# THE SOCIAL CONTEXT OF ACCOUNTABILITY: EFFECTS OF RATERS' EXPECTATIONS OF A SUPERVISORY REVIEW

by

Diana K. Corrigan

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

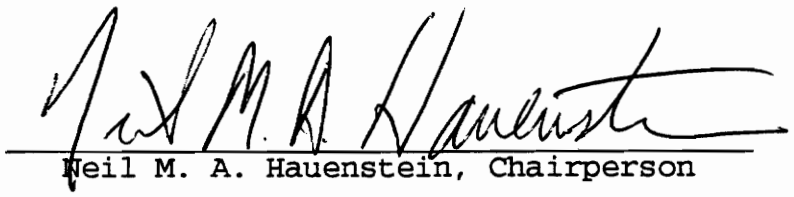
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

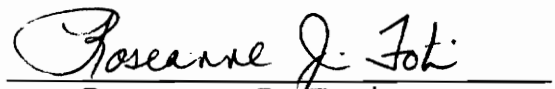
in

Psychology

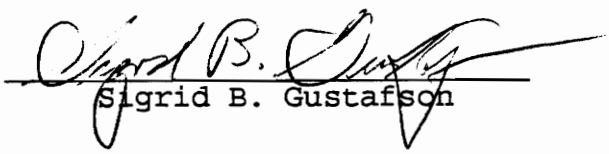
APPROVED:



Neil M. A. Hauenstein, Chairperson



Roseanne J. Foti



Sigrid B. Gustafson

January, 1994

Blacksburg, Virginia

THE SOCIAL CONTEXT OF ACCOUNTABILITY: EFFECTS OF RATERS' EXPECTATIONS OF A SUPERVISORY REVIEW

by

Diana K. Corrigan

Committee Chairperson: Neil M. A. Hauenstein

Department of Psychology

(Abstract)

Being held accountable for one's views in an ambiguous situation has been found to increase cognitive processing and reduce possible biases often pervasive in performance appraisals (Tetlock, 1983a). The present study examined the effects of making raters more accountable for their performance ratings through expectations of a supervisory review. Two types of accountability were manipulated. Interpersonal accountability induced subjects to believe that they would receive feedback based on their justifications of their ratings. Reward accountability subjects expected that they would receive a reward based on the quality of their ratings compared to true performance scores. Half of the subjects were also told that the supervisor gave the ratee a negative performance rating. 191 female subjects rated an interviewer's performance after receiving accountability and view manipulations. Results indicated that subjects who received accountability manipulations in conjunction with the view gave more unfavorable ratings and were less accurate on

Cronbach's (1955) elevation component of accuracy than no accountability and/or no view groups. Accountable subjects who did not receive the view were expected to increase cognitive processing. These hypotheses pertaining to increases in cognitive complexity and dimensional accuracy were not supported. Although unexpected, interpersonal accountability in conjunction with reward accountability prevented a decrement in processing when given the view. A biased versus unbiased vigilant processing hypothesis was proposed to explain these results. Possible limitations are also proposed as well as implications for future research.

Acknowledgements

I would like to sincerely thank several individuals whose support guided my research endeavors. Primarily, I would like to express sincere gratitude to Dr. Neil Hauenstein, the chair of this project, whose support, encouragement, and guidance provided me with the ability to plunge into research. I would also like to thank the other members of my committee, Dr. Roseanne Foti and Dr. Sigrid Gustafson, who were both unselfish with their time and whose questions encouraged me to think more critically.

I wouldn't have been able to finish this project without the friendship of many people I have met in Blacksburg. Specifically, I'd like to thank Dean Stamoulis, who not only provided his dissertation as the backbone for this thesis, but also time and assistance whenever I found it necessary to call on him. I appreciate his insight and confidence in making this a high quality study. I also want to thank Sharon Flinder, who continues to help keep things in perspective and whose friendship made this process as painless as possible. The other three fourths of the "Mill Mountain Crew" also deserve appreciation, for providing the necessary "positive reinforcement" throughout my entire time here at Virginia Tech. You know who you are. And lastly, I'd like to thank Gayle Kennedy and my research assistants, Chris Mattis, Wendy Privette, and Mandy McMurray for their

time and efforts in helping me with the data collection. I could not have completed this project without them.

I would like to recognize and thank my family for their continual support, especially my parents, J. Alan and Peggy Corrigan, who instilled in me the idea that I could achieve anything I wanted to in my career, and in my life. My sister, Maura, also requires much thanks for always listening to my complaints and being what only a sister could be, my best friend.

Table of Contents

Tables

Figures

## Introduction

Performance appraisal refers to the process by which each employee's performance is rated by an observer, usually the employee's supervisor. The ratings of each individual and the decisions that stem from those ratings can have a great impact on organizational effectiveness. Based on the results of performance evaluations, decisions are made in which employees are promoted, receive pay increases, chosen for training, demoted, or even fired. Because performance appraisals lead to decisions that have a wide effect on organizational practices, examining the process of performance appraisal is important to understand how raters make judgments from observations of behavior.

The rating process occurs within a context that has implications for how each individual is appraised. While the organization desires accuracy of the performance appraisal process, many factors can constrain the rater's ability to observe and evaluate the ratee. Motivational concerns of the rater also provide a context which often alters the evaluation from the "true" performance of the appraisee. Humans are often viewed as "cognitive misers" who place little effort into making decisions, thus errors are often made or biases interfere with accurate evaluations. If the organization could provide a context which minimizes such biases or evaluation errors, accurate performance evaluations are more likely to result. The following study proposes the

social context of accountability as one factor that has the potential to increase the rater's information processing efforts so accurate appraisals will be more likely to occur.

Several performance appraisal models have been proposed that place emphasis on the process of social cognitions in the appraisal rather than on rating outcomes (DeNisi, Cafferty, & Meglino, 1984; Ilgen & Feldman, 1983; Landy & Farr, 1983). More specifically, Ilgen and Feldman (1983) propose that the performance appraisal process involves three interacting systems which include the rater's information processing system, the organizational context, and the behavior of the ratee. They argue that factors from each system can introduce biases which constrain accurate appraisals. Much of the research on the rater's information processing system is based on Srull and Wyer's (1989) model of social cognition which explains how individuals form impressions of others. Although Srull and Wyer's (1989) model is not presented in the context of performance appraisal, the formation of impressions is implicit in the process and thus important in determining what information is observed by the rater.

Srull and Wyer (1989) suggest that subject's attempt to interpret a target's behavior in terms of a general trait concept. If expectations are available concerning the target's traits, then they will be used as the basis for

interpretations. If no expectations are evident, the individual will interpret behaviors according to the trait concept that most easily comes to mind. Subjects form a general evaluative conception of the person based on the trait concepts and interpret subsequent behaviors in terms of this conception. Srull and Wyer's (1989) model implies that situational aspects have an effect on which trait concepts are most readily available, which further has an effect on encoding of behaviors and the judgments that are made. The concept that most easily comes to mind often guides judgment without using the implications of specific observable behaviors to make a decision. Different decision making environments that individuals confront provoke different trait concepts, thus having implications for subsequent judgments.

Although Srull and Wyer's (1989) model provides important information about how individuals form impressions, their research does not directly take into account implications of differing organizational contexts that Ilgen and Feldman (1983) referred to as an integral part of the process. Srull and Wyer (1989) include information about how information is received, yet they do not specifically describe what occurs when the rater actively pursues information concerning particular objectives. DeNisi et al. (1984) propose a model which adds two important contributions

to previous social cognitive literature. Their model acknowledges that the rater is not a passive bystander, but the rater is an active seeker of the information required to complete evaluations. The second important feature provides the context of purpose that not only has the potential to motivate the rater to make more accurate appraisals, but also can determine what types of information are sought and encoded.

Because the rater is considered an active participant in the process, information is often sought in reference to prior expectations or biases which DeNisi et al. (1984) describe as preconceived notions. These preconceived notions can stem from earlier interactions with the ratee, prior evaluations, or information about the ratee obtained from others. These expectations provide a schema for interpreting incoming stimuli. As suggested by Srull and Wyer (1989), if no expectations are available, raters interpret behaviors in terms of the most applicable trait concepts that come to mind. The ambiguity and thus the availability of these expectations or preconceived notions provides for an important contextual factor of the appraisal process that could influence ratings.

Ambiguous situations (March & Olsen, 1979) are often confronted by organizational decision makers. The appraiser might be rating an individual for the first time and/or no

information is presented to the rater from other sources to aid in making an evaluation. The ambiguity of the situation has been shown to improve information processing. Tetlock (1983a) examined the manner in which subjects reacted when confronted with ambiguous situations. Subjects either knew in advance the opinion of the individual they were to confront or were in an ambiguous situation of not knowing the view of the individual. He found that subjects in situations of high ambiguity were more motivated to examine and evaluate information thoroughly and to consider arguments on both sides of an issue in order to prepare themselves for critical reactions about their decision.

Although ambiguity has been found to increase information processing, Porter, Allen, and Angle (1981) suggest that individuals may avoid or ignore certain information in ambiguous situations because those situations provide for more opportunities to engage in upward influence to promote self-interest. The unambiguous situation of knowing the opinion of the direct supervisor decreases the chance of manipulating information that the supervisor is already aware of. Although the effects of being in an ambiguous situation seem conflicting, these two interpretations may be reconciled.

Fandt and Ferris (1990) found that situations of low ambiguity influenced subjects to use defensive information

and concentrate on more positive aspects of their decision when justifying their decision. Although subjects selectively used certain information to justify their positions, they initially increased their information processing strategies when confronted with an ambiguous situation. The ambiguity of the situation may affect the encoding of information before an evaluation is made because the uncertainty involved increases the amount of information that is attended to. Once the evaluation is constructed, the ambiguity then might influence subjects to construct arguments to present themselves as rational and competent decision makers. Thus although ambiguity might influence a rater to examine more pieces of information, the individual still might use selective communication of that information. Although the research evidence has not resolved the effects that ambiguity brings to the situation, ambiguity is seen as an important contextual variable.

If a rater knows the opinion of upper-level management (e.g. his/her supervisor) about a prospective ratee's performance, an active rater is likely to use such information as one method of extracting performance cues from the environment and thus decreasing the ambiguity of the situation. DeNisi et al. (1984) included another contextual factor which has the potential to influence the cognitive and motivational aspects of the rater to seek certain performance

cues. They cite the purpose of the appraisal process as an important variable that may influence the selective attention, encoding, storage, and use of information. Landy and Farr (1983) also include purpose in their model which suggests that knowledge of the intended use of the ratings represents the "motivation to rate" by the active appraiser. They suggest that the organizational purpose of the appraisal does not always match the purpose of the rater so that the individual might not have the appropriate "motivation to rate". The rater often has hidden agendas indicating that purpose should be examined as perceived by the rater.

Numerous researchers have cited instances in which the overall purpose of the rating task has had important implications for what information is reported about the ratee. Tesser and Rosen (1975) found that subjects often distort or ignore negative information because of the reluctance to receive a negative reaction from giving a bad rating. Raters may be more lenient to avoid negative reaction from ratees, but they also may be more lenient under conditions where the purpose of appraisals are for administrative use rather than for research or feedback. Murphy, Balzer, Kellam, and Armstrong (1984) found a leniency bias in which observation of specific behaviors and performance evaluation ratings were less associated when the purpose was for personnel decisions versus for research.

Appraisers may give favorable evaluations to subordinates to whose performance they have contributed in order to enhance themselves (Baumeister, 1982). Raters might also rate leniently to embark confidence in their employees in hopes of motivating future performance (Villanova & Bernardin, 1989). Raters may also rate more severely than observed to impress their supervisors and in situations in which lenient ratings are admonished (Villanova & Bernardin, 1989).

Although most studies have focused on the overall purpose of the appraisal process, DeNisi et al. (1984) and Landy and Farr (1983) suggest that purpose may have a greater role in the rating process than previously examined. Beyond rater motivation of what information is reported, purpose may affect what information is sought and how that information is later encoded and stored. Zedeck and Cascio (1982) found that rater strategies for integrating information in making evaluations varied with the purpose of the rating. Lord and Smith (1983) proposed that the level of processing and motivated encoding of ratee information is influenced by purpose. They proposed that increased levels of overall attention and memory processes occur when the rater expects to have future interaction with the ratee and is more involved in the process. Thus, the purpose has implications not just for motivated ratings, but for the motivated use of performance cues and motivated cognitive processing.

The contextual aspects of purpose might also influence motivational processing when considering that raters are also ratees. If raters are evaluated by their supervisors on the accuracy of their evaluations, then the expectation of review might have consequences for the rater beyond what might result from giving lenient ratings. The orientation of raters towards various members of the organizational context could determine how purpose will affect the resulting evaluations (Landy & Farr, 1983). Several studies have examined the effect that expectations of validation against some performance measure will have on self-performance ratings (Boyle & Klimoski, 1992; Eder & Fedor, 1989; Farh & Werbel, 1986; Fox & Dinur, 1988). Farh and Werbel (1986) examined the effects of purpose of appraisal and expectation of validation against the tendency of subjects to be lenient on self-evaluations. They found that subjects were less lenient in conditions of high expectation of validation and when the evaluations were for research purposes. They emphasized the need for research determining which contexts will decrease the motivation of raters to be lenient.

Farh and Werbel (1986) found that the expectation of validation did decrease leniency, but other factors that make the rater more involved in the judgment process might have a greater influence. Hagafors and Brehmer (1983) proposed that the expectation of justifying one's views may influence

motivational processing. The authors found that subjects who had to justify their judgments used a more analytical mode of functioning in judgment behavior. Although their study involved justification of predictions as opposed to performance evaluation ratings, the results indicate that expecting to justify one's decision involves a context that influenced motivational processing.

Situations are prevalent in which raters do have to justify their performance evaluations. Hauenstein (1992) found that when raters must justify their ratings to the ratee they tend to provide more lenient performance evaluations. Although justifying appraisals to the ratee is an important contextual factor, expectations of justifying appraisals to a supervisor might have more extreme effects because the rater is put into the position of a ratee. When expecting to justify a decision to the supervisor, individuals are more likely to use self-presentational strategies to protect the self from looking bad when confronted by a superior. In order to appear competent when justifying a decision, raters will engage in more vigilant information processing (Schlenker & Weigold, 1989). Performance cues from the supervisor may reduce ambiguity, but they may also increase biases because of the need to present oneself in a favorable light. An ambiguous situation and the purpose of appraisal both provide a context that may

steer the rater away from being the traditional cognitive miser into being a more thorough, motivated information processor.

Tetlock's (e. g. 1985a) proposed politician research program on accountability in social cognition and decision theory is directly applicable to providing a context in which the rater is motivated to increase cognitive processing and increase the search for performance cues. Tetlock (1985a) proposed that the social context of the situation always entails consequences that the decision maker must consider when making an evaluation. He suggests that differing degrees of accountability, defined as the "pressure to justify one's opinions to others", can be seen as the link between the individual decision maker and the social system to which that individual belongs. Decision makers will be motivated to seek the approval and respect of those to which they are held accountable (accountability principles) (Tetlock, 1985a).

Tetlock (1985a) suggests that two main types of accountability provoke different levels of information processing. The difference between the two types concerns the coping mechanisms used in the different situations in which the view of the accountability principle is ambiguous or not. The first type of accountability concerns an unambiguous situation in which the socially acceptable option

is relatively obvious. Decision makers will choose the most salient, acceptable course of action which involves the avoidance of unnecessary cognitive work. Tetlock (1985a) labeled this shortcut as the "acceptability heuristic", which is compatible with cognitive miser perspectives that view the information processor as one who chooses the quickest route in order to avoid extra cognitive effort. When the decision environment is unambiguous, individuals will make a strategic attitude shift by using the available cues which are viewed as the most acceptable to the accountability principle.

The second type of accountability concerns the ambiguous situation of not knowing the view of those to which one is held accountable. In this situation, decision makers have no information to rely on so they will be motivated to consider as much information as possible to prepare for a variety of critical reactions. Tetlock (1985a) states that accountability to an unknown other can influence the "utilization of more cognitively complex decision strategies, provoke greater awareness of their own cognitive processes, and influence more data-driven processing of evidence in forming impressions". The cognitive reactions that result from being accountable to an unknown are viewed as an adaptive strategy to protect one's self-esteem and social image. Tetlock (1985a) describes this coping strategy as "preemptive self-criticism".

The level of complexity in cognitive processing that is motivated by accountability is operationalized as the level of differentiations and integrations in the structure of written thought protocols about decisions (Tetlock, 1983a). The coding system used to score responses was originally derived from the work of several researchers on the topic of integrative complexity (Schroder, Driver, & Streufert, 1967; Schroder & Suedfeld, 1971; Streufert & Streufert, 1978). Differentiation refers to the number of unique characteristics or performance dimensions that an individual takes into account (Tetlock, 1983a). An undifferentiated approach involves a response that offers only one view of an issue. A more differentiated view recognizes more than one perspective of an issue such as positive and negative views on different dimensions of the performance of a ratee. The number of complex connections among differentiated characteristics mentioned by the rater refers to integration. The individual may perceive the different dimensions existing in "isolation (low integration), in simple interactions (moderate integration), or in multiple, contingent patterns (high integration)" (Tetlock, 1983a).

Tetlock has provided support for his theoretical propositions about complex cognitive processing that results from accountability. The ambiguous situation of not knowing the view of the accountability principle has been found to

have desirable effects on information processing. Tetlock (1983b) found that preexposure accountability led to a reduction in the influence of early-formed impressions on final judgments of evidence concerning the guilt of a defendant. Accountability received prior to the judgment process also improved recall of the material more than postexposure or no accountability conditions suggesting that the preexposure condition affects how people initially encode and process information. Tetlock (1985b) also found that accountability motivates more data-driven processing by reducing the "fundamental attribution error" or the overreliance on dispositional cues as opposed to situational cues of behavior. Tetlock and Kim (1987) also found that subjects report more integratively complex thoughts in preexposure accountability conditions as well as more accurate behavioral predictions and more appropriate levels of confidence in their judgments.

Although accountability has been shown to increase complex cognitive processing in many different situations (Weldon & Gargano, 1988), the role of accountability in performance appraisal is limited (Klimoski & Inks, 1990). Most of the research on performance evaluation has examined accountability of the rater to the ratee rather than to the supervisor. In addition, these studies only examined biased shifts that occurred as a result of accountability

manipulations. Cognitive complexity was not tested in the studies. Judge and Ferris (in press) and Klimoski and Inks (1990) found that when faced with giving feedback to the subordinate ratee, rater's tended to distort information in line with ratee self-assessments. Raters accountable to the ratee avoided the negative event of giving inconsistent feedback by evaluating ratee's in a similar manner as their self-assessments. When raters are accountable to ratees, leniency seems to occur as a method to avoid negative confrontations because raters feel they do not have the means to justify such ratings (Judge & Ferris, in press).

As shown, raters that are held accountable to the ratee with a known view of the ratee's position tend to direct their ratings more leniently. As previously discussed, accountability of the rater to the supervisor may produce more desirable effects. If the view of the supervisor is not available to the rater, increases in complex cognitive processing are likely to ensue as well as decreases in biased shifts towards leniency because negative performance evaluations will be more accepted. Longenecker (1992) supports this view by reporting that immediate supervisors have the potential to greatly influence the rater. The author suggested that the quality of performance ratings will increase and rater biases will decline when raters are held accountable to supervisors for their rating practices.

Although the supervisor as the accountability principle seems to have an influence on ratings, not much research has used the supervisor in this manner. Although not dealing specifically with performance evaluation, Fandt and Ferris (1990) examined the opportunistic behavior that resulted when raters were held accountable for the decisions they made. The accountability principle was not specifically stated, but subjects in the accountability condition were told that their judgments would be assessed on their future job evaluations. This manipulation essentially places the rater in the position of being accountable to the supervisor because of the effects their decisions might have on their performance evaluations. Fandt and Ferris (1990) found that subjects in positions of high accountability used more defensive information and placed more emphasis on positive aspects of their decision as a means of impression management to the supervisor.

Stamoulis (1993) also used the supervisory role as the accountability principle as well as including measures of cognitive complexity and biased shift. Stamoulis (1993) also offered an extension on the typical operationalization and interpretation of accountability. Tetlock's (1985a) empirical studies have continued to emphasize that the decision maker is susceptible to the context of accountability to maintain the positive regard of the

accountability principle. Although this seems to be a psychological motive, Tetlock (1985a) indicated that both "symbolic psychological and tangible material rewards and punishments" are reasons why an individual might seek approval and respect when held accountable. Tetlock (1985a) also cited social exchange theorists such as Blau (1964) who focused on gaining desirable material resources (e.g. promotions, office space) as a source of motivation.

Stamoulis (1993) used a reward in his manipulation because no research had thus far attempted to incorporate this operationalization of accountability. The author used three levels of accountability, an interpersonal/reward accountability, interpersonal only accountability, and no accountability. The interpersonal/reward accountability incorporated an expected reward based on the raters performance evaluation as well as expected feedback concerning the ratings. He found no differences between the three groups in biased shift, but he did find that the conditions where the view of the supervisor were known produced biased shifts overall. Stamoulis (1993) also found that raters in the reward/interpersonal accountability condition were more conservative in their use of favorable performance information. The author predicted that those in the no view condition would exhibit complex cognitive processing, but this prediction was not supported over time

although those in the interpersonal/reward conditions did exhibit an increase in processing. No differences were found between the reward/interpersonal and interpersonal alone conditions, although differences from the no accountability condition did indicate some evidence of accountability overall.

The author partially attributed the lack of results to the strength of the manipulation of accountability. Although he did not obtain consistent results, his study did contribute to an understanding of the underlying mechanisms of accountability. This study proposes to replicate and expand on the Stamoulis (1993) study by modifying portions of the manipulation and adding a new one. The Stamoulis (1993) study did not include a reward only accountability manipulation although this manipulation might have different effects on cognitive processing. Reward is often seen as an important contextual variable as suggested by McCallum, McCallum, and Gurwitch (1987) who emphasize that reward and interpersonal value are the two most prominent characteristics of social interactions. Stamoulis (1993) suggests that the reward aspects of accountability focus the rater on the attributions of ratings, but interpersonal accountability focuses individuals on justifications. Reward accountability raters can attribute their negative ratings to the reward, while interpersonal accountability individuals

give more favorable ratings because negative ratings are more difficult to justify.

Reward accountability may also provoke different processing strategies. Individuals in a reward accountability situation are accountable for the ratings they make, but they do not have to justify the ratings as suggested by Tetlock's (1985a) definition pertaining to interpersonal accountability. Raters receiving the traditional view of Tetlock's (1985a) accountability are accountable to the supervisor and must therefore attempt to form impressions while also remembering behaviors to justify the evaluations they make. Although still responsible for their formed impressions and evaluations, reward accountability raters do not need to remember behaviors because they do not have to justify their decisions. The accuracy of ratings can receive more effort in reward accountability situations because integrations among dimensions do not need attention after the rater has evaluated the ratee.

Although the effects of each type of accountability are unresolved, interpersonal accountability and reward accountability are both important contexts which can potentially increase cognitive processing and increase the accuracy of ratings. Interpersonal accountability may provoke structural changes in information processing, but the

unstudied effects of reward accountability may be the missing link that promotes increased accuracy of performance evaluations. The following study included two levels of interpersonal accountability (interpersonal and no interpersonal) and two levels of reward accountability (reward and no reward) to assess the effects of a reward only manipulation against the other previously studied variables. The effects of reward accountability were assessed by measuring the accuracy of evaluations compared to true performance scores. Based on Cronbach's (1955) four deviational components of accuracy, Hauenstein and Alexander (1991) operationalized two components of accuracy that can be utilized in a single ratee situation. These two components, elevation and dimensional accuracy, were the basis for judging accuracy of evaluations.

Another modification involves the gender of the subjects. Stamoulis (1993) found the view effects to be especially evident in females, thus only females were used in this study to reduce possible rater biases from the effects of gender. Two alternative explanations exist concerning why females tended to be more affected by the view. First, females may be more motivated by the need for approval, so they attempt to present themselves in a favorable light more than males generally would. Females also may be generally more compliant than males, especially when under a male

authority figure. A meta-analyses of the conformity/compliance literature found that females are more likely to be influenced than males (Cooper, 1979). Eagly and Carli (1981) suggest that compliance in females is especially evident when a male researcher conducts the study, thus a male researcher was used in this study. While either explanation might explain the Stamoulis (1993) findings, females were found to be more affected by the manipulation, so females were used in this study in hopes of using subjects that were more attune to the accountability manipulation. Finally, other minor changes from the Stamoulis (1993) study were instituted to make the interpersonal accountability manipulation stronger and in attempts to alleviate hypothesis guessing based on recommendations made by Stamoulis (1993).

## Summary and Hypotheses

Decreasing biases in performance appraisal is a pervasive problem in organizations. Recent research (Ilgen & Feldman, 1983) in social cognition has focused on the need to study the cognitive processes involved in impression formation in order to understand why biases and inaccuracies occur in performance evaluation. The rater, viewed as an active seeker of information (DeNisi et al., 1984), is motivated to use available performance cues from the organizational context to aide in forming impressions. Differing organizational contexts have effects on what types of information are attended to and what trait concepts easily come to mind in this information acquisition process.

The ambiguity of the situation and the purpose of the appraisal are both seen as important contextual variables which contribute to the availability of performance cues. If the rater is confronted with an unambiguous situation of knowing the view of the supervisor, then he/she will use this information when forming an impression and will be more likely to put less cognitive effort into making an evaluation than when in an ambiguous situation (Tetlock, 1983a). Increases in cognitive processing can also be seen when the purpose of appraisal involves the expectation of review. When raters expect to justify their view, motivated cognitive

processing occurs so raters can prepare for criticism and appear more competent.

Tetlock's (1985a) research program concerning accountability is directly applicable to providing a context which will increase complex cognitive processing and accuracy levels. Tetlock (1985a) operationalized accountability as the the pressure to justify one's view to others. Accountability can be applied to performance appraisal by making raters accountable to their supervisors for their evaluations. In accountability situations where the view of the supervisor is unambiguous, raters will use the acceptability heuristic and make a biased shift towards the view of the supervisor. In interpersonal accountability situations where the view of the supervisor is ambiguous, raters will use motivated processing of performance cues in attempts at preemptive self-criticism. Raters will use more differentiations and integrate ratee behaviors in more complex ways to prepare to justify their decisions to an unknown view. Although mentioned but not studied by Tetlock (1985a), the context of reward accountability will provoke raters to focus on the accuracy of their ratings so that accuracy levels will increase. Both interpersonal and reward accountability have the potential to increase cognitive processing and accuracy as this study proposed to demonstrate.

This laboratory study was designed to provide a replication and extension of the Stamoulis (1993) study in order to assess the underlying mechanisms of accountability in relation to performance appraisal. Female subjects were randomly assigned to groups and viewed videotaped ratee behaviors. Ratee performance was operationalized as videotaped ratee vignettes that involved an interviewer of a job applicant. Although supervisors served as the accountability principles in performance appraisal situations, college professors were operationalized as the supervisor accountability principles because of the nature of this laboratory study. Interpersonal accountability to the supervisor possessed two levels: interpersonal accountability and no interpersonal accountability. Reward accountability to the supervisor also possessed two levels: reward accountability and no reward accountability. Finally, known view of the supervisor possessed two levels: view and no view. In contrast to previous studies, this study examined the independent effect of reward accountability. Dependent variables were biased shift, cognitive complexity, and accuracy. Cognitive complexity was based on the analysis of written thought protocols which Martin and Klimoski (1990) have suggested as an informative method in which to examine the cognitive processing of raters. A number of hypotheses

were made by integrating from Tetlock et al.'s (1989) empirical findings and from Stamoulis (1993):

1. An accountability by view interaction will occur for biased shift. The interpersonal/reward accountability with view group, the interpersonal accountability with view group, and the reward accountability with view group will exhibit the greatest biased shift over all other groups.

2. An interpersonal accountability by reward accountability by view interaction will occur for complex processing. The interpersonal accountability with no view group will exhibit the greatest cognitive complexity over all other groups.

3. An interpersonal accountability by reward accountability by view interaction will occur for accuracy. The reward accountability with no view group will exhibit the most accurate evaluations over all other groups.

## Method

### Experimental Design

Two levels of interpersonal accountability to the supervisor (interpersonal accountability and no interpersonal accountability), and two levels of reward accountability (reward accountability and no reward accountability) were crossed with two levels of the view of the supervisor (view, no view). Subjects in the view condition received the view manipulation after the accountability manipulation was read.

### Subjects

One hundred ninety-one female subjects from a pool of undergraduate volunteers from Virginia Polytechnic Institute and State University participated in the experiment. The subjects received extra credit points in exchange for participating. To help reduce possible rater bias effects due to gender, all the subjects were female. A male experimenter was used. Subjects were randomly assigned to one of the eight experimental groups. Twenty-one to twenty-six subjects were in each condition.

### Procedure

Experimental sessions were conducted in a classroom setting with a small group of subjects in each session. Performance evaluation and the performance evaluation dimensions were first defined. After receiving instructions, the subjects viewed the videotaped ratee behaviors. As in Tetlock (1983a), at the completion of the videotape, subjects

wrote their thought protocols, and then made evaluative ratings. Subjects were given five minutes to write down their thoughts about the performance of the interviewer in the previous vignettes. Similar to Tetlock (1983a), subjects were told "you should have the opportunity to collect your thoughts before giving an evaluation of the interviewer. All thoughts you report will be completely confidential. Please respond with complete candor and honesty." Subjects then rated the set of ratee behaviors as a whole. Subjects were then given manipulation check items and were debriefed. Each experimental session lasted approximately one hour.

Independent Variables

Interpersonal accountability. The interpersonal accountability manipulation was based on Tetlock (1983a), Tetlock (1983b) and specifically as an expansion and replication based on Stamoulis (1993). A pilot study was conducted by Stamoulis (1993) which aided in the development of the independent variable manipulations. Based on Stamoulis's (1993) results, the position of professor was chosen as the supervisor position. See Appendix A for a prepared script of the interpersonal accountability manipulation. An example of the sheet upon which subjects wrote meeting sign-up information as part of the interpersonal accountability manipulation is included in Appendix B.

Reward accountability. The reward accountability manipulation was based on Tetlock (1985a) and Stamoulis (1993). Reward was operationalized as additional extra credit points the subjects could receive as a result of the quality of their ratings compared to true dimension scores. A prepared script for the reward accountability manipulation can be found in Appendix C. An example of the sheet upon which subjects wrote envelope pick-up information as part of the reward accountability manipulation is included in Appendix D. A prepared script for the conditions receiving both the interpersonal and reward accountability manipulations is also included (see Appendix E) so replications in the two manipulations were not repeated when read. Subjects in the no accountability condition were told they would be asked to rate a set of job performances and told their performance evaluation rating would be "totally confidential and not traceable to them personally" (Tetlock, 1983b).

View. Subjects in the view condition were given the view of the professor after the accountability manipulation and before the set of ratee behaviors was presented. Stamoulis (1993) found that telling the subjects the view of the professor in a straightforward manner was the most effective manipulation of view. The view subjects silently read the prepared statement (see Appendix F) while the

experimenter read the statement aloud. This represents the supervisor's view or performance cue that was biased in relation to the average to good performance of the ratee. Subjects in the no view condition were not told the view of the professor at any stage of the experiment.

Dependent Variables

Biased Shift. Biased shift was measured by assessing group differences among the performance evaluation rating means to determine whether a particular group's performance rating means were in the direction of the known view or performance cue.

A 7-item, 7-dimension interviewer performance evaluation form from Hauenstein (1987) was used for rating the set of videotaped ratee behaviors (see Appendix G). The 7 job performance dimensions on the rating form include: Rapport Building, Organization of the Interview, Questioning Skill, Relevance of Questions, Company and Job Preview, Answering the Applicant, and an Overall evaluation. The rating format consists of a 7-point Likert-type scale. While the items themselves are behavioral in nature, the 7, 4, and 1 anchors on the rating scale represent three main levels of performance (effective, average, ineffective).

Cognitive Complexity. Cognitive complexity was measured in the same way Tetlock (1983a) has measured the variable. The overall rating of cognitive complexity was based on the

number of positive differentiations, negative differentiations, and integration statements that raters included in their thought protocols. See Appendix H for the protocol. Tetlock (1983a) drew from the work of several researchers in adopting his measure of complex processing (Schroder et al., 1967; Schroder & Suedfeld, 1971; Streufert & Streufert, 1978). The number of unique, functionally useful information dimensions was referred to as differentiation (Schroder et al., 1967). Integration referred to the number of complex connections among these unique dimensions. The measure addressed the structural aspects of written information and did not directly focus on the content of information used in decisions. The overall cognitive complexity rating represents the complexity of thoughts occurring in the evaluation process.

Schroder et al. (1967) described the methods by which content coders rated written thought protocols on differentiation and integration. Information concerning differentiation and integration were combined to yield an overall composite score of cognitive complexity. Responses were scored on a seven point scale. The scale represents a continuum of low to high levels of both differentiation and integration (see Appendix I). Two content coders were trained based on the information and examples presented in Schroder et al. book and by Tetlock (1983a) and Tetlock et

al. (1989). These coders were unaware of the hypotheses of the proposed experiment. Interrater reliability was assessed.

Accuracy. Accuracy was measured according to two components of accuracy operationalized to be computed in a single ratee study by Hauenstein and Alexander (1991) and based on Cronbach's (1955) four deviational accuracy components. The first component was labeled elevation because it is directly analogous to Cronbach's (1955) elevation component of accuracy. The second component was labeled dimensional accuracy. Conceptually, dimensional accuracy is equivalent to Cronbach's (1955) differential accuracy in that it measures the accuracy with which each rater evaluated a single ratee on each dimension. However, because there was only one ratee, the component was labeled dimensional accuracy instead of differential accuracy.

For elevation, perfect accuracy requires the rater's average observed rating to equal the average of the target scores. Perfect accuracy in terms of dimensional accuracy requires both a correlation of positive one between a rater's observed ratings and the target ratings, and that a rater's variance for her ratings equals the variance of the target scores.

Manipulation Check

At the end of the experimental session, all subjects responded to 11 manipulation check items (see Appendix J). Items 1 through 4 measure the perception of interpersonal accountability in writing thoughts and making performance evaluations. Items 5 through 8 assess the reward accountability manipulation in terms of the number of extra credit points received as a result of the quality of ratings. Items 9 through 11 measure the use of the known view in writing thoughts and making performance evaluations. Response format to the manipulation check consisted of a 7-point, Likert-type scale.

Development of Videotaped Ratee Behaviors

The identical set of videotaped ratee behaviors which Stamoulis (1993) used in his study were shown to raters. The videotaped ratee behaviors were initially constructed for a rater training study (Stamoulis & Hauenstein, 1992). Target ratee behaviors for the occupation of an applicant interviewer were derived from Hauenstein (1987). A range of these behaviors were chosen for inclusion into the videotape based upon the level of ratee performance they depicted and the ease in which the behaviors could be translated via video. Twelve behaviors were chosen from each of the good, average, and poor performance categories to represent a broad range of performance levels. To help control for any rating

scale or format biases which might affect the translation to this experiment, the same 7-item interviewer rating scale from Hauenstein (1987) was utilized in the development of this study's videotapes. This rating form was used in this study.

Scripts were carefully prepared based upon the interviewer behaviors. The interview vignettes were written so two behaviors of the same performance level were included. Therefore, 6 interview vignettes were composed for each of the good, average, and poor job performance categories (18 vignettes in total). Care was taken to preserve realism and match behaviors that might naturally covary, and to nest these critical behaviors amidst relatively innocuous interview content. Each vignette was in color, two to three minutes in duration, and contained the same two actors portraying the interviewer and interviewee on an office set.

A pilot study (Stamoulis, 1993) was then conducted to establish the target scores of the interview vignettes. The 18 videotaped vignettes were presented to 193 undergraduates within four different series of randomly ordered triads of the good, average, and poor performance levels. The mean ratings of each rating dimension from each interview vignette were used as vignette target scores. Ten vignettes with target scores above the scale midpoint on the overall evaluation dimension were used in this study. Therefore,

vignettes in this study portrayed average to good ratee performance.

The vignettes possess this slight skewness for three reasons. First, more of the vignette means on the overall evaluation dimension were above the scale midpoint rather than below the midpoint. In order to include as many vignettes as possible in the proposed study, the vignettes above the scale midpoint were used. Second, the proposed study includes the investigation of shifts in evaluations because of performance cues. In order to adequately measure the difference between raters evaluating "true" aspects of the vignette (of average to good ratee performance) and raters being affected by biased performance cues (of poor performance), it is desirable to include vignettes that are slightly skewed in their representation of "true" ratee performance. Third, only two studies have used Tetlock's (1985a) model to investigate the effects of accountability on performance evaluation ratings. Klimoski and Inks (1990) study used "moderately poor" work performance. Stamoulis (1993) used the same videos of "moderately good" work performance as this study. This study used the "moderately good" work performance in order to determine the effects of applying a reward only accountability manipulation to the same work performance levels as used in the Stamoulis (1993) study and as a replication to the other accountability

manipulations found in his study. Therefore, the use of the "moderately good" work performance expanded on both studies by widening the domain of work performance levels on different kinds of accountability manipulations.

## Results

### Manipulation Check

The 11 manipulation check items consisted of a 4-item scale measuring the interpersonal accountability manipulation, another 4-item scale measuring reward accountability, and a 3-item scale measuring the view manipulation. The 4 interpersonal accountability items, the 4 reward accountability items, and the 3 view scale items were intercorrelated respectively according to a priori expectations. The table of intercorrelations amongst the manipulation check items appears in Table 1. Considering their size, the scales possessed satisfactory internal consistency reliability. Cronbach's alpha for the interpersonal accountability scale was .88, Cronbach's alpha for the reward accountability scale was .73, and Cronbach's alpha for the view scale was .75. However, the intercorrelations among items on different scales were higher than expected. This raised the possible problem of discriminant validity of the manipulation check scales.

A 3-way, ANOVA was used to assess the effects of interpersonal accountability, reward accountability, and view on the interpersonal accountability scale. The source table appears in Table 2. A significant 2-way interaction was found between reward and interpersonal accountability on the interpersonal accountability scale, $F(1, 183) = 4.04$, $p <$

.05. Although the interpersonal/reward accountability interaction was unexpected, interpersonal accountability accounted for 63% of the variation, while the interaction only accounted for 2%. The interaction is depicted in Figure 1. The figure clearly suggested that subjects in the interpersonal accountability condition scored higher than no interpersonal accountability subjects on the manipulation check scale, regardless of the small difference that occurred between reward and no reward accountability subjects in the no interpersonal accountability condition. Interpersonal accountability subjects reported that they experienced interpersonal accountability to a greater extent ($M$ = 5.25) than the no-interpersonal accountability groups ($M$ = 2.30). Therefore, interpersonal accountability appeared to be manipulated successfully.

A 3-way, ANOVA was conducted to investigate the effects of interpersonal accountability, reward accountability, and the view manipulation on the reward accountability scale. The source table appears in Table 3. A significant main effect was found for reward accountability ($F(1, 183) = 38.63$, $p < .01$), but a significant 2-way interaction was also found between reward accountability and view on the reward accountability scale, $F(1, 183) = 9.71$, $p < .05$. The reward accountability manipulation accounted for more variance ($eta^2 = .174$) than the interaction ($eta^2 = .031$). Although

subjects in the reward condition ($M$ = 4.47) scored higher than no reward accountability subjects ($M$ = 3.31), subjects in the no view condition showed less difference than the view subjects from no reward to reward accountability (See Figure 2). This suggested that no view subjects did not perceive the reward accountability manipulation as strongly as the view subjects.

Because of this interaction, supplemental analyses were conducted on each dependent variable to compare the results when using the perceived reward (i.e. self-report on the reward manipulation check) to the results when using the reward manipulation as an independent variable. Biased shift and accuracy showed no differences in results when perceived reward was used as the independent variable of reward, but cognitive complexity results did change. These differences will be discussed in the cognitive complexity section.

A 3-way, ANOVA was conducted to assess the effects of interpersonal accountability, reward accountability, and view manipulations on the view scale. The source table appears in Table 4. A significant main effect was found for the view manipulation, $F(1, 183) = 94.53$, $p < .01$. The view groups' mean ($M$ = 4.71) indicated that they were aware of the performance view given to them, while the no-view groups ($M$ = 2.64) reported less awareness of any view of ratee performance.

Biased shift

An examination of the correlations between the rating dimensions revealed the six dimensions and the Overall performance dimension were significantly intercorrelated (Table 5). Therefore, a multivariate approach was used in the analysis of biased shift in the six specifically-oriented rating dimensions. Similar to other work in the performance evaluation literature, the Overall dimension was analyzed separately. The descriptive statistics for all six dimensions and overall performance were calculated for each experimental cell in Table 6.

A 3-way, MANOVA was conducted as an omnibus test of the effects of interpersonal accountability, reward accountability, and view on ratings of the six specific dimensions. The source table appears in Table 7. The significant effects included an interpersonal accountability by view interaction ($F(6, 178) = 2.48$, $p < .05$) and a main effect for view ($F(6, 178) = 4.78$, $p < .01$). Figure 3 displays the effects found using the mean of the six dimensions. This pattern of effects was consistent with what was expected from the hypotheses.

Between subjects one-way anova contrasts were used to investigate Hypothesis 1: the interpersonal/reward accountability with view group, the interpersonal accountability with view group, and the reward accountability

with view group will exhibit the greatest biased shift over all other groups. The between subjects contrast assessing hypothesis 1 found significant differences between the interpersonal/reward accountability with view group, the interpersonal accountability view group and the reward accountability view group compared to all other groups on all of the six rating dimensions. Table 8 displays the results of the t-tests. Therefore, the results of all six individual rating dimensions supported Hypothesis 1. The accountability with view groups exhibited biased shifts towards more unfavorable ratings more than all other groups.

Further analyses were used to examine the differences found in the hypothesized effects. Because univariate ANOVA's found significant effects for all six rating dimensions, the mean of the six dimensions were used for further analyses. T-tests were first conducted to examine differences among the accountability with view groups to insure the appropriateness of pooling these groups together to form one group. No significant differences were found between the interpersonal/reward accountability and the reward accountability with view groups ($t(50) = .02$, n.s.), or between the interpersonal/reward accountability and the interpersonal accountability with view groups, $t(45) = -.18$, n.s. Significant differences were also not found between the interpersonal accountability and reward accountability view

groups, $t(45) = -.22$, n.s., thus the type of accountability did not affect subjects differently on biased shift. Based on these results, the accountability/view groups were pooled together for further analyses.

T-tests were conducted with the collapsed accountability/ view group compared to each other experimental group individually. Significant differences were found between every other experimental group and the accountability/view group. The no accountability with view group was significantly different from the accountability/view group ($t(95) = 2.69$, $p < .01$), indicating that accountability was necessary for the view to effect subjects. The accountability/view group was also significant from the interpersonal/reward accountability with no view group ($t(94) = 3.49$, $p < .01$), the reward accountability/no view group ($t(96) = 3.68$, $p < .01$), the interpersonal accountability/no view group ($t(95) = 3.37$, $p < .01$), and the no accountability/no view group ($t(93) = 3.53$, $p < .01$). All no view groups were significantly different from the accountability/view group, but the no accountability/no view group was not significantly different from the no accountability group who received the view. These results indicated that accountability in conjunction with a negative performance view resulted in more unfavorable ratings than with no view and/or no accountability.

A 3-way, ANOVA was conducted as an omnibus test of the effects of interpersonal accountability, reward accountability, and view on ratings of the Overall performance dimension of the interviewer. The source table appears in Table 9. The significant between-subjects effects included a view main effect ($F(1, 183) = 20.23$, $p < .01$) and an interpersonal accountability main effect ($F(1, 183) = 3.85$, $p < .05$). View subjects produced less favorable ratings ($M = 3.19$) than no view subjects ($M = 4.04$). Interpersonal accountability subjects had less favorable ratings ($M = 3.43$) than no interpersonal accountability subjects ($M = 3.80$). This pattern of results was consistent with hypotheses, although the expected interactions of accountability by view were not found.

A one-way ANOVA contrast was conducted to investigate Hypotheses 1 for the Overall performance dimension. The difference between the interpersonal/reward accountability view group, the interpersonal accountability view group, and the reward accountability view group compared to all other groups was also significant for the Overall performance dimension, $t(183) = -5.08$, $p < .01$. Therefore, the results of the Overall performance rating dimension also supported Hypothesis 1.

Further analyses were used to examine the differences found in the hypothesized effects. As found for the six

rating dimensions, no differences were found between the three accountability with view groups on the overall performance dimension. Based on these results, the accountability/view groups were pooled together for further analyses. T-tests were conducted with the collapsed accountability/view group compared to each experimental group individually. Significant differences were found between each group and the accountability/view group, thus the overall performance dimension depicted the same pattern as found with the six rating dimensions. Accountability in conjunction with a negative view resulted in more unfavorable ratings than with no view and/or no accountability. An interpersonal accountability effect was found for the overall evaluative dimension, but the type of accountability did not result in differences. The interpersonal accountability effect could be attributed to the difference in ratings from the no accountability group in the view condition. No accountability with the view ($M$ = 3.88) did not result in more unfavorable ratings as found with other accountability/view groups ($M$ = 2.96).

In conclusion, any form of accountability in conjunction with the view caused increases in biased shift for all seven rating dimensions thus supporting hypothesis 1. Accountable subjects with the view provided more negative ratings than subjects in the no accountability and no view conditions.

More favorable ratings were found with the absence of a negative performance view. No differences were found between the different types of accountability (interpersonal/reward, interpersonal, and reward). Stamoulis (1993) found the performance view to be solely responsible for biased shifts in performance evaluations. The most likely explanation for the difference is that the current study used stronger accountability manipulations.

In this study, the lack of a view effect in the no accountability group is consistent with Tetlock's (e.g. 1985a) model and suggest that accountability is necessary for the view to have a strong effect. When accountable for ratings, subjects reverted to the acceptability heuristic by using available cues that were the most acceptable to the accountability principle. Subjects who were not accountable did not bias their ratings because they did not expect a review or have to justify their ratings.

For biased shift, the lack of significance found between the three types of accountability has implications for Tetlock's model. The results provide further empirical support for his theory, while also contributing additional information. The untested reward accountability manipulation and the conjunction of the two types of accountability influenced subjects to the same degree as the traditional interpersonal accountability. This study demonstrates that

Tetlock's model can apply to different types of accountability. As long as subjects are held accountable for their ratings, regardless of the type of accountability, they are affected by the performance view given.

Cognitive Complexity

Two content coders counted the frequency of positive differentiations, negative differentiations, and integration statements reported on the free thought protocols. The coders used these three variables in determining the cognitive complexity rating of thought protocols. Interrater reliability correlations for these four variables appear in Table 10. The interrater reliability correlations ranged from .55 to .89. The number of integrations counted in the thought protocols was the only component upon which the raters did not consistently agree ($r$ = .55). More importantly, the interrater reliability was high ($r$ = .84) on the overall cognitive complexity rating. Discrepancies between these two coders were resolved by the author. The descriptive statistics for these variables within each cell are in Table 11.

A 3-way, ANOVA was conducted to assess the effects of interpersonal accountability, reward accountability, and view on the overall cognitive complexity rating. The source table appears in Table 12. Significant between subjects effects included an interpersonal by reward by view interaction ($F$(1,

183) = 10.61, $p < .05$), an interpersonal by reward accountability interaction ($F(1, 183) = 4.03$, $p < .05$), and a reward accountability effect ($F(1, 183) = 7.53$, $p < .01$). The 3-way interaction is depicted in Figure 4. Although a 3-way interaction was hypothesized, the scores were not in the expected direction, thus the pattern of interactions were not consistent with what was expected.

A one-way ANOVA contrast investigated Hypothesis 2: the interpersonal accountability with no view group will have the highest cognitive complexity over all other groups. As expected based on Figure 4, the between subjects contrast assessing hypothesis 2 was not significant, $t(183) = -0.05$, n.s. Therefore, hypothesis 2 was not supported.

Although the hypothesis was not supported, the 3-way interaction between the 3 types of manipulations suggested differences not hypothesized. Multiple comparisons were used in a post hoc strategy to further investigate the effects of the manipulations on overall cognitive complexity. Comparisons using the Student-Newman-Keuls ($p < .05$) procedure showed that no significant differences were found between the view and no view for each accountability group. Significant differences were also not found between the different accountability groups in the no view condition. The only significant difference found between groups in the view condition was between the interpersonal accountability

group ($M$ = 1.90) and the interpersonal/reward accountability group ($M$ = 3.69). This effect accounted for the significant 3-way interaction found in the omnibus test. Although not expected, the interpersonal/reward accountability group had the highest cognitive complexity, regardless of view. The interpersonal/view group showed a decrement in complex cognitive processing, suggesting that interpersonal accountability subjects cannot overcome the tendency to accept the view and thus decrease processing, unless placed in conjunction with reward accountability.

In conclusion, the only significant difference was that the interpersonal/reward accountability with view group showed a higher level of cognitive complexity than the interpersonal accountability with view group. The low cognitive complexity scores for the interpersonal accountability/view group was consistent with Tetlock's (1983a) predictions and results, because the subjects who receive the view are expected to use the acceptability heuristic and thus show lower cognitive complexity. However, the high scores of interpersonal/reward accountability subjects who were under the same view conditions as their interpersonal accountability only counterparts were not consistent with Tetlock. These subjects biased their views in the same manner as the interpersonal only group, but they did not display low cognitive complexity when given the view.

The reward/interpersonal combined accountability group did not show a decrement in processing when given the view, which was an unexpected finding. Interpersonal and the untested reward accountability were seen as competing processing demands that would interfere with each other when received simultaneously. The competing nature of the two types of accountability were hypothesized to result in lower levels of cognitive complexity and accuracy, because subjects would not be able to concentrate on one processing strategy. As indicated by the results, these hypotheses did not hold. The conjunction of the two types of accountability produced a high level of cognitive complexity. The acquisition of some type of tangible reward based on the quality of ratings, combined with feedback based on justifications, prevented decrements in processing that were expected to occur when presented with a view.

Supplementary analyses were conducted to assess if differences were found due to the aforementioned problem with the reward accountability manipulation check. Each subject's score on the reward accountability manipulation scale was used as a measure of 'perceived reward'. 'Perceived reward' scores, the coded interpersonal accountability and view manipulations and their respective interactions were entered as independent variables in a regression analysis with the overall cognitive complexity rating as the dependent

variable. In comparison to the results from the reward manipulation, differences were found when using scores from perceived reward as an independent variable. A 3-way interaction was not found, although a significant 2-way interaction was found between perceived reward accountability and interpersonal accountability, $t(184) = 2.80$, $p < .01$. The main effects included an interpersonal accountability main effect ($t(184) = -2.63$, $p < .01$), and a perceived reward main effect ($t(184) = 2.17$, $p < .05$). An interpersonal accountability main effect was not found in the previous analyses with the coded reward accountability manipulation as the reward independent variable. The regression table appears in Table 13.

In order to more clearly see how the differing operational definitions affected results, Figure 5 depicts the graph with the coded reward manipulation with reward on the x-axis, and Figure 6 depicts the graph with perceived reward on the x-axis. Multiple comparisons were used in a post hoc strategy to further investigate the effects of the perceived reward on the overall cognitive complexity ratings. Comparisons using the Student-Newman-Keuls ($p < .05$) found that the major difference resulting from using perceived reward concerns the interpersonal accountability with no view group. This group was significantly different from low perceived reward to high perceived reward. This difference

was not found when the coded reward manipulation was used as the independent variable. The interpersonal accountability with view group displayed almost the same means as the interpersonal accountability no view group, with an increase in cognitive complexity when reward accountability was perceived by the subjects. The combination of reward and interpersonal accountability increased cognitive complexity regardless of the view, thus depicting the significant 2-way interaction found from the regression analyses. This affect was different from the initial 3-way interaction found, in which the view did affect results. Reward did not seem to increase cognitive complexity levels unless paired with interpersonal accountability.

To summarize the perceived reward results, the interpersonal accountability with no view did not have the highest cognitive complexity for either the no reward group from the coded reward manipulation or the low scorers on perceived reward. Therefore, hypothesis 2 was also not supported when using the self-report measure of reward. Although the hypothesis was not supported, the interpersonal/reward accountability with view group produced higher levels of cognitive complexity. The combination of reward and interpersonal accountability seemed to increase cognitive complexity levels beyond what the view provides. This finding was found for both the reward manipulation and

perceived reward using the reward manipulation check scale as the independent measure of reward.

Overall Cognitive Complexity Conclusions

Reward accountability combined with interpersonal accountability and a view produced the highest cognitive complexity, although not significantly different from all groups. This finding was contrary to hypothesis 2 in which the interpersonal only accountability with no view was hypothesized to have the highest cognitive complexity over all other groups, thus hypothesis 2 was not supported.

Although differences were found when using perceived reward scores from the reward accountability manipulation check scale, the major difference was that the interpersonal with perceived reward conditions had an increase in cognitive complexity, regardless of view. When the reward manipulation was used, the view did result in effects such that a significant difference between interpersonal/reward accountability and interpersonal accountability was found in the view condition, but not found in the no view condition. The interpersonal accountability only group did not increase cognitive complexity scores from view to no view when using perceived reward. The use of the reward manipulation did find this effect, which was expected from Tetlock's predictions and findings.

While this finding should be considered, results found when using the initial coded reward manipulation as the independent variable of reward were more consistent with hypothesized effects in concurrence with Tetlock' predictions. The view effects seen for reward only accountability for both independent variables of reward, and the interpersonal only accountability view effect found when using the reward manipulation were more consistent with increases in cognitive complexity from view to no view. The perceived reward findings were also not as significant when considering that the interpersonal/reward accountability view group produced high cognitive complexity ratings consistently across both types of the reward operationalized definition. The high cognitive complexity for the interpersonal/reward with view group did not support hypothesis 2, and was not consistent with Tetlock's findings.

Accuracy

The accuracy of ratings was compared to true performance scores for each dimension using elevation and dimensional accuracy. The two types of accuracy displayed a significant negative correlation, $r = -.22$, $p < .01$. The descriptive statistics for the two accuracy variables are displayed in Table 14.

A 3-way, ANOVA was conducted to investigate the effects of interpersonal accountability, reward accountability, and

view on elevation. The source table appears in Table 15. The significant effects included a view effect ($F(1, 183) = 31.81$, $p < .01$), and an interpersonal accountability effect, ($F(1, 183) = 6.14$, $p < .05$). Figure 7 depicts these effects. Although the hypothesized 3-way interaction was not found, the direction of the view effect was consistent with what was expected. Subjects in the no view conditions ($M = .84$) were more accurate than subjects in the view conditions ($M = 1.41$), as lower scores indicate higher accuracy levels. The direction of the interpersonal accountability effect was also consistent with expected patterns such that interpersonal accountability groups ($M = 1.25$) were less accurate than no interpersonal accountability groups ($M = 1.00$).

A one-way ANOVA contrast was used to assess Hypothesis 3: the reward accountability only, no view group will have the highest accuracy. The between subjects contrast assessing the difference between the reward accountability only, no view group and all other groups was significant, $t(183) = -2.98$, $p < .01$. Therefore, hypothesis 3 was supported by the elevation component of accuracy.

Although the hypothesis was supported and the hypothesized group did have the highest elevation accuracy ($M = .74$), Figure 7 clearly indicates that this result was due to the view effect, because of unexpected results that no differences were seen between groups in the no view

condition. T-tests were conducted as further analyses of the hypotheses. The reward only accountability with no view group was compared to each other group individually. As expected from the graph, the reward only accountability group was significantly different from each view group individually, but was not significantly different from other no view groups. Further analyses pooled the interpersonal/reward and interpersonal accountability groups in the view condition. The collapsed interpersonal accountability group was not significantly different from the reward only group in the view, but was significantly different from the no accountability with view group, $t(69) = -2.58$, $p < .05$. The difference found from the no accountability group shows the interpersonal accountability effect. Subjects in the view condition showed higher accuracy levels in the no accountability condition than in the accountability conditions.

Regardless of the type, accountability appears to make the view more salient to subjects, who exhibited lower elevation accuracy in the view condition than no accountability subjects. The no accountability group was the only group that was not affected by the view. These results are consistent with the biased shift measure in which accountability/view groups gave more severe ratings then no accountability and/or no view groups, thus leading to

inaccurate ratings as well. The lack of differences found between the types of accountability is inconsistent with the hypotheses that the two types of accountability provoke competing demands. Support was not found for the hypothesis that subjects would be more accurate if they focused on the quality of ratings rather than justifications, because significant differences were not found between the two types of accountability groups. The view influenced accountable subjects to adopt a negative stance which they probably perceived as accurate, instead of simply focusing on the behaviors to actually determine the true levels of performance. Subjects in the no accountability condition did not feel the pressure of accountability, so adopting the view of the supervisor was not necessary.

A 3-way, ANOVA was conducted to investigate the effects of interpersonal accountability, reward accountability, and view on dimensional accuracy. The source table appears in Table 16. No significant differences were found. A one-way ANOVA contrast was used to assess Hypothesis 3 based on dimensional accuracy. The contrast assessing the difference between the reward accountability with no interpersonal accountability with no view group was not significant, $t(183) = 1.42$, n.s. Hypothesis 3 was not supported by the dependent variable dimensional accuracy.

In conclusion, the view inflated ratings resulting in less accurate scores when combined with accountability conditions on the elevation component of accuracy. Although results were consistent with the biased shift measure, the results of elevation accuracy did not support hypothesis 3. Dimensional accuracy scores also did not support hypothesis 3.

## Discussion

The purpose of this study was to assess whether two types of accountability from Tetlock's (1985a) model could be provided as social contexts to bias the performance judgment process. Accountability has been shown to influence subjects to distort information in line with the opinions of another individual (Tetlock, 1983a), but this study draws attention to influences when the accountability principle has formal authority over the judge. This formal authority was represented by the expectation of a review by the subjects' professor. Results from the biased shift measure showed that accountability to the professor in conjunction with a negative performance view influenced raters to use the acceptability heuristic and subsequently give more negative ratings. These results are consistent with Tetlock's (1985a) theory. Although the type of accountability did not result in differences, accountability to a formal authority is necessary to influence subjects to bias their ratings towards the performance view.

The effects of accountability on biased shift were confirmed by the current study, but differences in two different types of accountability were also explored. Tetlock (1983a) has previously used interpersonal accountability, which influenced biased shifts towards the view and increased levels of cognitive complexity when the

raters expected to justify their ratings in a feedback session with the person to whom they were accountable. This study attempted to disentangle the mechanisms underlying accountability effects by also using reward accountability. Stamoulis (1993) used a reward/interpersonal accountability manipulation, but did not assess effects from a reward accountability alone manipulation. Interpersonal and reward accountability were hypothesized to provoke competing demands in the information processor, such that different effects were expected for cognitive complexity and accuracy. Interpersonal accountability with no view subjects were expected to have the highest cognitive complexity, but as the results indicated, this hypothesis was not supported.

A possible explanation for the lack of findings for cognitive complexity concerns the adequacy of the measure in assessing cognitive processes. Although cognitive complexity scale values above two represent the use of both positive and negative information simultaneously by the rater (an important characteristic of cognitive complexity), scale values of one and two represent the use of positive and negative information only. Because of the definition of the cognitive complexity dependent variable (Schroder et al., 1967), one rater can use many positive or negative differentiations and have a low cognitive complexity rating. Another rater can use a few positive and a few negative

differentiations and have a high cognitive complexity rating. The number of integration statements only affect the top two cognitive complexity scale values.

The distribution of cognitive complexity ratings and the low frequency of integration statements found in this study were similar to the findings of Schroder et al. (1967) and Stamoulis (1993). The inconsistency of correlations across the three components with the overall cognitive complexity rating (see Table 17) and the cognitive complexity definition put forth by Schroder et al. (1967) suggest that the three cognitive complexity components, when considered individually, do not represent the overall cognitive complexity construct. Future study needs to assess whether the cognitive complexity scale is an appropriate measure of cognitive complexity in performance evaluation settings. Other information processing measures, such as directed free recall or direct measures of encoding, may provide more insight than cognitive complexity.

The scale also did not produce many high scores of cognitive complexity. The measure may be more sensitive to changes on the low range of the cognitive complexity dependent variable. This is further substantiated by the fact that the components of cognitive complexity, (i.e., positive differentiations, negative differentiations, and integrations) do not appear to constitute the full construct

of cognitive complexity. The low levels of integration statements display that few people discussed cause and effect patterns of the interviewer's behavior. Tetlock (1983a) used the scale as a measure of cognitive complexity when discussing social issues. Because subjects in the current study were not told to present two sides of an issue, the manipulations may not have provoked high levels of cognitive complexity.

Although the lack of results for cognitive complexity could be attributed to measurement problems, the results for biased shift and the accuracy measures suggest a more complex conclusion. Neither type of accuracy supported the hypothesis that the reward accountability/ no view group would have the highest accuracy levels. However, the elevation accuracy measure reflected the findings for biased shift. Elevation accuracy was lower when accountable subjects were presented with the negative performance view, which corresponded to the more severe ratings found under the same circumstances for biased shift. The type of accountability did not have an effect, suggesting that any type of accountability received with a view will lead to the acceptability heuristic of accepting that view, whether accurate or not.

The elevation accuracy results support Tetlock's theory of accountability in unambiguous situations. However,

dimensional accuracy may be a more useful accuracy measure because dimensional accuracy shows how subjects rate patterns of ratee performance through comparisons of each dimension rating to a true score, instead of averaging across dimensions as done with elevation accuracy. However, significant differences were not found for dimensional accuracy. If it is assumed that cognitive complexity was a reasonable measure, then the lack of differences on dimensional accuracy is not surprising given the results for cognitive complexity. For groups to differ on dimensional accuracy, groups would also be expected to differ on cognitive complexity. That is, the group with the highest cognitive complexity would be the most accurate group.

However, the nature of the dimensional accuracy measure also compounds the problem. This study involved only one ratee which means that differences in dimensional accuracy were based on the comparisons of subjects' ratings to the true scores for each dimension. Rating differences on dimensions with only one ratee involves a situation that might not have been sensitive to accuracy differences. Differences in the methods by which raters' make evaluations are easier to recognize when many ratee's on several dimensions are rated as shown with Cronbach's (1955) differential accuracy. Furthermore, the true scores for each dimension were also very similar across dimensions, thus

limiting the variability that could have resulted from dimensional accuracy. Multiple ratees and more variability in the true scores across dimensions should be included for future study to assess the effects of accountability on dimensional accuracy.

To summarize, the findings for biased shift and elevation accuracy support Tetlock's theory. More specifically, in situations where the judge is accountable and the judge knows the view of the accountability principle, the judge will use the acceptability heuristic in the sense of biasing his/her ratings toward the view. In contrast, the cognitive complexity and dimensional accuracy results do not appear to support Tetlock's model. That is, no consistent decreases were found for cognitive complexity for accountable groups who received the view, and no differences were found in levels of dimensional accuracy.

One explanation for the cognitive complexity and dimensional accuracy results is that both measures were not sensitive enough to capture the accountability effects. However, an alternative explanation exists for the lack of findings for cognitive complexity and dimensional accuracy that is consistent with Tetlock's model. The key to invoking the acceptability heuristic is the lack of ambiguity in relation to the accountability principle's known view.

The view in the current study influenced subjects to provide more negative ratings, but subjects did not show lower cognitive complexity levels from no view to view. Although the view was presumed to be consistent with the views used in Tetlock's research, subsequent examination demonstrates that the views are not comparable in terms of ambiguity. For example, subjects in Tetlock (1983a) were told that they were accountable to an individual with known views, being either consistently conservative or liberal. Included in the view was information that the conservative viewpoint opposed policies such as affirmative action and supported policies such as increased defense spending and capital punishment, while the liberals held the opposite view. Given this type of information in the view, subjects easily produced normative explanations or justifications for why, for example, a liberal would oppose certain policies such as capital punishment because of the immorality of ever taking a human life.

The view presented in the current study simply stated that the accountability principle thought the interviewer performance set to be "poor or ineffective, overall, because the interviewer does not perform his job duties and responsibilities appropriately". The subjects received no explanations for the negative performance view, and normative reasons for why an interviewer performs poorly are not as

widely known as explanations concerning two sides of an issue. Because these subjects were students in a laboratory study, they were probably inexperienced and less likely to know the appropriate behaviors necessary for the successful performance of an interviewer.

The view presented in this study simply provided a negative view of the ratee's performance, but did not provide an explanation for that poor performance. The lack of an explanation possibly led to a higher level of ambiguity in the view condition for this study than in the previous studies by Tetlock and his colleagues (1983a, 1989). That is, the level of ambiguity necessary to provoke the full effect of the acceptability heuristic was probably not evident in the current study. The acceptability heuristic allows accountable subjects to avoid unnecessary cognitive work by accepting the position that he/she is most confident will be acceptable to others (i.e. the accountability principles).

In the present study, subjects were probably confident in the decision so they accepted the view, but were not as confident in the justifications they could make to support their decision. Thus biased shifts occurred, but subjects still needed to increase their cognitive effort in order to justify their ratings to the supervisor. Because the view condition did not completely reduce ambiguity, the lack of

effects for cognitive complexity and dimensional accuracy were not surprising.

These results have important implications for Tetlock's research, because the level of the ambiguity of the view has differing effects on cognitive processing. Subjects who received the view were probably looking for differences to justify the given view, thus they participated in biased vigilant processing. The no view subjects were presumably participating in unbiased vigilant processing, which Tetlock has shown to occur. Thus, this explanation is supportive of Tetlock's theory in that high levels of cognitive processing are likely to occur, but the nature of this study could not detect differences in biased versus unbiased vigilant processing.

Post hoc analysis of the cognitive complexity components support the unbiased versus biased vigilant processing explanation. As discussed previously, the cognitive complexity measure is an aggregate of positive and negative differentiations, and integration statements and thus might not be sensitive to the effects of the acceptability heuristic. The lack of findings in the Stamoulis (1993) study for the overall cognitive complexity rating also supports the examination of the components to explain the results, versus the overall rating.

Results from post hoc analyses of the components indicated that the view definitely contributed to a bias in information processing. Accountable subjects who received the view produced significantly less positive differentiations and significantly more negative differentiations. Unaccountable subjects were not affected by the view in terms of the number of positive or negative differentiations stated. Not only do accountable subjects rate the interviewer more severely as shown by the results for the biased shift measure, but they also reported more information that is consistent with the view. Thus, accountable subjects who received the view used biased vigilant processing to produce more negative information while decreasing the amount of positive information reported. Accountable subjects who did not receive the view used unbiased vigilant processing by reporting more information (positive or negative) than unaccountable subjects.

The number of positive and negative differentiations reported is useful in examining biased versus unbiased vigilant processing, but the number of integrations provides additional information about differences in subjects' cognitive processing. The number of integrations can be taken as a more direct measure of cognitive complexity, because integrations represent connections between positive and negative ideas. More integrations indicate that subjects

are able to connect different ideas in multiple, contingent patterns (Tetlock, 1983a).

Results from post hoc analyses of the integration component of cognitive complexity indicate that the amount of integrations reported by subjects was not affected by the view. Subjects in the interpersonal/reward accountability group produced more integrations than the other accountable groups and the no accountability group, regardless of view. The joint effect of interpersonal and reward accountability resulted in a high level of cognitive complexity that was not affected by biases induced by the view. The integration measure may not have the sensitivity to detect differences in biased versus unbiased processing. However, the results indicate that the conjunction of the two types of accountability is a powerful force to increase cognitive processing, whether that processing is biased or unbiased. Future research should attempt to break down these components to determine the extent of the usefulness of combining them into an overall cognitive complexity rating.

The unbiased versus biases information processing explanation is consistent with Tetlock as accountability and the no view would be more likely to increase unbiased vigilant processing, but this study potentially expands the boundaries of the level of ambiguity necessary to result in the occurrence of the full acceptability heuristic. The view

given in this study would be more likely to have an effect in performance appraisal in a field setting, because employees would have a better idea of the behaviors necessary for good or poor performance. Thus the view given in this study might be just as informative as the view presented in Tetlock's research if the view was presented in a field setting. The full extent of the acceptability heuristic might also be evident in a lab study if subjects were initially given a behavioral rating form to emulate the field by showing behaviors that contributed to effective or ineffective performance. The biased versus unbiased hypothesis could then be tested by determining if the reasons given in the behavioral rating form were presented more for view subjects than no view subjects when justifying their ratings. Future research is warranted to determine the level of ambiguity necessary for subjects to unconditionally accept the view.

Individuals appear to be motivated by the need for approval in performance evaluation situations, such that they are motivated to give biased ratings when they have knowledge of performance cues. Tetlock (1985a) suggested that there may be as many types of accountability situations as there are relationships between people. Reward in conjunction with interpersonal accountability provoked higher levels of cognitive processing, as seen by the number of integrations reported by this group. The appropriate type of

accountability needed for accurate performance appraisals warrants further study. Different types of accountability to different accountability principles should also be assessed further to determine if other types of accountability might have multiplicative affects.

Overall, ambiguity was clearly important in the performance evaluation setting. The negative performance view resulted in more unfavorable ratings such that biased shifts occurred, more negative and less positive differentiations were stated, and elevation accuracy levels were lower. Although ambiguity resulted in more vigilant information processing as Tetlock (1985a) suggested, the full extent of the acceptability heuristic was not tapped in this study because the level of ambiguity was higher in the view than was initially presumed. Accountability was necessary to provoke a biased shift when given the view, but the type of accountability did not make a difference. Differences found from the type of accountability might be more evident in the cognitive complexity or dimensional accuracy measure if the view was less ambiguous. Future study is needed to resolve these hypotheses.

If organizations want more accurate ratings, these results suggested providing an ambiguous situation for raters, regardless of accountability. But accountability might still be necessary, because the culture of the

organizations and relationships with the supervisor might provide low ambiguity situations. When environmental cues are more available, accountability may be necessary to influence subjects to be more accurate, specifically a high level of accountability. Tetlock (1985a) assumes that accountability is always naturally affecting individuals in their decision environments. However, the type of accountability that produces the most desired results in different situations still needs to be determined.

Four issues qualify the overall conclusions of this study. The first regards the use of only a negative performance view in this study. This study was a partial replication and expansion of the Stamoulis (1993) study, so the same negative performance view was utilized to compare the results across studies. A negative view produced the desired results, but the use of only a negative performance view limits the generalizations of this study to negative view effects. A positive view or a view not based on the sign of performance might produce different findings. Fandt and Ferris (1990) found a view effect by telling subjects how a supervisor handled a similar incident that subjects were to make decisions about. Future research should build a positive performance view factor into experimental designs, although a positive view may have little effect given the

natural inflation of ratings often found in the literature (Tesser & Rosen, 1975).

Second, only females were used in this study with a male experimenter. Because Stamoulis (1993) found gender effects in his study, females alone were used here to eliminate possible differences due to gender effects. Although the results are useful for comparing to the Stamoulis study, the use of only females limits the generalizations of this study to effects found with females. Future research should incorporate a fully crossed design with males and females used as subjects and experimenters, to more fully understand the effects of accountability on gender.

Third, although accountability effects were found, the manipulation of accountability in this laboratory study may not have been strong. In this study as with the Stamoulis (1993) study, the professor of the undergraduate subjects served as supervisor. The size of the professor's class was large (n = 1200), thus the professor has little one on one contact with students. Although the professor operationalization was affective given the accountability effects found, the preexisting supervisory relationship was probably not similar to that in most organizations. Future research should attempt to incorporate different accountability principles at different levels of

communication with the rater to fully understand overall accountability effects.

Fourth, results from the reward manipulation check scale indicated that reward accountability may not have been manipulated appropriately, although the reliability of the scale was high. When using perceived reward as an independent variable from the reward manipulation scale items, differences were only found for one dependent variable. Cognitive complexity displayed differences in results when using perceived reward versus the reward manipulation.

Although the differences should be examined, two of the items on the reward accountability manipulation check scale were not as specific to the reward manipulation as was originally expected. Discussions with subjects indicated that item 6 and item 7 might not have been specific to the reward manipulation versus the interpersonal manipulation. Although subjects in the reward group could receive more extra credit based on the quality of their ratings, subjects in the interpersonal group also received one extra credit point for participating. These questions both concerned extra credit, which interpersonal accountability subjects could have construed as applying to them as well as reward accountability subjects. If subjects in the interpersonal alone condition misconstrued these items, then their scores

might have contributed to the interaction found for the reward manipulation check scale. These 'bad' items might have resulted in the differences found in the cognitive complexity scores, rather than reward accountability not being manipulated appropriately. Perceived reward only produced a difference in results for one dependent variable, which further supports that the items were probably not specific enough to the distinct type of accountability manipulation. The results found for cognitive complexity when using perceived reward as the independent variable of reward were also not as logical as results found when using the coded reward manipulation. Future research is needed with more specific reward and interpersonal accountability items so that confidence can be placed in the strength of the reward accountability manipulation.

By using a reward only accountability manipulation, this study contributed to a further understanding of the underlying mechanisms of the social context of accountability. More research is needed to reconcile the two types of accountability and their effects on the dependent variables of cognitive complexity and accuracy. These results indicate that they work together to increase cognitive complexity. Tetlock (1985a) suggests that the protection of the self and social image, and the desire to gain material resources are the two major types of

accountability motivators. He also suggests that the two types of accountability are very closely intertwined, and thus may be difficult to disentangle. Accountability is a strong manipulation, but further study is needed to determine the different effects that these two major types of accountability can provide.

The investigation of impression management and influence factors in performance evaluation settings has increased recently (Villanova & Bernardin, 1989). More researchers are attempting to determine the specific situations that motivate an individual to manage their impressions to others. Tetlock's (1985a) model of accountability may provide one such social context that influences the motivations behind the ratings people give. When raters expect that they must justify their view, or they expect that their ratings will be compared to standards, subjects are more inclined to increase their information processing to prepare for critical reactions. As suggested by this study's findings, the level of ambiguity has a strong effect on whether vigilant information processing results in more accurate ratings. The effects of reward and interpersonal accountability may be centered on the type of information processing strategies that are provoked by the differing accountability situations. Studies of cognitive processing in organizations must continue to take into account the differing psychological and

tangible rewards that contribute to the social context and thus to the formation of evaluations.

## References


Baumeister, R. F. (1982). A self-presentational view of social phenomena. Psychological Bulletin, 91(1), 3-26.

Bieri, J., Atkins, A. L., Briar, S., Leaman, R. L., Miller, H., & Tripodi, T. (1966). Clinical and social judgment: The discrimination of behavioral information. NY: John Wiley & Sons.

Boyle, K. A., & Klimoski, R. J. (1992, May). Improving the quality of self-assessments: The role of accountability. Paper presented at the meeting of the Society for Industrial and Organizational Psychology, Montreal.

Cooper, H. M. (1979). Statistically combining independent studies: A meta-analysis of sex differences in conformity research. Journal of Personality and Social Psychology, 37, 131-146.

Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." Psychological Bulletin, 52, 177-193.

DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. (1984). A cognitive view of the performance appraisal process: A model and research propositions. Organizational Behavior and Human Performance, 33, 360-396.

Eagly, A. H., & Carli, L. L. (1981). Sex of researchers and sex-typed communications as determinants of sex

differences in influenceability: A meta-analysis of social influence studies. Psychological Bulletin, 90, 1-20.

Eder, R. W., & Fedor, D. B. (1989). Priming performance self-evaluations: Moderating effects of rating purpose and judgment confidence. Organizational Behavior and Human Decision Processes, 44, 474-493.

Fandt, P. M., & Ferris, G. R. (1990). The management of information and impressions: When employees behave opportunistically. Organizational Behavior and Human Decision Processes, 45, 140-158.

Farh, J., & Werbel, J. D. (1986). Effects of purpose of the appraisal and expectation of validation on self-appraisal leniency. Journal of Applied Psychology, 71, 527-529.

Fox, S., & Dinur, Y. (1988). Validity of self-assessment: A field evaluation. Personnel Psychology, 41, 581-592.

Hagafors, R., & Brehmer, B. (1983). Does having to justify one's judgments change the nature of the judgment process? Organizational Behavior and Human Performance, 31, 223-232.

Hauenstein, N. M. A. (1992). An information-processing approach to leniency in performance judgments. Journal of Applied Psychology, 77, 485-493.

Hauenstein, N. M. A., & Alexander, R. A. (1991). Rating ability in performance judgements: The joint influence of implicit theories and intelligence. Organizational Behavior and Human Decision Processes, 50, 300-323.

Ilgen, D. R., & Feldman, J. M. (1983). Performance appraisal: A process focus. In B. Staw (Ed.), Research in Organizational Behavior (Vol. 5, pp. 141-197). Greenwich CT: JAI Press.

Jones, E., & Pittman, T. (1982). Toward a general theory of strategic self-presentation. In J. Suls (Ed.), Psychological perspective on the self (pp. 231-263). Hillsdale, NJ: Lawrence Erlbaum.

Judge, T. A., & Ferris, G. R. (in press). The social context of performance appraisal evaluation decisions. Academy of Management Journal.

Kelly, G. A. (1955). The psychology of personal constructs. NY: W. W. Norton & Co.

Klimoski, R., & Inks, L. (1990). Accountability forces in performance appraisal. Organizational Behavior and Human Decision Processes, 45, 194-208.

Landy, F. L., & Farr, J. L. (1983). The measurement of work performance: Methods, theory, and applications. Orlando, FL: Academic Press.

Longenecker, C. O. (1992, May). More on the politics of employee appraisals. In R. J. Vance (Chair), Multiple

stakeholders and inevitable conflicts in performance appraisal. Symposium conducted at the meeting of the Society for Industrial and Organizational Psychology, Montreal.

Lord, R. G., & Smith, J. E. (1983). Theoretical, information processing, and situational factors affecting attribution theory models of organizational behavior. Academy of Management Review, 8, 50-60.

March, J. G., & Olsen, J. P. (1979). Ambiguity and Choice in Organizations. Bergen, Norway: Universitetsforlaget.

McCallum, D. M., McCallum, R., & Gurwitch, R. H. (1987). Dimensions of interpersonal events: Reward value and information value. Journal of Experimental Social Psychology, 23, 316-330.

Murphy, K. R., Balzer, W. K., Kellam, K. L., & Armstrong, J. G. (1984). Effects of the purpose of rating on accuracy in observing teach behavior and evaluating teaching performance. Journal of Educational Psychology, 76, 45-54.

Murphy, K. R., & Cleveland, J. N. (1991). Performance Appraisal. Needham Heights, MA: Allyn and Bacon.

Schlenker, B. R., & Weigold, M. F. (1989). Self-identification and accountability. In R. A. Giacalone & P. Rosenfeld (Eds.), Impression Management in the

Organization. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

Schroder, H. M., Driver, M. J., & Streufert, S. (1967). Human information processing: Individuals and groups functioning in complex social situations. New York: Holt, Rinehart, and Winston.

Schroder, H. M., & Seudfeld, P. (Eds.). (1971). Personality theory and information processing. New York: Ronald Press.

Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. Psychological Review, 96, 58-83.

Stamoulis, D. T. (1993). Making raters more accountable for their performance ratings: effects of expecting a supervisory review of ratings. Unpublished doctoral dissertation.

Streufert, S., & Streufert, S. C. (1978). Behavior in the complex environment. Washington, D. C.: H. Winston & Sons.

Sulsky, L. M., & Balzer, W. K. (1988). Meaning and measurement of performance rating accuracy: some methodological and theoretical concerns. Journal of Applied Psychology, 73(3), 497-506.

Tesser, A., & Rosen, S. (1975). The reluctance to transmit bad news. In L. Berkowitz (Ed.), Advances in

experimental social psychology (Vol. 8, pp. 193-232). NY: Academic Press.

Tetlock, P. E. (1983a). Accountability and complexity of thought. Journal of Personality and Social Psychology, 45, 74-83.

Tetlock, P. E. (1983b). Accountability and the perseverance of first impressions. Social Psychology Quarterly, 46, 285-292.

Tetlock, P. E. (1985a). Accountability: The neglected social context judgment and choice. In B. Staw & L. Cummings (Eds.), Research in organizational behavior (Vol. 1, pp. 297-332). Greenwich, CT: JAI Press.

Tetlock, P. E. (1985b). Accountability: A social check on the fundamental attribution error. Social Psychology Quarterly, 48, 227-236.

Tetlock, P. E., & Kim, J. I. (1987). Accountability and judgment processes in a personality prediction task. Journal of Personality and Social Psychology, 52, 700-709.

Tetlock, P. E., & Suedfeld, P. (1988). Integrative complexity coding of verbal behavior. In C. Antaki (Ed.), Analysing Everyday Explanation. London: Sage Publications.

Villanova, P. & Bernardin, H. J. (1989). Impression management in the context of performance appraisal. In

R. A. Giacalone & P. Rosenfeld (Eds.), Impression Management in the Organization. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

Weldon, E., & Gargano, G. M. (1985). Cognitive effort in additive task groups: The effects of shared responsibility on the quality of multiattribute judgments. Organizational Behavior and Human Decision Processes, 36, 348-361.

Williams, K. J., Cafferty, T. P., & DeNisi, A. S. (1990). The effect of performance appraisal salience on recall and ratings. Organizational Behavior and Human Decision Processes, 46, 217-239.

Zedeck, S., & Cascio, W. F. (1982). Performance appraisal decisions as a function of rater training and purpose of the appraisal. Journal of Applied Psychology, 67, 752-758.

Table 1

Correlations Amongst Manipulation Check Items

| | Int/1 | Int/2 | Int/3 | Int/4 | Rew/1 |
|---|---|---|---|---|---|
| Int/Item1 | | | | | |
| Int/Item2 | .77** | | | | |
| Int/Item3 | .59** | .67** | | | |
| Int/Item4 | .56** | .60** | .73** | | |
| Rew/Item1 | .20** | .19** | .24** | .38** | |
| Rew/Item2 | .14 | .10 | .15* | .26** | .38** |
| Rew/Item3 | .12 | .06 | .26** | .25** | .19** |
| Rew/Item4 | .22** | .26** | .32** | .36** | .43** |
| View/Item1 | .15* | .14 | .20** | .17* | -.03 |
| View/Item2 | .18* | .24** | .27** | .27** | .11 |
| View/Item3 | .08 | .15* | .18* | .23** | .12 |

| | Rew/2 | Rew/3 | Rew/4 | View/1 | View/2 | View/3 |
|---|---|---|---|---|---|---|
| Int/Item1 | | | | | | |
| Int/Item2 | | | | | | |
| Int/Item3 | | | | | | |
| Int/Item4 | | | | | | |
| Rew/Item1 | | | | | | |
| Rew/Item2 | | | | | | |
| Rew/Item3 | .58** | | | | | |
| Rew/Item4 | .39** | .44** | | | | |
| View/Item1 | .11 | .30** | .10 | | | |
| View/Item2 | .15* | .30** | .26** | .57** | | |
| View/Item3 | .21** | .30** | .32** | .43** | .51** | |

Note. *p < .05, **p < .01. Int=Interpersonal Accountability, Rew=Reward Accountability.

Table 2

Source Table for ANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on the Interpersonal Accountability Scale

| Source | df | Mean Square | F Value |
|---|---|---|---|
| Within | 183 | 1.35 | |
| Reward Acct | 1 | 10.43 | 7.74** |
| Interpersonal Acct | 1 | 414.83 | 308.07*** |
| View | 1 | 1.10 | .82 |
| Reward Acct X Interpersonal Acct | 1 | 5.44 | 4.04* |
| Reward Acct X View | 1 | .01 | .01 |
| Interpersonal Acct X View | 1 | .03 | .02 |
| Reward Acct X Interpersonal Acct X View | 1 | .05 | .04 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 3

Source Table for ANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on the Reward Accountability Scale

| Source | df | Mean Square | F Value |
|---|---|---|---|
| Within | 183 | 1.64 | |
| Reward Acct | 1 | 63.33 | 38.63*** |
| Interpersonal Acct | 1 | 2.99 | 1.82 |
| View | 1 | .05 | .03 |
| Reward Acct X Interpersonal Acct | 1 | .17 | .10 |
| Reward Acct X View | 1 | 9.71 | 5.92* |
| Interpersonal Acct X View | 1 | .87 | .53 |
| Reward Acct X Interpersonal Acct X View | 1 | 1.78 | 1.09 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 4

Source Table for ANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on the View Scale

| Source | df | Mean Square | F Value |
|---|---|---|---|
| Within | 183 | 2.15 | |
| Reward Acct | 1 | 5.76 | 2.69 |
| Interpersonal Acct | 1 | 8.13 | 3.79 |
| View | 1 | 202.49 | 94.35*** |
| Reward Acct X Interpersonal Acct | 1 | 3.06 | 1.43 |
| Reward Acct X View | 1 | .00 | .00 |
| Interpersonal Acct X View | 1 | .12 | .06 |
| Reward Acct X Interpersonal Acct X View | 1 | .70 | .33 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 5

Correlations Amongst Rating Dimensions

| | Rap. | Org. | Que. | Rel. | Job. | Ans. | Overall |
|---|---|---|---|---|---|---|---|
| Rap. | | | | | | | |
| Org. | .27 | | | | | | |
| Que. | .39 | .61 | | | | | |
| Rel. | .28 | .58 | .60 | | | | |
| Job. | .42 | .55 | .58 | .55 | | | |
| Ans. | .50 | .50 | .64 | .50 | .66 | | |
| Overall | .57 | .65 | .75 | .67 | .67 | .75 | |

Note. All correlations are significant at $p < .01$. Rap.= Rapport building, Org.= Organization of interview, Que.= Questioning skill, Rel.= Relevance of questions, Job.= Job and company preview, Ans.= Answering the applicant, Overall= Overall evaluation.

Table 6

Within-Cell Means for Rating Dimensions for View and No View

View

| | Rew/Int n=26 | Rew n=26 | Int n=21 | No Acct n=24 |
|---|---|---|---|---|
| Rap. | 3.50 | 3.46 | 4.38 | 4.17 |
| Org. | 2.92 | 2.50 | 2.71 | 3.67 |
| Que. | 2.92 | 3.00 | 3.29 | 3.75 |
| Rel. | 3.54 | 3.62 | 2.91 | 4.21 |
| Job. | 3.50 | 3.92 | 3.43 | 4.83 |
| Ans. | 3.69 | 3.54 | 3.81 | 4.33 |
| Overall | 3.04 | 3.00 | 2.86 | 3.88 |

No View

| | Rew/Int n=23 | Rew n=25 | Int n=24 | No Acct n=22 |
|---|---|---|---|---|
| Rap. | 4.30 | 4.48 | 4.17 | 4.55 |
| Org. | 3.91 | 3.48 | 3.75 | 3.64 |
| Que. | 3.96 | 4.00 | 3.92 | 3.64 |
| Rel. | 4.78 | 4.40 | 4.92 | 4.36 |
| Job. | 4.96 | 4.84 | 4.63 | 5.18 |
| Ans. | 4.35 | 5.00 | 4.63 | 5.05 |
| Overall | 4.04 | 4.16 | 3.79 | 4.18 |

Note. Rew=Reward Accountability, Int=Interpersonal Accountability, No Acct=No Accountability; Rap.=Rapport building, Org.= Organization of interview, Que.=Questioning skill, Rel.= Relevance of questions, Job.=Job and company preview, Ans.= Answering the applicant, Overall=Overall evaluation.

Table 7

Source Table for MANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on the Six Specific Rating Dimensions

| Source | df | F Value |
|---|---|---|
| Within | 178 | |
| Reward Acct | 6 | .77 |
| Interpersonal Acct | 6 | 1.57 |
| View | 6 | 4.78*** |
| Reward Acct X Interpersonal Acct | 6 | 1.39 |
| Reward Acct X View | 6 | 1.33 |
| Interpersonal Acct X View | 6 | 2.48* |
| Reward Acct X Interpersonal Acct X View | 6 | .73 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 8

A priori comparison of interpersonal/reward accountability, interpersonal accountability, and reward accountability view groups to other experimental groups on the Six Specific Rating Dimensions

| | df | t value |
|---|---|---|
| Rapport building | 183 | -2.44* |
| Organization of interview | 183 | -4.45*** |
| Questioning skill | 183 | -3.45*** |
| Relevance of questions | 183 | -5.01*** |
| Job and company preview | 183 | -5.35*** |
| Answering the applicant | 183 | -4.37*** |

Note. *p < .05, **p < .01, ***p < .001.

Table 9

Source Table for ANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on the Overall Rating Dimension

| Source | df | Mean Square | F Value |
|---|---|---|---|
| Within | 183 | 1.70 | |
| Reward Acct | 1 | .64 | .37 |
| Interpersonal Acct | 1 | 6.56 | 3.85* |
| View | 1 | 34.45 | 20.23*** |
| Reward Acct X Interpersonal Acct | 1 | 5.25 | 3.08 |
| Reward Acct X View | 1 | 2.53 | 1.49 |
| Interpersonal Acct X View | 1 | .66 | .39 |
| Reward Acct X Interpersonal Acct X View | 1 | 1.82 | 1.07 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 10

Interrater Correlations for Cognitive Complexity Rating and Its Components

| | |
|---|---|
| Cognitive Complexity Rating | .84 |
| Positive Differentiations | .89 |
| Negative Differentiations | .87 |
| Integration Statements | .55 |

Table 11

Within-Cell Means for Cognitive Complexity and Its Components for View and No View

View

| | Rew/Int n=26 | Rew n=26 | Int n=21 | No Acct n=24 |
|---|---|---|---|---|
| CCRat | 3.69 | 2.65 | 1.91 | 2.75 |
| PosDif | 1.12 | 0.85 | 0.67 | 1.33 |
| NegDif | 2.89 | 3.08 | 2.67 | 1.96 |
| Integ | 0.39 | 0.08 | 0.05 | 0.08 |

No View

| | Rew/Int n=23 | Rew n=25 | Int n=24 | No Acct n=22 |
|---|---|---|---|---|
| CCRat | 3.39 | 3.44 | 2.96 | 3.00 |
| PosDif | 1.57 | 1.84 | 1.29 | 1.41 |
| NegDif | 2.22 | 2.24 | 2.50 | 2.05 |
| Integ | 0.26 | 0.16 | 0.08 | 0.09 |

Note. Rew=Reward Accountability, Int=Interpersonal Accountability, No Acct=No Accountability; CCRat=Cognitive Complexity Rating, PosDif=Positive Differentiations, NegDif= Negative Differentiations, Integ=Integration Statements.

Table 12

Source Table for ANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on the Overall Cognitive Complexity Rating

| Source | df | Mean Square | F Value |
|---|---|---|---|
| Within | 183 | 2.59 | |
| Reward Acct | 1 | 19.52 | 7.53** |
| Interpersonal Acct | 1 | .03 | .01 |
| View | 1 | 9.50 | 3.67 |
| Reward Acct X Interpersonal Acct | 1 | 10.46 | 4.03* |
| Reward Acct X View | 1 | 1.99 | .77 |
| Interpersonal Acct X View | 1 | .24 | .09 |
| Reward Acct X Interpersonal Acct X View | 1 | 10.61 | 4.09* |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 13

Regression Analysis Assessing Interpersonal Accountability, Perceived Reward Accountability, and View Effects on the Overall Cognitive Complexity Rating

| Variable | B | T Value |
|---|---|---|
| Perceived Reward Acct | .18 | 2.17* |
| Interpersonal Acct | -.92 | -2.63** |
| View | -.36 | -1.02 |
| Perceived Reward Acct X Int Acct | .24 | 2.80** |
| Perceived Reward Acct X View | .04 | .48 |
| Interpersonal Acct X View | .06 | .49 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability, Int=Interpersonal.

Table 14

Within-Cell Means for Elevation and Dimensional Accuracy for View and No View

| | Rew/Int n=26 | Rew n=26 | Int n=21 | No Acct n=24 |
|---|---|---|---|---|
| **View** | | | | |
| Elevation | 1.55 | 1.37 | 1.63 | 1.08 |
| Dimensional | 1.32 | 1.29 | 1.32 | 1.33 |
| **No View** | | | | |
| | Rew/Int n=23 | Rew n=25 | Int n=24 | No Acct n=22 |
| Elevation | 0.92 | 0.74 | 0.90 | 0.81 |
| Dimensional | 1.27 | 1.40 | 1.32 | 1.37 |

Note. Rew=Reward Accountability, Int=Interpersonal Accountability, No Acct=No Accountability.

Table 15

Source Table for ANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on Elevation Accuracy

| Source | df | Mean Square | F Value |
|---|---|---|---|
| Within | 183 | .48 | |
| Reward Acct | 1 | .07 | .15 |
| Interpersonal Acct | 1 | 2.94 | 6.14* |
| View | 1 | 15.22 | 31.81*** |
| Reward Acct X Interpersonal Acct | 1 | .22 | .47 |
| Reward Acct X View | 1 | .20 | .41 |
| Interpersonal Acct X View | 1 | .64 | 1.34 |
| Reward Acct X Interpersonal Acct X View | 1 | .68 | 1.43 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 16

Source Table for ANOVA Assessing Interpersonal Accountability, Reward Accountability, and View Effects on Dimensional Accuracy

| Source | df | Mean Square | F Value |
|---|---|---|---|
| Within | 183 | .07 | |
| Reward Acct | 1 | .01 | .12 |
| Interpersonal Acct | 1 | .07 | .92 |
| View | 1 | .04 | .55 |
| Reward Acct X Interpersonal Acct | 1 | .00 | .06 |
| Reward Acct X View | 1 | .00 | .00 |
| Interpersonal Acct X View | 1 | .11 | 1.60 |
| Reward Acct X Interpersonal Acct X View | 1 | .04 | .57 |

Note. *p < .05, **p < .01, ***p < .001.
Acct=Accountability.

Table 17

Correlations Amongst the Cognitive Complexity Variables

| | PosDif | NegDif | Integ | CCRat |
|---|---|---|---|---|
| PosDif | | | | |
| NegDif | -.57** | | | |
| Integ | .09 | -.03 | | |
| CCRat | .38** | .03 | .75** | |

Note. *p < .05, **p < .01. PosDif=Positive Differentiations, NegDif=Negative Differentiations, Integ=Integration Statements, CCRat=Cognitive Complexity Rating.

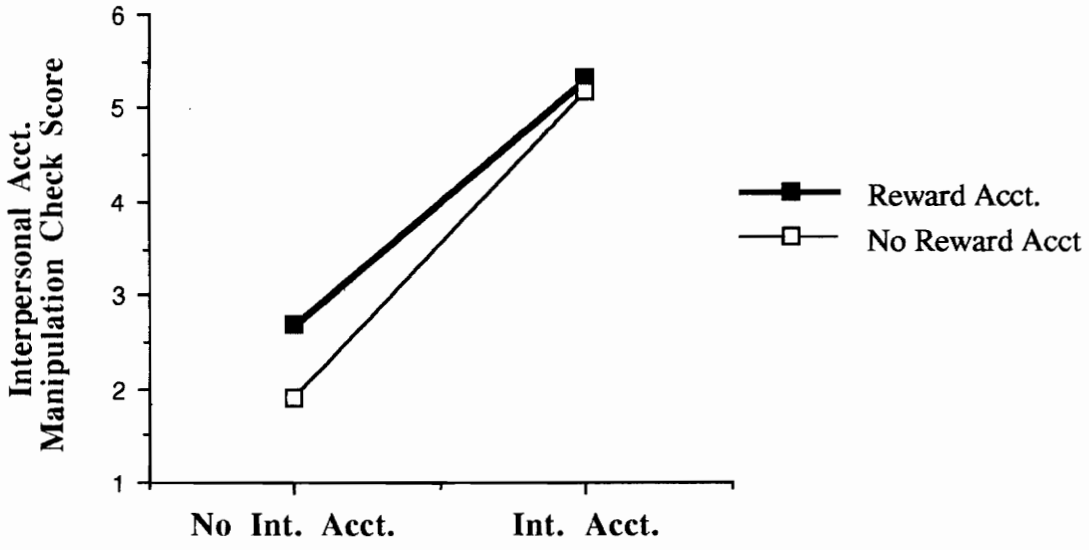



**Figure 1: Interpersonal accountability by reward accountability interaction on the interpersonal accountability manipulation check scale.**

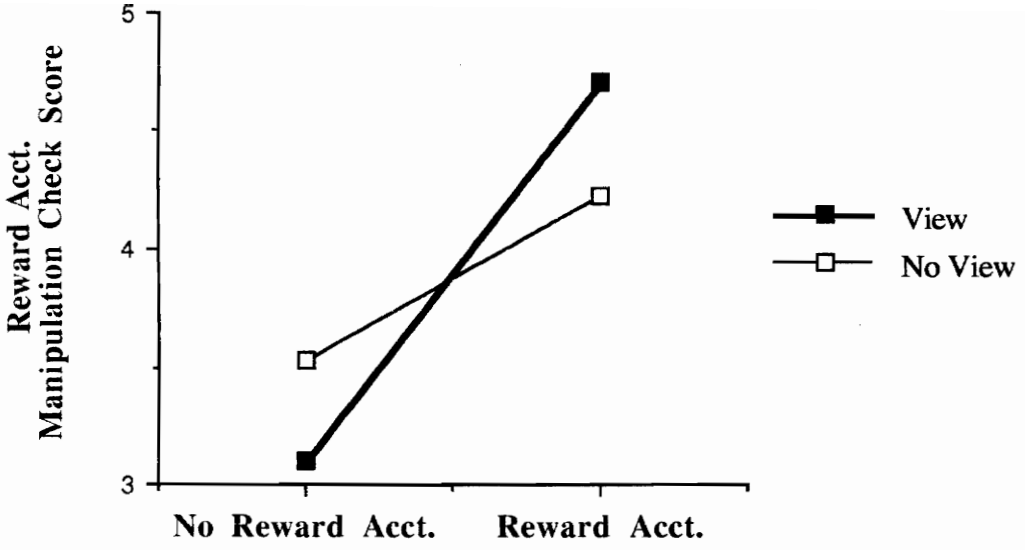



**Figure 2: Reward accountability by view interaction on the reward accountability manipulation check scale.**

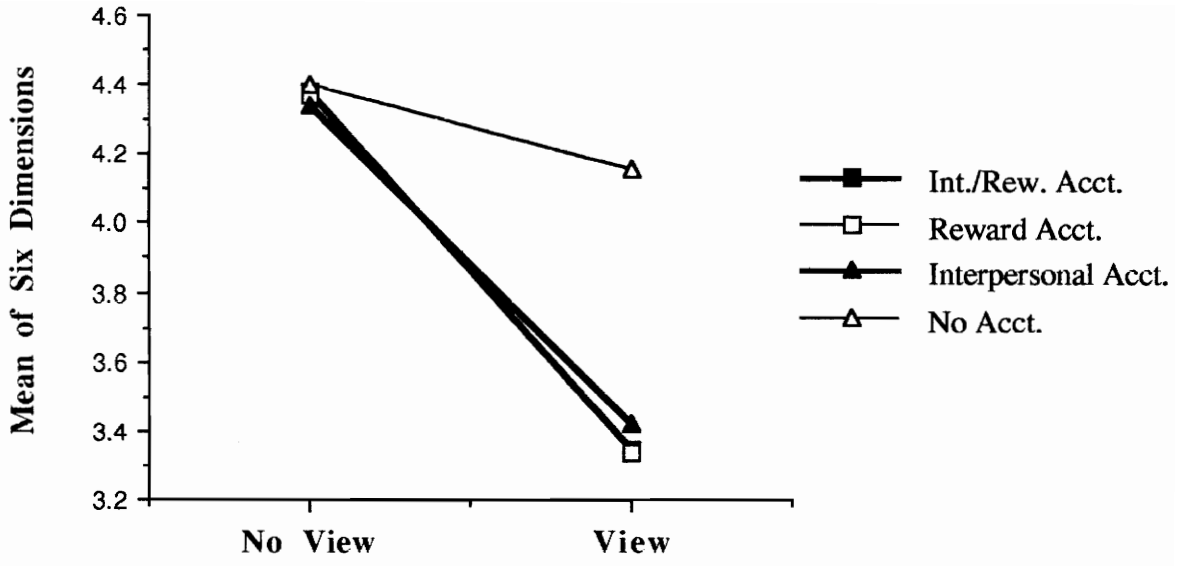



**Figure 3: Interpersonal accountability by view interaction on the mean of the six individual dimensions.**

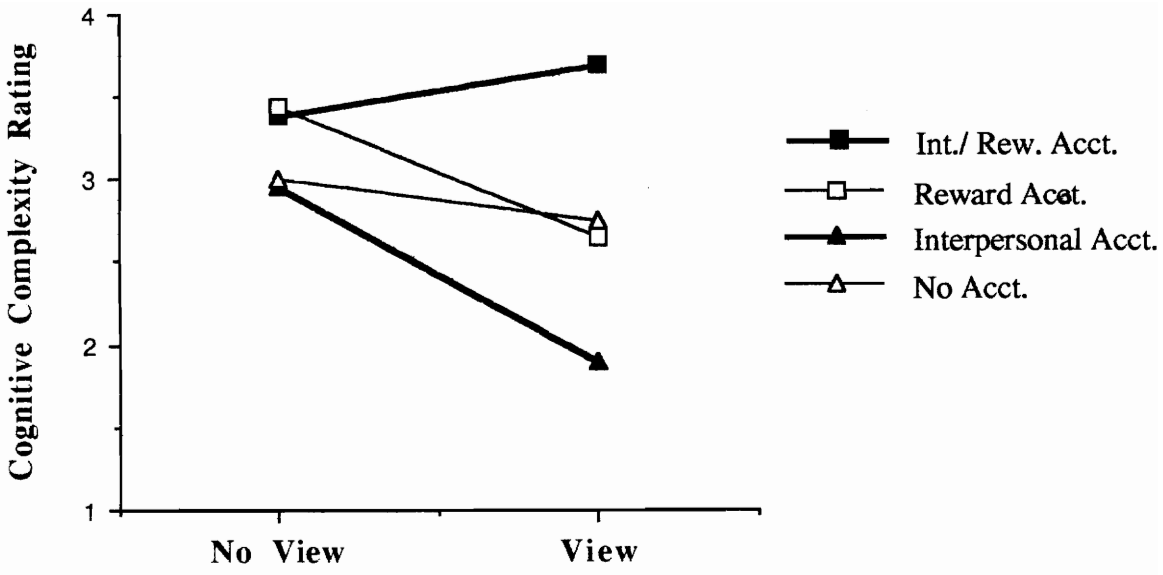



**Figure 4: Interpersonal accountability by reward accountability by view interaction on the overall cognitive complexity rating.**

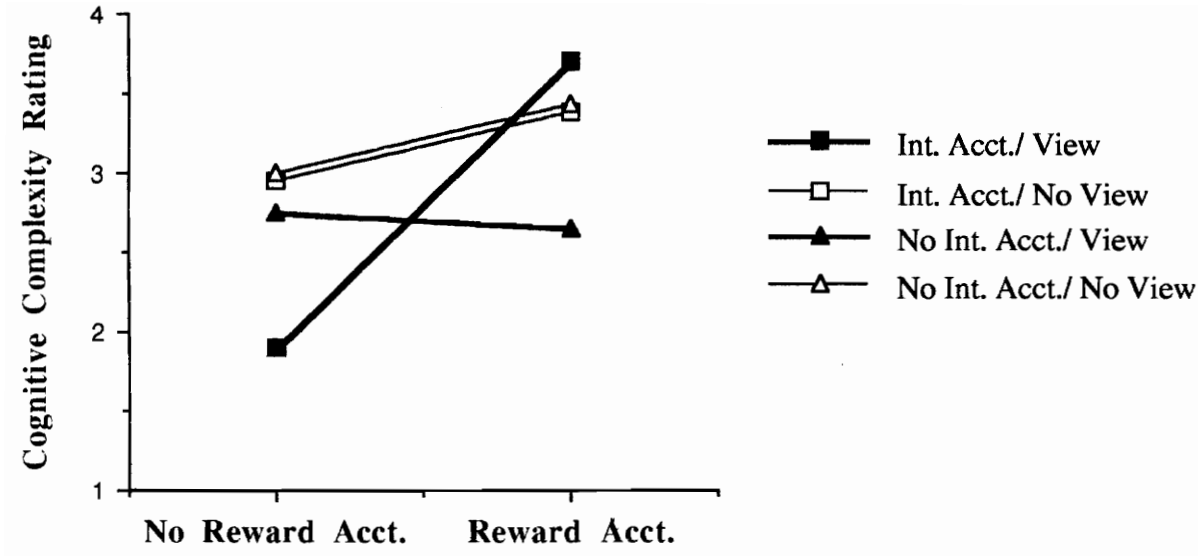



**Figure 5: Interpersonal accountability by reward accountability by view interaction on the overall cognitive complexity rating with reward as the independent variable on the x-axis.**

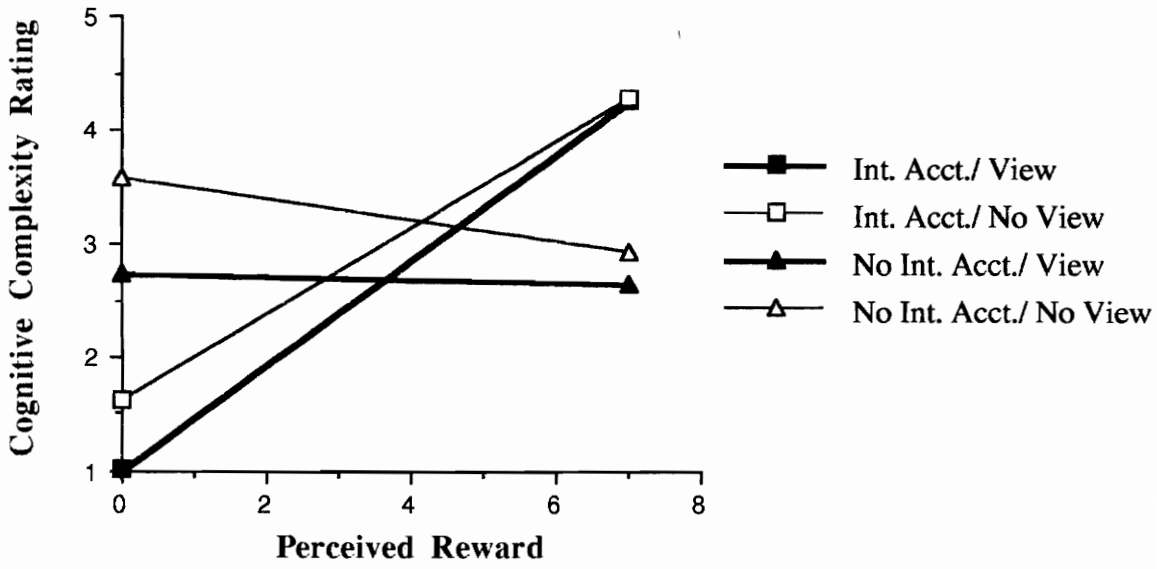



**Figure 6: Interpersonal accountability by perceived reward accountability interaction on the overall cognitive complexity rating with perceived reward as the independent variable on the x-axis.**

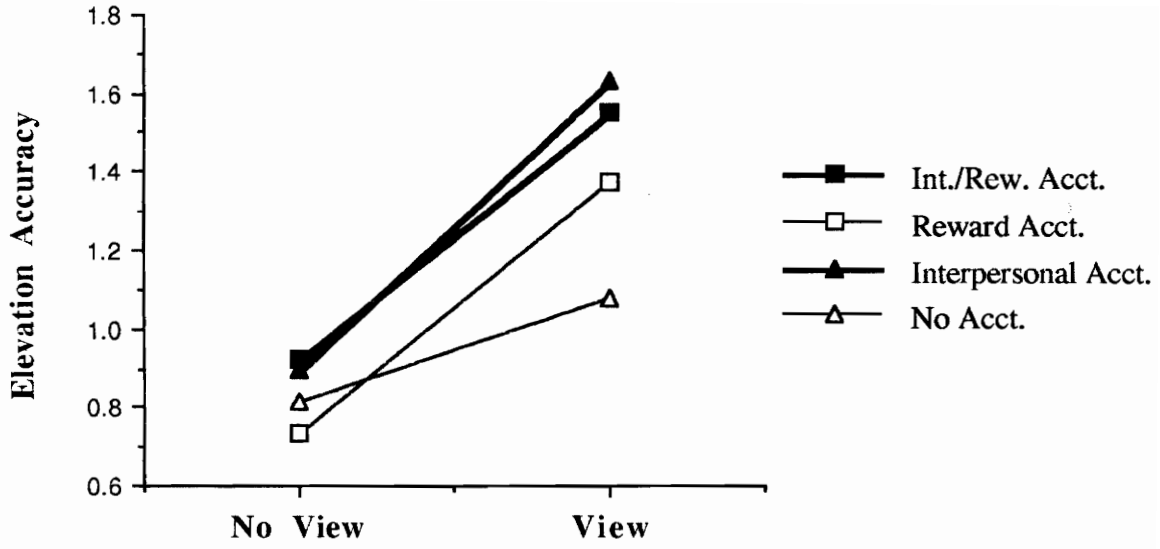



**Figure 7: View and interpersonal accountability effects on elevation accuracy.**

## Appendix A

**Interpersonal accountability**

Subjects in the interpersonal accountability condition will be told:

"This experiment is being conducted by Dr. ____, who you all may know as the instructor from your ____ class. Dr. ____ has written many journal articles about the cognitive aspects of performance evaluation, and has had much success with past research. As part of the study, you will be asked to meet with Dr. ____ for 10 minutes, at your convenience in the next 5 weekdays. In this meeting, you will be asked to explain and justify your performance ratings to Dr. ____. That is, this meeting will involve you telling Dr. ____ why you rated the interviewer the way you did. Your 1 extra credit point will only be given after the satisfactory completion of this meeting. He expects clear-cut statements from you concerning your reasons for rating the interviewer the way you did.

Also, Dr. ____ will be judging the quality of your ratings. Based on the quality of your ratings, Dr. ____ will give you feedback and tell you how well you rated the interviewer at the end of the meeting. Parts of this meeting may make you feel uncomfortable, because Dr. ____'s questions may require you to put a lot of effort in your answers. These meetings will be audiotaped so that Dr. ____ can

perform further analyses of your explanations. In order to make things easier for you, the meetings will be held next door in Derring 5077. I will hand out a sheet so that you can tell Dr. ____ when you are available to come in, and so you can give Dr. ____ your phone number so that he can reach you to set up a meeting."

**Appendix B**
**Subject Sign-Up Information-A**

FOR: REVIEW MEETING OF PERFORMANCE RATINGS

Your name: _____

Your Social Security Number: _____

Phone Number: (Home) _____

(Work phone, if applicable) _____

Potential Meeting Times (Indicate 6 10-minute slots in which you are available in the next 5 weekdays, Mon.- Fri., 7:30 a.m. - 5:30 p.m. Dr. _____ will call you and confirm 1 of the 6 meeting slots you requested.)

1. _____

2. _____

3. _____

4. _____

5. _____

6. _____

**Remember:** -Be on time for your review meeting.

-Review meetings will take place in Derring Hall. Rm. 5077

## Appendix C

**Reward accountability**

Subjects in the reward accountability condition will be told:

"This experiment is being conducted by Dr. ____, who you all know as the instructor from your ____ class. Dr. ____ has written many journal articles about the cognitive aspects of performance evaluation, and has had much success with past research. As part of the study, your ratings will be compared and evaluated in the next 5 weekdays according to a set of true performance dimension scores for the ratee. The true scores for the ratee were determined by Dr. ____ and in agreement with a number of other expert raters. Dr. ____ will view your scores compared to the standard. Your ratings will be reviewed according to the standards, and based on the quality of your ratings, Dr. ____ may award you 0, 1, or 2 additional extra credit points based on how well you rated the interviewer.

The number of extra credit points you receive will be strictly confidential. Because of the confidentiality, you will receive a sealed envelope concerning the number of extra credit points that you have been awarded based on the quality of your ratings. In order to make things easier for you, I will have the envelopes available next door in Derring 5077.

I will hand out a sheet so that you can tell me when you are available to come in to receive your extra credit points."

## Appendix D
## Subject Sign-Up Information-B

FOR: EXTRA CREDIT POINTS

Your name: _____

Your Social Security Number: _____

Phone Number: (Home) _____

(Work phone, if applicable) _____

Potential Meeting Times (Indicate 6 10-minute slots in which you can pick up your envelope in the next 5 weekdays, Mon.- Fri., 7:30 a.m. - 5:30 p.m. I will call you and confirm 1 of the 6 times you requested.)

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____

**Remember:**

-Be on time when you come to pick up the extra credit points that you might receive.

-A graduate student will have the envelopes in Derring Hall, Rm. 5077

## Appendix E

**Interpersonal/Reward Accountability**

Subjects in the interpersonal/reward accountability condition will be told:

"This experiment is being conducted by Dr. ____, who you all know as the instructor from your ____ class. Dr. ____ has written many journal articles about the cognitive aspects of performance evaluation, and has had much success with past research. As part of the study, you will be asked to meet with Dr. ____ for 10 minutes, at your convenience in the next 5 weekdays. In this meeting, you will be asked to explain and justify your performance ratings to Dr. ____. That is, this meeting will involve you telling Dr. ____ why you rated the interviewer the way you did. Your 1 extra credit point will only be given after the satisfactory completion of this meeting. He expects clear-cut statements from you concerning your reasons for rating the interviewer the way you did.

As part of the study, your ratings will also be compared and evaluated according to a set of true performance dimension scores for the ratee. The true scores for the ratee were determined by Dr. ____ and in agreement with a number of other expert raters. Your ratings will be reviewed according to the standards, and based on the quality of your ratings, Dr. ____ will give you feedback and tell you how well you rated the interviewer at the end of the meeting.

Dr. ____ may award you 0, 1, or 2 additional extra credit points based on how well you rated the interviewer.

Parts of this meeting may make you feel uncomfortable, because Dr. ____'s questions may require you to put a lot of effort in your answers. These meetings will be audiotaped so that Dr. ____ can perform further analyses of your explanations. The number of extra credit points you receive will be strictly confidential. Because of the confidentiality, you will receive a sealed envelope concerning the number of extra credit points that you have been awarded based on the quality of your ratings. In order to make things easier for you, the meetings will be held next door in Derring 5077 and the envelopes will also be distributed. I will hand out a sheet so that you can tell Dr. ____ when you are available to come in, and so you can give Dr. ____ your phone number so that he can reach you to set up a meeting."

## Appendix F

**View**

"I can tell you now, that several times, Dr. ____ has rated all of the interviewer performance you will see in today's session. Dr. ____ considers the performance set to be poor or ineffective, overall, because the interviewer does not perform his job duties and responsibilities appropriately. Dr. ____ rated the interviewer performance you are about to see as poor or ineffective."

## Appendix G
## PERFORMANCE RATING FORM

For each performance category, circle the number that most reflects your evaluation of the interviewer.

1. RAPPORT BUILDING: The extent to which the interviewer put the applicant at ease, built trust, and was socially desirable.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Poor | | | Average | | | Excellent |

2. ORGANIZATION OF INTERVIEW: The extent to which the interviewer had a clear format which was followed.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Poor | | | Average | | | Excellent |

3. QUESTIONING SKILL: The extent to which the questions were tactfully presented.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Poor | | | Average | | | Excellent |

4. RELEVANCE OF QUESTIONS: How useful the interviewer's questions were for assessing the knowledge, skills, abilities, and characteristics of the applicant.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Poor | | | Average | | | Excellent |

5. JOB AND COMPANY PREVIEW: The extent to which knowledge concerning the company and the job which the applicant has applied is appropriately conveyed.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Poor | | | Average | | | Excellent |

6. ANSWERING THE APPLICANT: The demonstrated willingness for answering the applicant's questions and the extent to which the responses provided meaningful information.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Poor | | | Average | | | Excellent |

7. OVERALL EVALUATION: Your general impression of the interviewer's overall performance in the interview.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Poor | | | Average | | | Excellent |

## Appendix H
## THOUGHTS SELF-REPORT

Write (neatly) your thoughts concerning the effectiveness of the interviewer's performances. If you need more space, use the back of this sheet.

## Appendix I
## EVALUATION COMPLEXITY SCALE

1. SOLELY positive-type OR negative-type of information about the interviewer, with 0 - 3 differentiations for that type of information

2. SOLELY positive-type OR negative-type of information about the interviewer, with 4 or more differentiations for that type of information

3. Positive-type AND negative-type of information about the interviewer, (some examples include contrast amongst vignettes, use of "although", "though", "however", etc.) with 1 type of information having 1 functionally differentiation and the other type of information having 1 or more differentiations

4. Positive-type AND negative-type of information about the interviewer, (some examples include contrast amongst vignettes, use of "although", "though", "however", etc.) with 1 type of information having 2 differentiations and the other type of information having 2 or more differentiations

5. Positive-type AND negative-type of information about the interviewer, (some examples include contrast amongst vignettes, use of "although", "though", "however", etc.) with 1 type of information having 3 or more differentiations and the other type of information having 3 or more differentiations

6. Positive-type AND negative-type of information about the interviewer, (some examples include contrast amongst vignettes, use of "although", "though", "however", etc.) with 1 type of information having 1 or more differentiations and the other type of information having 2 or more differentiations, AND subjects use 1 sentence to comment on THEIR theory of interviewing that helps to integrate their use of positive and negative types of information (some examples include use of "should", "need")

7. Positive-type AND negative-type of information about the interviewer, (some examples include contrast amongst vignettes, use of "although", "though", "however", etc.) with 1 type of information having 1 or more differentiations and the other type of information having 2 or more differentiations, AND subjects use 2 or more sentences to comment on THEIR theory of interviewing that helps to integrate their use of positive and negative types of information (some examples include use of "should", "need")

## Appendix J
## POST-EXPERIMENT QUESTIONNAIRE

Read each question carefully, and then please respond truthfully. Some questions may not apply to you. Circle your response to each question.

1. To what extent do you feel that a meeting at a later date will take place between you and the person who is the head of this research project?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

2. To what extent do you feel that you will have to explain and justify your performance evaluation ratings at a later date?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

3. To what extent did you think about an upcoming meeting with the person who is the head of this research project while you were observing the interviewer's performance?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

4. To what extent were you concerned about how good of an impression you would make in a future meeting with the person who is the head of this research project?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

5. To what extent were you concerned about the quality of the ratings you gave about the interviewer?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

6. To what extent were you concerned about how much extra credit you might earn as a result of the accuracy of the ratings you gave about the interviewer?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

7. To what extent did you think about the number of extra credit points you might receive while you were observing the interviewer's performance?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

8. To what extent were you concerned about the comparisons of your ratings to the true dimension standards while you were observing the interviewer's performance?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

9. How certain are you that you heard that the interviewer performance was poor or ineffective during the course of today's research session?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

10. To what extent do you believe that an opinion offered about the quality of the interviewer's performance was Dr. _____'s?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

11. To what extent did you consider knowledge that the interviewer performance was described as poor in your performance evaluation ratings?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| to no extent | | | to some extent | | | to a great extent |

**DIANA K. CORRIGAN**

**Home Address**
207 Jefferson Street
Blacksburg, Virginia 24060
(703) 552-9238

**Office Address**
Department of Psychology
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia 24061
(703) 231-6581

**EDUCATION**

**Master of Science** January 1994
Industrial/Organizational Psychology
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

**Bachelor of Arts** June 1991
Psychology
University of Georgia
Athens, Georgia

**PROFESSIONAL AND RESEARCH EXPERIENCE**

8/93 - present **Thesis Research**
Chairperson: Neil M. A. Hauenstein
*Virginia Polytechnic Institute and State University*, Blacksburg, VA
- Designed and conducted a laboratory study which investigated the effects of accountability on raters' expectations of a supervisory review.

3/92 - 5/92 **Research Coordinator**
*Virginia Polytechnic Institute and State University*, Blacksburg, VA
- Organized and administered laboratory study for Robert T. Brill's dissertation concerning the effects of job knowledge, observational purpose, and judgment complexity on recall and rating ability.

11/91 - 1/92 **Selection Test Administrator**
*Bell Atlantic*, Selection Research, Arlington, VA
- Administered computerized selection battery for concurrent validation study.

6/91 - 8/91 **Research Internship**
*Life Office Management Association,*
Atlanta, GA
- Analyzed, edited, and coded data using SAS, collected for the annual salary survey that I/O Psychologists at LOMA distribute as a non-profit service for insurance companies nationwide.

9/90 - 3/91 **Research Assistant**
*Dr. Karl Kuhnert,* University of Georgia,
Athens, GA
- Assisted in the analysis and interpretation of a survey concerning internship opportunities for I/O Psychology graduate students using SPSS.

**TEACHING EXPERIENCE**

8/93 - 12/93 **Instructor of Psychological Measurement Laboratory**
*Virginia Polytechnic Institute and State University*
- Taught basic statistical techniques, such as correlation and regression, on the PC Windows version of SAS.

8/92 - 5/93 **Graduate Teaching Assistant**
*Virginia Polytechnic Institute and State University*
- Assisted I/O Professors in conducting the undergraduate Introductory Psychology course. Graded tests and assignments, developed test questions, and aided students in understanding issues in I/O Psychology.

8/91 - 5/91 **Graduate Teaching Assistant**
*Virginia Polytechnic Institute and State University*
- Taught five undergraduate Introductory Psychology laboratory classes which are in conjunction with the Introductory Psychology lecture course. Prepared lecture, discussion and test questions each week. Guided and promoted discussion of issues in psychology through a planned outline of instruction.

**PROFESSIONAL ORGANIZATIONS AND ACTIVITIES**

Student Affiliate, American Psychological Association
Student Affiliate, Society of Industrial/Organization Psychology
Student Affiliate, Virginia Psychological Association
Student Affiliate, Atlanta Society of Applied Psychology
Member, Psi Chi

**PRESENTATIONS/POSTERS**

Brill, R. T., Hauenstein, N. M. A., & Corrigan, D. K. (April, 1994). The influence of job knowledge and complexity on free recall and performance ratings accuracy. Poster to be presented at the 65th annual meeting of the Eastern Psychological Association, Providence, Rhode Island.

**RELEVANT COURSEWORK**

Content Courses:
Industrial Psychology (I & II)
Organizational Psychology I (Motivation)
Organizational Psychology II (Leadership & Organizational Theory)
Social Cognition/Decision Making Seminar
Social Psychology
Learning

Research Methods and Statistics:
Research Methods
Quantitative Topics in Industrial/Organizational Psychology
Statistics for Social Science Research (I & II)
Regression Analysis
Multivariate Analysis

**REFERENCES** Available upon request.

Diana K. Corrigan