and abilities being measured in the interview, and thus legitimately belong to the true score variance of the knowledge, skills, and abilities being considered. If this were the case, nonverbal cues would actually be contributing to overall interview validity.

Research on Nonverbal Cues and Interview Ratings

Past research has demonstrated the influence of visual information on perception and decision-making processes in interviewing. Hellman and Saruwatari (1979) found physical attractiveness to affect interview evaluations. They found that physical appearance was consistently an advantage for males, only an advantage for females seeking nonmanagerial positions, and a hindrance for females seeking managerial positions. A study by Forsythe, Drake and Cox (1985) indicated a positive relationship between masculinity of female applicants' costume and favorability of hiring recommendations received by the applicants for managerial positions. Specifically, female applicants received more favorable hiring recommendations as costume masculinity increased. Bardack and McAndrew (1985) showed that both physical attractiveness and appropriateness of clothing influenced the hiring decision. Imada and Hakel (1977) demonstrated the salience of eye contact, gestures, smiling, and posture in the formation of impressions and decisions.

Research on Nonverbal Cues and Job Performance Ratings

If nonverbal cues do affect interview judgements, they could be either supressing or enhancing validity. Two studies have investigated the effect of nonverbal cues on interview validity by isolating the visual, aural, and verbal cues and comparing the accuracy of judgements made under three interview conditions designed to maximize each kind of cue. A transcript interview condition provides raters with only verbal cues. An audio interview condition makes both verbal and aural (e.g., vocal characteristics) available to raters. Finally, an audio-visual or face-to-face interview condition allows raters to process visual, aural, and verbal cues of the job candidates.

A study by Mair and Thurber (1968) showed that interviewer judgements were more accurate when the judgements were based on audio or transcript records of interviews than with direct observation. Their results suggest that the visual nonverbal cues of the interview serve mainly as distractors, lowering the proportion of accurate decisions. The authors explained that a larger proportion of more accurate judgements could be obtained if the interviewers based their decisions solely on what the applicants said.

Motowidlo et.al. (1992) conducted a study on a structured behavioral interview, which investigated whether valid judgements could be made from information about the content of applicants' responses, even when visual cues were not

available to raters. Results revealed that raters who did not have access to such nonverbal cues could successfully predict the job performance criterion. Additionally, a hierarchical regression analysis demonstrated that interviews in which raters had access to nonverbal cues (interviews conducted face-to-face) could not account for the variance beyond the variance in job performance accounted for by interviews in which raters did not have access to nonverbal cues (raters who read or listened to the interviews). Interestingly, however, the listened/read interviews accounted for variance in performance beyond the variance accounted for by the face-to-face interviews. These results also suggest that nonverbal cues may serve as irrelevant bits of information, uncorrelated with job performance.

It seems that when the mode of presentation does not make visual information available to raters, the judgements made about future job performance are more valid. One reason why ratings made from videotaped records of interviews were not as accurate as ratings made from audiotaped and transcribed records of interviews, may be that visually observed interviews contain a great deal of potentially irrelevant information which is not related to future job performance (e.g., rates attractiveness, behaviors not linked to performance). Written and audio records contain less information, and much of the missing information is

largely irrelevant (Murphy et. al., 1986). It is possible that the observer who has access to visual information becomes so involved in the pictures that it becomes difficult to grasp the verbal content (Furnham, Benson, Gunter, 1987).

Semantic and Episodic Memory

Another reason for differences in interview validity as a function of mode of presentation, could be that raters who have access to visual cues are encoding, storing, and retrieving information differently from raters who do not have access to visual cues. According to Tulving and Thompson (1973), information obtained by actually experiencing the event may be processed in episodic memory, and information one has read about may be processed in semantic memory.

Remembering information from episodic memory is recollection of past events. Episodic memory is a system that receives and stores information about temporally dated episodes or past events and the temporal spatial relations among them. Consequently, when remembering, the rater's mind looks back at a past event and recollects how things appeared, when they occurred, and their relationship to each other in time and space. Episodic memory is context dependent; such that the processing of a unit of information is influenced by other units of information. It has been suggested that the organization in episodic memory is relatively loose and can be easily changed or lost. In addi-

tion, information stored in episodic memory cannot be remembered readily and systematically, on command.

On the other hand, remembering information from semantic memory is recall of facts. Semantic memory is a mental thesaurus of organized knowledge a person possesses about words, verbal symbols, their meaning, and concepts and relations among them. Recall from semantic memory is recall of facts, independent of a particular time or place. Organization of knowledge in the semantic system is conceptual and tight. Individual facts and ideas, once assigned to a particular part of the memory system, do not wander around freely in the system. This permits efficient retrieval. Consequently, evaluations from transcribed or auditory records of an interview utilize semantic memory; and thus, recall is based on tightly organized facts that are easily recalled. Evaluations from face-to-face or audio-visual records utilize episodic memory; therefore, recall is dependent on loosely represented visual cues that are more easily distorted.

Research on Channel of Communication

Research on mode of presentation supports the episodic-semantic distinction. Studies have found that both learning and memory are better from print and audio-only than from audio-visual presentation. One study showed that subjects receiving information in print medium, a verbal-only channel of communication, remembered significantly more facts than

subjects receiving information in the audio-visual medium, a nonverbal/verbal channel of communication (Furnham, Gunter & Green, 1990). In addition, results have demonstrated that where additional sources of information (such as film footage) are absent, as in the transcript condition, memory performance does not deteriorate as much (Furnham, Benson, & Gunter, 1987). These studies suggest that reading generates a deeper and more effective level of cognitive information processing than does viewing or listening (Gunter, Furnham, & Leslie, 1986).

The Aims of This Study

Given the strong influence of nonverbal cues on perception, learning, and memory, this study focuses on the question of whether nonverbal cues facilitate or erode interview validity. Specifically, this study will investigate whether mode of presentation, and hence, nonverbal cues, affect the validity of the inferences made from the situational interview about job performance. Ratings based on transcriptions should not be affected either by visual cues, present in videotaped summaries of interviews, or by vocal or speech characteristics present in audiotaped summaries of interviews. By collecting one set of interviewee ratings based only on audiotape and another based only on written summaries of the interview, the design of this study attempts to filter out as many extraneous cues as possible and leave only information about the content of interviewee answers.

Based on results of past studies that suggest the presence of nonverbal cues will suppress the validity of rater judgements, it is predicted that ratings based only on information about the verbal content of interviewee answers will be more valid than ratings based on verbal, vocal, nonverbal, and visual information.

Hypothesis 1: Ratings made from transcribed situational interview records will be more correlated with job performance ratings than either interview ratings made from audiotaped or videotaped records.

By restricting raters to asking a standard set of questions and rating responses to the questions with behaviorally anchored scoring guides, the situational interview limits the amount of judgement made, on the part of the raters, to only considering the content of the interviewee responses. It has been suggested that the strict format of the situational interview is a source of its validity (Motowidlo et. al., 1992). Specifically, the tapping of behavioral intentions, the use of a behaviorally anchored scoring guide, and the job relevancy of the questions are all restrictive factors of the situational interview which are presumed to contribute to its validity. Gatewood and Field (1989) expressed a need for future research to investigate what specific design factors of the situational interview contribute to its validity. In response to Gatewood and Fields' request, the current study will also investigate the

impact of behaviorally anchored rating scales on the validity of the situational interview. If behavioral anchors are contributing to interview validity, then removing behavioral anchors from the rating scale should cause validity to decrease. Consequently, this study compares the validity of interview ratings made without behaviorally anchored scoring guides with interview ratings made with behaviorally anchored scoring guides.

Hypothesis 2: Ratings made from situational interview questions without behaviorally anchored rating scales are less predictive of job performance across the three modes of presentation (audio, video, transcript) than ratings made from situational interview questions with behaviorally anchored rating scales.

## METHOD

### Subjects

A one-hundred-dollar lottery and class credit was used as incentive to solicit undergraduate students at a state university in Southern California to participate in this study. Student subjects were between the ages of 19 and 56. Forty-five were female, eleven were male (seven student subjects did not indicate their sex). Twenty-two subjects were recruited from the personnel department of a large school district in Southern California. These subjects volunteered to participate in the study with the encouragement of the department director, and the incentive of the one-hundred-dollar lottery. Experiments were run on work time. These subjects were between the ages of 22 and 50. Seventeen were female and five were male. To make sure that the subjects recruited from the school district were comparable to the subjects recruited from the university, only individuals who had completed at least one year of college were allowed to participate in the study. A total of 85 subjects participated in the study.

### Development of the Situational Interview

The situational interview was used by subject raters to evaluate three recorded interviews for the job of Junior Cafeteria Manager of a large school district. Junior Cafeteria Managers are responsible for the operation of elementary school meal distribution kitchens (see Appendix A). The

interview content was developed from the results of a comprehensive job analysis, extracting critical incidents. The critical incidents were turned into twelve situational questions, and their corresponding behavioral anchors were developed with the assistance of job experts. The twelve interview questions were designed to measure four job dimensions: Supervision, Interpersonal Relations, Resourcefulness, and Attitude.

An example of a situational question and its corresponding behavioral anchors under the dimension of Attitude is shown below:

> The kitchen/cafeteria becomes unexpectedly busy one day and you feel that you are understaffed. What would you do?
>
> STRONG RESPONSE (5-7 POINTS):
> - Work with the employees in the kitchen all day; divide the work as evenly as possible among the staff.
>
> ACCEPTABLE RESPONSE (2-4 POINTS):
> -Call the supervisor and ask for assistance.
> -Try to get additional student help.
>
> POOR RESPONSE (0-1 POINTS):
> -Nothing.
> -Encourage employees to work faster and harder.

Each of the four dimensions measured contains three interview questions. For the purposes of this study, the behavioral anchors of one question from each job dimension was deleted from the rating scale. Therefore, eight of the interview questions had behavioral examples of what Strong, Acceptable, and Poor Responses are and four did not. This was to determine whether ratings made with behavioral an-

chors are more valid than ratings made without anchors. An example of a situational interview question and a rating scale without behavioral anchors is shown below:

> The kitchen/cafeteria becomes unexpectedly busy one day and you feel that you are understaffed. What would you do?
>
> STRONG RESPONSE (5-7 POINTS)
>
> ACCEPTABLE RESPONSE (2-4 POINTS)
>
> POOR RESPONSE (0-1 POINTS)

Choosing Three Stimulus Interviews

Approximately seventy interviews of candidates that were conducted for the position of Junior Cafeteria Manager at a large school district in Southern California were videotaped and audiotaped. Subsequently, forty of the seventy candidates who underwent the Situational Interview for Junior Cafeteria Manager at the school district were hired. Following, a multi-source (or 360- degree) job performance evaluation was conducted after two years on the job as a Junior Cafeteria Manager. Evaluations were collected from three sources: 1) Junior Cafeteria Managers, 2) their subordinates, 3) and their supervisors. The multi-source method of performance evaluation was used to produce a more thorough appraisal of the employees' job performance (Murphy and Cleveland, 1991). The evaluations were used as the measures of true job performance (see Appendix B).

Job performance was evaluated using the same four job dimensions measured in the situational interview (Supervi-

sion, Interpersonal Relations, Resourcefulness, and Attitude). Evaluators rated three to six specific work behaviors within each of the four dimensions. Each rating was made using a five-point scale (5=Excellent, 4=Very Good, 3=Good, 2=Satisfactory, 1=Needs Improvement). For each rating, evaluators were instructed to provide at least one specific example to justify their ratings. It was hoped that instructing evaluators to think about actual observations and experiences they had with the Junior Cafeteria Manager would improve the quality of their ratings (Bernardin & Beatty, 1984).

Evaluations were collected from the Junior Cafeteria Managers and their supervisors (Area Food Services Supervisors & School Principals). Their subordinates (Cafeteria Helpers and Ticket Clerks) also provided evaluations of the Junior Cafeteria Managers. However, because of the overall low education level of the subordinates, these individuals were interviewed by research assistants to collect the information necessary to fill out the job evaluations. All evaluators were ensured that the information they provided would be used for research purposes only, and would not affect the Junior Cafeteria Manager's career advancement.

Cafeteria Helpers and Ticket Clerks work with the Junior Cafeteria Managers every day, for about three hours. Cafeteria Helpers assist the Junior Cafeteria Managers in heating up the food, laying out the food, and serving food

to the children. Ticket Clerks also help with the same duties as the Helpers, however their primary responsibility is to collect meal tickets from the students. Junior Cafeteria Managers report to Area Food Services Supervisors (AFSSs). The AFSSs are responsible for ensuring that the Junior Cafeteria Managers complete their duties correctly. On average, AFSSs supervise ten to fifteen Junior Cafeteria Managers. Unless there is a problem with the kitchen, AFSSs do not visit the Junior Cafeteria Managers more than once every two months. However, they communicate with the Junior Cafeteria Managers every week, via telephone, to make sure paperwork is being completed and turned in, and that the overall running of the kitchen is being handled properly. School Principals interact with the Junior Cafeteria Managers on a daily basis, and are there to attend to immediate kitchen issues, such as an irate parent.

A full 360-degree performance evaluation was received for eighteen of the forty Junior Cafeteria Managers. That is, eighteen Junior Cafeteria Managers received a self-rating, a rating from one of the two types of subordinates, and a rating from one of the two types of supervisors. These evaluations were analyzed to choose a high, low, and average performer. Scores on all four job dimensions that were rated were averaged to produce one total score for each of the performance evaluations filled out for each Junior Cafeteria Manager. These total scores were then averaged

across the three levels of raters to produce one overall performance score for each Junior Cafeteria Manager. The eighteen Junior Cafeteria Managers were ranked based on their total performance score. The 1st-, 9th-, and 18th-ranked Junior Cafeteria Managers were chosen to represent the high, average, and low performers respectively. The standard deviations of the averaged performance ratings were examined to make sure that ratings were consistently high, average, or low.

Performance Criterion

The job performance scores used to choose the three stimulus interviews were also used as criterion scores for the three job candidates. These criterion scores were compared against the interview ratings of the job candidates made by the subject raters to determine the validity of the ratings.

Procedure

Rater Training: Before evaluating applicants, subject raters underwent a ten-minute training session which reviewed the Junior Cafeteria Manager position, the types of interview questions, and the scoring guidelines for the interview. First, subjects were given a Junior Cafeteria Manager Fact Sheet (see Appendix A). This informed subjects of the hourly pay rate Junior Cafeteria Managers received, the duties they performed, ideal knowledge, skills, and abilities they should possess, the subordinates they super-

vise, and superiors they report to. Second, subjects were given a copy of the twelve interview questions and Response Scoring Guide. After being informed how to use the Response Scoring Guide, subjects were given five minutes to review the twelve interview questions and their respective scoring guides (when available). The Response Scoring Guide was taken away from subjects before the presentation of each of the interviews. However, subjects were instructed to take notes during the presentation of the three interviews.

After being trained, subjects were presented interviews of the strong, average, and weak candidate in a separate communication mode for each candidate. Subjects watched and listened to a videotaped interview, listened to an audiotaped interview, and read a transcribed interview. The order of the communication channel and candidate performance level that was presented to subjects was randomly varied to counterbalance any order effects. For example, one group of subjects may have rated the videotape of the weak candidate, the audiotape of the average candidate, and the transcription of the strong candidate, and another group of subjects may have rated the videotape of the strong candidate, the transcription of the weak candidate, and the audiotape of the average candidate.

After the presentation of each interview, subjects were given back the Response Scoring Guide and asked to rate the candidates' interview performance using the notes they had

taken during the interview. Subjects were given as much time as they needed to evaluate. Subjects rated each candidate's response to each question individually by comparing it to the benchmarks (when available) of STRONG, ACCEPTABLE, and POOR, and assigning a point value for the response between the point range of the chosen benchmark.

Data Analysis

A factor analysis was conducted on all interview ratings. If ratings cluster into four factors, and factors seem to parallel the four interview dimensions (Supervisory Skills, Attitude, Resourcefulness, and Interpersonal Skills), then interview dimensions will be separately correlated with each of the job performance dimensions to determine validity of the interview ratings. A coefficient alpha was also conducted to evaluate the internal consistency of interview ratings, and to confirm the factor analysis.

In order to address the first hypothesis, three bivariate correlations were performed between interview ratings on each of the channels of communications and job performance ratings. Fisher's r to z transformations were computed to compare the relative magnitude of the three correlations with job performance. If ratings made in the transcript mode are more correlated with job performance than ratings made in audio or video mode, then making nonverbal cues accessible to raters is suppressing interview validity.

A hierarchical regression was performed to determine

whether ratings made with behavioral anchors could explain variance in job performance over and above ratings made without anchors. Ratings made without behavioral anchors were entered into the regression equations first, followed by ratings made with anchors. If the change in $R^2$ is significant, it provides evidence that the behaviorally anchored rating scales are a source of the interview's validity.

Finally, three more bivariate correlations were performed between the ratings made with behavioral anchors and job performance within each of the three communication modes. The same correlations were performed for the questions without behaviorally anchored rating scales. These r's were compared.

## RESULTS

Prior to analysis, five subject cases were deleted because of incomplete data. This left data from 80 subjects to be analyzed. It was also decided to delete ratings from one of the interview questions from analysis because of a severe inaccuracy in transcribing the question. This left 11 interview questions to be analyzed. Only one interview question was misrepresented in the transcriptions.

The data was then examined for accuracy of data entry, normality, and outliers. The data analyzed were 1) each of the eleven interview ratings, 2) the averaged ratings of all eleven interview questions, 3) the averaged ratings of the seven interview questions with behavioral anchors, and 4) the averaged ratings of the four questions without behavioral anchors. Data from the 80 subjects were divided into three groups according to the three levels of the independent variable: channel of presentation. All variables were normally distributed except for ratings from two interview questions made in the video mode. These distributions had kurtosis of 1.35 and 1.40 (see Appendix E). However, since subjects were rating three different candidate levels (high, average, and low), using both ends of the 0-7 point interview scale was expected.

Box plots revealed five univariate outliers in the interview question ratings (see Appendix F). These same outliers showed up in the box plots of mean interview rat-

ings, mean behaviorally anchored ratings, and mean unanchored ratings. The outlying ratings were visually examined to determine whether they made sense considering the candidate responses being scored. None of the outlying cases seemed unrealistic, given the responses being rated. It was determined that these ratings were thoughtfully given by the raters, and therefore, retained in the data set.

A factor analysis with principal component extraction was run on the all eleven interview questions to check for underlying factors and unidimensionality (see Appendix G). Kaiser's measure of sampling accuracy = .72. This indicates that the correlation matrix is factorable. Visual examination of the scree plot (see Appendix G) revealed that a line drawn through the points would change direction after the first two points, suggesting that two factors might be optimal to duplicate the correlation matrix. However, since the second factor only explains 14% more of the variance in the variables, and the goal of factor analysis is to summarize the patterns of correlations with the fewest factors, a one factor solution was attempted. This factor accounted for 33% of the variance in the interview ratings. Factor loadings ranged from .42 to .67.

A coefficient alpha was then conducted to estimate the internal consistency of the 11 interview questions across all three communication modes (see Appendix H). An alpha of .79 was computed. Item intercorrelations were all signifi-

cant at the .001 alpha level, with a mean of .26. The highest CORRECTED ITEM TOTAL CORRELATION was .53, and the lowest was .30. All ALPHA IF DELETED were equal or less than the overall alpha of .79. The high internal consistency and significant intercorrelations among the questions also suggest one common factor underlying the interview ratings.

It was hoped that interview ratings would cluster into four factors, representing each of the four interview dimensions; however, the presence of one factor does seem logical. First, the interview questions may be measuring exercise performance; that is, how well candidates answer situational questions. This phenomenon has been found to occur in assessment center ratings where the ratings cluster into factors of the different assessment center exercises instead of the different job dimensions they were intended to measure (Harris, Becker, Smith, 1993). Second, instead of predicting four separate constructs of performance, the situational interview ratings may be predicting one overall job performance construct. It is likely that supervisory skills, interpersonal skills, resourcefulness, and attitude have overlapping variance. For example, both supervisory skills and resourcefulness require assertiveness. Assertiveness is needed to discipline subordinates and to appropriately handle an emergency. Consequently, it seems logical that only one factor could be extracted through factor

analysis. This factor will be considered a predictor of overall job performance and will be correlated with the overall job performance ratings to determine the interview validity.

Bivariate correlations were performed with job performance ratings as the dependent variable and ratings made from videotaped, audiotaped, and transcribed interviews as the independent variable. Resulting r's and their significance are reported in Table 1. Ratings made from audiotaped and transcribed interviews did not explain a significant amount of variance in the job performance ratings. However, ratings made from videotaped interviews explained a significant 23% of the variance in job performance scores.

Table 1.

**CORRELATIONS BETWEEN INTERVIEW RATINGS AND JOB PERFORMANCE RATINGS**

| Channel of Communication | r | $r^2$ | F | p |
|---|---|---|---|---|
| **VIDEO** | .48418 | .23443 | 23.89* | .0000 |
| **AUDIO** | .15865 | .02517 | 2.01 | .1598 |
| **TRANSCRIPT** | .04000 | .00160 | .1202 | .7246 |

Fisher's r to z transformation was computed to compare these correlations. Resulting z's from the correlations are reported in Table 2. The validity of interview ratings made in the video mode was significantly higher than the validity of interview ratings made in the audio mode. The validity of interview ratings made in the video mode was also sig-

nificantly higher than the validity of ratings from the transcription mode. There was no difference in accuracy between interview ratings made from transcripts and interview ratings made from audiotapes.

Table 2.

**FISHER'S r TO z TRANSFORMATIONS COMPARING CORRELATIONS BETWEEN INTERVIEW RATING AND JOB PERFORMANCE**
(p = .05)

| Channel of Communication | VIDEO | AUDIO | TRANSCRIPT |
|---|---|---|---|
| **VIDEO** | **X** | **X** | **X** |
| **AUDIO** | z =2.26*, p=.0119 | **X** | **X** |
| **TRANSCRIPT** | z =3.02*, p=.0013 | z =.76, p=.2236 | **X** |

Hierarchical regression analysis was then performed to determine whether interview ratings made with behavioral anchors, across three modes of communication, could explain variance in job performance over and above the variance explained by ratings made without behavioral anchors. (See Table 3 for R's and change in $R^2$). The change in $R^2$ was significant; $R^2$ =.14067, F=39.04, p=.00.

Table 3.

**HIERARCHICAL REGRESSION PREDICTING JOB PERFORMANCE RATINGS**

| Rating Scale Format | R | $R^2$ | Change in $R^2$ | F | p |
|---|---|---|---|---|---|
| **First Step: Without Behavioral Anchors** | .07308 | .00534 | X | 1.28 | .2594 |
| **Second Step: With Behavioral Anchors** | .38211 | .14601 | .14601 | 39.04* | .0000 |

Bivariate correlations were also computed to compare how well the ratings made with behavioral anchors and without behavioral anchors correlated with job performance within each of the three communication modes. See Tables 4 & 5 for resulting r's. Ratings made with behaviorally anchored ratings scales significantly correlated with job performance in the video mode and audio mode. Ratings made without anchors successfully predicted job performance ratings in the video mode only.

Table 4.

**CORRELATIONS BETWEEN JOB PERFORMANCE RATINGS AND INTERVIEW RATINGS MADE WITH BEHAVIORALLY ANCHORED SCALES**

| Channel of Communication | r | $r^2$ | F | p |
|---|---|---|---|---|
| VIDEO | .51460 | .26481 | 28.09* | .0000 |
| AUDIO | .32167 | .10347 | 9.00* | .0036 |
| TRANSCRIPTION | .11955 | .01429 | 1.13 | .2908 |

Table 5.

**CORRELATIONS BETWEEN JOB PERFORMANCE RATINGS AND INTERVIEW RATINGS MADE WITHOUT BEHAVIORALLY ANCHORED SCALES**

| Channel of Communication | r | $r^2$ | F | p |
|---|---|---|---|---|
| VIDEO | .36631 | .13418 | 12.09* | .0008 |
| AUDIO | .15164 | .02300 | 1.84 | .1794 |
| TRANSCRIPTION | .10749 | .0115 | .91178 | .3426 |

## DISCUSSION

The present study found interview ratings made from videotaped summaries of situational interviews to be significantly correlated with job performance. Interview ratings made from audiotaped and transcribed situational interviews did not significantly correlate with job performance ratings. These findings are contrary to the hypothesis of this study, which predicted that ratings made in the video mode, where raters had access to interviewee nonverbal cues, would not be as valid as ratings made from transcription mode, where nonverbal cues were not accessible. It seems that when mode of communication did not offer nonverbal cues to raters, interview ratings could not accurately predict job performance. This implies that nonverbal cues exhibited by interviewees offer important, relevant information about potential to perform well on the job. In this study, at least, visual cues did not distract from the interview's validity, but instead, enhanced interview validity.

These findings are consistent with results from a recent study conducted by Burnett and Motowidlo (1993). This study quantified interviewee nonverbal cues and then correlated them with interview rating and job performance ratings. The nonverbal cues were significantly correlated with both interview ratings and job performance ratings, also suggesting that nonverbal cues exhibited in the interview are indicative of future job performance.

One explanation of why the only ratings that correlated with job performance were the ratings made in the video mode, where nonverbal cues were presented to raters, may be that nonverbal behaviors are not as easily manipulated as verbal cues. A job candidate may be able to easily alter or fake verbal responses to appear socially desirable; however, it might not be so easy to fake nonverbal behavior; especially, when there are so many facets to nonverbal information to alter (e.g., eye contact, hand gestures, dress, posture). Perhaps ratings that are influenced by a candidate's nonverbal behavior are more valid because nonverbal cues provide a more "pure" or truthful perspective of the candidate.

Another reason why nonverbal cues could be correlated with job performance may be that certain personality traits, related to the interview dimensions being measured, are manifested in the nonverbal cues. That is, certain nonverbal cues exhibited in the interview might reflect underlying personality traits that are related to the patterns of behavior that lead to effective supervisory skills, interpersonal skills, resourcefulness, and attitude (Burnett and Motowidlo, 1994). For example, nonverbal cues such as smiling, leaning forward, and nodding may be associated with the personality trait agreeableness. Accordingly, if agreeableness is an important aspect of effective interpersonal skills, then the nonverbal cues displayed during the inter-

view have provided job relevant information to raters. Past research has demonstrated relationships between certain personality traits and certain nonverbal cues (Albright, Kenny and Malloy, 1988; Borkenau and Liebleer, 1992; Kenny, Horner, Kashy and Chu, 1992). This research has shown nonverbal cues such as eye contact, smiling, hand gestures, posture, physical attractiveness, appropriateness of dress, fashionableness, and voice loudness to be correlated with personality traits such as extroversion, agreeableness, and conscientiousness. Consequently, nonverbal cues may reveal important information about personality traits which are related to the job interview constructs being measured.

Finally, nonverbal cues may indicate how interested and prepared a candidate is in a job. Candidates who are well groomed and dressed may be more serious about the job. Their efforts to physically "prepare" themselves for the interview implies a seriousness, interest, and effort that will be displayed on the job.

It is necessary to address how the episodic/semantic memory distinction, discussed earlier, applies to these results. It was theorized that raters utilize semantic memory when processing written information and episodic memory when processing face-to-face information. The results of this study are not relevant to the memory because subjects in this study were instructed to take notes during the presentations of the interviews. Allowing them to take

notes of the candidates' answers made it unnecessary for them to store and retrieve candidate responses to the interview questions. That is, raters had to encode candidate information; however, they did not have to store or retrieve candidate information. To score responses, raters simply had to review their notes. Previous research studying memory and comprehension as a function of channel found significant differences in comprehension and memory as a function of channel of communication. (Furnham, Gunter, Green, 1990; Furnham & Gunter, 1987; Gunter, Furnham, & Leese, 1986; Furnham, Benson, and Gunter, 1987). However, these studies did not allow subjects to take notes during the presentation of the stimulus.

Consequently, if raters in this study had been required to store and retrieve interviewee responses to make their ratings, their memory of interviewee responses presented in a verbal-only communication channel may have been more accurate than their memory of responses presented in a nonverbal/verbal channel. The results from this study do not conflict with past research which has demonstrated that memory and comprehension are not as strong in the presence of nonverbal cues. However, this study's findings do indicate that when memory (specifically, storage and retrieval of information) is not a factor, raters can successfully process and utilize interviewee information in a nonverbal/verbal channel of communication.

Finally, when the validity of ratings made with behavioral anchors and ratings made without behavioral anchors were compared across the three communication modes, ratings made with anchors explained variance in job performance beyond variance explained by ratings made without anchors. These results demonstrate that including behavioral anchors on rating scales has a positive impact on interview validity.

When interview ratings made with anchors and without anchors were compared to job performance within each communication mode, both types of ratings predicted job performance in the video mode, where nonverbal cues were available to raters. Additionally, ratings made with anchors predicted job performance in the audio mode. However, neither of the two types of ratings was significantly correlated with job performance in the transcription mode.

Tables 4 & 5 clearly illustrate that, in the situation where both behaviorally anchored ratings scales and nonverbal cues were provided to raters, ratings were most correlated with job performance. Conversely, interview ratings were least correlated with job performance in the transcript /unanchored situation, where neither nonverbal cues nor anchored scale were provided to raters. This pattern suggests that both nonverbal cues and behaviorally anchored rating scales are sources of situational interview validity.

It does not seem that ratings made with behavioral

anchors are more valid because of their strict structure, which was intended to restrict raters to only consider verbal content. If this were the case, then anchored ratings in the transcript mode, where only verbal content was presented, would have been able to predict job performance. Instead, it may be that ratings made with behavioral anchors enhance situational interview validity because they guide raters to maximize the usefulness of information obtained.

Limitations

One potential limitation of this study is that job performance ratings were collected from individuals who had access to the candidates' nonverbal information. It may have been that their ratings included variance due to the employees' nonverbal style on the job. If this were the case, then the interview ratings made from video records may have approximated the job performance ratings better than the ratings made from audio or transcription records because the video interview ratings shared systematic error variance with the job performance ratings. However, it is hoped that the evaluations collected from the immediate supervisors of the Junior Cafeteria Managers were free from the effects of nonverbal cues since most of the interactions between supervisor and subordinate are over the telephone.

Another limitation of the results of this study may be the thoroughness of the interview training. Subjects in this study were given only ten minutes of training on how to

use the situational interview. Research has the demonstrated that rater training can influence interview ratings (Latham, Wexley, and Pursell, 1975; Ivancevich, 1979). Although raters, in this study, were instructed to score candidates by matching responses to questions with behavioral anchors (when available), they were not specifically instructed to ignore nonverbal cues. Consequently, if more time and effort had been invested to thoroughly train raters to only attend to the content of candidate answers, it is possible ratings from made from transcripts, which only included verbal content, may have been significantly correlated with job performance.

Finally, the implications about rating scales with behavioral anchors and without behavioral anchors should be considered with certain limitations. First, it may have been that ratings from anchored scales were more correlated with job performance because there were seven questions that had behavior anchors and only four questions without anchors. Because there were more anchored than unanchored rating scales, the averaged ratings from questions with anchors may have been more reliable, and therefore more valid. Second, the questions that were anchored were different questions than the questions that were not anchored. Differences in validity between the two types of questions may have been due to the differences in the questions used, and not to whether questions had behavioral anchors or not.

Recommendations for Future Research

Before deciding on whether or not to train raters to attend to nonverbal cues, more research should be conducted to explore the interactions that nonverbal cues may have with other interview variables. Future research should study the relationship between nonverbal information and verbal information. What happens to interview validity when nonverbal information is not consistent with verbal information? For example, what happens when an interviewee's verbal answer to an interview question on leadership ability implies that he/she is an assertive, confident person, and yet his/her nonverbal cues are those of individual who is nervous and unconfident. Which cues do raters tend to include in their ratings; and which ones are actually related to job performance.

Future research should also investigate the validity contingencies of nonverbal cues. The validity of nonverbal cues is probably contingent on other factors, such as whether the interview dimensions being measured are job knowledge oriented or more construct oriented. For example, do nonverbal cues yield more, or less, job relevant information to raters about Cooking Knowledge, a job knowledge dimension, as opposed to Resourcefulness, a "softer" dimension?

Research should begin to explore how memory is affected by interviewee nonverbal cues. This is a practical question because interviewers in the applied settings do not always

take notes during interviews. More often, raters listen to the full set of interviewee responses to interview questions. Only when the candidate leaves the room do they write down notes and score interviewee responses. It would be interesting to investigate whether memory is affected by the presence of nonverbal cues, such that when nonverbal cues are present, memory of the interview responses is not as accurate. It could be that nonverbal cues exhibited in the employment interview are relevant bits of information that should be taken into account when predicting potential job performance. However, it could also be that when raters must store and retrieve candidate responses to interview questions, the nonverbal information inhibits proper retrieval of interviewee verbal information.

Another area that should be explored is the issue of rater training. Future research should study the effects of nonverbal cues on validity in relation to the amount of rater training provided to raters. Will in-depth rater training cause raters to focus more on verbal content and less on other factors, such as nonverbal cues? If this is so, what will the impact be on interview validity?

## CONCLUSIONS

The results of this study have demonstrated that the validity of situational interview ratings fluctuate as a function of channel of presentation. When interviewee nonverbal cues were made available to raters in a laboratory setting, interview ratings were more valid than when only verbal or transcribed information was offered to raters. These results suggest that nonverbal cues exhibited by candidates in the situational interview can provide raters with relevant information about the candidates' potential to succeed on the job. It seems that nonverbal cues do not suppress interview validity but, instead, enhance interview validity. This study also found that ratings made with behavioral anchors were more valid than ratings made without behavioral anchors. In sum, making interviewee nonverbal cues accessible to raters and including behavioral anchors on the rating scales both had a positive impact on the validity of the situational interview.

APPENDIX A: Junior Cafeteria Manager Fact Sheet

**SALARY:** 9.93-12.38/hour

**JOB DESCRIPTION:** A Junior Cafeteria Manager supervises, participates in and is responsible for the operation of an elementary school meal distribution kitchen, including the proper ordering, receiving, storing, heating, and distribution of meals prepared by the District Nutrition Center (headquarters).

**SOME OF THE KNOWLEDGE, SKILLS, AND ABILITIES NEEDED TO PERFORM THE JOB ARE:**

-knowledge of quantity food preparation
-record keeping skills
-supervisory skills
-dependability
-ability to work effectively with school personnel, students, and parents

**REPORTING RELATIONSHIPS:**

A Junior Cafeteria Manager reports to the Area Food Supervisor and School Principal. A Junior Cafeteria Manager supervises Cafeteria Helpers, Ticket Clerks, and Student Workers.

APPENDIX B: Sample Performance Evaluation Packet

Dear Principal:

The Personnel Selection is requesting your help in examination development. We are currently investigating the effectiveness of selection interviews. We need your help in providing us with information about your Junior Cafeteria Manager's job performance. The information you provide may have significant implications for interview programs already in place, as well as for interviews in general. The Deputy Branch Director of Food Services, has expressed her approval of this study and its goals, and encourages you participate.

Your responsibility in this project will be to provide information about your current or past Junior Cafeteria Manager based on your daily observations of and interactions with him or her. All the information that you will be providing will be used for RESEARCH PURPOSES ONLY. The information will be confidential and will not become part of the employee's personnel record. It will <u>not</u> be used, in <u>any</u> way, to make any decisions about the employee's career. Your evaluation will not have your name on it; therefore, you will remain anonymous.

Please do not show or discuss any of this information with anybody. If you have any questions, please contact me at (213)... Send back the completed questionnaire in the enclosed envelope via school mail. Your promptness will be greatly appreciated.

Sincerely,

Seema Thakur
Personnel Analyst

## Instructions

Before evaluating the employee try to remember specific behaviors/incidents you heard of, encountered, or observed the employee doing. You will be rating your employee on four areas of interest, including Supervisory Skills, Interpersonal Skills, Resourcefulness, and Attitude. For each category, please provide at least one example of your employees past behavior that justifies your rating(s). An example of a behavior to justify a rating made under Resourcefulness (A Junior Cafeteria Manager's ability to deal with any cafeteria related problem or situation effectively) may look like this:

"On Monday, all of the regular meals were bought before the last lunch period started. The Junior Cafeteria Manager quickly heated up back-ups so that all students who ordered served."

The following example would not qualify as a legitimately written example of a behavior:

"The employee completely lacked initiative in getting the job done. Even though there was plenty of opportunity, I couldn't count on him to deliver."

The example above is poorly written because the report mentions a trait (initiative), does not describe the situation in any detail, and is judgmental in nature.

When writing an example be sure to describe the situation in detail. Include the following information in your examples:

1. The circumstances that preceded the incident.
2. The setting in which it occurred.
3. The behavior and what made it effective or ineffective.
4. The consequences of the incident.

The example give above for Resourcefulness contains the appropriate detail of information.

1. The circumstances that preceded the incident: **all of the regular meals were bought before the last lunch period started**

2. The setting: **on Monday**

3. The behavior: **The Junior Cafeteria Manager quickly heated up back-ups**

4. The consequences: **all the students who ordered could be served**

**PLEASE EVALUATE YOUR EMPLOYEE AS HONESTLY AND OBJECTIVE AS POSSIBLE. REMEMBER, YOUR RATINGS WILL BE CONFIDENTIAL.**

***JUNIOR CAFETERIA MANAGER PERFORMANCE EVALUATION**

Employee's Name:_____

_____

How long has the employee worked under your supervision? ___

_____

How long has the employee worked as a Junior Cafeteria Manager?_____

How frequently did you observe this employee's work?

☐ more than once a day ☐ once a day ☐ once every 2-3 work days

☐ once a week ☐ once every two weeks ☐ once every three weeks

☐ once a month

---

Please rate the employee as honestly and objectively as possible on the knowledge, skills, and abilities that are listed under each area of interest. Provide at least one example under each area. **Use the following scale to make your ratings:**

| |
|---|
| **(1) MARGINAL (2) SATISFACTORY (3) GOOD (4) VERY GOOD**<br>**(5) OUTSTANDING (6) NO OPPORTUNITY TO OBSERVE** |

**I. <u>FOOD MANAGEMENT:</u> experience in quantity food preparation, recordkeeping, and food ordering.**

_____ Orders food accurately and promptly

_____ Knowledge of quantity food preparation

_____ Keeps accurate and up to date records

**<u>EXAMPLES</u>:**

| |
|---|
| **(1) MARGINAL (2) SATISFACTORY (3) GOOD (4) VERY GOOD**<br>**(5) OUTSTANDING (6) NO OPPORTUNITY TO OBSERVE** |

**II. SUPERVISORY SKILLS: planing, guiding, directing, delegating, training, evaluating, and disciplining employees**

_____ Ability to instruct other employees

_____ Ability to train and motivate other employees

_____ Ability to discipline employees

**Examples:**

| |
|---|
| (1) MARGINAL (2) SATISFACTORY (3) GOOD (4) VERY GOOD<br>(5) OUTSTANDING (6) NO OPPORTUNITY TO OBSERVE |

**III. INTERPERSONAL SKILL: effectively interacting with employees, other school personnel, teachers, parents, and/or students.**

_____ Establishes and maintains rapport with students and employees

_____ Cooperates with others and works to form harmonious work relationships

_____ Prioritizes group interests above individual inter ests

_____ Ability to deal with complaints from faculty or students

**EXAMPLES:**

| |
|---|
| (1) MARGINAL (2) SATISFACTORY (3) GOOD (4) VERY GOOD<br>(5) OUTSTANDING (6) NO OPPORTUNITY TO OBSERVE |

**IV. RESOURCEFULNESS: dealing with any cafeteria-related problem or situation effectively.**

_____ Ability to deal with any cafeteria related problem

_____ Ability to deal with stressful events, (e.g., short of staff or food)

_____ Ability to create and evaluate alternative solu tions when emergencies arise

_____ Ability to anticipate obstacles and develop plans to overcome them

_____ Ability to set priorities

**EXAMPLES:**

| |
|---|
| (1) MARGINAL (2) SATISFACTORY (3) GOOD (4) VERY GOOD<br>(5) OUTSTANDING (6) NO OPPORTUNITY TO OBSERVE |

**V. ATTITUDE: being reliable and trustworthy, having a positive attitude regarding attendance, punctuality and work.**

_____ Positive attitude regarding attendance, punctuali ty, and work

_____ Ability to be reliable and trustworthy

_____ Flexibility and willingness to change priorities and to have an "open mind"

**Examples:**

**Overall Performance** Please rate the employee's overall performance on a scale from one to five. Please circle only one rating.

**5** Performs job very well in all areas; exceeds standards and expectations for adequate job performance.

**4**

**3** Performs adequately in important areas of the job; meets standards and expectations for adequate job performance.

**2**

**1** Performs poorly in important areas of the job; does not meet standards and expectations for adequate job per formance.

**Comments:**

## APPENDIX C: Instructions for Data Collection

**READ ALOUD TO SUBJECTS:**

Today you will be evaluating the interview performance of 3 job candidates. The job that candidates are competing for is called Junior Cafeteria Manager **(Go over Junior Cafeteria Manager Fact Sheet).**

Job candidates were asked 12 job related questions. The 12 questions were designed to measure the candidates' skills in four job areas: Supervision, Interpersonal Relations, Resourcefulness, and Attitude. Each question asks candidates what they would do in a given situation. **(Allow subjects to review the 12 interview questions for about 5 minutes.)**

Your assignment will be to rate how well the candidates answered the 12 questions. You will classify candidates' responses to each of the interview questions as either as a Strong Response, Acceptable Response, or Poor Response. Strong Responses can receive a score of 5, 6, or 7; Acceptable Responses can receive a 2, 3, or 4; and, Poor responses can receive a score of 0 or 1.

The Response Scoring Guide will help you evaluate candidate responses. For example, the interview question reads, "The kitchen becomes unexpectedly busy one day, and you feel that you are understaffed. What would you do?" You observe an employee stealing food from the storeroom. What would you do?". If a candidate responded something like, "I would work with the employees in the kitchen all day.", he/she should receive anywhere from 5 to 7 points, according to the Response Scoring Guide. Follow the scoring guide to determine whether the answer is Strong, Acceptable, or Poor. However, you may use your judgement to determine how many points to give to the answer; as long as you stay within the point range of the category (Strong, Acceptable, Poor). If a candidate's answer to a question does not clearly fall into one of the three categories, place it in the category that it most closely resembles.

There are four questions that do not have examples of what Strong, Acceptable, and Poor Responses are. These questions are #s 3,6,8, and 12. For these four questions, you should use your own judgement to determine what category and point value to assign it.

You will not be allowed to rate the candidate's responses during the presentation of the interview; but, you make take notes on your scratch paper.

## APPENDIX D: Instructions to Subjects

Today, you will be evaluating the interview performance of 3 job candidates. Specifically, you will be asked to classify candidates' responses to each of twelve interview questions as either as a Strong Response, Acceptable Response, or Poor Response. Strong Responses can receive a score of 5, 6, or 7; Acceptable Responses can receive a 2, 3, or 4; and, Poor responses can receive a score of 0 or 1. The Response Scoring Guide will help you evaluate candidate responses.

The three interviews you will be evaluating today will each be presented in a different form of communication. One will be presented in video form, one in audio form, and one will be presented in written form. Before watching, listening or reading the interviews, please make sure that your Response Scoring Guide and answer sheet are put away. Leave your copy of the interview questions out so you can follow along with the interview. You may take notes during the interview.

After the interview is finished being presented, the proctor will instruct you to begin scoring. Use your notes and the Response Scoring Guide to score the candidate responses to the interview questions. You will be given approximately 10 minutes to score each interview.

Use 1-12 on the answer sheet to score the first job candidate's responses to the twelve questions, 13-24 to score the second candidate's responses, and 25-36 to score the third candidate's responses to the twelve questions.

Please do not hesitate to ask the proctor any questions you may have. Thank you.

APPENDIX E: Descriptives

## AVERAGE INTERVIEW RATINGS (ALL 11 QUESTIONS)

**Audio Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.267 | **Median** | 4.318 | **Mode** | 4.091 |
| **Std dev** | .850 | **Kurtosis** | 1.313 | **S E Kurt** | .532 |
| **Skewness** | -.617 | **S E Skew** | .269 | **Range** | 4.636 |
| **Minimum** | 1.364 | **Maximum** | 6.000 | | |

**Valid cases** 80 **Missing cases** 0

**Video Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.198 | **Median** | 4.455 | **Mode** | 3.727 |
| **Std dev** | 1.272 | **Kurtosis** | .391 | **S E Kurt** | .532 |
| **Skewness** | -.761 | **S E Skew** | .269 | **Range** | 5.909 |
| **Minimum** | .273 | **Maximum** | 6.182 | | |

**Valid cases** 80 **Missing cases** 0

**Transcript Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.297 | **Median** | 4.273 | **Mode** | 4.455 |
| **Std dev** | .986 | **Kurtosis** | -.342 | **S E Kurt** | .532 |
| **Skewness** | .151 | **S E Skew** | .269 | **Range** | 4.273 |
| **Minimum** | 2.091 | **Maximum** | 6.364 | | |

**Valid cases** 80 **Missing cases** 0

## INTERVIEW RATINGS FOR 11 INTERVIEW QUESTIONS
(across all 3 communication modes)

**Supervisory Skills:**

### Interview Rating #1

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.308 | **Median** | 4.000 | **Mode** | 6.000 |
| **Std dev** | 1.801 | **Kurtosis** | -.972 | **S E Kurt** | .313 |
| **Skewness** | -.201 | **S E Skew** | .157 | **Range** | 7.000 |
| **Minimum** | .000 | **Maximum** | 7.000 | | |

**Valid cases** 240 **Missing cases** 0

### Interview Rating #2

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.058 | **Median** | 4.000 | **Mode** | 3.000 |
| **Std dev** | 1.851 | **Kurtosis** | -.801 | **S E Kurt** | .313 |
| **Skewness** | -.054 | **S E Skew** | .157 | **Range** | 7.000 |
| **Minimum** | .000 | **Maximum** | 7.000 | | |

**Valid cases** 240 **Missing cases** 0

**Interpersonal Skills:**

### Interview Rating #3

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.196 | **Median** | 4.000 | **Mode** | 5.000 |
| **Std dev** | 2.058 | **Kurtosis** | -.894 | **S E Kurt** | .313 |
| **Skewness** | -.301 | **S E Skew** | .157 | **Range** | 7.000 |
| **Minimum** | .000 | **Maximum** | 7.000 | | |

**Valid cases** 240 **Missing cases** 0

### Interview Rating #4

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 3.504 | **Median** | 4.000 | **Mode** | 5.000 |
| **Std dev** | 2.004 | **Kurtosis** | -1.094 | **S E Kurt** | .313 |
| **Skewness** | -.160 | **S E Skew** | .157 | **Range** | 7.000 |
| **Minimum** | .000 | **Maximum** | 7.000 | | |

**Valid cases** 240 **Missing cases** 0

## Interview Rating #5

| | | | | | |
|---|---|---|---|---|---|
| Mean | 5.371 | Median | 6.000 | Mode | 6.000 |
| Std dev | 1.503 | Kurtosis | .803 | S E Kurt | .313 |
| Skewness | -1.001 | S E Skew | .157 | Range | 7.000 |
| Minimum | .000 | Maximum | 7.000 | | |

Valid cases 240 Missing cases 0

Attitude:

## Interview Rating #6

| | | | | | |
|---|---|---|---|---|---|
| Mean | 3.592 | Median | 4.000 | Mode | 4.000 |
| Std dev | 1.552 | Kurtosis | -.219 | S E Kurt | .313 |
| Skewness | .027 | S E Skew | .157 | Range | 7.000 |
| Minimum | .000 | Maximum | 7.000 | | |

Valid cases 240 Missing cases 0

## Interview Rating #7

| | | | | | |
|---|---|---|---|---|---|
| Mean | 4.646 | Median | 5.000 | Mode | 5.000 |
| Std dev | 1.708 | Kurtosis | -.189 | S E Kurt | .313 |
| Skewness | -.539 | S E Skew | .157 | Range | 7.000 |
| Minimum | .000 | Maximum | 7.000 | | |

Valid cases 240 Missing cases 0

## Interview Rating #8

| | | | | | |
|---|---|---|---|---|---|
| Mean | 4.125 | Median | 5.000 | Mode | 6.000 |
| Std dev | 2.092 | Kurtosis | -.954 | S E Kurt | .313 |
| Skewness | -.502 | S E Skew | .157 | Range | 7.000 |
| Minimum | .000 | Maximum | 7.000 | | |

Valid cases 240 Missing cases 0

Resourcefulness:

## Interview Rating #9

| | | | | | |
|---|---|---|---|---|---|
| Mean | 3.983 | Median | 4.000 | Mode | 4.000 |
| Std dev | 1.771 | Kurtosis | -.676 | S E Kurt | .313 |
| Skewness | -.225 | S E Skew | .157 | Range | 7.000 |
| Minimum | .000 | Maximum | 7.000 | | |

Valid cases 240 Missing cases 0

## Interview Rating #10

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.642 | **Median** | 5.000 | **Mode** | 6.000 |
| **Std dev** | 1.942 | **Kurtosis** | -.482 | **S E Kurt** | .313 |
| **Skewness** | -.578 | **S E Skew** | .157 | **Range** | 7.000 |
| **Minimum** | .000 | **Maximum** | 7.000 | | |

**Valid cases** 240 **Missing cases** 0

## Interview Rating #11

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.367 | **Median** | 5.000 | **Mode** | 5.000 |
| **Std dev** | 1.716 | **Kurtosis** | -.233 | **S E Kurt** | .313 |
| **Skewness** | -.483 | **S E Skew** | .157 | **Range** | 7.000 |
| **Minimum** | .000 | **Maximum** | 7.000 | | |

**Valid cases** 240 **Missing cases** 0

## INTERVIEW RATINGS MADE WITH BEHAVIORAL ANCHORS

**Audio Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.043 | **Median** | 4.143 | **Mode** | 3.143 |
| **Std dev** | .944 | **Kurtosis** | .219 | **S E Kurt** | .532 |
| **Skewness** | -.368 | **S E Skew** | .269 | **Range** | 4.714 |
| **Minimum** | 1.286 | **Maximum** | 6.000 | | |

* Multiple modes exist. The smallest value is shown.

| | | | |
|---|---|---|---|
| **Valid cases** | 80 | **Missing cases** | 0 |

**Video Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.045 | **Median** | 4.286 | **Mode** | 3.857 |
| **Std dev** | 1.317 | **Kurtosis** | -.039 | **S E Kurt** | .532 |
| **Skewness** | -.538 | **S E Skew** | .269 | **Range** | 6.000 |
| **Minimum** | .429 | **Maximum** | 6.429 | | |

* Multiple modes exist. The smallest value is shown.

| | | | |
|---|---|---|---|
| **Valid cases** | 80 | **Missing cases** | 0 |

**Transcript Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.063 | **Median** | 4.000 | **Mode** | 3.143 |
| **Std dev** | 1.073 | **Kurtosis** | -.579 | **S E Kurt** | .532 |
| **Skewness** | .338 | **S E Skew** | .269 | **Range** | 4.429 |
| **Minimum** | 2.000 | **Maximum** | 6.429 | | |

| | | | |
|---|---|---|---|
| **Valid cases** | 80 | **Missing cases** | 0 |

## INTERVIEW RATINGS MADE WITHOUT BEHAVIORAL ANCHORS

**Audio Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.659 | **Median** | 4.500 | **Mode** | 4.250 |
| **Std dev** | 1.059 | **Kurtosis** | .151 | **S E Kurt** | .532 |
| **Skewness** | -.234 | **S E Skew** | .269 | **Range** | 5.250 |
| **Minimum** | 1.500 | **Maximum** | 6.750 | | |

* Multiple modes exist. The smallest value is shown.

**Valid cases** 80 **Missing cases** 0

**Video Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.466 | **Median** | 4.625 | **Mode** | 5.250 |
| **Std dev** | 1.385 | **Kurtosis** | .690 | **S E Kurt** | .532 |
| **Skewness** | -.880 | **S E Skew** | .269 | **Range** | 6.750 |
| **Minimum** | .000 | **Maximum** | 6.750 | | |

**Valid cases** 80 **Missing cases** 0

**Transcript Mode:**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.706 | **Median** | 4.750 | **Mode** | 4.250 |
| **Std dev** | 1.080 | **Kurtosis** | -.310 | **S E Kurt** | .532 |
| **Skewness** | -.177 | **S E Skew** | .269 | **Range** | 4.750 |
| **Minimum** | 2.250 | **Maximum** | 7.000 | | |

**Valid cases** 80 **Missing cases** 0

**INTERVIEW RATINGS MADE WITH BEHAVIORAL ANCHORS**
**(across all 3 communication modes)**

| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.050 | **Median** | 4.143 | **Mode** | 4.286 |
| **Std dev** | 1.117 | **Kurtosis** | .052 | **S E Kurt** | .313 |
| **Skewness** | -.262 | **S E Skew** | .157 | **Range** | 6.000 |
| **Minimum** | .429 | **Maximum** | 6.429 | | |

*Multiple modes exist. The smallest value is shown.

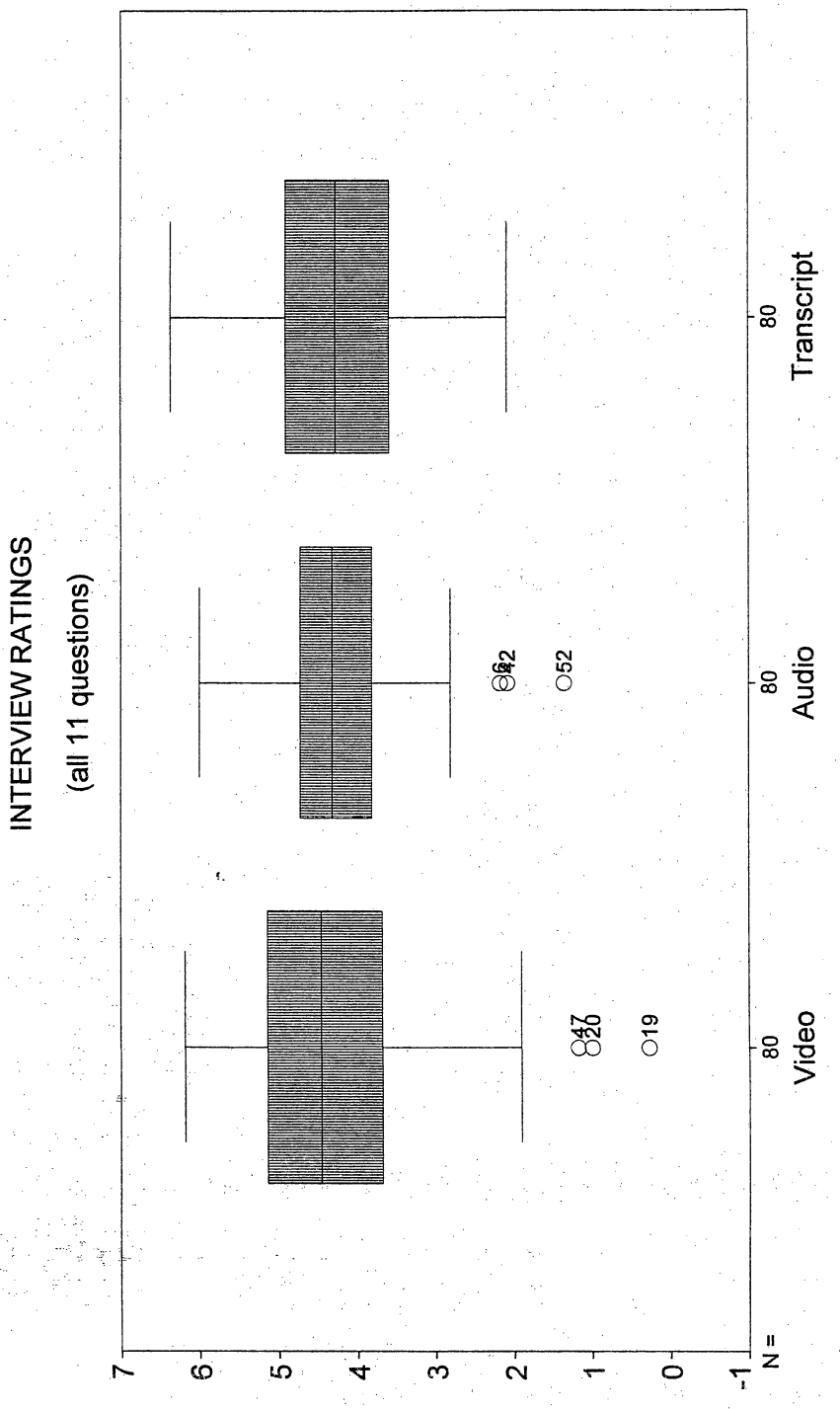
| | | | |
|---|---|---|---|
| **Valid cases** | 240 | **Missing cases** | 0 |

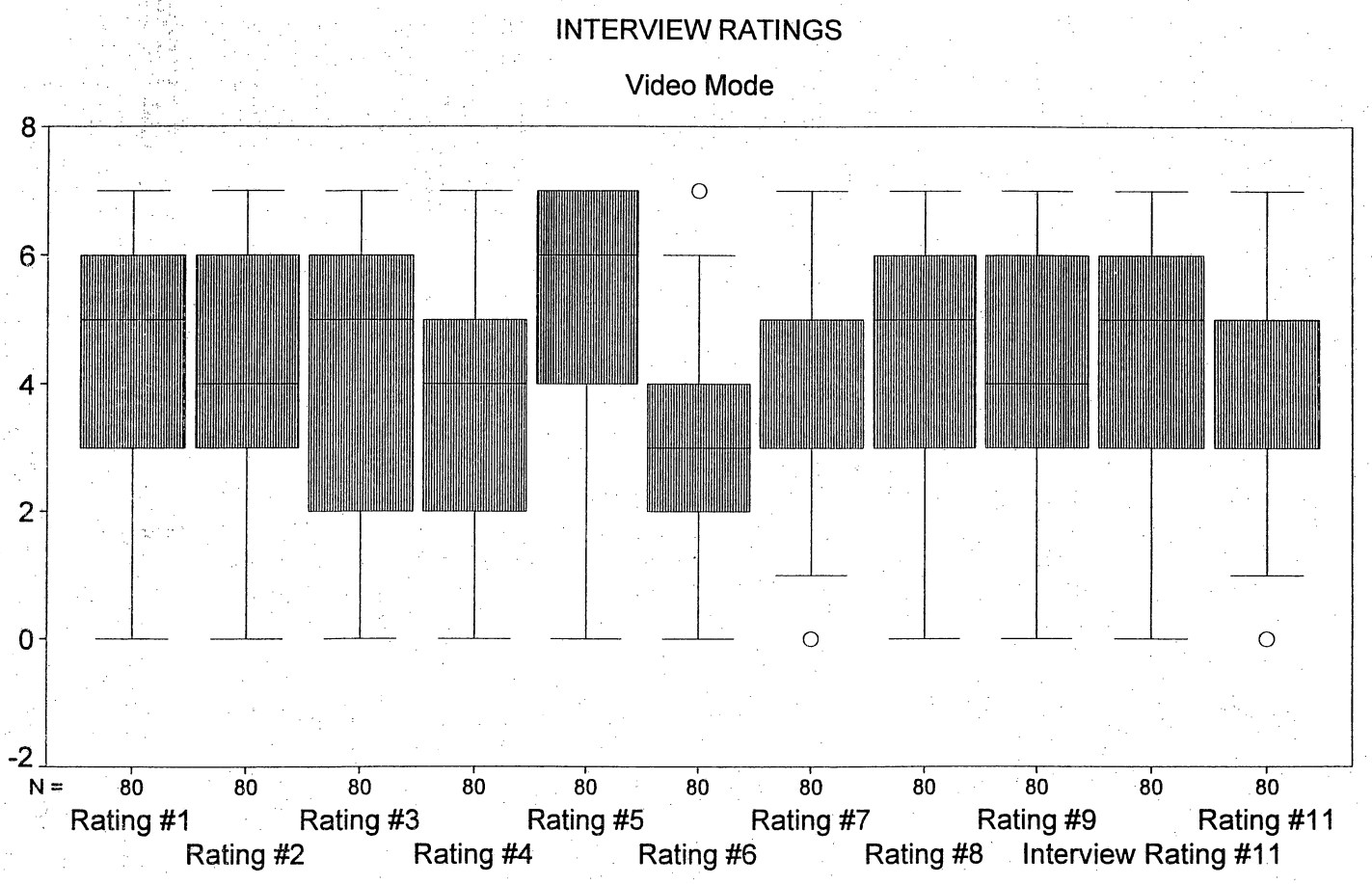
**INTERVIEW RATINGS MADE WITHOUT BEHAVIORAL ANCHORS**
**across all 3 communication modes)**

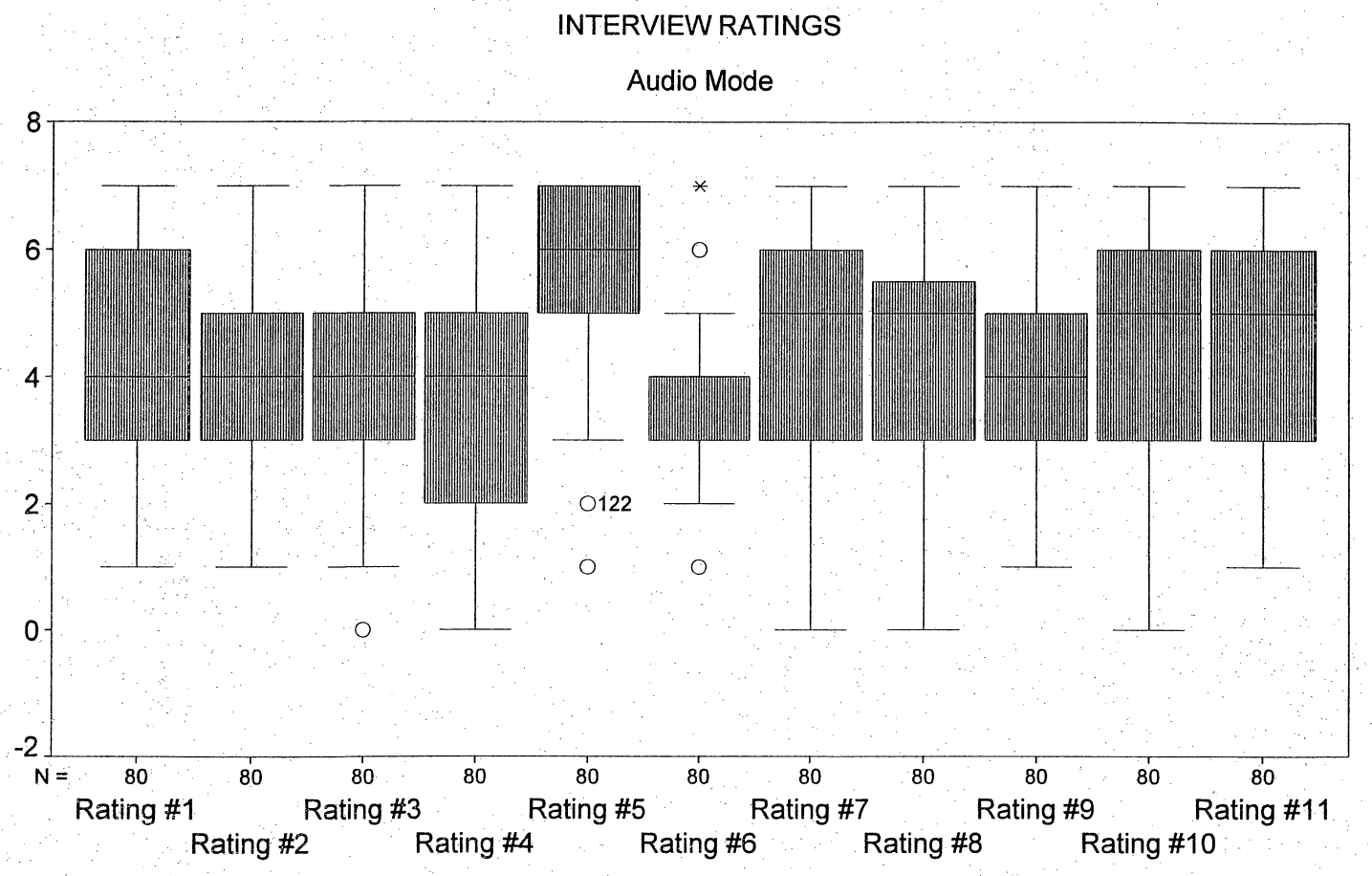
| | | | | | |
|---|---|---|---|---|---|
| **Mean** | 4.610 | **Median** | 4.750 | **Mode** | 4.250 |
| **Std dev** | 1.184 | **Kurtosis** | .734 | **S E Kurt** | .313 |
| **Skewness** | -.624 | **S E Skew** | .157 | **Range** | 7.000 |
| **Minimum** | .000 | **Maximum** | 7.000 | | |

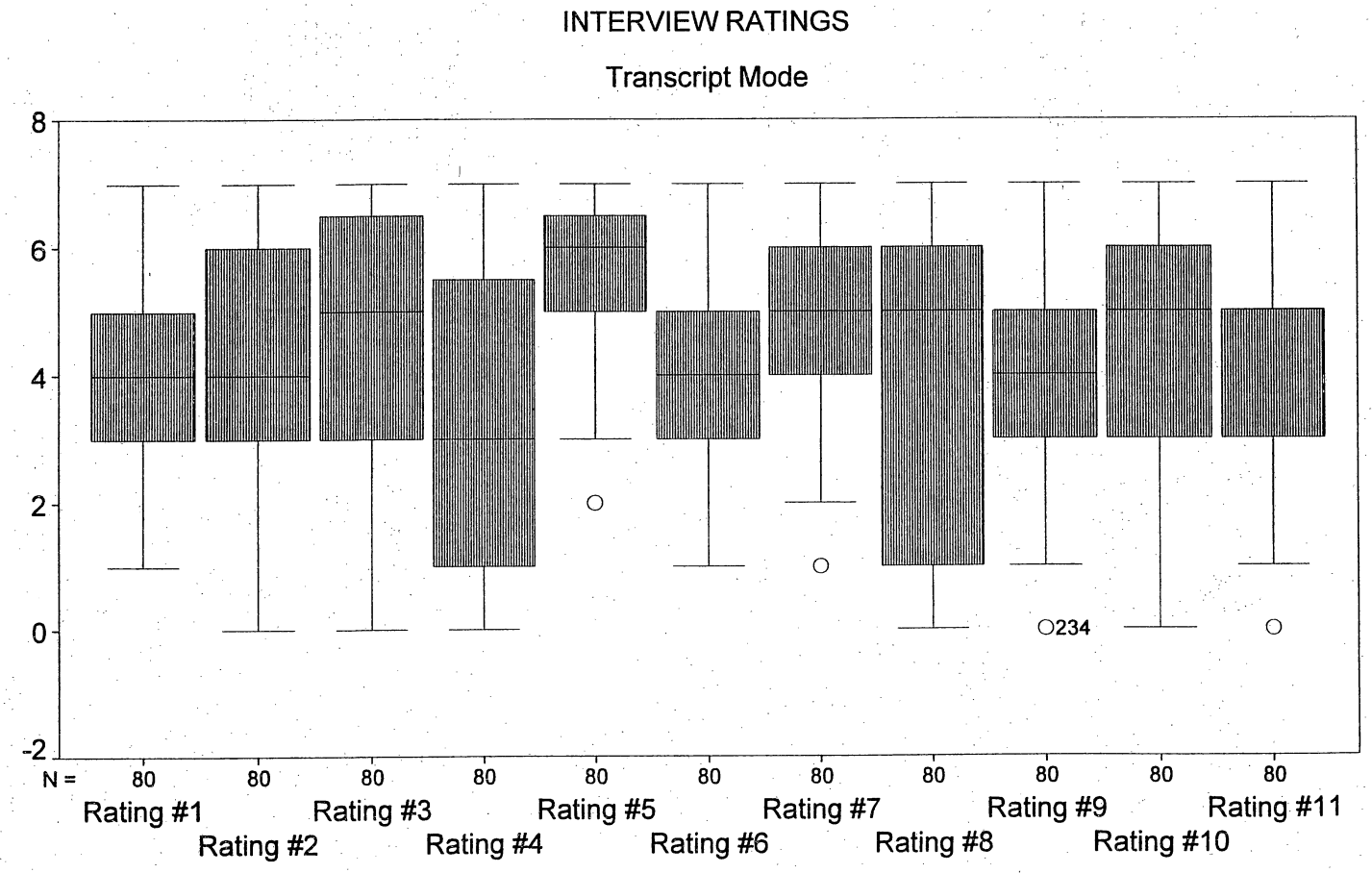
**Valid cases** 240 **Missing cases** 0

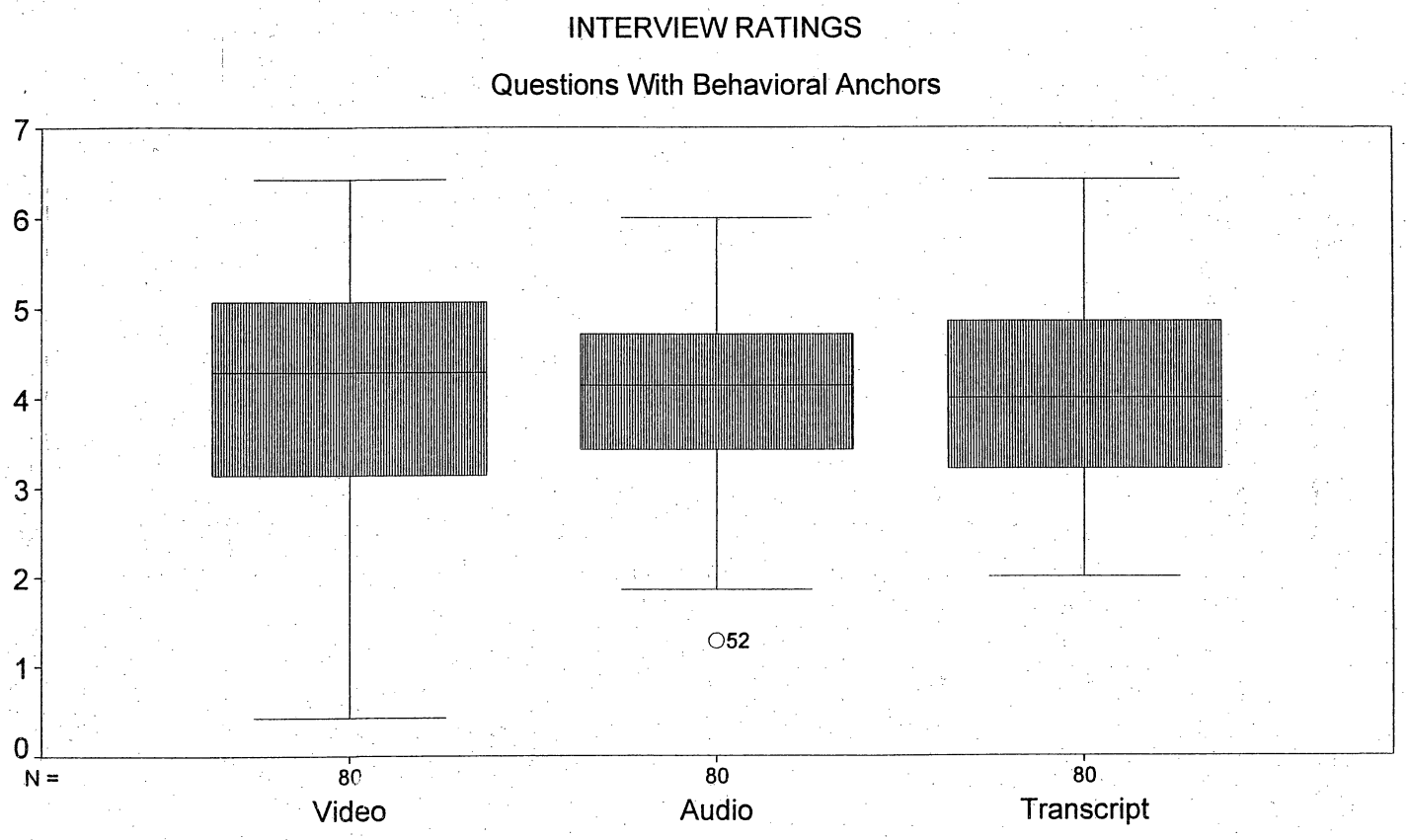
APPENDIX F: Graphs

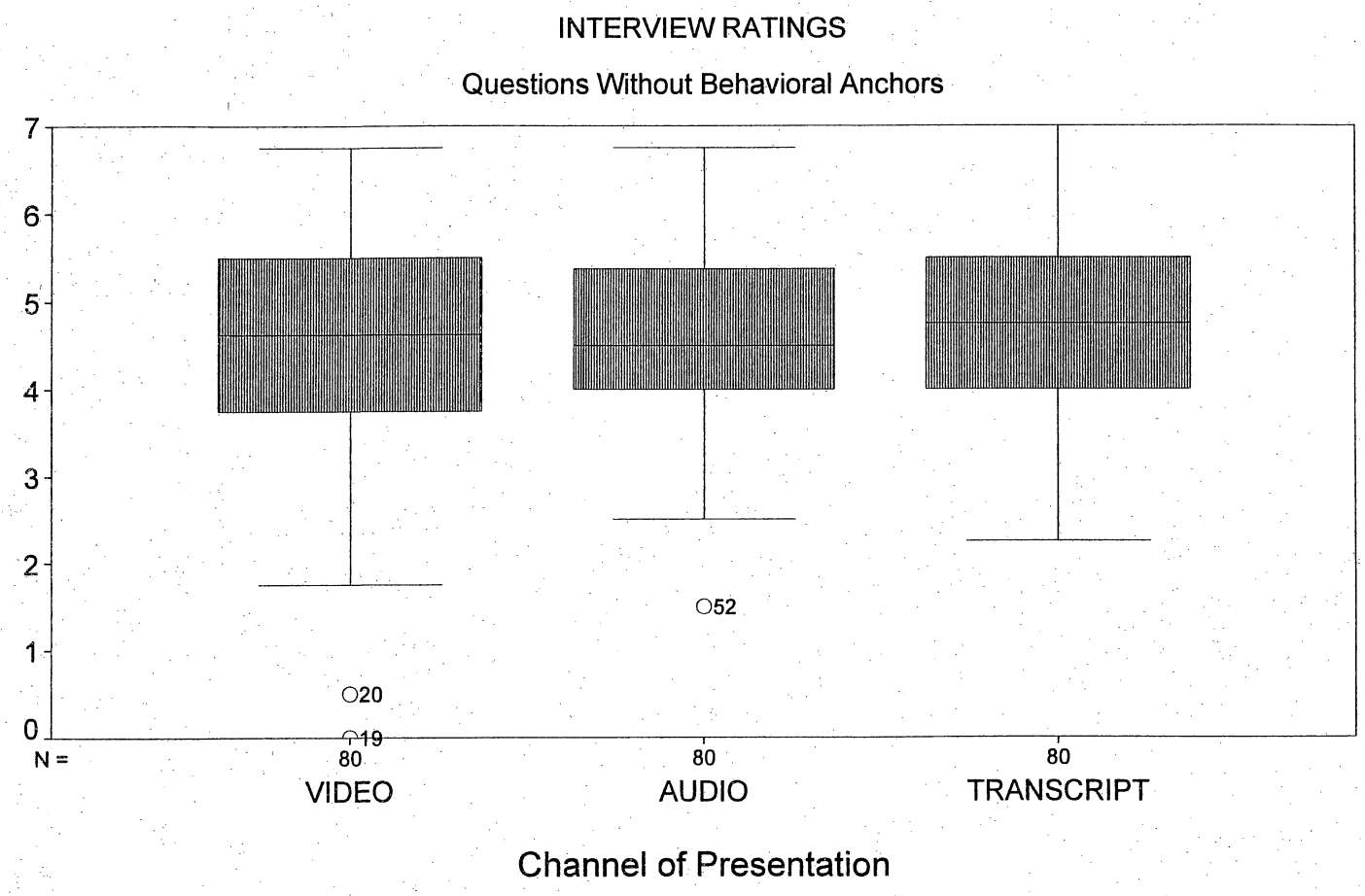



INTERVIEW RATINGS
Video Mode
8
6
4
2
0
-2
N =
80
Rating #1
Rating #2
Rating #3
Rating #4
Rating #5
Rating #6
Rating #7
Rating #8
Rating #9
Interview Rating #11
Rating #11


INTERVIEW RATINGS
Audio Mode
8
6
4
2
0
-2
122
N =
80
80
80
80
80
80
80
80
80
80
80
Rating #1
Rating #2
Rating #3
Rating #4
Rating #5
Rating #6
Rating #7
Rating #8
Rating #9
Rating #10
Rating #11


INTERVIEW RATINGS
Transcript Mode
8
6
4
2
0
-2
234
N =
80
Rating #1
Rating #2
Rating #3
Rating #4
Rating #5
Rating #6
Rating #7
Rating #8
Rating #9
Rating #10
Rating #11


INTERVIEW RATINGS
Questions With Behavioral Anchors
7
6
5
4
3
2
1
0
52
N =
80
80
80
Video
Audio
Transcript
Channel of Presentation


INTERVIEW RATINGS
Questions Without Behavioral Anchors
7
6
5
4
3
2
1
0
52
20
19
N =
80
80
80
VIDEO
AUDIO
TRANSCRIPT
Channel of Presentation

## APPENDIX G: Factor Analysis

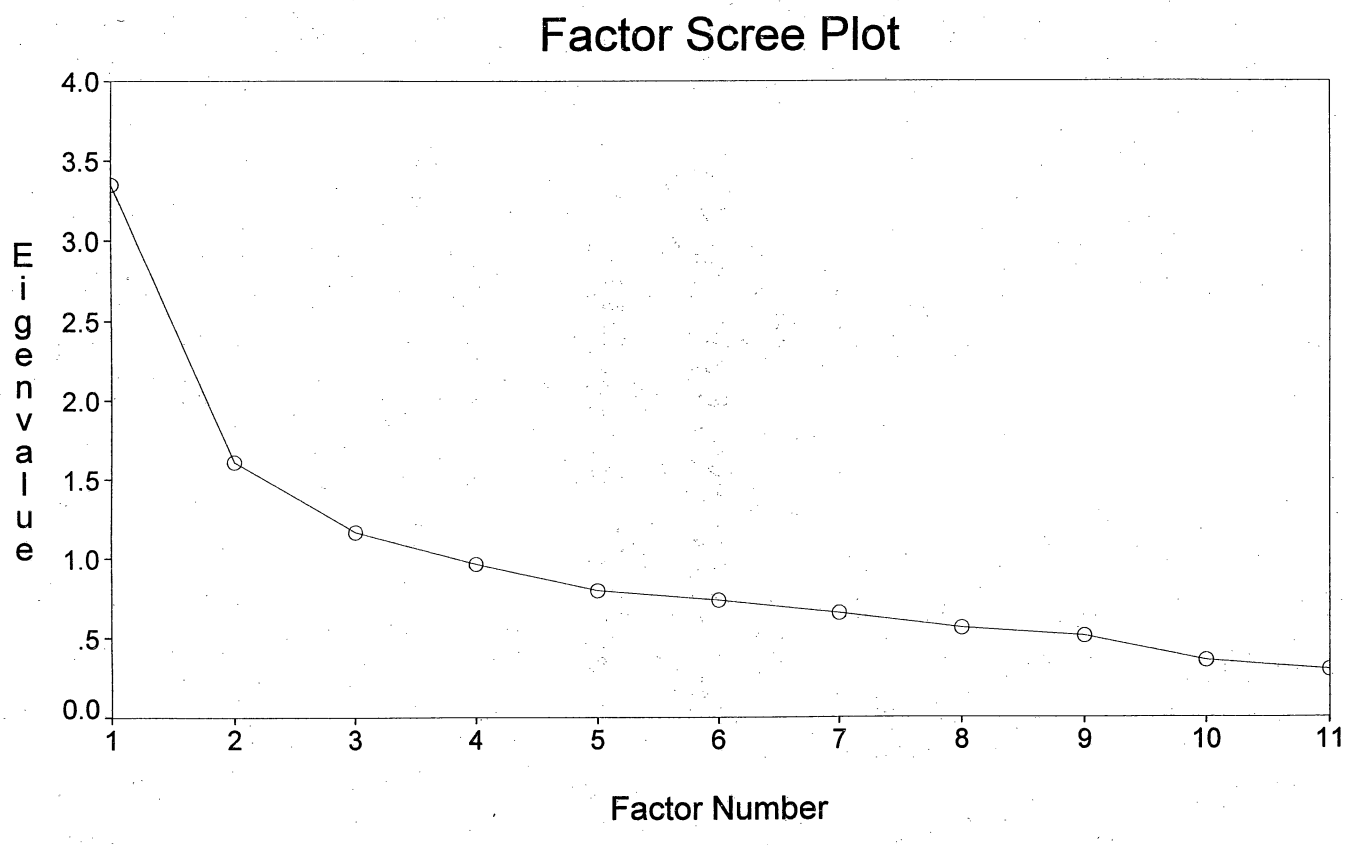
Kaiser-Meyer-Olkin Measure of Sampling Adequacy = .75619

**Factor Matrix:**

| | Factor 1 |
|---|---|
| Rating11 | .66489 |
| Rating5 | .65344 |
| Rating2 | .62884 |
| Rating6 | .60107 |
| Rating9 | .59897 |
| Rating4 | .59410 |
| Rating3 | .58595 |
| Rating1 | .55433 |
| Rating10 | .50441 |
| Rating7 | .48655 |
| Rating8 | .42116 |

**Final Statistics:**

| Variable | Communality | | Factor | Eigenvalue | Pct of Var |
|---|---|---|---|---|---|
| Rating1 | .30728 | * | 1 | 3.65665 | 33.2 |
| Rating2 | .39544 | * | | | |
| Rating3 | .34333 | * | | | |
| Rating4 | .35295 | * | | | |
| Rating5 | .42698 | * | | | |
| Rating6 | .36129 | * | | | |
| Rating7 | .23673 | * | | | |
| Rating8 | .17737 | * | | | |
| Rating9 | .35876 | * | | | |
| Rating10 | .25442 | * | | | |
| Rating11 | .44208 | * | | | |


Factor Scree Plot
Eigenvalue
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5
0.0
1
2
3
4
5
6
7
8
9
10
11
Factor Number

APPENDIX H: Reliability Analysis

| Statistics for Scale | Mean | Variance | Std Dev | No of Variables |
|---|---|---|---|---|
| | 46.9247 | 128.9943 | 11.3576 | 11 |

| Item Means | Mean | Minimum | Maximum | Range | Max/Min | Variance |
|---|---|---|---|---|---|---|
| | 4.2659 | 3.5146 | 5.3891 | 1.8745 | 1.5333 | .2733 |

| Inter-item Correlations | Mean | Minimum | Maximum | Range | Max/Min | Variance |
|---|---|---|---|---|---|---|
| | 609 | -.2436 | .5071 | .7507 | -2.0816 | .0150 |

**Item-total Statistics**

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Alpha if Item Deleted |
|---|---|---|---|---|---|
| Rating1 | 42.6025 | 109.5346 | .4332 | .2577 | .7738 |
| Rating2 | 42.8619 | 105.8422 | .5168 | .3297 | .7644 |
| Rating3 | 42.7155 | 105.4145 | .4599 | .4309 | .7711 |
| Rating4 | 43.4100 | 105.1085 | .4843 | .3652 | .7680 |
| Rating5 | 41.5356 | 110.2834 | .5318 | .3438 | .7658 |
| Rating6 | 43.3264 | 111.1199 | .4726 | .3239 | .7706 |
| Rating7 | 42.2636 | 113.1697 | .3592 | .3094 | .7813 |
| Rating8 | 42.7866 | 111.3618 | .3015 | .4815 | .7908 |
| Rating9 | 42.9331 | 107.7182 | .4937 | .3484 | .7673 |
| Rating10 | 42.2678 | 110.7515 | .3570 | .4618 | .7827 |
| Rating11 | 42.5439 | 107.1063 | .5376 | .3313 | .7630 |

Reliability Coefficients 11 items

**Alpha = .7891** Standardized item alpha = .7952

## REFERENCES


Albright, L., Kemy, D.A., Malloy, T.A. (1988). Consensus in personality judgments of zero acquaintance. Journal of Personality and Social Psychology, 55, 387-395.

Arvey, R.D. and Campion, J.E. (1982). The employment interview: a summary and review of recent research. Personnel Psychology, 35, 281-323.

Bardack, N.R. and McAndrew, F.T. (1985). The influence of physical attractiveness and manner of dress on success in a simulated personnel decision. The Journal of Social Psychology, 165 (6), 777-778.

Baron, R.A. (1983). "Sweet smell of success"? The impact of pleasant artifical scents on evaluations of job applicants. Journal of Applied Psychology, 68, 709-713.

Borkenau, P., and Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. Journal of Personality and Social Psychology, 62, 645-657.

Burnett, J.R., and Motowidlo, S.J. (1993). Utilization and validity of nonverbal cues in the structured interview. Dissertation.

Eder, R.W. and Ferris, G.R. (1989). The employment interview: Theory, research, and practice. Sage Publications. P169-182.

Forsythe, S., Drake, M.F. and Cox, C.E. (1985). Influence of applicant's dress on interviewer's selection decisions. Journal of Applied Psychology, 70, 2, 374-378.

Furnham, A., and Gunter, B. (1987). Effects of time of day and medium of presentation on immediate recall of violent and nonviolent news. Applied Cognitive Psychology, 1, 255-262.

Furnham, A., Benson, I. and Gunter, B. (1987). Memory for television commercials as a function of the channel of communication. Social Behaviour, 2, 105-112.

Furnham, A. Gunter, B. and Green, A. (1990). Remembering science: The recall of factual information as a function of presentation mode. Applied Cognitive Psychology, 4, 203-212.

Gifford, R., Ng, C.F., and Wilkinson, M. (1985). Nonverbal cues in the employment interview: Links between qualities and interviewer judgments. Journal of Applied Psychology, 4, 729-736.

Gunter, B., Furnham, A. and Leese, J. (1986). Memory of information from a party political broadcast as a function of the channel of communication. Social Behaviour, 1, 135-142.

Hellman, M.E., Saruwatari, L.R. (1979). When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and nonmanagerial jobs. Organizational Behavior and Human Performance, 23, 360-372.

Imada, A.S. and Hakel, M.D. (1977). Influence of nonverbal communication and rater proximity on impressions and decisions in simulated employment interviews. Journal of Applied Psychology, 62, 3, 295-300.

Ivancivich, J.M. (1979). Longitudinal study of the effects of rater training on psychometric error in ratings. Journal of Applied Psychology, 64, 502-508.

Kenny, D.A., Horner, C., Kashy, D.A., and Chu, L. (1992). Consensus of Zero acquaintance: Replication, behavioral cues, and stability. Journal of Personality and Social Psychology. 62, 88-97.

Latham, G.P., Saari, L.M., Pursell, E.D. and Campion, M.A. (1980). The situational interview. Journal of Applied Psychology. 65 (4), 422-427.

Latham, G.P. and Saari, L.M. (1984). Do People do what they say? Further studies on the situational interview. Journal of Applied Psychology. 65 (4), 569-573.

Latham, G.P., Wexley, K.N., and Pursell, F.D. (1975) Training managers to minimize rating errors in the observation of behavior. Journal of Applied Psychology, 70, 550-555.

Lin, T.R., Dobbins, G.H. and Farh, J. (1992). A field study of race and age similarity effects on interview ratings in conventional and situational interviews. Journal of Applied Psychology, 77 (3), 363-377.

Mair, N.R.F. and Thurber, J.A. (1968). Accuracy of judgments of deception when an interview is watched, heard, and read. Personnel Psychology, 21, 23-30.

Motowidlo, S.J., Carter, G.W., Dunnette, M.D., Tippins, N., Werner, S., Burnett, J.R., Vaughan, M.J. (1992) Studies of the structured behavioral interview. Journal of Applied Psychology, 77 (5), 571-587.

Murphy, K.R. and Cleavland, J.N., (1991). Performance appraisal: An organizational perspective. Allyn and Bacon.

Murphy, K.R., Herr, M.B., Lockhard, M.C., Maguire, E., (1986). Evaluating the performance of paper people. Journal of Applied Psychology, 71, 2, 654-661.

Rasmussen Jr., K.G. (1984). Nonverbal behavior, verbal behavior, resume credentials and selection interview outcomes. Journal of Applied Psychology, 69, 551-556.

Tulving, E. and Thompson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 5, 352-373.

Washburn, P.V. and Hakel, M.D. (1973). Visual cues and verbal content as influenced on impressions formed after simulated employment interviews. Journal of Applied Psychology, 58 (1), 137-141.

Weekley, J.A. and Gier, J.A. (1987). Reliability and validity of the situational interview for a sales position. Journal of Applied Psychology, 72(3) 484-487.

Wexley, K.N., Fugita, S.S. and Malone, M.P. (1975). An applicants nonverbal behavior and student evaluators' judgements in a structured interview setting. Psychological Reports, 36, 391-394.