7-1-2007

The Effect of Evaluator's Mood and Type of Accountability on Performance Appraisal Evaluations: A Study of the Affect Infusion Model

Jason E. Gerlt
University of Nebraska at Omaha

The Effect of Evaluator's Mood and Type of Accountability on Performance Appraisal Evaluations: A Study of the Affect Infusion Model

A Thesis

Presented to the

Department of Psychology

and the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment

of the requirements for the Degree

Master of Arts in Psychology

by

Jason E. Gerlt

July, 2007

Thesis Acceptance

Acceptance for the faculty of the Graduate College,
University of Nebraska, in partial fulfillment of the
requirements for the degree Master of Arts,
University of Nebraska at Omaha.

Committee

Jessica Anderson, Ph.D.

James M. Thomas

William T. Cluto

Chairperson Susa Scherer, Ph.D

Date July 11, 2007

Dedication

I would like to thank Lisa Scherer, my advisor, for her continual support, patience, advice, and feedback; without her, this thesis would not have been possible. I would also like to thank the members of Jim Thomas, Jessiline Anderson, and William Clute, my thesis committee, for their patience and feedback. I would like to thank Arlene, my wife, for enduring with me for the duration of this project. Finally, I would like to thank Kelvin VanManen for developing the computerized information board used in this investigation.

The Effect of Evaluator's Mood and Type of Accountability on Performance Appraisal Evaluations: A Study of the Affect Infusion Model

Jason E. Gerlt, MA

University of Nebraska, 2007

Advisor: Lisa L. Scherer, Ph. D.

Abstract

The purpose of this study was to test the theoretical framework of Forgas' (1995) Affect Infusion Model on the extent to which mood and type of accountability (no, process, and outcome) influenced information search strategies and judgment outcomes. Information boards (e.g. Billings & Scherer, 1991; Payne, 1976) were utilized to examine the amount of information searched and the performance ratings made of hypothetical teaching assistants. A 2 (mood) X 3 (accountability) between-factors design was used to examine the data. Seventy four undergraduate students were randomly assigned to one of six groups: positive mood/no-accountability, positive mood/ outcome accountability, positive mood/ process accountability, negative mood/ no- accountability, negative mood/ outcome accountability, and negative mood/ process accountability. Participants in the outcome accountability condition, regardless of mood, were expected to utilize a motivated processing strategy; participants in the process accountability, regardless of

mood, condition were expected to utilize a substantive processing strategy; participants in the no accountability condition, regardless of mood, were expected to utilize a heuristic processing strategy. Participants in the outcome accountability and process accountability conditions were expected to search significantly more information compared to participants in the no-accountability condition. However, for the process accountability and the no accountability conditions, participants in the positive mood condition were expected to rate teaching assistants more positively compared to participants in the negative mood condition, but for participants in the outcome accountability condition, no difference in performance ratings were expected. The results of the investigation do not support the predictions made.

List of Figures

List of Appendices

List of Tables

The Effect of Evaluator's Mood and Type of Accountability on Performance Appraisal Evaluations: A Study of the Affect Infusion Model

Organizations are interested in the decision-making processes utilized by its employees, or more precisely, how to help employees make better decisions. One frequently researched decision-making process is the performance appraisal process. A performance appraisal is a process in which a rater (a supervisor, peer, or subordinate) evaluates the work performance of another employee. Landy and Farr (1980), and Murphy and Cleveland (1995) conducted extensive reviews of the performance appraisal literature and have indicated that performance appraisal ratings are not always reflective of actual job performance. The apparent inconsistency between performance appraisal ratings and actual levels of performance has led researchers to investigate possible causes of the discrepancy.

Numerous researchers have discussed some of the possible factors that could influence a rater's evaluation of an employee. One frequently researched factor believed to promote inaccuracy in ratings is the purpose of the performance appraisal evaluation. Researchers suggest that performance appraisal ratings given for administrative purposes – such as promotion or pay increases – were more lenient than what the actual job performance warranted, whereas performance appraisal ratings given for developmental purposes – such as the identification of people who need training – tended to be harsher than what the actual job performance warranted (Aleamoni, & Hexner, 1980; Gmelch, & Glasman, 1977; Hobson, Mendel, & Gibson, 1981; Jawahar, & Williams, 1997; McIntyre, Smith, & Hassett, 1984; Murphy, Balzer, Kellam, & Armstrong, 1984; Zedeck,

& Cascio, 1982). Another factor thought to influence performance appraisal ratings is the interpersonal affect the rater feels toward the ratee. Researchers suggest that evaluators who like the ratee tend to evaluate the ratee more favorably compared to evaluators who have neutral or negative feelings about the ratee (Cardy, & Dobbins, 1994; Robbins, & DeNisi, 1998; and Tsui, & Barry, 1986). One factor thought to influence performance appraisal ratings, which has not received much research attention, is mood. Studies examining the influence of mood on the performance appraisal processes have found mood effects for the amount and type of information remembered, but have found mixed results on performance appraisal ratings (Robbins, & DeNisi, 1998; Sinclair, 1988). Although research examining the influence of mood on the performance appraisal process has been limited, many researchers have examined mood effects on cognitive processes and in other workplace contexts.

To understand how moods influence the performance appraisal process, it is important to differentiate a mood from an emotion. Once the difference between mood and emotion is established, a theoretical model of mood and emotion will be presented to explore the influence of mood on performance appraisals. Next, an examination of the research on influence of mood on judgment will be reviewed. I will then examine accountability as a situational variable that could influence the affect of mood on performance appraisal evaluations. After defining accountability, I will discuss how accountability affects performance appraisal evaluations. Finally, I will present my hypotheses regarding the interactive effect of mood and accountability on performance appraisal evaluations.

Mood and Emotion

Within the last few years several books have been published describing the impact moods and emotions have in the workplace (e.g. Ashkanasy, Zerbe, & Hartel, 2002; Lord, Klimoski, & Kanfer, 2002; Payne, & Cooper, 2001). Most people are capable of experiencing a wide range of moods, such as happiness or sadness, and emotions, such as anger or elation. Moods and emotions appear to be very similar in nature, but there are subtle differences between the two affective states. According to Forgas (1995), a mood is an affective state that is low in intensity, has no specified target, and is relatively enduring, whereas an emotion is an affective state that is high in intensity, has a specific target, and lasts a relatively short-period of time in comparison to a mood. However, the intensity of a mood is not always less intense than a discrete emotion, nor is an emotion always less enduring than a mood. Several researchers have concluded that the presence of a specific target is the primary factor that distinguishes an emotion from a mood (e.g. Forgas, 1995). That is, emotions are directed at a target, I am angry with the student, whereas a mood usually does not have a specific target, I am upset right now. The consensus among researchers is that the workplace is replete with moods and emotions and that these affective states have a major influence on an individual's behaviors and decision making. Several theories have been developed to explain the effects of moods and emotions on decision-making.

Two prominent theories are Schwarz and Clore's (1983) Affect-as-Information and Martin, Ward, Achee, and Wyer's (1993) Mood-as-Input theory. Schwarz and Clore (1983) conceptualized the Affect-as-Information theory, which states that people use

their current mood state as evaluative information when making decisions. According to this model, people will make mood congruent judgments when making decisions regarding a target. That is, people in a negative mood are more likely to view the target more negatively and make a more negative judgment. Conversely, individuals in a positive mood are more likely to view the target more positively and make positive judgments about the target. Similar to Schwarz and Clore (1983), Martin et al's. (1993) Mood-as-Input theory states that an individual's mood influences their judgments about specific targets resulting in mood congruent judgments. Unlike Schwarz and Clore (1983), Martin et al. (1993) also suggest that individuals' interpretation of their moods, not simply because they are in a particular mood, influences the types of processing strategies utilized (heuristic or elaborative). That is, individuals in negative moods will utilize either an elaborative processing strategy or a heuristic processing strategy depending on their understanding of their moods and their processing goals. The Affect-as-Information theory differs from the Mood-as-Input theory in that an underlying assumption is that people misattribute their current mood state as to how they feel about the target, whereas in the Mood-as-Input theory individuals use their mood as a source of information, which might or might not be caused by or attributed to the target. However, both the Affect-as-Information and the Mood-as-Input theories state that individuals will make judgments regarding a target that is congruent with their current mood state.

Most theories regarding the influence of mood on judgments have the same basic assumption; individuals make evaluations about a target that are congruent with their current mood state (e.g. Martin et al, 1993; Russell, 2003; Schwarz, & Clore, 1983).

However, researchers have found mixed results. Some researchers have found mood congruent judgments (e.g. Martin et al. 1993; Sinclair, 1988), whereas other researchers have not found mood congruent judgments (Erber & Erber, 1994; Robbins, & DeNisi, 1998). Forgas (1995) has attempted to integrate the mixed results into a comprehensive theory of the effects of moods and emotions on social judgments.

Forgas (1995) developed the Affect Infusion Model (AIM) as an attempt to explain the conflicting results regarding the effects of mood on social judgments. The main premise of the model is that individuals choose among four processing strategies when making social judgments regarding another person or target. The processing strategy chosen depends on characteristics of the target, the situation, and the decision maker. Forgas (1995) theorized when a person will utilize a specific processing strategy, described next.

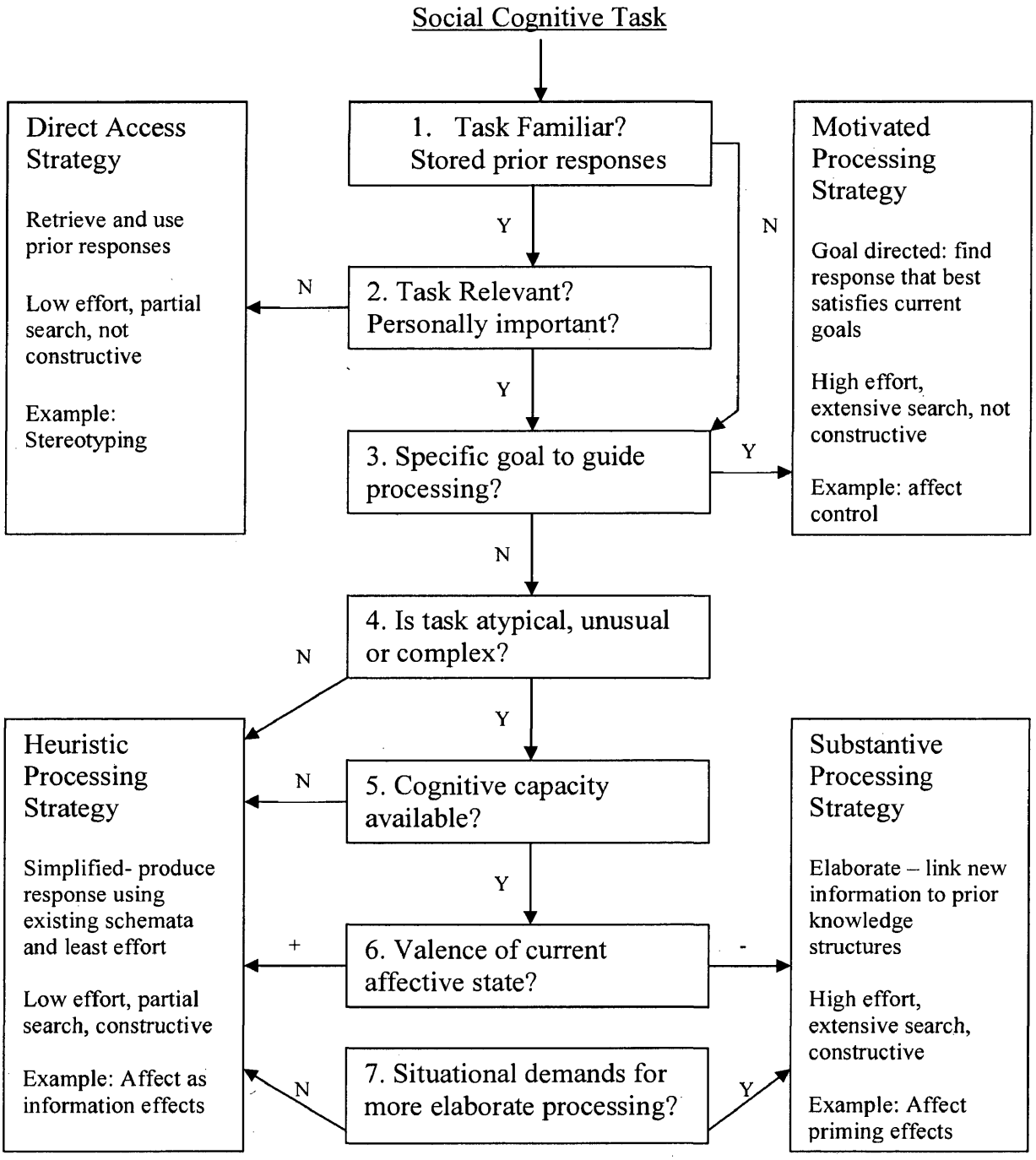
Affect Infusion Model

According to Forgas (1995), people utilize one of four different information-processing strategies when making social judgments. The information processing strategies, direct access, motivated, heuristic, and substantive (described below) utilized by an individual are influenced by characteristics of the target, the decision maker, and the situational context (task characteristics, person characteristics, and situational characteristics, respectively). Task features include the familiarity of the target, such as the target is a close, personal friend of the decision maker, and thus the decision maker is likely to have preconceived judgments regarding the target's ability or performance. Another task feature is the typicality or complexity of the target (targets that are unusual

or complex require more processing). According to Forgas (1994), person features are characteristics of the decision maker that influence his/her social judgment. Typical person characteristics include: (a) personal relevance of the judgment, how will the decision affect the individual making the decision; (b) motivational goals, is there something guiding the decision makers information search; (c) cognitive capacity, is the decision maker able to think about and process complex information; and (d) affective state, the current mood state of the decision maker. Situation characteristics include the perceived need for the decision to be accurate according to a specific standard or ideal and will the decision be evaluated by others. Although the presence of task, person, and situational characteristics serve as a guide to determining processing strategy, individuals do not explicitly consider these characteristics when making evaluative judgments. The AIM serves as a guide to explaining differences in evaluative judgments, but is not a direct representation of the cognitive processes used by individuals when making evaluative judgments. Forgas created a decision flow chart (see Figure 1) to represent how information-processing strategies are selected and when the strategies are influenced by the decision maker's mood. It is important to note that a person's mood influences the processing strategy he/she uses to make a decision, and a person's mood influences the valance of the outcome of the decision or judgment (Forgas, 1995).

Figure 1.

Flowchart of Forgas' (1995) Affect Infusion Model.



Note. From Theories of mood and cognition (p. 107), by L. L. Martin & G. L. Clore (Eds.), 2002, Mahaw, NJ: LEA. Copyright 2002 by LEA.Adapted

The information processing strategies are divided into high vs. low infusion processes (processing strategies that are and are not influenced by the decision maker's mood), and simple vs. elaborative processing strategies (the amount of information processed, and how the information is processed). Direct access and motivated processing strategies are considered low infusion processes because the decision outcomes are not influenced by an individual's mood. In contrast, heuristic and substantive processing strategies are considered high infusion processes because the decision outcomes are influenced by an individual's mood.

Direct access processing, considered the easiest processing strategy due to its reliance on past responses to make current judgments, is used when the target is highly familiar, the judgment is not personally relevant, and when situational factors do not call for more elaborate processing (Forgas, 1995). Direct access processing relies on memories or stored responses about a target, and evaluations of the target will be not influenced by an individual's mood (Forgas, 1995). For example, if a professor is in a negative mood state and is asked for his/her opinion regarding a student's performance with whom he/she has been working for the past year, the professor is likely to give an accurate assessment of the student's performance. The evaluation will not be influenced by the professor's mood because the professor is knowledgeable about the student's performance and will not use his/her current mood state as information when making the evaluation.

A motivated-processing strategy is used when an individual experiences a specific motivational pressure to achieve a specific goal or outcome. When a person is trying to

achieve a certain goal, he or she uses a highly selective and guided information search and is less influenced by his or her current mood. That is, a person searches for information that will support his or her goal, and the person does not use his or her current mood state as a source of information (Forgas, 1995). Forgas also states that motivated processing is evoked from more than just a perceived pressure to be careful or accurate and can include several motivational forces such as the need for affiliation, mood repair, and ego enhancement. According to Forgas (2001) the influence of mood on motivated processing is indirect if it has any influence.

According for Forgas (1995), heuristic and substantive-processing strategies are considered high infusion processes because the judgments or decisions made by the decision maker are influenced by the decision maker's current mood state. For example, if the task is atypical or complex and the decision maker has the cognitive capacity to make thorough information searches, then the decision maker's mood will determine the information processing strategy utilized by the individual. That is, a person in a negative mood will use a substantive processing strategy, whereas a person in a positive mood will use a heuristic processing strategy. However, when the situation calls for a more elaborative processing the decision maker will use a substantive processing strategy regardless of his or her mood. Conversely, when the situation does not explicitly demand more elaborative processing the decision maker will use a heuristic processing strategy regardless of his or her mood. Regardless of the processing strategy chosen under the high infusion processes, the person's mood will influence his or her judgment or decision

outcomes. That is, people in positive moods will evaluate targets more positively compared to individuals in negative moods and vice versa.

Heuristic processing is used when the target is highly typical, there is no motivation for a particular response, and the person has no stored responses regarding the target. People use this strategy when they want to use the least amount of effort to make a decision. That is, the person uses whatever information is available to make a quick decision. One source of information is the individual's current mood state. Thus, heuristic processing is highly influenced by current mood state.

Substantive processing is characterized as a systematic processing strategy, which is used when the target is highly atypical and situational demands call for more elaborative processing. According to Forgas (2001), substantive processing is a constructive processing strategy in that individuals form judgments about the information they are receiving. Judgments about the information received are influenced by the individual's current mood state. That is, an individual's mood influences how incoming information is interpreted. An important note is that past literature has found that people in positive moods use heuristic processing, whereas people in negative moods use substantive processing (e.g. Martin, & Clore, 2001). Forgas (1995) suggests that mood can influence the processing strategy chosen, such that positive affective states will lead to heuristic processing, whereas negative affective states will lead to substantive processing. However, Forgas also suggests that cognitive factors, such as characteristics of the situation that demand more elaborative processing, play a role in determining the processing strategy chosen under high affect infusion processing. That is, when the

situation demands more elaborative processing, individuals will use a substantive processing strategy regardless of affective state, but when the situation does not explicitly demand more elaborative processing, individuals will use a heuristic processing strategy regardless of mood. One specific situation in the workplace that calls for one person to make a judgment regarding another person is during a performance appraisal evaluation. Next, I will discuss the influence of mood on social judgments and on the performance appraisal process.

Mood and judgment

Forgas and Moylan (1987) examined the influence of transient mood states on social judgments. The researchers administered a questionnaire to individuals who had just finished watching a happy, sad, or aggressive movie. The questionnaire asked questions that pertained to; (a) political judgments, (b) the likelihood of future events, (c) satisfaction with personal life, and (d) judgments regarding responsibility and guilt. Forgas and Moylan predicted that participants in positive moods would make judgments that were more positive compared to individuals in negative or aggressive moods. The results indicated that individuals in positive moods were more optimistic in their judgments on all four areas of the survey compared to individuals in negative and aggressive moods. Forgas and Moylan suggested that mood states might activate mood-congruent cognitive categories, which guide an individual's interpretation of information. For example, for participants in negative moods, the negative aspects of a category are more salient and more easily recalled compared to non-mood-congruent cognitive categories. Because individuals attend to the negative aspects and recall more negative

information, they are more likely perceiving things as negative. Thus, mood influences an individual's cognitive processing of information. According to Murphy and Cleveland (1995), performance appraisal is a cognitive process, and it is important to explore how mood can influence the performance appraisal process.

Mood and performance appraisal

Murphy and Cleveland (1995) presented a basic model of performance evaluations in which the rater observes, encodes information, stores information, retrieves information, and integrates information about an employee's behavior before evaluating the employee's performance. Previous research has stated that affective states influence cognitive processes (Forgas, 1995; Schwarz, 2001).

Although a considerable amount of research has focused on various sources of bias in performance appraisal evaluations, few studies have explicitly examined the influence of mood on performance appraisal judgments (Murphy, & Cleveland, 1995). Two studies have examined the influence of current affective states on performance appraisal judgments, but have found mixed results.

Robert Sinclair (1988) examined the influence of mood and the order of information presentation on the amount of error in performance appraisal judgments. Sinclair speculated that there would be a linear trend for error exhibited by participants. Halo error occurs when a person evaluates the performance of another person on some dimension and then assumes that the person is likely to perform as well on other performance dimensions. For example, a professor rates a teaching assistant high on teaching ability and assumes the teaching assistant is likely to perform equally well on

organizing class material and, thus, evaluates the teaching assistant high on organization. Sinclair (1988) speculated that individuals in negative moods would display the least amount of halo error, whereas participants in positive moods would display the most halo error, and participants in neutral moods would display a moderate amount of halo error. For Sinclair's study, halo error was defined as the amount of the inter-dimension correlation for each participant. Sinclair also predicted that participants in negative moods would be the most accurate in their performance appraisal judgments. Finally, Sinclair predicted that mood congruency effects would be the greatest for participants that received mood congruent information first. Mood was manipulated by having the participants complete the Velten (1968) mood measure, and asking participants to read and write a statement of a past event in their lives, and concentrate on that statement and think about how they felt about the event. Order of information was manipulated by presenting participants with either eight pieces of information that had positive or negative valence followed by 24 statements, in random order, of negative and positive valence.

Participants were instructed that they would be participating in two experiments. In the first experiment, participants were told that they would be helping to develop a new behaviorally based teacher performance rating scale. Participants read 32 behavioral statements regarding the teacher's behavior and then completed a questionnaire asking about the use of behaviors to rate teacher performances. Next, the participants were directed to a second experimenter whom they were told was interested in validating the Velten (1968) mood measure. Upon completion of the mood measure, the first

experimenter returned and stated that he was interested in time-delayed effects on memory. Participants then completed a 12-item questionnaire consisting of four overall measures of teaching effectiveness, and eight items that focused on specific categories of behavior. Following administration of the questionnaire, participants were given 5 minutes to write down as many behaviors as they could remember regarding the teacher.

According to Sinclair (1988), the mood manipulation was successful, such that participants in the positive mood, negative mood, and neutral mood conditions were significantly different from each other in the amount of affect and activity measured. According to Sinclair, the results supported the predictions. Sinclair found larger inter-dimension correlations for participants in the positive mood condition compared to participants in the negative mood condition. Participants in the neutral mood condition showed inter-dimension correlations that fell between the positive and negative mood conditions, but the differences between the neutral and the positive conditions were not significant, nor were the differences between the neutral and negative mood conditions significantly different. Sinclair suggested that participants in the positive mood condition tended to group behaviors into fewer categories compared to participants in the negative condition, which resulted in greater halo error. That is, participants in the negative mood condition appeared to show great differentiation among behaviors compared to participants in the positive mood condition. Sinclair (1988) then examined the correlations between the number of positive behaviors within a category and the participant's performance appraisal rating of the teacher to determine the accuracy of the performance appraisal judgments.

Sinclair (1988) found an effect of mood on accuracy of ratings, such that, participants in negative moods examined more information and showed greater differentiation among behaviors compared to participants in the positive and neutral mood conditions. Specifically, results showed lower correlations between a participant's rating of teaching performance on a dimension and number of positive behaviors per category of specific behavior. As such, participants in negative moods were more accurate in their performance appraisal ratings. Sinclair then examined participants' ratings of the two measures of global evaluations (four-item questionnaire and the open-ended evaluation question) and found that participants in positive moods made more positive evaluations of the teacher's performance than did participants in either the negative or neutral mood conditions (Sinclair, 1988). Sinclair also found that the order in which the behavioral information was presented influenced the subsequent evaluations of the teacher's performance, such that, participants receiving positive information first rated the teachers performance more positively than participants who read negative behaviors first. However, there was no interaction between mood and order of information (Sinclair, 1988). In a final task of the experiment, participants were instructed to write down as many behaviors as they could remember.

Sinclair (1988) found a mood by valence of information interaction on recall of information. That is, participants in positive moods retrieved more positively valenced and less negatively valenced information compared to the neutral and negative mood conditions. In contrast, participants in the neutral and negative mood conditions retrieved less positive information compared to negative information, and retrieved more negative

information than did participants in the positive mood condition. Sinclair also found an interaction between order of information and valence of information on recall. Participants who received positive information first tended to recall more positive information than participants who received negative information first, but there was no difference in the amount of negative information recalled regardless of which information was viewed first (Sinclair, 1988).

According to Sinclair, the results indicated that participants in negative moods were more accurate in their ratings compared to individuals in positive moods who appeared to display more halo error in their ratings. Sinclair (1988) speculated that participants in positive moods used an automatic or a heuristic processing strategy when making decisions, whereas participants in negative moods used a controlled or analytic (a more thorough) processing strategy. Because participants in negative moods used a controlled processing strategy, they reviewed more information before they made their decisions, which led to accurate appraisals of performance.

Sinclair's results can be interpreted using the Affect Infusion Model (AIM). The automatic processing strategy is similar to what Forgas termed heuristic-processing strategy, whereas the controlled processing strategy is similar to substantive-processing strategy. Both processing strategies should lead to mood congruent appraisals (Forgas, 1995), which were found in this study. However, Sinclair's results do not support his claim that participants in negative moods processed more information. If participants in negative moods processed more information or read more compared to those in positive moods, then they should have recalled more information. This did not occur. Although

participants in negative moods recalled more negative behaviors than did participants in positive moods, participants in positive moods recalled more information overall. If participants in positive moods recalled more information, they may have perceived the negative valence information as less negative, which might have led to the more positive appraisals. In contrast, participants in negative moods remembered less positive valence information, and perhaps perceived the positive behaviors that they did recall as more negative. Other researchers have found mood effects on information retrieved, but did not find mood effects on performance appraisal ratings (Robbins, & DeNisi, 1998).

Robbins and DeNisi (1998) examined the influence of interpersonal affect and mood on the performance appraisal process. Robbins and DeNisi speculated that interpersonal affect felt towards another person would influence performance appraisal ratings and that mood would not have an effect on the ratings. Robbins and DeNisi also speculated that interpersonal affect, rather than mood, would influence the information recalled and the weighting of the information recalled. Undergraduate business majors from three classes evaluated the performance of the professor teaching their class. That is, the participants evaluated the performance of a professor who was well known. Interpersonal affect was defined as how well the participant liked the professor, which resulted in one of three types of affect toward the professor: positive affect, neutral affect, and negative affect. Interpersonal affect was measured four weeks prior to the beginning of the study and again ten weeks after the semester had started. Mood was manipulated by having the participants focus on happy, sad, or neutral events from their past and to concentrate on the feelings associated with that event. After the mood manipulation,

participants watched video clips of the professors' teaching performance from a previous year. The videos contained nine behaviors encompassing three positive, three neutral, and three negative levels of performance. The dependent variables were recall (the number of behaviors remembered), weighting (participants' ranking of the behaviors in terms of how much each should be considered when making overall evaluations), and ratings (overall evaluation of teaching effectiveness evaluated on a 7-point scale). The ratings of effectiveness were then compared to the professor's true score of effectiveness, which was determined in an earlier study by Hom and associates (Hom, DeNisi, Kinicki, & Bannister (1982). Using the ratee as a control variable, Robbins and DeNisi (1998) used a 3 (interpersonal affect) X 3 (mood) X 3 (true score) factorial design to analyze the data.

The results of the analysis indicated that there was a main effect for interpersonal affect on performance appraisal ratings, such that participants who liked the professor tended to give higher ratings of effectiveness. The results also indicated that there was no effect of mood on performance appraisal ratings, which supported the hypothesis that interpersonal affect and not mood influenced the participant's rating of the professor's effectiveness. Robbins and DeNisi (1998) also suggested that interpersonal affect, as opposed to mood, would influence the information recalled. Contrary to the prediction, participants in negative moods recalled the most information, and there was no influence of interpersonal affect on the amount of information recalled. Additionally, Robbins and DeNisi (1998) found an interaction between interpersonal affect and true score, such that participants gave the most weight to items that were similar to their level of liking toward the professor. That is, participants with positive affect toward the professor gave more

weight to positive behaviors compared to negative behaviors, but participants with negative affect toward the professor did not give significantly more weight to negative behaviors compared to positive behaviors. Finally, Robbins and Denisi found an interaction between mood and true score on subjective weights assigned to behaviors, such that weight assigned to the behaviors was incongruent with their current mood state. That is, participants in negative moods put more weight on positive behaviors than negative behavior, but participants in positive moods assigned similar weight to all behaviors.

The overall finding of the study was that although mood influences the information recalled, interpersonal affect has more of an influence on the performance appraisal ratings (Robbins & DeNisi, 1998). Robbins and DeNisi's (1998) results can be explained using the AIM. Robbins and DeNisi did not find mood effects in their study, but did find, albeit marginal, effects for interpersonal affect such that raters who liked the professor rated him more favorably compared to raters who did not like the professor as much. According to the AIM, a person will use direct access processing when the target is well known. For a person to like or dislike someone else, the person would have to be familiar with the other person. Thus, in Robbins and DeNisi's study, one would not expect to find mood effects on performance appraisal because the raters already had schemata and prepared responses for evaluating the professor. If participants had used a substantive processing strategy rather than a direct access or motivated processing strategy, Robbins and DeNisi might have found mood effects on performance appraisal ratings. Additionally, when using a direct access processing strategy individuals do not

process all of the available information. One technique found to increase the amount of information a person utilizes when making evaluations is to make the person accountable for his or her decision.

Accountability

Tetlock (1992) defined accountability as "the implicit or explicit expectation that one may be called on to justify one's beliefs, feelings, and actions to others" (p. 256). Additionally, when individuals are held accountable for their decisions there is the implication that if they do not provide satisfactory justification for their decisions or actions they will suffer negative consequences, particularly social repercussions (Weigold & Schlenker, 1991). That is, people are worried about how others will view them based on their decisions and actions. Researchers who have examined the effects of accountability on decision making have found that individuals who were held accountable for their decisions used more analytical decision strategies (Chaiken, 1980; Hagafors & Brehmer, 1983; Hattrup & Ford, 1995; McAllister, Mitchell, & Beach, 1979; Mero & Motowidlo, 1995), recalled more information (Mero & Motowidlo, 1995; Tetlock, 1983), and made more complex judgments (Simonson & Staw, 1992; Tetlock & Kim, 1987) compared to individuals who were not held accountable for their decisions. That is, searching through more information, being able to recall more information about a target, and being able to make more complex judgments should improve the quality of the individual's decision. However, other researchers have questioned the notion that making individuals accountable for their decision always has positive consequences.

Adelberg and Batson (1978) examined the impact of making financial aid

advisors accountable for their decision to distribute financial aid to prospective college students. Adelberg and Batson found that when individuals were accountable for the distribution of financial aid, the decision makers tended to distribute money to all candidates, such that no single candidate received enough financial aid. However, individuals not held accountable for their decisions tended to provide sufficient funding to only a few candidates (Adelberg & Batson, 1978). Adelberg and Batson (1978) reasoned that making individuals accountable for their decisions created a sense of anxiety, which caused individuals to make a decision that was more defensible (i.e. fund everyone applying for financial aid). That is, when held accountable for their decisions, individuals are likely to make decisions that cause the least anxiety and are easy to justify.

Another drawback of making individuals accountable for their decisions is that individuals may report the most easily defensible position, which may or may not reflect their actual opinion (Lerner & Tetlock, 1999). Cialdini, Levy, Herman, Kozlowski, and Petty (1976) examined the influence of an individual's attitude (whether the attitude was important or not important to the individual) and accountability (to an individual with an opposing view) on an individual's propensity to temporarily alter their view. Cialdini et al. expected that individuals would express a moderate view of their attitude to an individual with an opposing view when the attitude was of low personal relevance. However, individuals would express a more polarized view of their attitude to an individual with an opposing view when the attitude was of high personal relevance. That

is, individuals would express a stronger view of their attitude than was actually held, when discussing their view with a person with an opposite attitude.

Cialdini et al. (1976) determined the participants' personal relevance of issues, indicated that participants would be discussing their opinion to an individual with an opposing view, and manipulated when the individual would meet with the other person (delayed vs. immediate). In the immediate condition, the results were as expected. Individuals with strong views evinced more polarized views (stronger towards their own belief) compared to the control group, whereas individuals with less important views evinced moderate views compared to the control group. However, when the discussion was delayed individuals evinced more polarized views compared to the control group regardless of the personal importance of their view (Cialdini et al, 1976). Cialdini et al. suggested that individuals engaged in moderation of their views for strategic reasons. That is, in the immediate condition people who had low important views moderated their position because it would be a more defensible position. However if the participant's view was personally important, or the individual had time to prepare an argument, then the expressed views were more polarized in the direction of the personal view (Cialdini et al, 1976). That is, the person would be able to justify and defend their position to an individual with an opposing viewpoint. Within the broader context of the affect infusion model, the Cialdini et al study may exhibit differences in processing strategies based on personal relevance and time of accountability. When an attitude is low in personal relevance and individuals have little time to think about the topic, the individuals may engage in direct access processing and state an opinion they think the person with an

opposing view wants to hear. However, when the attitude is high in personal relevance, the individuals may feel passionate about their view and engage in heuristic processing about the topic and thus expresses a more polarized view of their attitude. Finally, when individuals have time to prepare their discussion or argument about a topic the individuals may engage in motivated processing when the topic is of low personal relevance, and substantive processing when the topic is of high personal relevance. Thus, accountability can influence the processing strategy an individual might use when preparing for a discussion with a person who holds an opposing view. However, accountability could also influence an individual's choice of processing strategies when the views of the other person are not known.

Tetlock, Skitka, and Boettger (1989) examined the influence of accountability to a person with an unknown view, accountability to a person with a known view (conservative or liberal), and no accountability on an individual's reported stance on a controversial public policy. Half of the participants wrote down their thoughts and feelings about the public policies and then completed a survey that measured their attitudes about each public policy. The other half of the participants completed the attitude survey first and then wrote down their thoughts and feelings about the policies. Before completing the attitude survey and writing down their thoughts, some participants were told that they would be discussing their thoughts, feelings, and attitudes with an individual who was either a conservative or a liberal, whereas others were told that they would be discussing their thoughts, feelings, and attitudes with a person whose views were not known. The remaining participants were told that their thoughts, feelings, and

attitudes regarding the public policies would remain confidential. Tetlock et al. found that when individuals were asked to discuss their thoughts, feelings, and attitudes with a person whose view was known, the participants tended to report a view that was more compatible with the other person's view compared to individuals who were not held accountable, or who did not know the view of the person with whom they would be speaking. However, individuals tended to shift their attitudes only when they wrote down their thoughts and feelings before making an attitudinal stance. When individuals reported their attitudinal stance before writing their thoughts on the issue, they tended to report views that were consistent with their original attitudes (Tetlock et al., 1989). Additionally, Tetlock et al. (1989) found that individuals who reported their thoughts about an issue before taking an attitudinal stance generated more integratively complex views when they did not know the opinion of the person to whom they were accountable compared to individuals who knew the opinions of the people to whom they were accountable. Individuals who reported their attitudes before writing down their thoughts and feelings tended to generate less integratively complex views compared to individuals who reported their thoughts first (Tetlock et al., 1989). Thus, the research on accountability has generated mixed results.

When individuals are accountable for their decisions, they tend to use more complex processing strategies and generate more complex decisions compared to individuals who are not accountable. However, when individuals know the opinions of the person to whom they are accountable, they tend to espouse views that are easily defensible or similar to the known view; especially when the task is not personally

relevant, or when they are asked to write their thoughts about the task before stating their attitude toward the issue. Thus, research on accountability has found both benefits and problems with holding people accountable for their decisions. One reason the research on accountability has found mixed results is that researchers tend to focus only on outcome accountability without distinguishing between outcome and process accountability.

Process and Outcome Accountability

According to Siegel-Jacobs and Yates (1996), process accountability occurs when individuals' judgments are evaluated by how the judgments were reached and not on the quality of the judgment outcomes; whereas outcome accountability is when individuals' judgments are compared to judgments made by experts (i.e. the quality of the judgment is assessed). Using Siegel-Jacobs and Yates definitions of accountability, researchers have conducted studies that have examined the influence of process and outcome accountability on judgments made in the workplace.

Brtek and Motowidlo (2002) examined the influence of procedure and outcome accountability on the validity of interview judgments. Brtek and Motowidlo speculated that participants would make more valid evaluations in the procedure accountability condition compared to participants in the no accountability condition because participants would be more attentive to relevant information regarding an interviewee. Brtek and Motowidlo also speculated that the effects of procedure accountability on interview decision validity would be mediated by attentiveness. The results of the study indicate that participants who were held accountable for the procedure they followed when making interview decisions made more valid interview decision based on the interview

criteria compared to participants who were not held accountable for the procedure they followed or for the outcome of the interview decision. Similarly, participants who were accountable for their outcomes when making interview decisions made more valid interview decisions based on the interview criteria compared to participants who were not held accountable for the procedure they followed or the outcome of the interview decision. The results of the study also indicated that the effect of procedure accountability on interview validity was fully mediated by attentiveness. Brtek and Motowidlo speculated that participants in the procedure accountability condition were motivated to pay more attention to the information in the interview and were thus able to make valid interview decisions. The results of the Brtek and Motowidlo (2002) study lend support to the concept that participants held accountable, either for the procedure in which judgments are made or for the outcome of the judgments, examine more information before making a decision. That is, accountability is a situational factor that, according to Forgas (1995), should elicit more elaborative processing strategies.

For the purpose of this study, accountability is defined as the implicit or explicit expectation that one will be asked to justify either the quality of the process used to arrive at a final decision or evaluation, or to justify the quality of the final decision or evaluation. Distinguishing between process and outcome accountability is important for understanding how each can influence the performance appraisal process.

Accountability and Performance Appraisal

Mero and Motowidlo (1995), using an elaborate in-basket simulation, examined the influence of accountability on performance appraisal ratings under four different

motivational contexts. The motivational contexts provided information about previous performance appraisal ratings. In the first context, participants did not receive any information about previous performance appraisal ratings. In the inflationary context, participants were told that previous performance appraisal ratings were consistently lower compared to other people in the company. In the accuracy context, participants were told that performance appraisal ratings were inflated and that it was impossible to discriminate between actual levels of performance. In the equity context, participants were told that women were consistently rated lower than men were rated. Accountability was manipulated through the instructions given to the participant. Participants in the accountability condition were told that at the completion of the experiment they would be required to justify their ratings to the researcher via a written statement. Participants in the no accountability condition were told that their ratings would remain anonymous and that their written assignment due at the end of the experiment was to critique the simulation. Mero and Motowidlo (1995) predicted that when there is no specific motivation to achieve a specific outcome, the first motivational context, participants who were held accountable would rate performance more accurately than would non-accountable raters. That is, performance evaluations would be consistent with the evaluations of experts. When the motivational context creates pressure to achieve a desired outcome, Mero and Motowidlo predicted that high accountable raters would administer ratings that were more congruent with the motivation context than would non-accountable raters. That is, in the leniency condition high-accountable raters were expected to rate people more leniently compared to non-accountable raters; in the

accuracy condition, high-accountable raters were expected to discriminate between different levels of performance more so than non-accountable raters; and in the equity condition, high-accountable raters were expected to rate women more leniently than were the non-accountable raters.

According to Mero and Motowidlo (1995), when there was no motivation to achieve a specific outcome, high-accountable raters were more accurate in their performance appraisal evaluations than were non-accountable raters. For the inflationary and the accuracy conditions, high-accountable raters made evaluations that were more consistent with the motivational context compared to non-accountable raters. That is, for the inflationary condition, high accountable raters were more lenient in their evaluations than were non-accountable raters. In the accuracy condition, high accountable raters displayed less discrepancy between the ratings given and the true score of the employee compared to non-accountable raters. However, in the equitable treatment condition, high-accountable raters did not rate women's performance more leniently than did non-accountable raters (Mero, & Motowidlo, 1995).

Mero and Motowidlo (1995) concluded that participants in the no motivation condition, leniency condition, and the accuracy condition gave performance evaluations that were easier to justify when they were held accountable for their decision compared to participants who were not held accountable. However, for the equity condition the authors reasoned that high-accountable raters failed to give lenient ratings to females because lenient ratings for a select group of individuals would be harder to justify than accurate appraisals; conversely non-accountable raters did not feel the need to justify

their ratings and may have been more socially motivated to rate females more leniently (Mero, & Motowidlo, 1995).

The Affect Infusion Model can be used to interpret the results of the Mero and Motowidlo (1995) study. According to the AIM, participants in the three motivational contexts probably utilized a motivated processing strategy. That is, participants examined relevant information before making an evaluation, but the evaluation was consistent with an implied goal in two of the three motivational context conditions. High accountable participants who were not asked to achieve a specific outcome probably utilized a substantive processing strategy, but because there was no mood manipulation, the varying moods might have canceled each other out resulting in no effects of mood on performance appraisal outcomes. Non-accountable raters probably utilized a heuristic processing strategy in the motivated and unmotivated conditions. Use of the direct processing strategy is improbable because participants were unlikely to have stored prior responses regarding the employees whom they evaluated. One of the major limitations of the Mero and Motowidlo (1995) study is that the researchers failed to examine the effects of different types of accountability on performance appraisal judgments. The previous discussion of the effects of mood on performance appraisal judgments and the influence of accountability on judgments will help support the hypotheses of this investigation.

This Investigation

The purpose of this study was to examine the extent to which individuals' current mood states influence the evaluations they make under varying types of accountability. More specifically, this study examined the influence of positive and negative mood

states and no accountability, process accountability, and outcome accountability on the amount of information searched and on performance appraisal evaluations. The premise of this investigation was that the processing strategy utilized by participants would be influenced by the type of accountability to which the participant is exposed, such that participants in the outcome accountability condition would use a motivated processing strategy, participants in the process accountability would use a substantive processing strategy, and participants in the no accountability condition would use a heuristic processing strategy. I also predicted an interaction between mood and accountability on the performance appraisal outcomes. The rationale for the hypotheses falls out of the framework of the AIM (see Figure 1).

According to Forgas' model (1995), individuals ask themselves a series of questions when determining what type of processing strategy they will utilize when making judgments. The individual first determines if the task or the target is highly familiar. For this investigation the participants evaluated the performance of teaching assistants with whom they were unfamiliar. As such, participants were unable to utilize a direct access processing strategy. The next question individuals will ask themselves is whether or not they have a specific objective that they wish to attain. For this investigation, only participants in the outcome accountability condition could satisfy this criterion and engage in motivated processing, which is not influenced by the individuals' moods. Individuals then determined if the task is typical or complex. Although participants were somewhat familiar with evaluating the performance of teaching assistants, the participants were not familiar with the method used in this investigation.

Because the task is atypical, participants will then determine if they have the cognitive capacity available to process the information. For this investigation, it is expected that participants will have the cognitive capacity available to process information in a thorough manner. Although cognitive capacity is one factor that can influence the processing strategy utilized by individuals, the individuals' moods will also influence the processing strategy utilized. According to the model, the participants' mood will influence the processing strategy people will want to use, such that participants in positive moods will use a heuristic processing strategy, whereas participants in negative moods will use a substantive processing strategy. However, if there are situational demands that call for more elaborate processing (accountability for process or outcome), then participants will utilize a more elaborative processing strategy regardless of their current mood state. Similarly, if there are no situational demands for more processing, then individuals in positive and negative moods will utilize a heuristic processing strategy because, according to Forgas (1995), mood effects for determining processing strategy are secondary to situational factors that demand more elaborative processing. For this investigation it is believed that process accountability will act as a situational condition that will elicit a more substantive processing strategy, whereas the no-accountability condition will elicit a heuristic processing strategy for participants in positive and negative moods. The previous rationale leads to the following hypotheses:

Hypothesis 1: The amount of information searched will vary depending on the accountability condition of the participant (see Figure 2).

1a) Participants in the process accountability condition will examine a similar amount of information as participants in the outcome accountability condition.

1b) Participants in the no accountability condition will examine less information compared to individuals in the process and outcome accountability conditions.

Hypothesis 2: There will be an interaction between mood and accountability on participant's evaluation of performance (see Figure 3).

2a) In the no accountability condition, participants in a positive mood will give higher performance appraisal evaluations compared to those in a negative mood.

2b) In the process accountability condition, participants in a positive mood will give higher performance appraisal evaluations compared to individuals in a negative mood.

2c) In the outcome accountability condition, there will be no mood effects on performance appraisal evaluations.

Hypothesis 3: There will be a main effect of mood on performance appraisal ratings, such that participants in positive moods will give significantly higher ratings than participants in negative moods.

Figure 2.

Predicted relationship between number of behaviors examined and accountability condition.

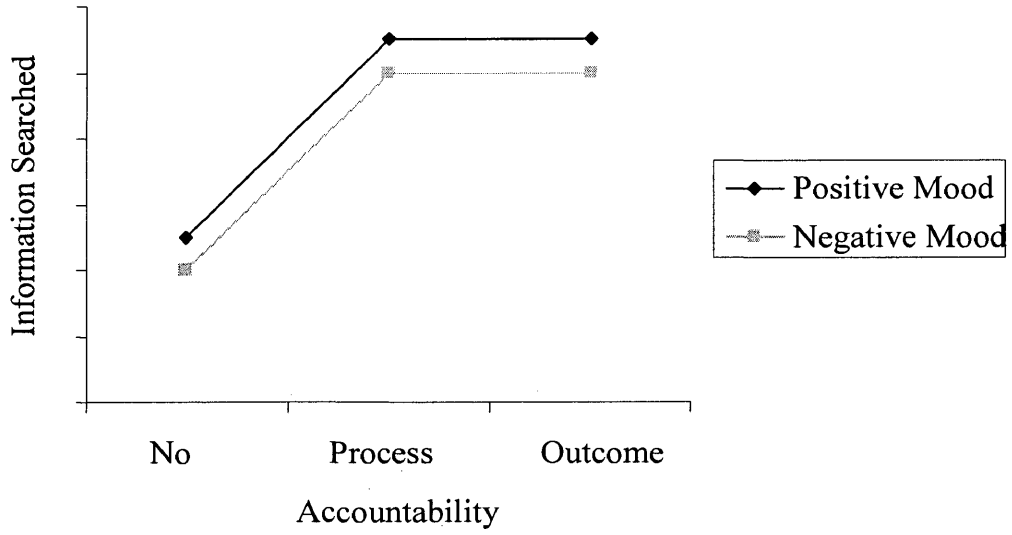
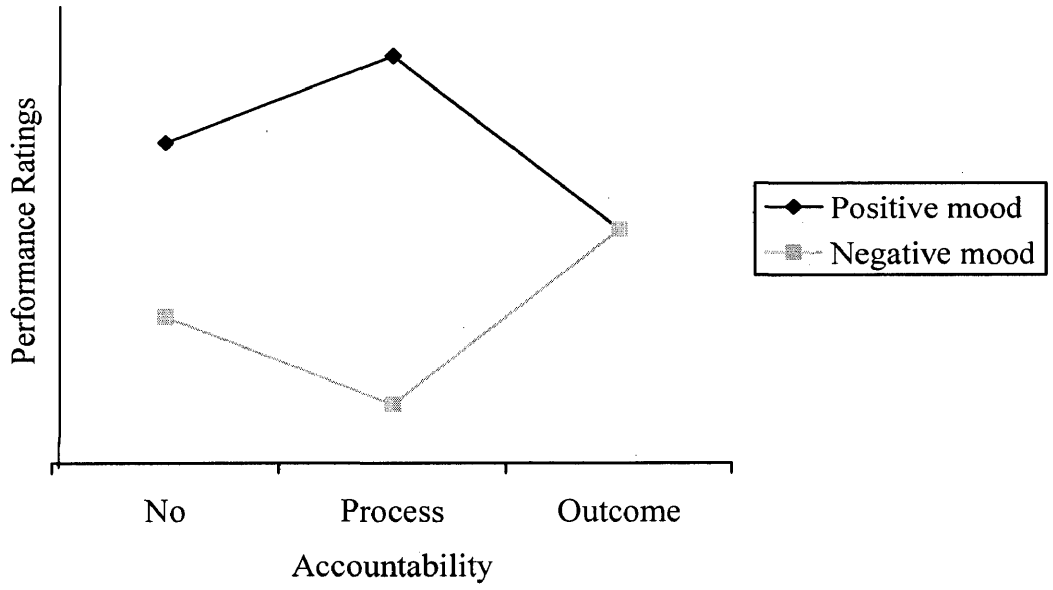


Figure 3.

Predicted Interaction between Mood and Accountability on Performance Ratings.



Method

Participants

Ninety-four undergraduate college students from a medium sized Mid-western university participated in this investigation. Participants ranged in age between 18 and 57 years (M = 21.81 years, SD = 5.80 years). The participants consisted of 60 females, 32 Males, and 2 people who chose not to respond to the question. Of the 94 participants, 32% were freshmen, 16% were sophomores, 26% were juniors, 26% were seniors, one person chose not to respond to the question. The participants were 84% White, 3% Black, 9% Hispanic, and 2% Asian, and two individuals chose not to respond to the question.

Design

Participants were randomly assigned to one of six conditions: negative affect/no-accountability, positive affect/no-accountability, negative affect/outcome accountability, positive affect/outcome accountability, negative affect/process accountability, or positive affect/process accountability. The data were analyzed using a 2 (mood) X 3 (accountability) between-subjects design.

Past research examining the effects of accountability on performance appraisal judgments have found moderate effect sizes. For example, Mero and Motowidlo (1995) have found effect sizes ranging from $d = .46$ to $d = .75$, where d equals the difference between the means of the treatment group (M*t*) and the control group (M*c*) divided by the pooled standard deviation (SD) ($d = (Mt - Mc)/SD$). Prior research that has examined the effects of mood on performance appraisal ratings have found low to moderate effect sizes. Given the range of possible effect sizes, and a PA-PC of 5, I would have needed

about 500 participants to have an 80% chance of finding a significant effect of mood on performance appraisal judgments. However, due to financial and time constraints, and the feasibility of recruiting enough participants, I intended to use 180 participants, which would have given me a probability of .89 to find an effect of accountability on performance appraisal ratings with a moderate effect size.

Independent Variables

The independent variables for this study were mood (positive or negative) and accountability (no, process, and outcome). Participants were randomly assigned into either the positive mood or negative mood condition. Within each mood condition, participants were randomly assigned to one of the three accountability conditions.

Mood. To elicit a positive mood, participants watched a 10-minute video clip of Ferris Bueller's Day Off, whereas individuals in the negative mood condition watched a 10-minute video clip of Midnight Express. Both movies had been found to elicit the desired mood (Weiss, 1996). A pilot study was also conducted to ensure the videos elicited the appropriate mood state. To ensure the mood manipulation worked, participants completed the Affective Reactions Scale shown in Appendix A (Scherer, Reiter-Palmon, Butler, & Weiss, 1994).

Accountability. The three types of accountability (no, process, and outcome) were manipulated through the instructions given to the participants at the beginning of the performance appraisal task. Accountability is defined as the implicit or explicit expectation that one will be called on to justify either the quality of the process used to arrive at a final decision or to justify the quality of the final decision or evaluation of a

teaching assistant's performance. Prior research has successfully manipulated no-accountability by assuring the participants that their performance evaluations would remain anonymous (Brtek, & Motowidlo, 2002; Mero, & Motowidlo, 1995) and that their evaluations would be averaged with other individuals who were also participating in the study (Brtek, & Motowidlo, 2002); additionally, Tetlock et al. (1989) told participants that their views and evaluations would remain confidential. For this investigation, participants in the no-accountability condition were told that their evaluations of the teaching assistants' performance would remain confidential and would be averaged with the evaluations of other individuals who are also evaluating the performances.

Brtek and Motowidlo (2002) manipulated process accountability by telling participants that they would be meeting with the Dean of the college at the end of the experiment, at which time they would be asked to justify the procedure they used when evaluating interviewees. Tetlock and Kim (1997) informed participants that at the end of the experiment they would be meeting with the researcher to discuss the types of information that were used in forming impressions of other people. Finally, Simonson and Staw (1992) told participants that an evaluation of their decision making ability would be based on their use of effective decision making strategies and not on the outcome of their decisions. For this study, participants in the process accountability condition were told that they would be meeting with the Dean of the college at the end of the experiment to discuss information they used to arrive at their evaluations.

Outcome accountability has been manipulated by telling the participants that their evaluations will be compared to expert evaluations, and any discrepancies between the

evaluations would need to be justified (Brtek, & Motowidlo, 2002), or told that they will discuss their attitudes regarding an issue with someone who's views were known (Tetlock et al., 1989). For this investigation, participants were told that they would be meeting with the Dean of the college to discuss their ratings of the teaching assistants.

Stimulus Materials

Information board. I used a modified information board (Payne, 1976) to assess performance appraisal ratings (see Appendix B). Rather than have many boards that represent the same information for many different candidates, I used the same teaching assistants across several boards and varied the information presented on the boards. Each board consisted of behavioral statements that varied in perceived effectiveness, and that represented six dimensions of teaching. The names of teaching assistants were listed along the vertical axis of the board with the behaviors listed across the top horizontal axis. There was a seventh column on the board in which participants recorded an overall effectiveness rating of the teaching assistant's performance.

Mood manipulation. At the beginning of the experiment, participants read an excerpt from 'On comparison meaningfulness of aggregation functions' (Marichal, & Mathonet, 2000). The purpose of this article was to generate a neutral mood state in the participants so that the manipulation of mood would not be influenced by the participants' current mood states. Herman (2004) found that the Marichal and Mathonet (2000) article does elicit a neutral mood state. A pilot study was conducted to ensure the article still placed individuals into a neutral mood state.

To elicit a positive mood, participants watched a 10-minute video clip of Ferris Bueller's Day Off, whereas individuals in the negative mood condition watched a 10-minute video clip of Midnight Express. The movies were pilot tested to ensure they elicited the appropriate mood states.

Mood measurement scale. Participants completed the Scherer et al. (1994) Affective Reactions Scale three times (see Appendix A) after the reading the article, after watching the video, and after completing their evaluations of the teaching assistants. The scale consists of 42 bipolar adjective pairs measured on a 6-point continuum. That is, participants indicated which adjective in the pair was more representative of how they felt after completing each task. The scale contains five sub-dimensions. A reliability analysis of each sub-dimension indicated that the scales are reliable. The first dimension is negative arousal ($\alpha = .91$), and consists of thirteen adjective pairs such as apprehensive vs. relieved, and composed to nervous. The second dimension, termed positive arousal ($\alpha = .74$), was moderately reliable, and consists of four adjective pairs like tired vs. energetic and pacified vs. riled. Results of the reliability analysis suggest that the reliability of the scale could be improved from $\alpha = .74$ to $\alpha = .87$ by removing item 3 (riled vs. pacified) from the scale. As such, item 3 of the positive arousal sub-dimension, was removed for the calculation of the participants positive arousal score. Fear is the third dimension ($\alpha = .90$), and consists of five adjectives pairs, such as serene vs. jittery and unafraid vs. afraid. Elations is the fourth dimension ($\alpha = .97$), and consists of eight adjective pairs (e.g. repulsed vs. attracted and somber vs. cheerful). The fifth dimension is termed boredom and consists of five adjective pairs such as interested vs. bored and concerned

vs. unconcerned. The initial reliability of the boredom scale was $\alpha = .63$, but results indicated that removing item 11 (concerned vs. unconcerned) would improve the reliability of the scale to $\alpha = .88$. Thus, item 11 of the boredom scale was dropped for the calculation of boredom for the primary investigation. Factor analysis, either confirmatory or exploratory, was not conducted because the number of participants, 94, in relation to the number of scale items (35) was less than the recommended ratio proposed by Stevens (2002) of 5 to 1 to ensure the stability of the results.

Dependent Measures

For this study the two dependent measures were processing strategy and performance appraisal rating. Participants were expected to engage in a heuristic, substantive, or motivated processing strategy, regardless of their mood condition. Performance appraisal evaluations were the average of the two overall ratings per teaching assistant across the two information boards, and the overall measure of teaching effectiveness.

Processing strategy. Payne (1976) suggests that there are two general types of processing strategies, compensatory and non-compensatory. According to Payne, in the compensatory processing strategy individuals assign weights or importance values to the individual aspects of the target (in this case the teaching behaviors of the teaching assistants). The individual would then examine each piece of information and then make a judgment regarding the target. That is, individuals will search across dimensions for a single teaching assistant and then make a rating or judgment regarding the teaching assistant's performance.

According to Payne (1976) A common non-compensatory processing strategy used by individuals is the elimination by aspects (EBA) strategy. In EBA individuals search information within a dimension (i.e. across targets) and make comparisons between targets based on the attractiveness of the information. For example, when searching for a car individuals do not compare all cars on all features, rather, individuals eliminate cars that do not have important characteristics such as color. Individuals keep eliminating cars based on importance of characteristics until one option remains. For this investigation, participants may rate teaching assistants based on the importance of the dimension and thus compare teaching assistants based on a single or a few dimensions, which would be evinced by accessing less information before making an evaluation.

The type of processing strategy used by the participants was determined by counting the number of behaviors per teaching assistant the participant accessed and the order in which information was searched before giving an overall rating of the teaching assistant's performance.

Performance Ratings. Participants gave an overall rating per cluster for a total of two overall ratings. An average performance rating was calculated from the two overall ratings. Participants were instructed to give an overall rating of teaching effectiveness for each teaching assistant. Ratings for each dimension and the overall rating were assessed on a 7-point scale, such that, 1 = highly ineffective teaching behavior to 7 = highly effective teaching behavior.

Design Constants

Interpersonal affect. To control for interpersonal affect, the teaching assistants being rated were not known by the participants. The teaching assistants being rated were paper people, but the participants were told that the teaching assistants were graduate students at another university.

Depression and anxiety. To control for the possible effects of depression and anxiety on performance appraisal evaluations, the Costello-Comrey Depression and Anxiety Scales (CCDAS) (Costello, & Comrey, 1967) (see Appendix D) was used to identify individuals who displayed high levels of depression or anxiety. The CCDAS depression scale is an 11-item inventory measured using a 9-point scale. Item responses range from 1 (absolutely not) to 9 (absolutely), or from 1 (never) to 9 (always), with higher scores indicating higher levels of depression. The anxiety scale of the CCDAS is a 12-item scale measured using the same 9-point response scheme. For this investigation, the Depression and Anxiety Scales displayed high levels of reliability ($\alpha = .90$; and $\alpha = .84$, respectively). Again, due to the low number of participants, a factor analysis of the two scales was not conducted.

Participants mean scores on the CCDAS depression scale ranged between 1.00 and 4.58 ($M = 2.54$, $SD = .97$), and mean scores on the CCDAS anxiety scale ranged between 1.90 and 7.00 ($M = 3.68$, $SD = 1.05$). Participants were identified as high in depression or anxiety if their standardized score on the scale was greater than 2.33 standard deviations above the mean. One participant was identified as high in depression ($z = 2.36$) and anxiety ($z = 3.16$).

Accountability Manipulation

After completing the video rating task, participants proceeded to individual testing rooms where they began the performance appraisal task. Participants read the instructions for the task on a computer. The instructions presented to the participants varied depending on the type of accountability to which the person was exposed (see Appendix C). A manipulation check of accountability was conducted by asking the participants the extent to which they believed the Dean would be here at the end of the study to discuss their performance ratings. Belief was measured on a 5-point scale ranging from 1 – strongly disagree, to 5 – strongly agree.

Procedure

Participants reported to the experiment in groups of four to six. Upon entering the room, participants were randomly given participant numbers which assigned the participant to one of the three accountability conditions. Participants were informed that they were participating in two studies. First participants were told that the primary investigator was working with a professor from a local area community college who was interesting in collecting information on how students react to different types of presented information. Participants were told that they would read an article for about 10 minutes and then complete a questionnaire. After reading the article participants would then watch a video for about 15 minutes and complete another questionnaire. The questionnaires were the Scherer et al. (1994) Affective Reactions Scale, and were used to determine how the participants felt after reading the article and watching the video clip.

Participants were told that the professor intended to use the results as part of his film appreciation class he teaches.

Participants were also told that the primary investigator was working with the Dean of the college who was interested in developing a new performance appraisal system to evaluate the teaching performance of teaching assistants. Participants were told that the Dean was interested in developing a computer based evaluation system so that students would not have to spend time in class filling out evaluation forms. Participants were told that the second study would be conducted in a separate room across the hall and would be explained in more detail later.

After completing the article and video rating tasks, participants moved to another room and were shown to individual computer rooms. Participants were told that they would be evaluating the teaching performance of 30 teaching assistants on the teaching assistants' planning for lecture and presentation of lecture using an information board. The information board was described to the participants. Participants were told to click on a box and read the information completely before selecting another box, and once they felt they had enough information to rate the performance of the teaching assistant they should use the 7-point scale to evaluate the teaching assistant's performance. Participants were then instructed on how to access the computer program where they were given more detailed instructions about the task and were presented with the accountability instructions.

Upon the completion of the performance appraisal task, participants completed a demographic questionnaire (see Appendix E), the Affective Reactions Scale (Scherer et.

al., 1994), and the CCDAS (Costello & Comrey, 1967). Participants were then debriefed as to the true nature of the experiment, thanked for their time, and asked not to talk about the experiment with anyone who might participate in the experiment in the future.

Results

Pilot Study I

The purpose of pilot study 1 was to determine the perceived effectiveness of various teaching behaviors. Teaching behaviors were categorized as either planning for class, or presentation of class material. Each category of teaching behavior contained six dimensions of behavior, and each dimension contained five statements for a total of 60 behavioral statements, which were thought to vary between highly ineffective and highly effective. Once the perceived effectiveness of the teaching behaviors were determined, the statements were used to create the information boards for the main study. Additionally, once all of the statements were rated for effectiveness, the overall teaching ability of the fictitious teaching assistants was determined.

Method

Participants. Forty-five undergraduate students from a medium sized Midwestern university were instructed to rate the perceived effectiveness of twelve teaching behaviors.

Stimulus material and task. The twelve teaching behaviors were categorized as either planning for class (six statements), or presentation of class material (six statements). The six planning statements were, a) understanding, b) organization, c) knows the material, d) time spent preparing, e) punctuality, and f) current grade point

average. The six dimensions of presentation were, a) clarity, b) stimulates thinking, c) enthusiasm, d) engagement, e) answers questions, and f) year in school.

Participants evaluated the effectiveness of 12 behavioral statements (one statement per dimension of teaching behavior) using a 7-point Likert type scale. The scale ranged from 1 (not at all effective) to 7 (extremely effective). To ensure that participants were not influenced by the set of items, the 60 behavioral statements were randomized among ten questionnaires. That is, nine participants rated each statement, and of the nine, only four or five of the participants responded to the same questionnaire.

Results and Discussion

To determine the effectiveness of each behavioral statement, the means and standard deviations of each statement were calculated (see Appendix F). Results suggest that the five statements of each dimension vary in the perceived effectiveness of the teaching behavior.

Once the effectiveness of the teaching behaviors was established, the statements were used to create the information boards. The statements were entered into the information boards such that all of the fictitious teaching assistants would have similar average level of overall teaching effectiveness (see Appendix G and H).

Pilot II

The purpose of pilot study two was twofold. The first goal of the study was to determine student perceptions of the relative importance of outcomes that would result from a negative performance appraisal evaluation. The second goal of the study was to determine, given specific negative consequences of poor performance appraisal

evaluations, the likelihood of a student giving a teaching assistant a rating that was higher than warranted given the actual level of performance. Once the performance appraisal outcome rated highest in importance and highest in the likelihood to elicit inflated ratings was determined, the scenario would be used as the basis for the accountability instructions. A single scenario will be used to control for the effects of importance of the outcome.

Method

Participants. Nine undergraduate students from a Midwestern university participated in the second pilot study.

Stimulus materials and task. Participants were presented with eight possible outcomes of a poor performance appraisal rating. Four of the outcomes related to group outcomes (e.g. all teaching assistants with below average teaching evaluations will lose their assistantships), and four outcomes related to individual outcomes (e.g. the teaching assistant with the lowest teaching evaluation will lose his or her assistantship) (see Appendix I and J). Statements related to individual and group outcomes were counterbalanced to prevent order effects. Participants evaluated the importance of the eight outcomes using a 6-point Likert type scale. Responses ranged from 1, "very unimportant to me", to 6 "very important to me". Participants then evaluated the same statements on the likelihood that they would give a higher evaluation of a teaching assistant's performance than was warranted, given the probable outcome of a poor evaluation. Participants rated their likelihood to inflate ratings using a Likert type scale that ranged from 1 "very unlikely", to 6 "very likely."

Results and Discussion

Importance and likelihood ratings were established by averaging the ratings across all participants. Results indicate that participants perceived all of the individual and group outcomes as important (see Table 4 and 5). Additionally, participants indicated that they were likely to inflate their ratings given the probable consequences of poor evaluations. Because all of the statements were similar in importance, the group statement relating to "all low performing teaching assistants losing their positions, resulting in the elimination of the graduate program and the reduction in the number of undergraduate courses offered at the university," was selected as the scenario which will be used as part of the main study.

Pilot III

Prior research by Herman (2004) found the Marichal and Mathonet (2000) article titled "On comparison meaningfulness of aggregation functions" elicited a neutral mood state in participants. The purpose of this pilot study was to ensure that the aforementioned article would still elicit a neutral mood state. The article would then be used for the current research to elicit a neutral mood state in participants before exposing participants to positive or negative mood inducing stimuli.

Method

Participants. Twenty-one undergraduate students enrolled in a junior level psychology course at a Midwestern university participated in the third pilot study.

Stimulus material and task. For this pilot study, participants read the Marichal and Mathonet (2000) article for about seven to ten minutes and then completed the Scherer et.

al. (1994) Affective Reactions Scale. The scale consists of 42 adjective pairs that comprise five scales (negative arousal, positive arousal, boredom, fear, and elation) and a 6-point scale used to evaluate which adjective best describes how the participants felt after reading the article. Higher scores indicate higher levels of affective reactions.

Results and Discussion

A one-sample t-test was conducted for each sub-dimension. Sub-dimension scores were then compared to the median of the scales (*Mdn* =. 3.5) to test for significant affective reactions. Results of the one-sample t-tests indicated that participants had generally neutral reactions to the article. That is, participants evinced significantly high levels of boredom, low levels of negative arousal, and significantly low levels of positive arousal. Additionally, participants did not evince significant levels of elation (see Table 1.). Thus, the article appears to elicit high levels of boredom and fails to elicit any type of arousal (positive and negative arousal were actually lowered), or other positive (elation) and negative (fear) reactions in participants. As such, the article was used to elicit a neutral affective state and thereby control for current affective states of participants.

Table 1.

Affective Reaction after Reading the Mood Neutralizing Article

Dimension	*M*	*SD*	*difference*	*df*	*t*
Negative Arousal	3.03	.69	-.47	20	3.11*
Elation	3.22	.72	-.18	20	-1.14
Fear	3.50	.88	.00	20	.03
Boredom	4.07	1.09	.57	20	2.38*
Positive Arousal	2.83	.73	-.67	20	4.21*

Note. N = 21, median = 3.50

* $p < .05$

Pilot Study IV

As part of the main investigation, mood was manipulated by having the participants watch a 10 to 15 min. video clip. The purpose of this pilot study was to determine which video clips elicited positive or negative affective reactions.

Method

Participants. Fifty-three undergraduate students from a midwestern university participated in the fourth pilot study.

Stimulus materials and tasks. At the beginning of the study, all participants read the Marichal and Mathonet (2000) mood-neutralizing article to ensure that current affective states would not interfere with the mood video manipulation task. After reading the mood-neutralizing article, participants completed the Scherer et al. (1994) Affective Reactions Scale. Upon completion of the questionnaire, participants watched a 10- to 15-minute video clip of a suspected negative mood eliciting video (Mad Max or Midnight Express), or a 10 to 15 minute video clip of a suspected positive mood eliciting video (Ferris Bueler's Day Off or Stripes). Upon completion of the video, participants completed the Scherer et al. (1994) Affective Reactions Scale. The five sub-dimensions of the Scherer et al. (1994) Affective Reactions Scale were then analyzed to determine the affective nature of each video.

Results and Discussion

For each video, the means for each of the sub-dimensions of the Scherer et al. (1994) Affective Reactions Scale were compared to the median of the scales (*Mdn = 3.5*). Results indicated that the video clip from the movie Mad Max elicited low levels of

boredom and high levels of positive arousal. The video did not elicit affective reactions related to negative arousal, elation, or fear (see Table 2). That is, Mad Max is not boring, seemed to elicit positive arousal, but does not appear to elicit a negative mood state as anticipated.

The video clip from Ferris Bueller's Day Off elicited low levels of negative arousal, fear, and boredom (see Table 2.). Ferris Bueller's Day Off also elicited high levels of elation and positive arousal (see Table 2.). That is, Ferris Bueller's Day Off appeared to elicit a positive mood state in participants.

Participants' affective reactions to Stripes was similar to participants' reactions to Ferris Bueler's Day Off. That is, the Stripes video clip elicited high levels of elation and positive arousal, and low levels of negative arousal, fear, and boredom (see Table 2.). Thus, Stripes appeared to elicit a positive mood state.

For the video Midnight Express, the affective reactions of the participants were in the expected direction, but were not significantly different from the median of the scale (see Table 2.). The lack of significant difference could indicate that the video placed participants in a neutral mood state. However, prior research by Weiss (1996) found that Midnight Express elicited a negative mood state. One explanation for the present results could be the presence of the researcher in the room during the presentation of the video, and while participants completed the affective reactions scale. That is, the presence of the researcher in the room during the experiment could have influenced the participants' reactions to the video. Rosenthal (1976) states that expectations of the researcher can influence the performance of the participant in a study. During the experiment, the

researcher tried to remain passive (without expression) during the presentation of the videos. However, the lack of expression could have unintentionally elicited a neutral response to videos. Perhaps participants focused more on the experimenter's reaction for a guide on how to react rather than how the video made them feel.

Given the results of pilot study four, the limitation of the presence of the researcher, and prior research, Midnight Express and Ferris Bueller's Day Off were selected as the mood manipulation videos for the primary study.

Table 2.

Means and Standard Deviations for Each Sub-Dimension of the Affective Reactions Scale for each Video

Reaction	*n*	*M*	*SD*	*difference*	*t*
			Mad Max		
Negative Arousal	15	3.43	.67	-.07	-.40
Elation	14	3.59	.54	.09	.62
Fear	15	3.68	.64	.18	1.10
Boredom	15	2.69	.88	-.81	-3.53*
Positive Arousal	15	4.22	.90	.72	3.10*
			Ferris Bueller's Day Off		
Negative Arousal	12	1.98	.36	-1.52	-14.58*
Elation	12	5.31	.46	1.81	13.74*
Fear	12	2.22	.59	-1.28	-7.49*
Boredom	12	2.67	.51	-.83	-5.69*
Positive Arousal	12	4.67	.54	1.17	7.53*
			Midnight Express		
Negative Arousal	12	3.50	1.15	.00	.00
Elation	12	3.11	.93	-.39	-1.44
Fear	12	3.52	1.22	.02	.05
Boredom	12	3.68	1.36	.18	.47
Positive Arousal	12	3.21	1.24	-.29	-.81
			Stripes		
Negative Arousal	14	2.55	.76	-.95	-4.64*
Elation	14	4.96	.57	1.46	9.57*
Fear	14	2.59	1.02	-.91	-3.37*
Boredom	14	2.44	.67	-1.06	-5.89*
Positive Arousal	14	4.68	.58	1.18	7.66*

* $p < .05$.

Pilot Study V

Researchers have suggested that individuals in positive moods tend to use a heuristic information processing strategy when evaluating information about a target, whereas as individuals in negative moods tend to use a substantive information processing strategy (Forgas, 1994; Martin, Ward, Achee, & Wyer,1993). That is, individuals in positive moods are likely to use less information when making a judgment regarding a target, whereas individuals in negative moods are likely to search through more information before making a judgment. Prior research on the impact of affective state on the decision making process has found that affective state influences the type of information evaluated (Mischel, Ebbesen, & Zeiss, 1973), and the type of information recalled about a target (Isen, Shalkar, Clark, & Karp, 1978; Sinclair, 1988). Research by Elsbach and Barr (1999) found that individuals in negative moods were more likely than individuals in positive moods to correctly follow a detailed decision protocol. However, Elsbach and Barr did not examine the pattern in which the information was searched.

Research by Isen and Means (1983) examined the amount of information searched and search strategy utilized by participants in positive affective states compared to control participants. The results indicated that participants in positive affective states searched less information when making decisions to purchase a fictitious car compared to control subjects (Isen & Means, 1983). However, Isen and Means (1983) failed to examine the impact of negative affective state on the amount information searched and the type of search strategy employed. The goal of the current pilot study was to determine

if affective states (positive vs. negative) influence the amount of information searched, and the pattern in which information was searched.

Method

Participants. Thirty undergraduate students from a Midwestern university participated in pilot study five. However, due to technical problems with the computer program used to measure the amount of information searched and the type of search pattern utilized by participants, only 22 participants are included in this investigation. Of the twenty two participants included in this investigation, only 12 completed the information search task.

Stimulus material and task. At the beginning of the study, participants were told that they would be participating in two studies. The first study was to gather student reactions to different types of presented information (article vs. video). Participants were informed that a professor at local area college was colleting this information and planned to use the results for a presentation in his Film Appreciation Class. Participants were also told that they would be participating in a computerized evaluation task in another room.

For the first part of the investigation, all participants read the Marichal and Mathonet (2000) mood neutralizing article to ensure that current affective states would not interfere with the mood video manipulation task. After reading the mood-neutralizing article, participants completed the Scherer et al. (1994) Affective Reactions Scale. Upon completion of the questionnaire, participants watch a 10 to 15 minute video clip of either Midnight Express (negative mood elicitor), or Ferris Bueler's Day Off (positive mood elicitor). The experimenter left the room during the viewing of the video to prevent the

presence of the researcher from influencing the participants' affective reactions. After watching the video clip, participants completed the Scherer et al. (1994) Affective Reactions Scale indicating how they felt after watching the video.

Upon completion of the video rating task, participants moved to another room to complete the computer based evaluation task. For the evaluation task, participants were instructed that they would be evaluating the teaching performance of 30 fictitious teaching assistants (TA's) on the teaching assistants planning for lecture and presentation of material in class. Participants were told that the information regarding the TA's performance would be displayed on an information board. After a brief description of an information board and how to use the board, participants accessed the computer program and read the no-accountability instructions on the purpose of the information board. The no accountability instructions (Appendix D) promise confidentiality of ratings and state the possible consequences of poor evaluations: a) all TA's with low performance will be terminated, b) the graduate program will be eliminated, and c) fewer undergraduate classes will be offered. Participants were debriefed at the end of the study as to the purpose of the video rating task and the evaluation task.

Results and Discussion

Article manipulation check. Results indicate that participants were in a neutral mood state after reading the article (see Table 3.). Regardless of mood condition, participants did not differ in their affective reactions to the article. Additionally, all participants reported significantly high levels of boredom and significantly low levels of positive arousal compared to the median of the scale (*Mdn* = 3.50).

Video manipulation check. Results indicate that participants who viewed the positive mood inducing video compared to participants who viewed the negative mood inducing video, evinced significantly lower levels of negative arousal (F [1,20] = 15.88, p < .05) and fear (F [1,20] = 12.78, p < .05), and significantly higher levels of elation (F [1,20] = 16.16, p < .05). Participants did not differ in their affective reactions to the videos on dimensions of boredom or positive arousal (see Table 4.). Thus, the participants were experiencing different mood states depending on the video watched.

Mood and information search. For this investigation, ten of the twenty-two participants did not make all 60 ratings of the teaching assistants' teaching performance. As such, all analyses were conducted using an ANCOVA controlling for the number of ratings made. That is, people who made fewer ratings spent less time on the experiment and searched less information because they did not complete the rating task.

For the total amount of information searched, results indicate no significant difference between mood conditions, but the results were in the expected direction. That is, participants in the positive mood condition searched less information per teaching assistant (M = 5.00, SD = 1.64) compared to participants in the negative mood condition (M = 5.63, SD = 2.71).

Results also suggest that participants, regardless of mood, search information in an inter-dimensional fashion. That is, participants look at each teaching dimension for a teaching assistant before searching for information about the next teaching assistant.

All other results were also not significant, but were in expected directions. That is, on average participants in a positive mood spent less time searching information (M =

18.74 min, *SD* = 2.04 min) than participants in a negative mood (*M* = 19.54 min, *SD* = 1.86 min), and participants in a positive mood rated the overall performance of the teaching assistant higher (*M* = 3.76, *SD* = .51) than participants in a negative mood (*M* = 3.66, *SD* = .68). Several limitations could have influenced the results.

Limitations. One limitation of the pilot study was the presence of the researcher in the room during the evaluation task. The presence of the researcher could have prompted the participants to search more information. Additionally, Participants were observed clicking on all of the information boxes, which influences the information actually used and the time spent on the task.

Another limitation of this investigation was the presence of other participants in the room during the evaluation task. That is, participants were aware when other participants were having computer problems, and when other participants finished. Participants seemed to rush when they felt they were the last ones to complete the study, which would influence the time spent on task, and could influence the amount of information searched and the rating of the teaching assistants.

These limitations were corrected for the main study. That is, the larger number of participants increased the power of the study. Additionally, participants were shown to individual computer rooms, thus, removing the influence of the experimenter and the presence of other participants.

Table 3.

Means and Standard Deviations for each Sub-Dimension of the Affective Reactions Scale after Reading the Article

	Positive		Negative	
Affective Reaction	*M*	*SD*	*M*	*SD*
Negative Arousal	3.45	.80	3.82	.65
Elation	3.47	.65	3.3	.84
Fear	3.20	.79	3.54	1.39
Boredom	5.00	.83	4.6	1.14
Positive Arousal	2.42	.41	2.43	.79

Note. for positive mood, n = 10; for negative mood, n = 7

Table 4.

Means and Standard Deviations for each Sub-Dimension of the Affective Reactions Scale after Watching a Positive Mood Inducing and a Negative Mood Inducing Video Clip

	Ferris Bueler's Day Off		Midnight Express	
Affective Reaction	*M*	*SD*	*M*	*SD*
Negative Arousal	2.46	.39	3.93	1.11
Elation	4.58	.76	3.25	.78
Fear	2.30	.45	3.93	1.38
Boredom	3.36	1.09	2.93	.95
Positive Arousal	3.70	1.27	3.29	1.23

Note: for Ferris Bueler's Day Off, n = 10; for Midnight Express, n = 12

Primary Investigation

Article Manipulation Check

At the beginning of the experiment, participants read a mood-neutralizing article to ensure that their current mood state would not interfere with the effect of the mood manipulation task. Results of a one-way ANOVA indicate that regardless of mood condition there were no significant differences in the participants' affective reactions to the article. Additionally, when compared to the median of the scale (*Mdn* = 3.50), all participants evinced significantly high levels of boredom ($M = 4.56$, $t[93] = 10.64$, $p < .01$), significantly low levels of positive arousal ($M = 2.62$, $t[93] = -9.90$, $p < .01$), and marginally low levels of negative arousal ($M = 3.33$, $t\ [93] = -2.01$, $p < .05$). Thus, the article was effective in neutralizing the participants' mood prior to watching the mood manipulation video.

Video Manipulation Check

Prior research found that watching a video clip is an effective method of eliciting either a positive or negative mood in participants (Isen, Daubman, & Nowicki, 1987). After watching a 15-minute video clip, participants completed the Scherer et al. (1994) Affective Reactions Scale. A t-test was conducted to compare dimension scores to the median of the scale (*Mdn* = 3.50) to determine how the participants reacted to the video they watched. Results indicate that the positive mood inducing video elicited low levels of negative arousal, fear, and boredom, and high levels of elation and positive arousal (See Table 5). Results also indicate that the negative mood inducing video elicited

significantly high levels of negative arousal and fear, significantly low levels of elation and positive arousal, and neutral levels of boredom in the participants (see Table 6).

A one-way ANOVA was then conducted to verify that the videos elicited different affective reactions. Results indicate that participants differed in their affective reactions to the videos on all dimension of the Affective Reactions Scale except boredom (see Table 7). The results indicate that there was a significant difference in the homogeneity of variance between the two groups on the negative arousal scale (*Levene* $[1,92] = 7.90, p < .01$). However, according to Stevens (2002), ANOVA is generally robust to the violation of equal variances when the cell sizes are equal; both the positive and negative conditions had 47 participants. Thus, participants who watched the positive mood eliciting video appeared to be in a positive mood, whereas participants who watched the negative eliciting video appeared to be in a negative mood. As such, the mood manipulation task was effective in altering the participants' current mood state.

Accountability Manipulation Check

Accountability was manipulated by the written instructions given to the participants before engaging in the information board rating tasks. Non-accountable participants were informed that their ratings would be kept confidential, whereas process and outcome accountability participants were informed that "Time permitting" the dean would be here to discuss the information they used to determine the performance ratings given to teaching assistants.

Table 5.

Affective Reactions of Participants after Watching the Positive Mood Video Clip

Dimension	*M*	*SD*	*difference*	*df*	*t*
Negative Arousal	2.48	.61	-1.02	46	-11.56*
Elation	5.05	1.03	1.55	46	10.27*
Fear	2.09	.82	-1.41	46	-11.77*
Boredom	2.91	.86	-.59	46	-6.68*
Positive Arousal	4.29	1.19	.79	46	4.57*

Note. n = 47; median = 3.50

* $p < .01$

Table 6.

Affective Reactions of Participants after Watching the Negative Mood Video Clip

Dimension	*M*	*SD*	*difference*	*df*	*t*
Negative					
Arousal	3.95	.91	.45	46	3.41*
Elation	2.93	.88	-.57	46	-4.43*
Fear	4.03	.95	.53	46	3.78*
Boredom	3.04	1.03	-.46	46	-3.05*
Positive					
Arousal	3.35	.96	-.15	46	-1.04

Note. n = 47, median = 3.50

* $p < .01$

Table 7.

ANOVA Results Comparing Affective Reactions between Mood Conditions

	SS	*df*	*MSE*	*F*	*p*
Negative Arousal					
Between	51.21	1	51.21	85.51	0.000
Error	55.10	92	.60		
Total	106.31	93			
Elation					
Between	105.44	1	105.44	114.22	0.000
Error	84.93	92	0.92		
Total	190.36	93			
Fear					
Between	87.90	1	87.90	110.99	0.000
Error	72.87	92	0.79		
Total	160.77	93			
Boredom					
Between	.41	1	.41	.45	0.503
Error	83.02	92	.90		
Total	83.43	93			
Positive Arousal					
Between	20.67	1	20.67	17.72	0.000
Error	107.32	92	1.17		
Total	127.99	93			

Note. N = 94

To determine if participants felt accountable during the information board rating tasks, participants were asked, at the end of the study, the extent to which they believed the dean would be there to discuss their ratings.

There were 33 participants in the no-accountability condition, 16 in the positive mood condition and 17 in the negative mood condition. The mean score of the accountability check question was compared to the median of the scale (*Mdn* = 3) to determine if participants felt accountable. Results suggest that participants in the no-accountability condition did not feel accountable ($M = 2.61$, t [32] = -2.14, $p < .05$) with the majority of participants indicating that they were not sure if the dean was coming (i.e. answered 3 to the accountability check question).

In the process-accountability condition, there were 29 participants, 16 in the positive mood condition and 13 in the negative mood condition. On average, participants in the process accountability condition tended to believe or were at least uncertain as to whether the dean would be coming to speak with them at the end of the study ($M = 2.79$, t [28] = -1.09, *ns*). Results also indicate that six participants in the positive mood condition and three in the negative mood condition did not believe the dean was coming. That is, the accountability manipulation did not work on these individuals. As such, these nine participants were removed from the data set and the accountability manipulation check was recalculated. The results indicate that participants believed the dean might be coming to speak with them at the end of the study ($M = 3.35$, t [19] = 2.67, $p < .05$).

Finally, for the outcome-accountability condition there were 31 participants, 15 in the positive mood condition and 16 in the negative mood condition. Results of the

manipulation check indicate that participants were not certain if the dean was coming to speak with them at the end of the study ($M = 2.87$, t [30] = -.64, *ns*). However, the results also indicate that three participants in the positive mood condition and seven participants in the negative mood condition did not believe the dean would be coming to speak with them at the end of the study. That is, the accountability manipulation did not work. As such, these ten participants were removed from the investigation. Once these ten participants were removed from the study, participants in the outcome accountability condition tended to believe that the dean would be coming to speak with them at the end of the study ($M = 3.48$, t [19] = 2.91, $p < .05$).

After removing the nine participants in the process accountability condition who did not feel accountable, the ten participants in the outcome accountability condition who did not feel accountable, and the one participant who scored high in anxiety and depression, a total of 20 participants were removed from the study leaving a total sample size of 74 participants.

Test for Normality and Extreme Cases

Before proceeding with hypothesis testing, it is important to ensure that the dependent variables of amount of information searched and performance ratings are normally distributed. According to Stevens (2002) excessive skewness and Kurtosis can attenuate power. In addition to testing the normality of the distribution, it is also important to test for outliers and influential cases as these participants can impact the results of the investigation.

Descriptive statistics were collected for the dependent variables of total amount of information searched by participants, the average amount of information per teaching assistant searched by participants, and the overall average performance rating given to teaching assistants. The examination of the assumption of a normal distribution indicates that the three dependent variables have non-problematic levels of skewness and kurtosis (see Table 8.). That is, the dependent variables have normal distributions.

In addition to examining the normalcy of the distribution, standard scores were calculated to test for possible outlier and influential cases. According to Stevens (2002) an outlier or influential case could exist if the standardized score on the variable of interest was greater than the absolute value of three standard deviations from the mean. Based on the criteria espoused by Stevens (2002), two participants have extreme scores on the total amount of information searched. One participant in the no accountability/ positive mood condition had a z-score of -3.47, and one participant in the no accountability/ negative mood condition had a z-score of -3.31. All other z-scores fell between three standard deviations above and below the mean.

Further examination of the two cases suggested that two participants were members of the sample of interest. However, the cases could be influential; as such, hypothesis testing was conducted once with the two participants included in the analysis and once with the participants excluded.

Table 8.

Descriptive Statistics for the Total Amount of Information Searched, The Average Amount ofInformation Searched, and Average Performance Rating Given to Teaching Assistants

					Skewness		Kurtosis	
	min	max	*M*	*SD*	statistic	S. E.	statistic	S. E.
Total info Searched	124	360	323.76	57.56	-1.76	.28	2.71	.55
Avg. Info Searched	4.13	12	10.79	1.92	-1.76	.28	2.71	.55
Avg. Pefr. Rate	2.57	5.42	3.89	.58	.19	.28	0.24	.55

Note. N = 74

Correlations

Correlations were calculated between demographic information of age, gender, class standing, the CCDAS depression and anxiety scales, and the outcome variables of average performance rating, amount of information searched, total amount of time spent on the information board task, and the average amount of time individuals spent examining each piece of information.

Results indicated a significant negative correlation between gender and overall average performance rating ($r = -.25, p < .05$) suggesting that women tended to evaluate the teaching performance of teaching assistants less favorably than males. The significant positive correlation between age and class standing ($r = .49, p < .05$) suggested that older individuals tend to be upperclassmen. The significant correlation between anxiety and total amount of information searched ($r = .23, p < .05$) suggested that anxious individuals tend to search through more information compared to less anxious individuals. The significant relationship between age and the amount of time participants spent on the task ($r = .23, p < .05$) suggested that older participants tended to spend more time on the task compared to younger individuals.

Additionally, the significant positive correlation between the total time spent on the task and the average amount of time participants examined each piece of information ($r = .60, p < .05$) suggests that people who spent more time on the information search task tended to spend more time examining each piece of information. However, there was a significant negative relationship between average amount of time spent examining each piece of information and the total amount of information that was examined ($r = -.58, p <$

.05). That is, participants who searched less information tended to spend more time examining the information they accessed.

Hypothesis Testing

Recall the methodology; participants searched for information on six information boards, and each board contained the initial of ten fictitious teaching assistants. Additionally, each board contained six statements related to teaching performance on planning and presentation for a total of 60 statements per board, and 360 total pieces of information.

Hypothesis 1. Hypothesis one predicted that accountability would influence the amount of information searched by participants. That is, participants in the no-accountability condition were expected to search less information compared to participants in the process and outcome accountability conditions, whereas no difference was expected in the amount of information searched by participants in the process and outcome accountability conditions.

Hypothesis one was evaluated using a 2 (mood: positive vs negative) X 3 (accountability: no, process, and outcome) by 2 (order: planning boards first vs presentation boards first) between subjects design. Although participants were randomly assigned to receive either the three planning boards first or the three presentation boards first, order was included in the model to ensure the order in which the information boards were presented did not affect the amount of information searched by participants.

The results indicate that there was no significant main effect for order, nor were any of the interactions with order significant. Additionally, there were no other

significant main effects or interactions on the amount of information searched by participants. As such, order was removed from the model and the results were rerun using a 2 (mood) X 3 (accountability) between subjects design.

On average, participants in the no accountability condition examined 327 pieces of information, participants in the process accountability condition examined about 313 pieces of information, and participants in the outcome accountability condition examined about 328 pieces of information (see Table 9.). The results indicate that there were no significant difference in the amount of information searched based on the type of accountability to which a person was exposed ($\eta^2 = .011$, F [2,71] = .36, *ns*), for mood of the participants ($\eta^2 = .019$, F [1,72] = 1.33, *ns*), nor for the interaction between mood and accountability ($\eta^2 = .008$, F [2,71] = .77, *ns*). Thus, hypothesis one, predicting differences in the amount of information searched as evidence of different processing strategies used to evaluate information was not supported. Finally, removing the two participants with extreme scores did not change the results of the analysis.

*Hypothesis 2. H*ypothesis two predicted that there would be a significant interaction between mood and accountability on the average performance rating given to teaching assistants. Specifically, it was expected that participants in the positive mood condition would give more positive evaluations of performance compared to participants in the negative mood condition for the no-accountability and process accountability conditions. However, for the outcome accountability no difference in average performance rating was expected between participants in positive and negative moods.

Table 9.

Means and Standard Deviations for the Total Amount of Information Searched

Accountability Condition	*M*	*SD*	*N*
	Positive Mood		
No	333.06	64.12	16
Process	330.20	51.64	10
Outcome	330.83	41.30	12
	Negative Mood		
No	321.76	61.21	17
Process	299.09	70.69	11
Outcome	324.63	51.20	8
	Total		
No	327.24	61.92	33
Process	313.90	62.86	21
Outcome	328.35	44.31	20

The data were analyzed using a 2 (mood: positive vs negative) by 3 (accountability: no, process, and outcome) by 2 (order: planning boards first vs presentation boards first) between subjects design on the overall average performance rating score across the six information boards. Results indicated a marginal interaction between mood and accountability on the average performance rating given to the teaching assistants ($\eta^2 = .077$, F [2, 71] = 2.57, $p = .085$). However, the interaction was not in the predicted form. It was predicted that for participants in the no accountability and the process accountability conditions an effect of mood on performance rating would occur , but for participants in the outcome accountability condition no effect for mood on performance rating was expected.

A probe of the interaction indicates that there was no effect of mood on performance rating for participants in the no accountability condition. Participants in the positive mood/ no accountability condition and the negative mood/ no accountability condition rated the teaching assistants about the same on their overall teaching performance (see Table 10.). There was a significant difference in the average performance ratings given by participants in the process accountability condition (partial $\eta^2 = .26$, F [1,19] = 6.32, $p = .022$), but the effects were in the opposite direction. That is, participants in the negative mood/ process accountability condition rated the overall average performance of the teaching assistants more favorably compared to participants in the positive mood/ process accountability condition (see Table 10.). Finally, there was no significant effect for mood on average performance rating for participants in the outcome accountability conditions (see Table 10.).

Table 10.

Means and Standard Deviations for Average Performance Rating

Accountability Condition	*M*	*SD*	*N*
	Positive Mood		
No	4.05	.61	16
Process	3.45	.61	10
Outcome	3.64	.53	12
	Negative Mood		
No	4.07	.42	17
Process	4.06	.59	11
Outcome	3.85	.60	8
	Total		
No	4.06	.51	33
Process	3.77	.66	21
Outcome	3.73	.55	20

Hypothesis 2, unexpected findings. Order was included in the model as a check to make sure that the order in which the boards were presented did not influence the average performance rating given to teaching assistants. The results indicate that there was a marginal interaction between board order and accountability on the average performance rating given to teaching assistants ($\eta^2 = .083$, F [2, 71] = 2.82, $p = .067$). A probe of the interaction indicates that, for participants in the no accountability and the process accountability conditions, there was no significant difference in average performance rating based on the order in which the boards were presented (see Table 11). However, the difference in the average performance rating for participants in the outcome accountability condition was approaching significance ($\eta^2 = .177$, F [1,18] = 3.86, $p = .065$) (see Table 11.) Because the interaction between accountability and order was marginal, and the probe of the interaction did not evince significant differences at conventional levels ($p < .05$), order was removed from the model and the analysis was rerun.

When order was removed from the model, resulting in a 2 (mood) by 3 (accountability) between subjects design, the results indicate that there was no significant interaction between mood and accountability on the average performance rating given to the teaching assistants ($\eta^2 = .051$, F [2,71] = 1.82, $p = .17$). Thus, hypothesis two was not supported.

Hypothesis 3. Hypothesis three predicted that there would be a main effect of mood on the average performance appraisal rating given to the teaching assistants, such that, participants in the positive mood condition were expected to give more favorable

performance ratings compared to individuals in the negative mood condition. Hypothesis three was not supported. There was a significant main effect for mood on the average performance rating given to teaching assistants, but it was in the opposite direction (η^2 = .062, F [1,72] = 4.47, $p < .05$). That is, participants in the negative mood condition tended to rate the teaching assistants more favorably compared to participants in the positive mood condition. The overall average performance rating given by participants in the positive mood condition (M = 3.76, SD = .63) was significantly lower than the average performance rating given by participants in the negative mood condition (M = 4.02, SD = .51). However, it appears as though both groups evaluated the teaching assistants' performance harsher than was actually warranted. To determine if participants actually rated the teaching assistants teaching performance harsher than was actually warranted, additional analyses were conducted.

The results of pilot study I were used to create the information boards and serve as a guide to determine the overall perceived teaching effectiveness of each teaching assistant, from which the overall average effectiveness of all the teaching assistants was calculated (M = 4.42). Separate T-tests were conducted for the positive and negative mood conditions to compare the average of each group's overall average performance rating to the predetermined overall average performance rating. There was a significant difference in the performance rating given by participants in the positive mood condition and the performance rating actually warranted based on the perceived effectiveness of the teaching behaviors. (M = 3.76, t [37] = -6.50, $p < .05$), as was the average rating given by participants in the negative mood condition (M = 4.02, t [35] = -4.78, $p < .05$).

Table 11.

Means and Standard Deviations of Average Performance Rating for Mood by Accountability by Order of Board Presentation

	Planning First			Presentation First		
	n	*M*	*SD*	*n*	*M*	*SD*
			Positive Mood			
No	10	3.92	.53	6	4.27	.72
Process	7	3.60	.48	3	3.11	.86
Outcome	8	3.47	.47	4	3.99	.52
			Negative Mood			
No	9	4.02	.51	8	4.13	.30
Process	6	4.16	.68	5	3.93	.50
Outcome	5	3.71	.73	3	4.09	.14
			Total			
No	19	3.96	.51	14	4.19	.50
Process	13	3.86	.63	8	3.62	.73
Outcome	13	3.56	.57	7	4.03	.38

Note. Total N = 74

Additional results indicate there is a strong trend for a main effect for accountability on the average performance rating given to teaching assistants ($\eta^2 = .077$, F [2,71] = 2.85, $p = .065$). Post hoc analysis using Fishers least significant difference, related that the average performance ratings between participants in the no-accountability condition ($M = 4.06$, $SD = .51$) were systematically more positive than the performance ratings provided by participants in the outcome accountability condition ($M = 3.73$, $SD = .55$) ($diff = .33$, $p < .05$). Participants in the process accountability condition provided average ratings ($M = 3.77$, $SD = .66$) that fell between, but were not statistically different from, the performance ratings provided by participants in the no-accountability and outcome accountability conditions.

Exploratory Analyses

For the primary analyses, the total amount of information searched and the average performance ratings of teaching assistants were calculated by collapsing information across six information boards. However, by collapsing across boards some information is lost. For example, there might be significant differences in the amount of information participants accessed across similar types of boards, either across the three planning boards, or across the three presentation boards. Additionally, there could be differences in the amount of information searched based on the type of board, comparing information searched on the planning vs. the presentation boards. These differences could extend to the average performance ratings. As such, several exploratory analyses adding a within-subjects variable were conducted. Before the analyses were conducted, additional analyses for extreme scores and influential cases were conducted.

Test for Outliers and Influential Cases

For the current study there are several sets of dependent variables. For example, participants rated the teaching performance of 30 teaching assistants on two dimension of teaching behavior across six information boards. That is, three information boards with ten teaching assistants per board were used to evaluate the teaching behavior of planning for lecture, and three boards with the same ten teaching assistants per board were used to assess the teaching behavior of presentation of lecture. Additionally, the amount of information a participant searched before making a performance evaluation for each board was calculated. Thus, for performance ratings and amount of information searched each participant had a total of 18 data points; three data points per type of board (planning vs. presentation) and three sets of measures. Additionally, average scores (corrected for non-independence) were calculated for the average performance ratings across the three planning and presentation boards and the average amount of information searched across the three planning and presentation boards. The six previously mentioned variables allow one to check for differences in performance ratings and information searched based on the type of board a participant was viewing.

According to Stevens (2002) the best procedure for determining the presence of outliers or influential cases on a set of variables is the use of Mahalanobis Distance (D^2). Mahalanobis Distance is used to calculate how discrepant a participant's scores are from the centroid of the scores for the entire population. Thus, D^2 was calculated for performance rating and amount of information searched across the three planning and the three presentation boards, resulting in four calculations of D^2. Because average

performance ratings and average amount of information searched were calculated across each board so that scores could be compared based on type of board examined, two additional measures of Mahalanobis distance were calculated.

According to Stevens (2002) a case could be considered an influential data point if the calculated value of D^2 is greater than the critical value of D^2, which is determined by the number of participants and the number of variables measured. For this investigation, N equals 74, and for the first six D^2 calculations there are three variables (one for each board). Extrapolating from the table presented by Stevens (2002, p. 133), D^2 values greater than 15.32 should be closely examined to determine if the case is an outlier or influential. For the remaining three D^2 calculations there are two variables; so extrapolating from the table presented by Stevens (2002, p. 133), D^2 values greater than 13.22 should be closely examined to determine if the case is an outlier or influential.

Results of the outlier and influential cases analyses on the performance ratings on the three planning boards indicates the existence of one possible influential case (case 102036, $D^2_{calc} = 17.48 > D^2_{crit}$ 15.32). For the performance ratings on the three presentation boards, the results indicated the presence of one influential case (case 302095, $D^2_{calc} = 32.42 > D^2_{crit}$ 15.32).

Next, outlier and influential case analyses for the amount of information searched on the three planning boards and three presentation boards were conducted. Results for the three planning boards indicate that there are possibly five influential cases (cases 101007, 302092, 102036, 501125, and 101003; $D^2_{calc} =$ 15.04, 23.61, 24.46, 24.68, and 40.13 > D^2_{crit} 15.32, respectively). Case 101007 was included as a possible influential

case for two reasons. First, the calculated value is close to the critical value, and the critical value was extrapolated from a table which serves as a guide to determining influential cases. Finally, the calculated D^2 was much larger than the next lowest calculated value of $D^2 = 7.89$. Results for the three presentation boards indicate the presence of three possible influential cases (cases 101003, 301007, and 501136; $D^2_{calc} =$ 20.05, 29.11 and 44.22 > D^2_{crit} 15.32, respectively)

Finally, Mahalanobis distances were calculated for the average performance rating given on the planning and presentation boards, and the average amount of information searched on the planning and presentation boards. For the average performance ratings given on the planning and presentation boards, there appear to be no influential cases (all $D^2_{calc} < D^2_{crit}$ 13.22). However, the analysis indicates the presence of four possible influential cases on the amount of information searched (cases 101007, 302093, 101003, and 102036; D^2_{calc} = 12.88, 17.23, 17.82, and 20.77 > D^2_{crit} 13.22, respectively). Case 101007 was include for the same two reasons as before; a) the calculated value is close to the critical value, and b) the calculated value is much larger than the next lowest value of $D^2 = 6.74$.

Further examination of all the cases suggests that none of the cases are outliers as all participants were drawn from the population of interest (i.e. undergraduate psychology students). However, some of the cases could still be influential and thus alter the results of the analysis; as such all analyses were conducted with and without the possible influential cases.

Exploratory Hypothesis Testing

Hypothesis I. Hypothesis one predicted that the amount of information searched by participants would differ depending on the accountability condition to which they were exposed. That is, participants in the no-accountability condition were expected to search less information compared to the process and outcome accountability conditions, and no difference was expected in the amount of information searched by participants in the process and outcome accountability conditions. However, because participants are required to search for information on three boards related to the same type of teaching behavior, one cannot assume that the amount of information searched on each board is independent of the other two boards, Thus, in one analysis, the average amount of information searched was analyzed using a 2 (mood: positive vs negative) X 3 (accountability: no, process, and outcome) X (3) (planning set: boards 1, 2, and 3). In the second analysis, the amount of information searched was analyzed using a 2 (mood: positive vs negative) by 3 (accountability: no, process, and outcome) by (3) (presentation set: boards 4, 5, and 6) mixed designs. Because order did not have any significant effects for the initial analyses, it was left out of the current exploratory analyses.

The results of the within-subjects analysis indicated a significant linear trend in the amount of information searched on the planning boards (partial $\eta^2 = .198$, $F\,[1,62] = 16.80$, $p < .001$), such that participants tended to search through the most information on the first board and less information on subsequent boards (see Table 12.). Within-subjects analysis indicated a significant linear trend in the amount of information searched across the three presentation boards (partial $\eta^2 = .083$, $F\,[1,62] = 6.12$, $p < .02$)

(see Table 13.). However, there were no other significant within-subject effects or between-subject effects on the amount of information searched. Thus, hypothesis one. predicting differences in the amount of information searched as evidence of different processing strategies used to evaluate information, was not supported.

Prior analysis indicated the presence of four possible influential cases on the amount of information searched on the planning boards. As such, the four cases were removed from the sample, and the analysis was rerun. Removing the four cases did not change the results of the analysis.

Prior analysis also indicated the presence of three possible influential cases on the amount of information searched on the three presentation boards. As such, the three cases were removed from the sample, and the analysis was rerun. When the three cases were removed there was a significant within-subjects interaction between mood and board on the linear trend in the amount of information searched (partial $\eta^2 = .076$, $F\,[1,65] = 5.37$, $p < .024$). The nature of the interaction is such that participants in the positive mood condition searched a similar amount of information on all three presentation boards, whereas participants in the negative mood condition tended to search through less information on subsequent presentation boards (see Table 14.) (see Figure 4). The remaining within-subjects and between-group effects for mood and accountability on the amount of information searched did not change with the removal of the three influential cases.

Table 12.

Means and Standard Deviations for the Total Amount of Information Search on Each Planning Information Board Collapsing Across Mood and Accountability

	M	*SD*
Board 1	55.38	10.31
Board 2	53.43	10.84
Board 3	51.16	13.69

Note. N = 74

Table 13.

Means and Standard Deviations for the Total Amount of Information Search on Each Presentation Information Board

	M	*SD*
Board 4	55.59	8.99
Board 5	54.30	9.80
Board 6	53.89	10.19

Note. Total N = 74

Table 14.

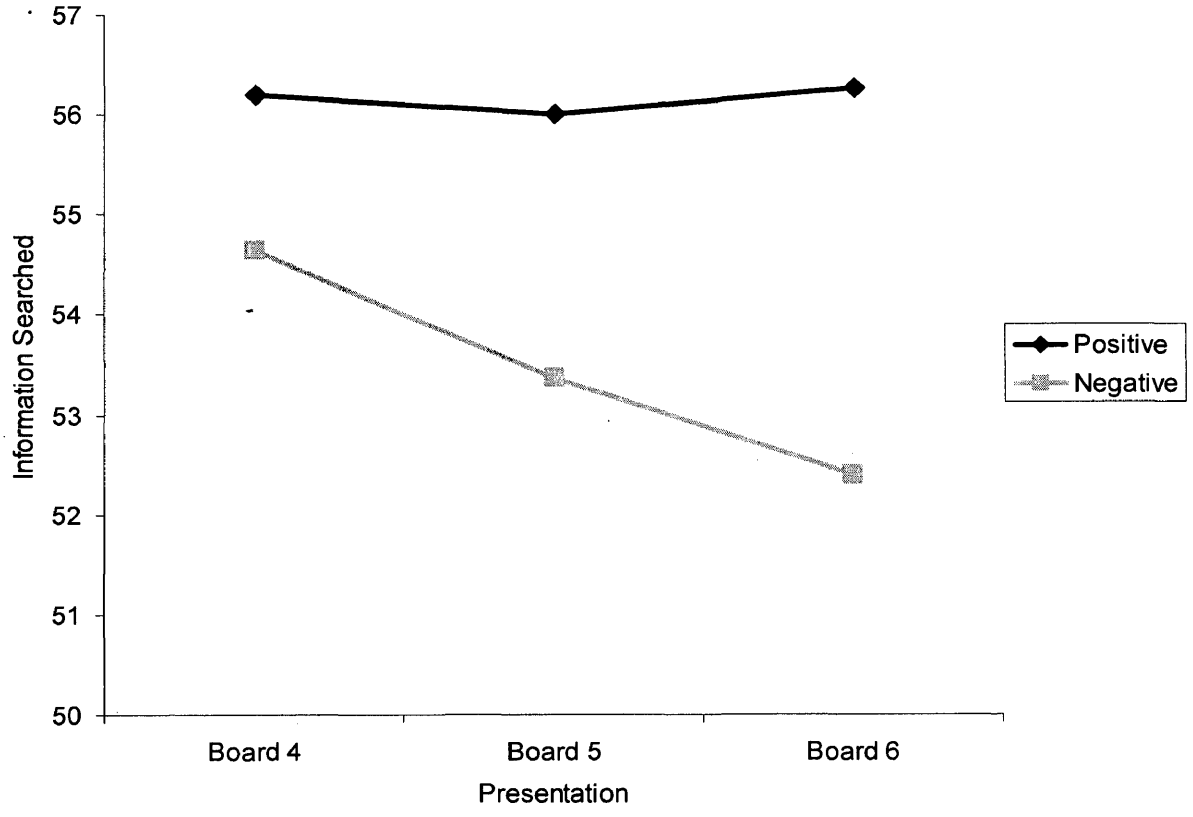
Overall Means and Standard Deviations for the Total Amount of Information Search by Mood on Each Presentation Information Board after Removing Three Influential Cases

	Positive		Negative	
Presentation	*M*	*SD*	*M*	*SD*
Board 4	56.20	8.95	54.64	9.37
Board 5	56.00	9.05	53.36	10.06
Board 6	56.26	8.51	52.39	10.84

Note. N = 74

Figure 4.

Within-Subjects Interaction between Mood and Board on the Amount of Information Searched.



The initial analyses only looked at differences for each type of board (i.e. on the set of planning boards and the set of presentation boards) and not at differences between the types of boards. As such, a secondary analysis using a 2 (mood) by 3 (accountability) by (2) (board type, planning vs. presentation) mixed analysis was conducted. The results of the analysis indicate that there were no significant within- or between-subject effects for mood or accountability on the amount of information searched by participants. That is, participants searched a similar amount of information on the three planning boards as they did on the three presentation boards. However, the extreme case analysis indicated the presence of three possible influential cases. As such, the three cases were removed, and the analysis was conducted again. The results of the new analysis were not significantly different from the previous analysis. That is, the three extreme scores were not influential cases. The overall conclusion is that participants searched a similar amount of information on all information boards regardless of the mood or accountability condition to which they were exposed. Thus, the current study does not support the hypothesis that mood and accountability influenced the processing strategy used by participants. More specifically, the amount of information searched by participants was ineffective at distinguishing between high and low effort information processing strategies.

Hypothesis 2. Hypothesis two predicted a significant interaction between mood and accountability on the average performance rating given to teaching assistants. Specifically, it was predicted that participants in the positive mood condition would give

more positive evaluations of performance compared to participants in the negative mood condition for the no-accountability and process accountability conditions. However, no effect of mood on the average performance rating for participants in the outcome accountability condition was expected. The data were analyzed using a 2 (mood: positive vs. negative) by 3 (accountability: no, process, and outcome) by 3 (planning set: boards 1, 2, and 3) mixed design. A second analysis was conducted using a 2 (mood: positive vs. negative) by 3 (accountability: no, process, and outcome) by (3) (presentation set: boards 4, 5, and 6) mixed design.

Results of the mixed analysis on the average performance ratings of planning set indicate that there were no significant effects either within-, or between-subjects (see Table 15.). Additionally, when the one previously identified extreme case was removed from the sample, the results did not change.

Results of the mixed design examining the average performance rating on presentation set indicate that there was a significant within-subjects quadratic trend for board on the average performance rating of teaching assistants on presentation (partial η^2 = .083, F [1, 68] = 6.16, p = .016). That is, participants tended to evaluate the teaching assistants on presentation board 5 more favorable compared to presentation boards 4 and 6 (see Table 16.).

The between-subjects results indicate that there was a significant interaction between mood and accountability on the average performance rating given to teaching assistants across the three presentation boards (partial η^2 = .085, F [2, 68] = 3.15, p = .049). A probe of the interaction indicates that, for the no accountability and the outcome

accountability conditions, there was no significant difference in the average performance rating of presentation between the positive mood and negative mood conditions (see Table 17). However, for the process accountability condition, there was a significant difference in the average performance rating of presentation between participants in the positive mood and negative mood conditions (partial $\eta^2 = .341$, $F\,[1, 19] = 9.84$, $p = .005$), such that participants in the negative mood condition tended to evaluate the teaching performance more favorably compared to participants in the positive mood condition (see Table 17.).

Prior screening for extreme and influential cases indicated the presence of one possible influential case. As such, the within-subjects analysis on the average performance rating on presentation was rerun after removing the possible influential case. The results of the analysis did not change as a result of removing the one extreme case from the sample. Thus, the overall conclusion is that hypothesis two was not supported. That is, there was no interaction between mood and accountability on the average performance rating of planning, and the significant interaction between mood and accountability on the average performance of presentation was not in the predicted pattern.

The previous analyses examined the within and between differences in average performance ratings for the planning set and the presentation set separately. As such, an additional analysis to examine differences in ratings based on type of board was conducted. That is, the data were analyzed using a 2 (mood) by 3 (accountability) by 2 (planning vs. presentation) mixed design. Average performance rating scores across the

three planning and three presentation boards were calculated using the procedure suggested by Judd and McClelland (2001) for determining mean scores when independence between scores can not be assumed. Results of the analysis indicate that there was a significant within-subject effect for type of board on the average performance rating given to teaching assistants (partial $\eta^2 = .335$, F [1, 68] = 34.26, $p > .001$), such that participants gave more favorable ratings on planning ($M = 4.05$, $SD = .58$) compared to presentation ($M = 3.73$, $SD = .67$). Results of the between-subjects analysis on the overall average rating across all six boards indicates that there was no significant interaction between mood and accountability on performance ratings (partial $\eta^2 = .051$, F [2, 72] = 1.82, *ns*). Thus, again hypothesis two was not supported.

Table 15.

Means and Standard Deviations for Average Performance Rating of Planning on each Information Board

		Positive			Negative		
Planning	Condition	*n*	*M*	*SD*	*n*	*M*	*SD*
Board 1							
	No	16	4.18	.56	17	4.31	.62
	Process	10	3.72	.69	11	4.07	.73
	Outcome	12	3.95	.46	8	4.24	.46
Board 2							
	No	16	4.21	.69	17	4.17	.53
	Process	10	3.64	.68	11	4.03	.84
	Outcome	12	3.82	.64	8	3.99	.53
Board 3							
	No	16	4.09	.64	17	4.25	.78
	Process	10	3.67	.72	11	4.09	.62
	Outcome	12	3.79	.63	8	4.14	.69

Note. Total N = 74

Table 16.

Quadratic Trend in Average Performance Rating of Presentation Across Three Information Boards

Presentation	Board 4		Board 5		Board 6	
	M	*SD*	*M*	*SD*	*M*	*SD*
	3.70	0.68	3.80	0.72	3.68	0.80

Note. N = 74

Table 17.

Means and Standard Deviation of Performance Rating of Presentation Averaged Across Three Information Boards

	Positive			Negative		
Condition	*n*	*M*	*SD*	*n*	*M*	*SD*
No	16	3.94	.70	17	3.89	.44
Process	10	3.22	.63	11	4.05	.58
Outcome	12	3.43	.68	8	3.58	.79

Note. Total N = 74

Hypothesis 3. Hypothesis three predicted a main effect for mood on performance ratings, such that participants in the positive mood condition would give more favorable evaluations of teaching performance compared to participants in the negative mood condition. There were main effects for mood on the average performance rating of presentation (partial $\eta^2 = .06$, $F\,[1, 72] = 4.32$, $p = .041$), and overall average rating (partial $\eta^2 = .062$, $F\,[1, 72] = 4.47$, $p = .038$), but the effects were in the opposite direction than expected (see Table 18.). That is, participants in the negative mood condition tended to give more favorable performance ratings than participants in the positive mood condition. Thus, hypothesis three was not supported.

Table 18.

Means and Standard Deviations of Average Performance Rating of Planning, Presentation, and Overall by Mood Condition

	Positive		Negative	
Ratings	*M*	*SD*	*M*	*SD*
Planning	3.94	.58	4.16	.58
Presentation	3.59	.73	3.87	.58
Overall	3.76	.63	4.02	.51

Note. for Positive mood n = 38; for Negative mood n = 36

Post Experimental Task Affective Reactions

Upon completion of the performance rating task, an exploratory analysis of mood condition on participant affective reactions was conducted in order to get some sense as to how affective reactions changed from the initial manipulation to the end of the experiment. Results indicated no significant difference between participants in the positive and negative mood conditions on negative arousal or elation (see Table 19 and 20). The results indicated as significant difference between individuals in a positive versus a negative mood on boredom and fear, such that, individuals in a positive mood evinced lower levels of boredom and higher levels of fear compared to individuals in the negative mood condition. Participants in the positive mood condition reported marginally higher levels of positive arousal compared to individuals in the negative mood condition. Initially, individuals in the positive and negative mood conditions differed significantly on all dimensions of the affective reactions scale except boredom. By the end of the study, the participants only differed on the dimensions of boredom and fear. Thus, it appears as though the affective reactions of the participants changed throughout the course of the study.

To more fully explore the change in affective reactions, participants' scores were compared to the midpoint of the scale. Consistent with initial affective reactions, participants in the positive mood condition demonstrated significantly low levels of negative arousal. Inconsistent with initial affective reactions, participants in the positive mood condition reported elation, fear, positive arousal, and boredom scores that did not differ significantly from the midpoint of the scale (see Tables 21 and 22). Consistent with

initial affective reactions, participants in the negative mood condition evinced significantly low levels of positive arousal. Inconsistent with initial affective reactions, participants in the negative mood condition reported significantly low levels of negative arousal and fear; significantly high levels of elation; and marginally high levels of boredom (see Tables 21 and 22). Thus, participants in the positive mood condition appear to demonstrate a neutral mood state, whereas participants in the negative mood condition appear to be in a more positive mood.

Table 19.

Means and Standard Deviations of Post Experimental Affective Reactions

	Positive Mood		Negative Mood	
Affective Reaction	*M*	*SD*	*M*	*SD*
Negative Arousal	2.67	.63	2.62	.64
Elation	4.04	1.19	3.92	.65
Fear	3.40	.65	2.88	.43
Boredom	3.09	.76	3.95	.97
Positive Arousal	3.75	1.10	2.97	.89

Note. Positive mood n = 7; Negative mood n = 19

Table 20.

Post Experimental Comparison of Affective Reactions.

	SS	*df*	*MSE*	*F*	*p*
Negative Arousal					
Between	.01	1	.01	.027	0.870
Error	9.82	24	.41		
Total	9.83	25			
Elation					
Between	.07	1	.07	.10	0.755
Error	16.15	24	.67		
Total	16.22	25			
Fear					
Between	1.36	1	1.36	5.49	0.028
Error	5.95	24	.25		
Total	7.31	25			
Boredom					
Between	3.80	1	3.80	4.44	0.046
Error	20.54	24	.86		
Total	24.33	25			
Positive Arousal					
Between	3.08	1	3.08	3.44	0.076
Error	21.49	24	.90		
Total	24.57	25			

Table 21.

Post Experimental Affective Reactions of Participants in the Positive Mood Condition.

Dimension	*M*	*SD*	*difference*	*df*	*t*
Negative					
Arousal	2.67	.63	-.83	6	-3.48*
Elation	4.04	1.19	.54	6	1.19
Fear	3.40	.65	-.10	6	-.41
Boredom	3.09	.76	-.41	6	-1.43
Positive					
Arousal	3.75	1.10	.25	6	.62

Note. n = 7

Table 22.

Post Experimental Affective Reactions of Participants in the Negative Mood Condition.

Dimension	*M*	*SD*	*difference*	*df*	*t*
Negative					
Arousal	2.62	.64	-.87	18	-5.95*
Elation	3.92	.65	.42	18	2.82*
Fear	2.88	.43	-.62	18	-6.20*
Boredom	3.95	.97	.45	18	2.01
Positive					
Arousal	2.97	.89	-.53	18	-2.58*

Note. n = 19

Discussion

General Overview

The goal of this investigation was to test and extend Forgas' (1995) Affect Infusion Model (AIM). First, the model was tested by examining the interaction between evaluator mood (positive vs. negative) and type of accountability (no, process, and outcome) on the amount of information used when making an evaluative judgment. Finally, the model was tested by examining the influence of mood and accountability on performance rating outcomes of teaching assistants.

The following discussion will begin with a brief review of the AIM model and a rationale for the hypotheses of the study. Following the rationale of the hypotheses will be a discussion of the results of each hypothesis. The discussion will conclude with limitations of the current research and directions for future research.

Review of the Affect Infusion Model and Rationale for Hypotheses

Forgas (1995) developed the AIM to help explain when an individual would engage in high effortful (motivated, or substantive) vs. a less effortful (direct access, or heuristic) search of information, and when either of those processing strategies would be influenced by the individual's current affective state, high infusion (heuristic, or substantive) vs. low infusion (direct access, or motivated) processes. Specifically, Forgas states that an individual will use direct access processing (low infusion/ low effort) when the person has a stored response regarding a target and when the judgment made is not personally relevant. For this investigation, none of the participants were expected to use a direct access processing strategy as all of the targets were unknown, and the task was

viewed as moderately important. Forgas (1995) claims that individuals should use a motivated processing strategy (low infusion/ high effort) when a task is important, and when there are specific goals to guide their processing. For this investigation, individuals in the outcome accountability condition were expected to utilize a motivated processing strategy as the requirement to justify one's ratings was believed to elicit a guide to processing.

According to Forgas (1995), individuals in a positive mood usually engage in heuristic processing, whereas individuals in a negative mood usually engage in substantive processing. However, when individuals do not have the cognitive capacity to evaluate information, individuals will use a heuristic processing strategy, regardless of mood. Additionally, when situational demands explicitly call for more effortful processing of information, individuals will utilize a substantive processing strategy regardless of mood (Forgas, 1995). Conversely, when the situational demands explicitly suggest that more effortful processing is not necessary, then individuals, regardless of mood, will utilize a heuristic processing strategy (Forgas, 1995). Thus, Forgas' makes specific predictions of when individuals will use specific processing strategies. For this investigation, individuals in the no accountability condition were expected to use a heuristic processing strategy as no goal existed to guide processing and no situational demands existed that require more elaborative processing. Finally, individuals in the process accountability were expected to utilize a substantive processing strategy as no goal existed to guide processing and situational demands existed that require more effortful processing of information.

The Effect of Accountability on the Amount of Information Searched

Hypothesis one predicted that participants in the no-accountability condition would search less information compared to participants in the process and outcome accountability conditions. However, the results indicated that participants searched similar amounts of information regardless of the accountability condition to which they were exposed. Thus, the current results do not support hypothesis one and appear to be inconsistent with AIM. That is, based on the amount of information searched, it appears that all individuals utilized a high-effort, motivated or substantive, processing strategy whether they were held accountable or not (see Figure 5 & 6). Several factors could explain the results of this investigation, such as the importance of the decision task and the response mode of the decision task.

Task importance. According to the AIM, the personal relevance or importance of decision is a necessary but not sufficient reason to elicit a high-effort decision strategy (Forgas, 1995). However, Beach and Mitchell (1978) claim that when a decision is important to the individual, the individual will use a complex decision strategy and examine more information before making a decision. For this investigation, individuals were informed that any teaching assistant who received a below average rating would lose his or her assistantship, which could result in the elimination of a graduate program and ultimately decrease the number of undergraduate courses offered by the department. The purpose of providing all individuals with the same rationale was to control for task importance and prevent the use of direct access processing, but doing so could have prevented the use of the heuristic processing strategy. That is, although participants were

told that the overall goal of the investigation was to develop a new performance appraisal system, which should only be viewed as personally relevant by participants who were held accountable for either their ratings or the information used, the task may have been viewed as very important by all participants and thus elicited high information search strategies.

Figure 5.

Predicted Processing Strategy Based on the Amount of Information Searched and Outcome of Performance Ratings.

	Outcome Accountability	Process Accountability	No Accountability
Positive Mood	Motivated Processing High Info. Search Neutral Rating	Substantive Processing High Info. Search High Rating	Heuristic Processing Low Info. search High Rating
Negative Mood	Motivated Processing High Info. Search Neutral Rating	Substantive Processing High Info. Search Low Rating	Heuristic Processing Low Info. search Low Rating

Figure 6.

Actual Processing Strategy Based on the Amount of Information Searched and Outcome of Performance Ratings.

	Outcome Accountability	Process Accountability	No Accountability
Positive Mood	Motivated Processing High Info. Search Low Rating	Motivated Processing High Info. Search Low Rating	Motivated Processing High Info. Search High Rating
Negative Mood	Motivated Processing High Info. Search High Rating	Motivated Processing High Info. Search High Rating	Motivated Processing High Info. Search High Rating

According to Beach and Mitchell (1978), another way to make a decision important to an individual is to hold the individual accountable for his or her decision. One could make a person accountable for the outcome of his or her decision (Chaiken, 1980), or hold the individual accountable for the process he or she used to make the decision (Hagafors & Brehmer, 1983). In the current study, participants were told one of three things; (a) that they were accountable for the performance ratings provided, (b) they were accountable for the information they accessed, or (c) they were not accountable for the performance ratings they provided or the information they accessed. As such, the non-accountable raters should have searched through less information before making a performance rating compared to participants held accountable for either the performance ratings made or the information used; but, as previously stated, all individuals searched a similar amount of information despite the possible difference in task importance. However, most of the decision making research equates decision choice with decision judgment even though there are distinct differences between the two response modes (Billings & Scherer, 1988, 1991; Einhorn & Hogarth, 1981; Einhorn, Kleinmuntz, & Kleinmuntz, 1979, Payne, 1976; 1982).

Response mode: Judgment vs. choice. Although judgments and choices may occur as part of the same decision process, the two constructs not the same. According to Einhorn et al. (1979), evaluative judgments often precede and facilitate choices, but evaluative judgments can occur with out making choices, and choices can be made without first making evaluative judgments. Einhorn et al. also stated that the two decision motivations, judgment vs. choice, can lead to different decision strategies. Einhorn et al.

suggested that individuals who make evaluative judgments or choices preceded by evaluative judgments utilize a compensatory decision strategy, whereas individuals who only make choices use a non-compensatory decision strategy. Payne (1976) indicated that in compensatory search strategies, described as a linear or additive search patterns, individuals searched across dimensions of information for each alternative before making an evaluative judgment; individuals then choose the highest rated alternative. Non-compensatory search strategies, described as a satisficing search patterns, occurred when individuals searched within dimensions of information and retained alternatives that met some minimum requirement (Payne, 1976). The satisficing procedure was repeated until one alternative remained. Payne (1982) conducted a review of the decision-making literature on judgment and choice and suggested that when individuals make judgments, or judgments followed by choice, they tend to search more information than when decisions only require choosing between alternatives. However, Payne (1982) also indicated several factors that could influence the decision strategy used, such as, the cost or benefit of accessing information, time pressure, task complexity, and task importance. For example, Billings and Scherer (1988) found that the importance of the decision task influenced the amount of information searched for decision choices, but not for decision judgments.

Billings and Scherer (1988) examined differences between decisions of choice vs. judgment, and between decisions of high vs. low importance. Billings and Scherer found that when individuals were required to make evaluative judgments about resident-hall advisors, they searched more information compared to individuals who were required to

choose the best resident hall advisor. Additionally, individuals required to make evaluative judgments search similar amounts of information regardless of the importance of the task, whereas individuals required to make a choice searched more information when the task was important compared to when the task was unimportant (Billings & Scherer, 1988).

For this investigation, all individuals were required to make an evaluative judgment. As such, regardless of the accountability condition to which an individual was exposed, individuals might access the majority of the available information before making their judgments of the teachings assistants' teaching performance. Thus, because individuals were asked to make an evaluative judgment rather than a choice, no difference in amount of information searched, or any apparent difference in the type of information processing strategy utilized, was found. However, this rationale is inconsistent with the AIM, and points out a potential omission of the AIM in that it fails to account for differences between judgment and choice.

Summary of Hypothesis 1

Thus, the AIM model (Forgas, 1995) appears to be inconsistent with at least some the decision-making literature. Forgas suggested that individuals should engage in high effort processing strategies when specific goals to guide processing existed. Forgas also suggested that if no specific goals to guide processing existed, individuals would engage in high effort processing strategies when; (a) they have the cognitive ability to process additional information, (b) when they are in a negative mood, and (c) when situational demands exist that call for more elaborative processing. However, the decision making

literature suggested that individuals would use complex decision strategies when the task is important (e.g. Beach & Mitchel, 1978; Billings & Scherer, 1988) or when individuals were required to make evaluative judgments versus choices (Billings & Marcus, 1983; Billings & Scherer, 1988; 1991). In an effort to understand the relationship between mood and accountability on the amount of information searched, exploratory analyses were conducted.

Exploratory Analyses of the Amount of Information Searched

Although participants tended to examine a majority of the information before making an evaluative judgment, an exploratory analysis indicated a significant within-subjects linear trend in the amount of information searched across the three planning and the three presentation information boards. The results indicated that participants tended to examine the most information on the first board of the series and less information on subsequent boards. The results of these analyses are consistent with the results of Billings and Scherer (1988) who found that individuals searched through the most information on the first information board and less information on subsequent boards. One possible explanation for the results is that on the first information board of the series (planning or presentation) individuals were unfamiliar with the task and the information they would be reading. However, on subsequent boards related to the same aspect of teaching, individuals were selective in the information they accessed and might have focused on information they felt was more relevant to the evaluation of teaching performance.

Interestingly, when the three individuals with extreme scores on the amount of information searched on the set of presentation information boards were removed from

the sample, a significant within-subject interaction between mood and board on the amount of information searched was found. Examination of the three influential cases indicated that all three participants were in a positive mood condition. Additionally, one individual was from the no-accountability condition, one was from the outcome accountability condition, and one was from the process accountability condition. After removing the three participants from the sample, the results indicated that individuals in the positive mood condition tended to search a similar amount of information across the three presentation boards, whereas individuals in the negative mood condition tended to search less information on subsequent boards. A probe of the interaction indicated that individuals in a positive mood searched significantly more information on the last presentation board compared to individuals in a negative mood. However, the results of this investigation appear to contradict the findings of Martin et al. (1993).

Rationale for Stopping Information Search

Martin et al. (1993) suggested that an individual's motivation to continue accessing information was influenced by the individual's reason for engaging in the search task and the individual's current mood state. Martin et al. (1993) found that when individuals were told to stop searching for information when the task became uninteresting, individuals in a negative mood stopped sooner and read less information, than individuals in a positive mood. However, when individuals were told to stop searching for information when they felt they had enough information to make a judgment about a target, individuals in a positive mood stopped sooner and read less information than individuals in a negative mood.

For this investigation, individuals were instructed to access as much information as they felt was necessary to make an evaluation of the teaching assistants performance; as such, one would have expected individuals in the positive mood to access less information on the third presentation board compared to individuals in the negative mood condition. However, the current study differed from the Martin et al (1993) study in several ways. In the Martin et al study, individuals made an evaluative judgment on a single target, whereas this investigation required individuals to make evaluative judgments on thirty teaching assistants. Additionally, Martin et al. provided individuals with a stack of cards describing various aspects of the target and instructed them to read the cards one at a time. After reading a card, individuals were told to turn the card face down and not reread it. For the current investigation, individuals were instructed to access information one box at a time, but the information remained visible. Thus, individuals could reread information before making a judgment. Thus, individuals in a negative mood should have stopped accessing information sooner than individuals in a positive mood based on the decision rule to read until you feel you have enough information. Additionally, because there were only six pieces of information per teaching assistant, individuals may have ignored the decision rule and stopped accessing information when they were no longer interested in the task. However, these conclusions are only speculative as the results were found after removing the three influential cases from the data set.

Despite the interesting within-subject findings, there were no between-groups effects on the amount of information searched on the planning or on the presentation

information boards. Thus, one could conclude that individuals, regardless of mood or accountability condition, utilized a high-effort information search strategy when making evaluative judgments about the teaching performance of teaching assistants. The question remains as to which high-effort information search strategy individuals used.

Overall Summary of Hypothesis 1

According to Forgas (1995), processing strategies are divided into low vs. high effort decision processes and low vs. high infusion decision processes. Thus, based on the results, participants appeared to have used a high effort processing strategy. However, examining the amount of information searched is not sufficient to determine which high effort processing strategy an individual utilized; rather the combination of the amount of information searched and the presence of mood congruent or incongruent judgments, meaning no effect of mood on processing outcomes, could indicate the processing strategy utilized.

The Interaction between Mood and Accountability on Performance Ratings

Hypothesis two predicted an interaction between mood and accountability on the performance ratings made regarding the teaching performance of teaching assistants. Specifically, mood congruent judgments were expected for participants in the no accountability and the process accountability conditions, such that, individuals in a positive mood were expected to provide more favorable performance ratings compared to individuals in a negative mood. Regardless of mood, no difference in performance ratings was expected for individuals in the outcome accountability condition. The results indicated that for participants in the no accountability and outcome accountability

conditions mood did not influence the performance ratings provided, whereas for individuals in the process accountability condition mood did influence the performance ratings provided. However, the influence of mood on performance ratings for individuals held accountable for the information they accessed was in the opposite direction than predicted. That is, individuals in a negative mood provided more favorable performance ratings than individuals in a positive mood. In fact, regardless of accountability, individuals in a negative mood provided more favorable performance ratings than individuals in a positive mood, which is contrary to hypothesis three. The results of this investigation appear to contradict the majority of the literature examining affective influences on decision making, which tends to find mood congruent decision outcomes (e.g. Martin et al., 1993; Schwarz & Clore, 1983; Sinclair, 1988). However, the current investigation differed from previous research on one very important aspect.

Previous research exploring the impact of mood on perceptions of others (Martin et al, 1993; Schwarz & Clore, 1983) and on performance appraisal evaluations (Sinclair, 1988) examined judgments based on memory recall. That is, individuals read information about a target and then made an evaluative judgment based what information they remembered. For the current investigation, individuals accessed information and made immediate evaluations. Additionally, all information accessed remained visible until individuals finished rating all of the teaching assistants on the board. That is, for this investigation, individuals made what Hastie and Park (1986) referred to as on-line judgments, or judgments that do not rely on memory retrieval. Thus, when individuals make evaluative judgments that do not rely on memory retrieval, their current mood state

is not used as information about how they feel about the target, nor does mood influence their interpretation of, or bias their impression of the information they are accessing. However, this rationale does not fully explain the main effect for mood in which individuals in a negative mood provided more favorable evaluations compared to individuals in a positive mood.

Main Effect for Mood on Performance Ratings

The results indicated that individuals positive moods provided less favorable performance ratings compared to individuals negative moods. General theories exploring the impact of mood on decision making suggested that individuals in a negative mood engage in more effortful or piecemeal examination of information, whereas individuals in a positive mood tended to use less effortful processing strategies (Forgas, 1995; Martin, 2001; Schwarz, 2001). Conversely, Isen (2001) suggested that positive mood enhanced cognitive processing by enabling individuals to integrate information into fewer related categories of information. For the current investigation, all participants searched a similar amount of information regardless of mood and accountability, but there is an effect of mood on the performance ratings provided by individuals. Perhaps individuals in a negative mood examined the information in a systematic and piecemeal method and calculated an average performance rating based on the perceived effectiveness of each statement. The results showed that individuals in a negative mood provided performance ratings that were less discrepant than performance ratings provided by individuals in a positive mood. These results are similar to the Sinclair (1988) results showing that individuals in a negative mood were more accurate in their performance appraisal ratings

compared to individuals in a positive mood. Unfortunately, Sinclair found that individuals in a positive mood were more lenient in their ratings, inflated performance ratings, whereas in the current investigation individuals in a positive mood provided harsher ratings. Perhaps individuals in a positive mood formed an overall impression of the teaching assistant based on the information provided, and the overall impression was influenced by the presence of less effective teaching behaviors, resulting in a harsher rating than would be expected based on the mathematical calculation of teaching performance. Hastie and Park (1986) suggested that behaviors which are incongruent with overall impressions of individuals are the easiest to recall in memory tasks. Although the current study was not an evaluation task based on memory recall, less effective teaching behaviors may have stood out and influenced an individual's overall impression of the teaching assistant. According to Forgas (1995), the presence of mood incongruent judgments would indicate that individuals utilized a motivated processing strategy.

Forgas (1995) stated that individuals would utilize a motivated processing strategy when a task was important and a specific goal to guide processing existed. The current investigation seems to run counter to the predictions made by Forgas' AIM. As predicted, for individuals in the outcome accountability condition, regardless of mood, no difference in performance ratings were found. Apparently, having to justify one's ratings is sufficient to elicit a goal directed information search. However, one would not expect individuals held accountable for the information they used, or not held accountable at all, to utilize a motivated processing strategy. When not held accountable for decisions,

research suggested that individuals would make judgments or take stances that were congruent with their attitudes and perceptions (e.g. Cialdini et al., 1976; Tetlock, 1993). Additionally, being held accountable for the information accessed, despite eliciting a more thorough search, should not have influenced the appraisal of the information as individuals were explicitly told that they would not be asked about the ratings they provided. However, mood incongruent judgments were made by all participants regardless of the mood or accountability condition to which they were exposed. Thus, other factors could influence an individual's decision to use a motivated processing strategy beyond the presence of a specific goal to guide processing. Perhaps the requirement of making an immediate evaluation rather than an evaluation that relied on information stored in memory was sufficient to elicit a motivated processing strategy. A more plausible explanation is that on-line judgments are less susceptible to mood effects because information coding was not based on valence. The results of this investigation could have some interesting implications.

Implications

Forgas developed the Affect Infusion Model (AIM) to reconcile conflicting findings within the mood and emotion literature. However the results of this investigation do not coincide with what was expected based on the AIM. One reason for the apparent difference was the use of on-line vs. memory based judgment. When individuals make on-line judgment, individuals in a positive mood appear to be negatively influenced by the presence of negative information, more so compared to individuals in a negative

mood. Thus, Forgas Affect Infusion Model may need to be revised to account for the influence of on-line judgments.

The results of this investigation could also have some interesting applied implications. For example, the process of conducting a performance appraisal is infused with human error (Landy & Farr, 1980; Murphy & Cleveland, 1991). For instance, Sinclair (1988) found that individuals tended to recall information that was congruent with his or her current mood state, which in turn influenced the final evaluation. One way to correct for the biased recall of information is to ensure the decision maker has the necessary performance information available before evaluating a subordinated performance. However, one must use caution when recording the information regarding performance as negative behavior tend to have a negative effect on performance for individuals in positive moods.

Finally, Semmler and Brewer (2002), found that jurors who where in a sad mood tended to process trial information more thoroughly and indicated more inconsistencies in information compared to jurors in a neutral mood. However, in the Semmler and Brewer (2002) study, individuals relied on memory retrieval when evaluating trial information consistency. Allowing jurors to take notes, or making the trial notes available during the decision making process could eliminate the disparity between mood conditions on the evaluation of the consistency of information presented during trial. One must use caution to ensure the information presented to the jurors is affect neutral, or risk the chance that individuals in positive moods will evaluate negative information more critically. Despite

the interesting results and implication of this investigation, the interpretation of the results and implications must be tempered by the limitations of the study.

Limitations

Affect Infusion Model. According to Forgas (1995) several characteristics of the decision maker, the task, and the target influence an individual's choice of information processing strategies. The current investigation was designed with consideration of Forgas' assumptions, but the assumptions were not explicitly tested (see Figure 1).

According to the AIM, individuals use the direct access processing strategy when the target is familiar and the decision is not personally relevant or important. In the current investigation, individuals evaluated the teaching performance of teaching assistants, a task that the majority of the participants have performed at one point in their college career. Because the participants have evaluated the teaching performance of teaching assistants, one could argue that the task was familiar. One could also argue that individuals might also have preconceived ideas about teaching assistants. However, when the decision task is important, individuals will not use the direct access processing strategy when evaluating targets (Forgas, 1995). That is, task familiarity and having stereotypes about a target are not sufficient to elicit the direct access processing strategy, and individuals could still engage in high effort and high infusion decision processes. In the current investigation, task importance was not explicitly measured, but one of the goals of pilot study II was to determine the perceived importance of various performance appraisal outcomes. The performance appraisal outcome used in the current investigation was rated as very important by undergraduates in pilot study II. Thus, one could reason

that the performance rating task was viewed as important by the participants. As such, none of the participants should have used the direct access processing strategy when making performance ratings. However, although the task was designed to be important enough to prevent the use of direct access processing, the task might have been viewed as very important which would limit the use of heuristic processing. According to Beach and Mitchel (1978), when a decision task is viewed as important, individuals tend to use more complex and thorough decision strategies.

According to Forgas (1995), when there is a specific goal to guide processing individuals will utilize a motivated processing strategy, but in the absence of a specific goal individuals will engage one of the high infusion decision strategies. For the current investigation, individuals in the outcome accountability condition were expected to engage in motivated processing. That is, being called on to justify the performance ratings provided should have elicited a goal directed behavior of generating defensible performance ratings. Additionally, because individuals in the no accountability and the process accountability conditions were not required to justify their ratings, they should not have felt pressure to engage in goal directed behavior. However, individuals were not explicitly asked if they felt they had a specific goal, or their rationale for the performance ratings they provided. Asking individuals for their rationale for the performance ratings provided could have added valuable insight. Additionally, the importance of the task might have elicited a specific goal to guide processing, such as to evaluate the information in an objective manner so as to provide valuable feedback, which may account for the results that all individuals utilized a high effort processing strategy.

According to Forgas (1995), when no specific goal to guide processing is present, the complexity of the task and the cognitive capacity of the individuals are the next factors that influence the processing strategy chosen. For this investigation individuals were not asked if they felt the task was complex or unusual. However, although information boards are common in research, the use of information boards is not common in everyday life. Additionally, the information boards of the current investigation were set up differently from other information boards. Instead of having each board represent new targets and new dimensions, the information boards for this study were established as sets, three boards related to presentation of lecture and three boards related to preparation for class. Thus, the rating task was not common, and, individuals were expected to find the task atypical. When the task is atypical the individuals' cognitive capacity is instrumental in determining the processing strategy chosen.

The final assumption of the AIM which was not tested was the cognitive capacity of the individuals. The original proposal called for four information boards with six teaching assistants and five behaviors on each board. The original proposal was rejected because the information task was deemed to have insufficient cognitive demand (i.e. too easy to access all of the information). Due to limited space constraints of using a computer screen, cognitive load was increased by adding four teaching assistants and one additional behavior to each board, and increasing the number of boards from four to six. However, the ease of accessing information on a computer screen, as opposed to physically turning over a card on a conventional information board, could have limited the expected increase in cognitive load experienced by the participants. Finally, the lack

of time constraint and the process of making immediate evaluations, rather than making evaluations based on memory recall of information, could have eliminated any cognitive demand experienced by the individuals. The lack of cognitive load makes it impossible to discern, based on the amount of information accessed, whether individuals used a substantive, heuristic, or motivated processing strategy when evaluating the performance of teaching assistants.

Methodological limitations. Although the task was unfamiliar to the participants, the task might have been easier than expected, which might have caused a ceiling effect on the amount of information searched (i.e. participants searched through most of the information) preventing one's ability to distinguish between type of processing strategies utilized by participants. Increasing the cost associated with the information search could have increased the complexity or cognitive load of the decision making task. For example, Payne, Bettman, and Johnson (1988) found that creating time pressure and offering monetary rewards for accurate decisions influenced the decision strategies individuals utilized. Similarly, Gilliland, Schmitt, and Wood (1993) found that forcing individuals to pay for the information they accessed impacted the information search strategy utilized.

Another limitation of this investigation was the successful manipulation of accountability. For this investigation, 19 of the 61 participants who were held accountable did not feel accountable, or did not believe the Dean of the college would be speaking with them at the end of the study. One way to improve the accountability manipulation could have been to suggest that a representative of the Deans office would

show up at the end of the investigation to discuss the performance ratings provided or the information used when making performance ratings. However, although 19 individuals did not feel accountable, 41 individuals believed that the Dean of the college would show up at the end of the investigation to discuss the performance ratings they provided. As such, the manipulation of accountability was successful.

Although the accountability manipulation was successful, individuals who were not held accountable may have still felt accountable, or were highly engaged in the task, which resulted in high information search. Weiss (1996), while not manipulating accountability, suggested that participants were more engage in a selection decision task than expected because individuals searched through more information than was necessary to reach and informed decision. Herman (2004) also found evidence that participants were more engaged, or exhibited more effort than expected, during a solution generation task. Herman (2004) instructed participants to generate multiple solutions to an ill-structured problem and found that although participants generated several solutions to the problem, the overall quality of solutions was low. Herman suggested that participants tended to focus on the instruction to generate multiple solutions at the expense of generating high quality solutions. Thus, is appears as though participants from this particular university tend to be highly engaged in experimental tasks and put forth more effort than required, perhaps in an attempt to please the experimenter. For this investigation, individuals might have searched more information, even though they did not feel accountable, because they were engaged in the task, or in an effort to please the

experimenter, which could explain the no effect of accountability on information search, or on the amount of time spent on the task.

Another methodological limitation of this investigation was the successful manipulation of mood vs. manipulation of an emotion. Forgas (1994) suggests that an emotion is directed a specific target which is attributed as eliciting the emotion, whereas for a mood, the feeling is not directed at, or attributed to a specific target. In the current investigation, participants watched a video designed to elicit either a positive, or negative mood. Prior research using videos as elicitation aids have claimed that videos successfully alter a participants mood (e.g. Isen & Daubman, 1984; Isen, Daubman, & Nowiki, 1987; Weiss, 1996). In fact, the videos used in this investigation were the same videos used by Weiss (1996), and although the mood effect was subtle, particularly for the negative mood condition, the effect was in the expected direction and significant. One explanation for the subtle negative mood effect could by caused by recent events. The scene in Midnight Express depicted a prisoner being beaten by prison guards. Although the scene is quite disturbing, the presence of, and easy access to violent images, such as pictures of the Abu Ghairb incident, could have limited the effectiveness of the video. Other videos or other techniques might have elicited a stronger effect.

Although videos alter one's current mood state, researchers have questioned the duration of the mood effect. Isen and Gorgoglione (1983) found that positive moods tended to endure thorough an experiment, whereas negative moods tended to dissipate. For this investigation, at the conclusion of the performance evaluation task, participants in the negative mood condition reported low levels of negative arousal and fear, and high

levels of elation and boredom, which are opposite reactions of what were reported after watching the negative mood inducing video. Additionally, participants in the negative mood condition tended to report low levels of positive arousal before and after the performance evaluation task. Thus, the duration of the negative mood is questionable, as is the influence of negative mood on the evaluation task. The results of this investigation could show differences between positive mood and neutral mood rather than differences between positive and negative moods.

Another limitation of this investigation was that performance rating information was condensed into a single number of overall average performance rating across all teaching assistants and across all boards. For this investigation, individuals made 30 performance ratings on planning and 30 ratings on presentation; by collapsing the ratings across boards and dimensions, some information was lost. For example, the exploratory analyses on average performance rating indicated that individuals tended to rate the performance of teaching assistants harsher on presentation than on planning. However, looking at overall scores is appropriate to do before proceeding to more complex research models. Despite the limitations and null findings of this investigation, it demonstrates a need for more research in the area of mood and decision making, especially as it relates to performance appraisal.

Future Research

According to Forgas (1995), the affect infusion model (AIM) serves as guide to understanding what factors influence how social judgments are made. Forgas also suggested that there are four types of decision processes characterized as either high vs.

low effort, and as high vs. low infusion, and that individuals make judgments using one of the four processing strategies. However, Forgas, like other researchers, might be equating judgment with choice, whereas other researchers indicate that there is a clear distinction between the two constructs (e.g. Billings & Scherer, 1988; Einhorn et al, 1979; Payne, 1982). Thus, the AIM may not account for situations that only require evaluative judgments.

Additionally, Forgas is somewhat vague on what constitutes a specific goal to guide processing. Forgas (1995) suggested that mood repair and need to affiliate were goals that could elicit motivated processing, whereas the instructions to make accurate ratings were not sufficient. However, in the current investigation, individuals who were not held accountable for the performance ratings they provided and were required to make evaluative judgments, appeared to use a motivated processing strategy. That is, the need to make evaluative judgments without making a choice could elicit goal directed processing of information. Thus, future research should examine the influence of mood, accountability, and response mode on the amount of information searched and on the valance of the decision outcomes.

The results of this investigation indicated that individuals in positive moods tended to rate the teaching performance of the teaching assistants less favorably than individuals in negative moods. These results conflict with results of other research that examined the influence of moods on judgments (e.g. Martin et al., 1983; Sinclair, 1988). However, the current investigation examined judgments made immediately after accessing information, whereas most research examined judgments based on recall of

information. Researchers have suggested the mood influences how information is encoded and the type of information recalled (e.g. Martin et al., 1983; Schwarz, 2001). Thus, individuals who are required to make immediate judgments may not have encoded the information according to valence and, thus, were not influenced by their current mood state, whereas individuals who were required to make judgments based on information stored in memory might have encoded and retrieved information that was congruent with their current affective state.

Additionally, some researchers suggested that individuals in a negative mood examined information in a more thorough manner compared to individuals in a positive mood who tended to search information more superficially (e.g. Forgas, 1995, Martin et al., 1983; Schwarz, 2001). However, other researchers suggested that individuals in a positive mood integrated information into fewer categories and evaluated information more critically compared to individuals in a negative mood (e.g. Isen, 2001; Isen and Means, 1983). Perhaps when individuals make evaluative judgments about another person, with all available information immediately at hand, individuals in positive moods think more critically about information and form an overall impression based on the perceived effectiveness of teaching behaviors, whereas individuals in a negative mood may focus on individual behaviors and form an average impression based on the perceived effectiveness of each statement. Thus, future research should explore the impact of mood on immediate judgments and memory based judgments.

In addition to exploring mood incongruent effects, future research should explore factors that influence performance rating error, or the presence of HALO error. Sinclair

(1988) found that individuals in a positive mood provided lenient performance ratings, whereas individuals in a negative mood provided relatively accurate ratings. In the current investigation, individuals in positive and negative moods provided performance ratings that were harsher than expected, but individuals in a positive mood provided the harshest ratings. Previous research on performance appraisal evaluations suggested that raters evaluate the performance of others harsher when the purpose of the appraisal is for developmental reasons (e.g. Jawahar & Williams, 1997; McIntyre, Smith, & Hassett, 1984; Murphy, Balzer, Kellam, & Armstrong, 1984). Perhaps the purpose of the appraisal eliminated the influence of mood on performance ratings. As such, future research should examine the impact of mood, accountability, and purpose of appraisal on performance ratings.

References

Adelberg, S., & Batson, C. D. (1978). Accountability and helping: When needs exceed resources. *Journal of Personality and Social Psychology, 36,* 343 – 350.

Aleamoni, L. M., & Hexner, P. Z. (1980). A review of the research on student evaluations and a report on the effect of different sets of instructions on student course and instructor evaluation. *Instructional Science, 9,* 67-84.

Ashkanasy, N M., Zerbe, W. J., & Hartel, C. E. J. (2002). *Managing emotions in the workplace.* New York: M. E. Sharp.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review, 3,* 439 – 449.

Billings, R. S. & Marcus, S. A. (1983). Measures of compensatory and noncompensatory models of decision behavior: Process tracing vs. policy capturing. *Organizational Behavior and Human Performance, 31,* 331 – 352.

Billings, R. S., & Scherer, L. L. (1988). The effects of response mode and importance on decision-making strategies: Judgment vs. choice. *Organizational Behavior and Human Decision Processes, 41,* 1 – 19.

Billings, R. S., & Scherer, L. L. (1991). Response mode effects in decision making: Process, outcomes and postdecision regret. *Advances in Information Processing in Organizations, 4,* 1 – 17.

Brtek, M. D., & Motowidlo, S. J. (2002). Effects of procedure and outcome accountability on interview validity. *Journal of Applied Psychology, 87,* 185-191.

Cardy, R. L., & Dobbins, G. H. (1994). *Performance Appraisal: Alternative perspectives.* Cincinnati, OH: South-Western Publishing Co.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39,* 752 – 766.

Cialdini, R. B., Levy, A., Herman, C. P., Kozlowski, L. T., & Petty, R. E. (1976). Elastic shifts of opinion: Determinants of direction and durability. *Journal of Personality and Social Psychology, 34,* 663 – 672.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. New York, NY: Academic Press.

Costello, C. G., & Comrey, A. L. (1967). Scales for measuring depression and anxiety. *The Journal of psychology, 66,* 303 – 313.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32,* 53 – 88.

Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process tracing models of judgment. *Psychological Review, 86,* 465 – 485.

Elsbach, K. D., & Barr, P. S. (1999). The effects of mood on individual's use of structured decision protocols. *Organization Science, 10,* 181 – 198.

Erber, R. & Erber, M. W. (1994). Beyond mood and social judgment: Mood incongruent recall and mood regulation. *European Journal of Social Psychology, 24,* 79 – 88.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117,* 39-66.

Forgas, J. P. (2001). The affect infusion model (AIM): An integrative theory of mood effects on cognition and judgment. In L. L. Martin, & G. L. Clore (Eds.), *Theories of mood and cognition: A users guidebook* (99 – 134). Mahwah, NJ: Lawrence Earlbaum Associates.

Forgas, J. P., & Moylan, S. (1987). At the movies: transient mood and social judgments. *Personality and Social Psychology Bulletin, 13,* 467-477.

Gilliland, S. W., Schmitt, N., & Wood, L. (1993). Cost-benefit determinants of decision process and accuracy. *Organizational Behavior and Human Decision Processes, 56,* 308 – 330.

Gmelch, W. H., & Glasman, N. S. (1977). The effect of purposes on student evaluation of college instructors. *Educational Research Quarterly, 2,* 45-55.

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.

Hattrup, K., & Ford, K. J. (1995). The roles of information characteristics and accountability in moderating stereotype-driven processes during social decision making. *Organizational Behavior and Human Decision Processes, 63,* 73 – 86.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93,* 258 – 268.

Herman, A. (2004). *Emotional intelligence as a moderator of problem based negative arousal on solution quantity and quality*. Unpublished Master's thesis, University of Nebraska, Omaha.

Hobson, C. J., Mendel, R. M., & Gibson, F. W. (1981). Clarifying performance appraisal criteria. *Organizational Behavior and Human Performance, 28,* 164-188.

Hom, P. W., DeNisi, A. S., Kinicki, A. J., & Bannister, B. D. (1982). Effectiveness of performance feedback from behaviorally anchored rating scales. *Journal of Applied Psychology, 67,* 568 – 578.

Isen, A. M. (2001). An influence of positive affect on decision making in complex situation: Theoretical implications with practical applications. *Journal of Consumer Psychology, 11,* 75 – 85.

Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology, 47,* 1206 – 1217.

Isen, A. M., Daubman, K. A, & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology, 52,* 1122 – 1131.

Isen, A. M., & Gorgoglione, J. M. (1983). Some specific effects of four affect-induction procedures. *Personality and Social Psychology Bulletin, 9,* 136 – 143.

Isen, A. M., & Means, B. (1983). The influence of positive affect on decision-making strategy. *Social Cognition, 2,* 18 – 31.

Isen, A. M., Shalker, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop. *Journal of Personality and Social Psychology, 36,* 1 -12.

Jawahar, I. M., & Williams, C. R. (1997). Where all the children are above average: The performance appraisal purpose effect. *Personnel Psychology, 50*, 905-923.

Judd, C. M., & McClelland C. H. (2001). *Data analysis: A model comparison approach.* Belmont, CA: Thomson Learning.

Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87,* 72 – 107.

Lerner, J. S., & Tetlock, P.E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125,* 255 – 275.

Lord, R.G., Kilmoski, R. J., & Kanfer, R. (2002). *Emotions in the workplace: Understanding the structure and role of emotions in organizational behavior.* San Francisco, CA: Josey-Bass.

Marichal, j. & Mathonet, P. (2000). On comparison meaningfulness of aggregation functions. *Journal of Mathematical Psychology, 45*, 213-223.

Martin, L. L., & Clore, G. L. (2001). *Theories of mood and cognitions: A users guide.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishing.

Martin, L. L., Ward, D. W., Achee, J. W., & Wyer, R. S. (1993). Mood as input: people have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology, 64,* 317 – 326.

McAllister, D. W., Mitchell, T. R., &, Beech, L. R. (1979). The contingency model of the selection of decision strategies: An empirical test of the effects of significance, accountability, and reversibility. *Organizational Behavior and Human Performance, 24,* 228 – 244.

McIntyre, R. M., Smith, D. E., & Hassett, C. E. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of rating. *Journal of Applied Psychology, 69*, 147-156.

Mischel, W. Wbbesen, E. B., & Zeiss, A. R. (1973). Selective attention to the self. Situational and dispositional determinants. *Journal of Personality and Social Psychology, 27,* 129 – 142.

Mero, N.P., & Motowidlo, S. J. (1995). Effects of rater accountability on the accuracy and the favorability of performance ratings. *Journal of Applied Psychology, 80*, 517-524.

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

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: social, organizational, and goal-based perspectives.* Thousand Oaks, CA: Sage Publications.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational and Human Performance, 16,* 366 – 387.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin, 92,* 382– 402.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 366 – 387.

Payne, R. L., & Cooper, C. L. (2001). *Emotions at work: theory, research, and applications for management*. New York: John Wiley & Sons, LTD.

Robbins, T. L., & DeNisi, A. S. (1998). Mood vs. interpersonal affect: Identifying Process and rating distortions in performance appraisal. *Journal of Business and Psychology, 12*, 313-325.

Rosenthal, R. (1976). *Experimenter effects in behavioral research* (enlarged ed.). New York, NY: Irvington Press.

Russel, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145 – 172.

Scherer, L. L., Reiter-Palmon, R. Butler, A. & Weiss, R. J. (1994). *Toward a taxonomy of reactions to ill-defined problems*. Paper presented at the annual meeting of the Society for Judgment and Decision Making, St. Louis, MO.

Schwarz, N. (2001). Feelings as information: Implications for affective influences on information processing. In L. L. Martin, and G. L. Clore (Eds.), *Theories of mood and cognition: A users guide* (pp. 159-176). New Jersey: Lawrence Erlbaum Associates.

Schwarz, N., & Clore, G. L. (1983). Mood, misattributions, and judgments of well-being: informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513 – 523.

Siegel-Jacobs, K., & Yates, J. F. (1996). Effects of procedural and outcome accountability on judgment quality. *Organizational Behavior and Human Decision Processes, 65*, 1-17.

Simonson, I., & Staw, B. M. (1992). Deescalation strategies: A comparison of techniques for reducing commitment to losing courses of action. *Journal of Applied Psychology, 77,* 419 – 426.

Sinclair, R. C. (1988). Mood, categorization breadth, and performance appraisal: The effects of order of information acquisition and affective state on halo, accuracy, information retrieval, and evaluations. *Organizational Behavior and Human Decision Processes, 42*, 22-46.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Tetlock, P. E. (1992). The impact of accountability on judgment and choice: toward a social contingency model. *Experimental Social Psychology, 25,* 331 – 376.

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., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology, 57,* 632 – 640.

Tsui, A. S., & Barry, B. (1986). Interpersonal affect and rating errors. *Academy of Management Journal, 29,* 586-599.

Velten, E. Jr. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy, 6*, 473 – 482.

Weigold, M. F., & Schlenker, B. R. (1991). Accountability and risk taking. *Personality and Social Psychology Bulletin, 17,* 25 – 29.

Weiss, R. J. (1996). *The effect of response mode and affective state on multiattribute decision making.* Unpublished Master's thesis, University of Nebraska, Omaha.

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.

Appendix A

Affective Reactions Scale

Appendix A

Affective Reactions Scale

Affective Reactions Scale. Please read the following sets of adjectives pairs and circle the number that most closely resembles what you are feeling right now. For example, if you are feeling scared you might circle the number 1; conversely, if you are feeling reassured you might circle the number 6. However, if you are feeling neither scared nor reassured you might circle the number 3. If you have any questions regarding the instructions, please let me know.

1	scared	1	2	3	4	5	6	reassured
2	energetic	1	2	3	4	5	6	tired
3	riled	1	2	3	4	5	6	pacified
4	relieved	1	2	3	4	5	6	apprehensive
5	admiration	1	2	3	4	5	6	contempt
6	afraid	1	2	3	4	5	6	unafraid
7	tranquil	1	2	3	4	5	6	agitated
8	repulsed	1	2	3	4	5	6	attracted
9	interested	1	2	3	4	5	6	bored
10	passive	1	2	3	4	5	6	uptight
11	concerned	1	2	3	4	5	6	unconcerned
12	undisturbed	1	2	3	4	5	6	mad
13	apathetic	1	2	3	4	5	6	enthusiastic
14	cranky	1	2	3	4	5	6	good-humored
15	unperturbed	1	2	3	4	5	6	anxious
16	somber	1	2	3	4	5	6	cheerful
17	calm	1	2	3	4	5	6	excited
18	disappointed	1	2	3	4	5	6	delighted
19	detached	1	2	3	4	5	6	engrossed
20	relaxed	1	2	3	4	5	6	tense
21	jittery	1	2	3	4	5	6	serene
22	fearful	1	2	3	4	5	6	fearless
23	hopeful	1	2	3	4	5	6	hopeless
24	offended	1	2	3	4	5	6	unoffended
25	composed	1	2	3	4	5	6	nervous
26	captivated	1	2	3	4	5	6	disinterested
27	sluggish	1	2	3	4	5	6	alert
28	placated	1	2	3	4	5	6	angry
29	depressed	1	2	3	4	5	6	elated
30	sedate	1	2	3	4	5	6	jumpy
31	unruffled	1	2	3	4	5	6	irritated
32	pleased	1	2	3	4	5	6	displeased
33	alarmed	1	2	3	4	5	6	unalarmed

34	dejected	1	2	3	4	5	6	exhilarated
35	peppy	1	2	3	4	5	6	drained
36	objectionable	1	2	3	4	5	6	unobjectionable
37	unbothered	1	2	3	4	5	6	disgusted
38	threatened	1	2	3	4	5	6	secure
39	sad	1	2	3	4	5	6	happy
40	lively	1	2	3	4	5	6	quiet
41	distressed	1	2	3	4	5	6	comforted
42	passionate	1	2	3	4	5	6	dispassionate

Appendix B

Example of the Information Board

Appendix B

Example of the Information Board

Area of Evaluation

Dimension: Planning

	Comes Prepared for Class	Time spent preparing	Class Material is Organized	Lacks Materials Needed for Class	Number of Late starts	Lecture Strays from the Syllabus	Rating
J. S.							
C. S.							
E. J.							
M. S.							
S. J.							
B. R.							

Appendix C

Accountability Instructions

Appendix C

Accountability Instructions

No Accountability

Dean Hendricks has asked me to validate a computer-based method for evaluating the teaching performance of the colleges' teaching assistants. As part of the validation process you will be asked to assess the teaching performance of 30 teaching assistants in the college of Arts and Sciences. You will be asked to evaluate the teaching assistants on two dimensions of teaching behavior; preparation and presentation.

The purpose of the new performance evaluation system is to help facilitate the decision making process of the Department Chairs in determining which teaching assistants should have their contracts renewed and which should have their contracts terminated. The new system will also assist the Dean in determining which graduate programs will continue to receive funding and which will be eliminated.

The determination of which teaching assistants to retain and which to terminate, and which graduate programs will continue to receive funding or be eliminated will be determined using the following criteria. All teaching assistants who receive a below average rating will NOT have their teaching contracts renewed for the next semester. Additionally, the teaching assistant position will be eliminated, which could result in the elimination of the graduate program for that department. The elimination of the graduate program will force the department to offer fewer undergraduate courses during the semester.

To evaluate the teaching performance of the teaching assistants you will be presented with an information board containing information about the teaching assistants' behavior on several aspects of preparation and presentation. In the last column of the information board you will be asked to rate the teaching assistants' overall performance on the relevant dimension (see example below).

Your evaluations of the teaching assistants' performance will remain anonymous.

If you have any questions about what you are to do, please ask now.

Appendix C cont.

Process Accountability.

Dean Hendricks has asked me to validate a computer-based method for evaluating the teaching performance of the colleges' teaching assistants. Dean Hendricks is especially interested in what information students use when evaluating the performance of teaching assistants, and how important the information was in determining the performance rating. As part of the validation process, you will be asked to assess the teaching performance of 30 teaching assistants in the college of Arts and Sciences on two dimensions of teaching behavior; preparation and presentation.

The purpose of the new performance evaluation system is to help facilitate the decision making process of the Department Chairs in determining which teaching assistants should have their contracts renewed and which should have their contracts terminated. The new system will also assist the Dean in determining which graduate programs will continue to receive funding and which programs will be eliminated.

The determination of which teaching assistants to retain and which to terminate, and which graduate programs will continue to receive funding or be eliminated will be determined using the following criteria. All teaching assistants who receive a below average rating will NOT have their teaching contracts renewed for the next semester. Additionally, the teaching assistant position will be eliminated, which could result in the elimination of the graduate program for that department. The elimination of the graduate program will force the department to offer fewer undergraduate courses during the semester.

To evaluate the teaching performance of the teaching assistants you will be presented with an information board containing information about the teaching assistants' behavior on several aspects of preparation and presentation. In the last column of the information board you will be asked to rate the teaching assistants' overall performance on the relevant dimension.

Dean Hendricks will be here at the end of the evaluation period to discuss with you what information you used, and how important the information was, when evaluating the performance of the teaching assistants. Dean Hendricks is not interested in the actual performance evaluations, he is only interested in the information that was used to make the evaluations.

Appendix C cont.

Outcome Accountability

Dean Hendricks has asked me to validate a computer-based method for evaluating the teaching performance of the colleges' teaching assistants. Dean Hendricks is especially interested in whether student evaluators will generate performance ratings that are similar to the ratings generated by professors. As part of the validation process, you will be asked to assess the teaching performance of 30 teaching assistants in the college of Arts and Sciences on two dimensions of teaching behavior; preparation and presentation.

The purpose of the new performance evaluation system is to help facilitate the decision making process of the Department Chairs in determining which teaching assistants should have their contracts renewed and which should have their contracts terminated. The new system will also assist the Dean in determining which graduate programs will continue to receive funding and which programs will be eliminated.

The determination of which teaching assistants to retain and which to terminate, and which graduate programs will continue to receive funding or be eliminated will be determined using the following criteria. All teaching assistants who receive a below average rating will NOT have their teaching contracts renewed for the next semester. Additionally, the teaching assistant position will be eliminated, which could result in the elimination of the graduate program for that department. The elimination of the graduate program will force the department to offer fewer undergraduate courses during the semester.

To evaluate the teaching performance of the teaching assistants you will be presented with an information board containing information about the teaching assistants' behavior on several aspects of preparation and presentation. In the last column of the information board you will be asked to rate the teaching assistants' overall performance on the relevant dimension

Dean Hendricks will be here at the end of the evaluation period to discuss with you your ratings and compare your ratings to the evaluations made by professors. Dean Hendricks is not interested in the information that you used when evaluating the teaching assistants, he is only interested in how close your ratings match the ratings given by professors.

Appendix D

Costello-Comrey Depression and Anxiety Scales

Appendix D

Costello-Comrey Depression and Anxiety Scales

For the Following 23 questions, please circle the number that best describes your response to each item.

1 I feel that life is worthwhile.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

2 When I wake up in the morning I expect to have a miserable day.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

3 I wish I had never been born.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

4 I feel that there is more disappointment in life than satisfaction.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

5 I want to run away from everything

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

6 My future looks hopeful and promising.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

7 When I get up in the morning I expect to have an interesting day.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

8 Living is a wonderful adventure for me

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

9 I am a happy person.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

10 Things have worked out well for me.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

11 The future looks so gloomy that I wonder if I should go on.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

12 I feel that life is drudgery and boredom.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

13 I feel blue and depressed.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

14 When I look back I think life has been good to me.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

15 I get rattled easily.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

16 When faced with excitement or unexpected situations, I become nervous and jumpy.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

17 I am calm and not easily upset.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

18 When things go wrong I get nervous and upset instead of calmly thinking out a solution.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

19 It makes me nervous when I have to wait.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

20 I am a tense "high strung" person.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

21 I am more sensitive than most other people.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

22 My hand shakes when I try to do something.

Always	almost always	very frequently	frequently	fairly often	occasionally	rarely	almost never	never
9	8	7	6	5	4	3	2	1

23 I am a very nervous person.

absolutely	very definitely	definitely	probably	possibly	probably not	definitely not	very definitely not	absolutely not
9	8	7	6	5	4	3	2	1

Appendix E

Demographic Questionnaire

Appendix E

Demographic Questionnaire

Please answer the following 7 questions.

What is your gender (circle one)?

Male Female

What is your ethnicity (circle one)?

Caucasian	Black	Hispanic	Asian	Native American	No Response

How old are you? _____

How many credit hours have you completed (circle one)?

0 - 26 27 - 57 58 - 90 91 -130 131 +

What was the purpose of the video rating task?

Dean Hendricks will be here to discuss my ratings at the end of the study?

Strongly Disagree	Disagree	not sure	Agree	Strongly Agree
1	2	3	4	5

The video rating task and the information search task are related?

Strongly Disagree	Disagree	not sure	Agree	Strongly Agree
1	2	3	4	5

Appendix F

Perceived Effectiveness of Teaching Behaviors

Appendix F

Perceived Effectiveness of Teaching Behaviors

Dimension	Item	Statement	*M*	*SD*
Comprehension	1	The TA demonstrates a clear understanding of course material	6.00	0.82
	2	The TA appears to understand the course material	6.43	0.79
	3	The TA repeatedly corrects him/herself during the discussion	3.71	2.06
	4	The TA pauses frequently during the discussion as if trying to remember his/her line of thought.	2.44	1.51
	5	The TA speaks confidently about the material	6.14	0.90
Organization	1	Discussions seem well organized and are easy to follow	6.00	1.53
	2	Discussions seem disorganized and are difficult to follow	1.29	0.49
	3	Discussions seem haphazard, but are enjoyable	4.14	1.46
	4	Discussion frequently strays from the topic	3.86	1.57
	5	Discussion is focused and rarely strays from the topic	5.89	0.78
Readiness	1	The TA usually has all the material necessary for discussion	6.57	0.53
	2	The TA has been known to go back to his/her office to retrieve class material	3.29	2.29
	3	The TA frequently looks down at his/her notes during the discussion	2.71	1.60
	4	The TA hardly ever looks at his/her notes during discussion	4.71	1.11
	5	The TA repeatedly shuffles his/her notes during the discussion	2.33	1.12
Preparation	1	The TA spends at least 1 hour before class preparing for the discussion	6.00	0.82
	2	The TA spends at least 2 hours per week preparing for the discussion	4.71	0.76
	3	The TA spends at least a 1/2 hour per week preparing for the discussion	4.44	1.59
	4	The TA spends at least a 1/2 hour every day preparing for the discussion	5.43	0.98

Appendix F cont.

Dimension	Item	Statement	*M*	*SD*
Punctuality	1	Discussion usually starts on time	6.44	0.53
	2	Discussions frequently run late	3.29	2.36
	3	Discussion rarely starts on time	2.14	1.07
	4	Discussion usually ends on time	5.71	0.76
	5	Discussion frequently ends early	4.14	1.57
GPA	1	The TA's grade point average is 3.345	4.29	1.38
	2	The TA's grade point average is 3.1	4.56	1.01
	3	The TA's grade point average is 3.949	4.57	2.15
	4	The TA's grade point average is 3.31	4.57	0.98
	5	The TA's grade point average is 3.549	3.86	2.27
Clarity	1	In an attempt to clarify course material, the TA uses interesting examples during the discussion	6.57	0.53
	2	In an attempt to clarify course material, the TA uses abstract examples during the discussion	5.00	1.53
	3	In an attempt to clarify course material, the TA uses real-world examples during the discussion	6.71	0.49
	4	In an attempt to clarify course material, the TA uses the same examples that are in the text	4.14	2.12
	5	In an attempt to clarify course material, the TA uses obscure examples during the discussion	3.00	1.12
Stimulates thinking	1	The TA assigns materials that have been effective aids to learning	6.33	0.71
	2	The TA only lectures during the discussion	4.00	1.83
	3	The TA continually reads from the book and his/her notes	4.43	1.51
	4	The TA engages the students with questions during the discussion	6.14	1.07
	5	The TA spends more time talking about the weekend than on course material	1.14	0.38

Appendix F cont.

Dimension	Item	Statement	*M*	*SD*
Enthusiasm	1	The TA continually moves around the class during the discussion	4.00	1.63
	2	The TA speaks in a monotone voice during discussion	3.14	1.86
	3	The TA speaks at a rapid pace, but with enthusiasm	4.57	1.40
	4	Discussion comes across dry and monotonous	1.56	0.53
	5	The TA generally seems exited about the discussion material	5.71	1.38
Engagement	1	The TA tries to call on all of the students to answer questions	6.00	0.82
	2	The TA tends to ignore questions from students	1.22	0.44
	3	The TA tries to ask questions that the students can answer	6.00	1.29
	4	The TA asks questions that are beyond the level of the students	3.30	1.70
	5	The TA calls on students to answer questions who appear not to know the answer	2.00	1.41
Answers questions	1	The TA responds to questions with enthusiasm	6.14	1.07
	2	The TA answers questions with a concise answer	6.29	1.11
	3	When the TA can not answer the question, he/she often states that he/she will get the answer	4.57	1.90
	4	The TA tends to answer a question with another question	2.44	1.33
	5	The TA asks the question to the rest of the class	5.57	0.98
Year in School	1	The TA is a first year graduate student	3.89	1.17
	2	The TA is a second year graduate student	4.86	0.69
	3	The TA is a third year graduate student	5.14	1.46
	4	The TA has recently returned to school after a 3 year break	3.43	1.62
	5	The TA took a year off from school before beginning graduate school	3.43	2.23

Appendix G

Three Information Boards on Planning for Discussion, Boards 1, 2, and 3.

Appendix G

Three Information Boards on Planning for Discussion, Boards 1, 2, and 3.

Teaching Assistant	Understanding	Organization	Knows the Material	Time Spent Preparing	Punctuality	Current GPA
J. S.	The TA demonstrates a clear understanding of course material	Discussion is focused and rarely strays from the topic	The TA hardly ever looks at his/her notes during discussion	Spends at least a 1/2 before class preparing for the discussion	Discussion rarely starts on time	The TA's grade point average is 3.001
C. S.	The TA understands course material	Discussions frequently stray from the topic	The TA frequently looks down at his/her notes during the discussion	Spends at least 1 hour before class preparing for the discussion	Discussion frequently run late	The TA's grade point average is 3.949
E. J.	The TA repeatedly corrects self during the discussion	Discussions seem haphazard but are enjoyable	The TA has been known to go back to his/her office to retrieve class materials	Spends at least a 1/2 hour every day preparing for the discussion	Discussion usually starts on time	The TA's grade point average is 3.549
M. S.	The TA pauses frequently during the discussion	Discussions seem disorganized and are difficult to follow	The TA usually has all the material necessary for discussion	Spends at least 2 hours per week preparing for the discussion	Discussion usually ends on time	The TA's grade point average is 3.301
S. J.	The TA speaks confidently about the material	Discussions seem well organized and are easy to follow	The TA repeatedly shuffles his/her notes during the discussion	Spends at least a 1/2 hour per week preparing for the discussion	Discussion frequently ends early	The TA's grade point average is 3.345
B. R.	The TA understands course material	Discussion is focused and rarely strays from the topic	The TA hardly ever looks at his/her notes during discussion	Spends at least a 1/2 before class preparing for the discussion	Discussion frequently ends early	The TA's grade point average is 3.001
C.T.	The TA repeatedly corrects self during the discussion	Discussions frequently stray from the topic	The TA frequently looks down at his/her notes during the discussion	Spends at least 1 hour before class preparing for the discussion	Discussion usually ends on time	The TA's grade point average is 3.345
A.G.	The TA pauses frequently during the discussion	Discussions seem haphazard but are enjoyable	The TA has been known to go back to his/her office to retrieve class materials	Spends at least a 1/2 hour every day preparing for the discussion	Discussion rarely starts on time	The TA's grade point average is 3.301
C.M.	The TA speaks confidently about the material	Discussions seem disorganized and are difficult to follow	The TA usually has all the material necessary for discussion	Spends at least a 1/2 hour per week preparing for the discussion	Discussion frequently run late	The TA's grade point average is 3.549
C.F.	The TA demonstrates a clear understanding of course material	Discussions seem well organized and are easy to follow	The TA repeatedly shuffles his/her notes during the discussion	Spends at least 2 hours per week preparing for the discussion	Discussion usually starts on time	The TA's grade point average is 3.949

Appendix G cont.

Teaching Assistant	Understanding	Organization	Knows the Material	Time Spent Preparing	Punctuality	Current GPA
A.H.	The TA repeatedly corrects self during the discussion	Discussion is focused and rarely strays from the topic	The TA has been known to go back to his/her office to retrieve class materials	Spends at least a 1/2 hour every day preparing for the discussion	Discussion usually ends on time	The TA's grade point average is 3.949
S.P.	The TA pauses frequently during the discussion	Discussions frequently stray from the topic	The TA frequently looks down at his/her notes during the discussion	Spends at least a 1/2 before class preparing for the discussion	Discussion frequently ends early	The TA's grade point average is 3.301
J.J.	The TA speaks confidently about the material	Discussions seem haphazard but are enjoyable	The TA repeatedly shuffles his/her notes during the discussion	Spends at least a 1/2 hour per week preparing for the discussion	Discussion frequently run late	The TA's grade point average is 3.549
J.Y.	The TA demonstrates a clear understanding of course material	Discussions seem disorganized and are difficult to follow	The TA hardly ever looks at his/her notes during discussion	Spends at least 2 hours per week preparing for the discussion	Discussion rarely starts on time	The TA's grade point average is 3.001
D.B.	The TA understands course material	Discussions seem well organized and are easy to follow	The TA usually has all the material necessary for discussion	Spends at least 1 hour before class preparing for the discussion	Discussion usually starts on time	The TA's grade point average is 3.345
E.F.	The TA pauses frequently during the discussion	Discussion is focused and rarely strays from the topic	The TA frequently looks down at his/her notes during the discussion	Spends at least a 1/2 hour every day preparing for the discussion	Discussion frequently run late	The TA's grade point average is 3.345
J.G.	The TA speaks confidently about the material	Discussions frequently stray from the topic	The TA repeatedly shuffles his/her notes during the discussion	Spends at least 2 hours per week preparing for the discussion	Discussion rarely starts on time	The TA's grade point average is 3.301
R.S.	The TA demonstrates a clear understanding of course material	Discussions seem haphazard but are enjoyable	The TA hardly ever looks at his/her notes during discussion	Spends at least a 1/2 hour per week preparing for the discussion	Discussion usually ends on time	The TA's grade point average is 3.549
K.T.	The TA understands course material	Discussions seem disorganized and are difficult to follow	The TA has been known to go back to his/her office to retrieve class materials	Spends at least a 1/2 before class preparing for the discussion	Discussion usually starts on time	The TA's grade point average is 3.001
B.H.	The TA repeatedly corrects self during the discussion	Discussions seem well organized and are easy to follow	The TA usually has all the material necessary for discussion	Spends at least a 1/2 hour every day preparing for the discussion	Discussion frequently ends early	The TA's grade point average is 3.949

Appendix G cont.

Teaching Assistant	Understanding	Organization	Knows the Material	Time Spent Preparing	Punctuality	Current GPA
P.R.	The TA speaks confidently about the material	Discussion is focused and rarely strays from the topic	The TA has been known to go back to his/her office to retrieve class materials	Spends at least a 1/2 before class preparing for the discussion	Discussion frequently run late	The TA's grade point average is 3.949
S.S.	The TA demonstrates a clear understanding of course material	Discussions frequently stray from the topic	The TA hardly ever looks at his/her notes during discussion	Spends at least a 1/2 hour every day preparing for the discussion	Discussion usually starts on time	The TA's grade point average is 3.301
T.R.	The TA understands course material	Discussions seem haphazard but are enjoyable	The TA repeatedly shuffles his/her notes during the discussion	Spends at least a 1/2 hour per week preparing for the discussion	Discussion frequently ends early	The TA's grade point average is 3.001
K.S.	The TA repeatedly corrects self during the discussion	Discussions seem disorganized and are difficult to follow	The TA usually has all the material necessary for discussion	Spends at least 1 hour before class preparing for the discussion	Discussion usually ends on time	The TA's grade point average is 3.549
L.S.	The TA pauses frequently during the discussion	Discussions seem well organized and are easy to follow	The TA usually has all the material necessary for discussion	Spends at least 2 hours per week preparing for the discussion	Discussion rarely starts on time	The TA's grade point average is 3.345
M.B.	The TA speaks confidently about the material	Discussion is focused and rarely strays from the topic	The TA frequently looks down at his/her notes during the discussion	Spends at least 1 hour before class preparing for the discussion	Discussion frequently ends early	The TA's grade point average is 3.549
E.J.	The TA pauses frequently during the discussion	Discussions frequently stray from the topic	The TA repeatedly shuffles his/her notes during the discussion	Spends at least a 1/2 before class preparing for the discussion	Discussion usually starts on time	The TA's grade point average is 3.001
K.R.	The TA repeatedly corrects self during the discussion	Discussions seem haphazard but are enjoyable	The TA hardly ever looks at his/her notes during discussion	Spends at least a 1/2 hour per week preparing for the discussion	Discussion usually ends on time	The TA's grade point average is 3.345
G.D.	The TA understands course material	Discussions seem disorganized and are difficult to follow	The TA has been known to go back to his/her office to retrieve class materials	Spends at least 2 hours per week preparing for the discussion	Discussion frequently run late	The TA's grade point average is 3.301
M.M.	The TA demonstrates a clear understanding of course material	Discussions seem well organized and are easy to follow	The TA frequently looks down at his/her notes during the discussion	Spends at least a 1/2 hour every day preparing for the discussion	Discussion rarely starts on time	The TA's grade point average is 3.949

Appendix H.

Three Information Boards on Presentation of Class Material Boards 4, 5, and 6

Appendix H.

Three Information Boards on Presentation of Class Material Boards 4, 5, and 6

Teaching Assistant	Clarity	Stimulates Thinking	Enthusiasm	Engagement	Answers Questions	Year in School
J. S.	The TA uses interesting examples during the discussion	The TA spends more time talking about the weekend than on course material	The discussion comes across dry and monotonous	The TA calls on students to answer questions who appear not to know the answer	The TA tends to answer a question with another question	The TA is a third year graduate student
C. S.	The TA uses abstract examples during the discussion	The TA engages the students with questions during the discussion	The TA speaks in a monotone voice during discussion	The TA asks questions that are beyond the level of the students	The TA responds to questions with enthusiasm	The TA is a first year graduate student
E. J.	The TA uses real-world examples during the discussion	The TA continually reads from the book and his/her notes	The TA continually moves around the class during discussion	The TA tends to ignore questions from students	When students asks questions, the TA often states that he/she will get the answer	The TA is a second year graduate student
M. S.	The TA uses the same examples that are in the text	The TA only lectures during the discussion	The TA speaks at a rapid pace and with enthusiasm	The TA tries to ask questions that the students can answer	When a student asks a question, the TA asks the question to the rest of the class	The TA has recently returned to school after a 3 year break
S. J.	The TA uses obscure examples during the discussion	The TA assigned materials that were effective aids to learning	The TA generally seems exited about the discussion material	The TA tries to call on all of the students to answer questions	The TA responds to questions with a concise answer	The TA took a year off from school before beginning graduate school
B. R.	The TA uses the same examples that are in the text	The TA assigned materials that were effective aids to learning	The TA speaks in a monotone voice during discussion	The TA calls on students to answer questions who appear not to know the answer	The TA responds to questions with a concise answer	The TA has recently returned to school after a 3 year break
C.T.	The TA uses real-world examples during the discussion	The TA spends more time talking about the weekend than on course material	The TA speaks at a rapid pace and with enthusiasm	The TA tries to call on all of the students to answer questions	When a student asks a question, the TA asks the question to the rest of the class	The TA is a third year graduate student
A.G.	The TA uses obscure examples during the discussion	The TA continually reads from the book and his/her notes	The TA continually moves around the class during discussion	The TA tends to ignore questions from students	The TA tends to answer a question with another question	The TA took a year off from school before beginning graduate school
C.M.	The TA uses abstract examples during the discussion	The TA only lectures during the discussion	The discussion comes across dry and monotonous	The TA asks questions that are beyond the level of the students	When students asks questions, the TA often states that he/she will get the answer	The TA is a second year graduate student
C.F.	The TA uses interesting examples during the discussion	The TA engages the students with questions during the discussion	The TA generally seems exited about the discussion material	The TA tries to ask questions that the students can answer	The TA responds to questions with enthusiasm	The TA is a first year graduate student

Appendix H cont.

Teaching Assistant	Clarity	Stimulates Thinking	Enthusiasm	Engagement	Answers Questions	Year in School
A.H.	The TA uses abstract examples during the discussion	The TA spends more time talking about the weekend than on course material	The TA generally seems exited about the discussion material	The TA calls on students to answer questions who appear not to know the answer	The TA responds to questions with enthusiasm	The TA has recently returned to school after a 3 year break
S.P.	The TA uses real-world examples during the discussion	The TA continually reads from the book and his/her notes	The TA speaks at a rapid pace and with enthusiasm	The TA asks questions that are beyond the level of the students	When students asks questions, the TA often states that he/she will get the answer	The TA is a first year graduate student
J.J.	The TA uses the same examples that are in the text	The TA only lectures during the discussion	The TA speaks in a monotone voice during discussion	The TA tries to ask questions that the students can answer	When a student asks a question, the TA asks the question to the rest of the class	The TA is a second year graduate student
J.Y.	The TA uses obscure examples during the discussion	The TA engages the students with questions during the discussion	The discussion comes across dry and monotonous	The TA tends to ignore questions from students	The TA tends to answer a question with another question	The TA is a third year graduate student
D.B.	The TA uses interesting examples during the discussion	The TA assigned materials that were effective aids to learning	The TA continually moves around the class during discussion	The TA tries to call on all of the students to answer questions	The TA responds to questions with a concise answer	The TA took a year off from school before beginning graduate school
E.F.	The TA uses abstract examples during the discussion	The TA assigned materials that were effective aids to learning	The TA speaks at a rapid pace and with enthusiasm	The TA tries to call on all of the students to answer questions	When students asks questions, the TA often states that he/she will get the answer	The TA is a third year graduate student
J.G.	The TA uses real-world examples during the discussion	The TA continually reads from the book and his/her notes	The TA speaks in a monotone voice during discussion	The TA asks questions that are beyond the level of the students	The TA tends to answer a question with another question	The TA is a second year graduate student
R.S.	The TA uses interesting examples during the discussion	The TA spends more time talking about the weekend than on course material	The TA generally seems exited about the discussion material	The TA calls on students to answer questions who appear not to know the answer	When a student asks a question, the TA asks the question to the rest of the class	The TA has recently returned to school after a 3 year break
K.T.	The TA uses the same examples that are in the text	The TA engages the students with questions during the discussion	The discussion comes across dry and monotonous	The TA tries to ask questions that the students can answer	The TA responds to questions with a concise answer	The TA took a year off from school before beginning graduate school
B.H.	The TA uses obscure examples during the discussion	The TA only lectures during the discussion	The TA continually moves around the class during discussion	The TA tends to ignore questions from students	The TA responds to questions with enthusiasm	The TA is a first year graduate student

Appendix H cont.

Teaching Assistant	Clarity	Stimulates Thinking	Enthusiasm	Engagement	Answers Questions	Year in School
P.R.	The TA uses the same examples that are in the text	The TA spends more time talking about the weekend than on course material	The TA continually moves around the class during discussion	The TA tries to call on all of the students to answer questions	The TA responds to questions with a concise answer	The TA is a second year graduate student
S.S.	The TA uses abstract examples during the discussion	The TA assigned materials that were effective aids to learning	The TA speaks in a monotone voice during discussion	The TA tends to ignore questions from students	The TA tends to answer a question with another question	The TA is a third year graduate student
T.R.	The TA uses real-world examples during the discussion	The TA only lectures during the discussion	The TA speaks at a rapid pace and with enthusiasm	The TA calls on students to answer questions who appear not to know the answer	When students asks questions, the TA often states that he/she will get the answer	The TA has recently returned to school after a 3 year break
K.S.	The TA uses obscure examples during the discussion	The TA engages the students with questions during the discussion	The discussion comes across dry and monotonous	The TA asks questions that are beyond the level of the students	When a student asks a question, the TA asks the question to the rest of the class	The TA took a year off from school before beginning graduate school
L.S.	The TA uses interesting examples during the discussion	The TA continually reads from the book and his/her notes	The TA generally seems exited about the discussion material	The TA tries to ask questions that the students can answer	The TA responds to questions with enthusiasm	The TA is a first year graduate student
M.B.	The TA uses abstract examples during the discussion	The TA spends more time talking about the weekend than on course material	The discussion comes across dry and monotonous	The TA tries to call on all of the students to answer questions	The TA responds to questions with enthusiasm	The TA is a second year graduate student
E.J.	The TA uses obscure examples during the discussion	The TA engages the students with questions during the discussion	The TA speaks at a rapid pace and with enthusiasm	The TA tries to ask questions that the students can answer	The TA tends to answer a question with another question	The TA is a third year graduate student
K.R.	The TA uses the same examples that are in the text	The TA continually reads from the book and his/her notes	The TA generally seems exited about the discussion material	The TA asks questions that are beyond the level of the students	When students asks questions, the TA often states that he/she will get the answer	The TA has recently returned to school after a 3 year break
G.D.	The TA uses real-world examples during the discussion	The TA only lectures during the discussion	The TA speaks in a monotone voice during discussion	The TA tends to ignore questions from students	The TA responds to questions with a concise answer	The TA took a year off from school before beginning graduate school
M.M.	The TA uses interesting examples during the discussion	The TA assigned materials that were effective aids to learning	The TA continually moves around the class during discussion	The TA calls on students to answer questions who appear not to know the answer	When a student asks a question, the TA asks the question to the rest of the class	The TA is a first year graduate student

Appendix I

Importance and Likelihood Ratings of Individual Performance Appraisal Outcomes

Appendix I

Importance and Likelihood Ratings of Individual Performance Appraisal Outcomes

	Importance		Likelihood	
Statement	*M*	*SD*	*M*	*SD*
The teaching assistant who receives the lowest rating will NOT have his or her teaching contract renewed for the next semester and will be replaced by another graduate student.	4.11	0.51	3.44	0.47
The teaching assistant who receives the lowest rating will NOT have his or her teaching contracts renewed for the next semester. The teaching assistant will not be replaced by another graduate student, which could result in a reduction in the number of courses offered by the department.	4.78	0.36	3.78	0.46
The teaching assistant who receives the lowest rating will NOT have his or her teaching contracts renewed for the next semester. Additionally, the teaching assistant position will be eliminated, which could result in the elimination of the graduate program for that department. The elimination of the graduate program will force the department to offer fewer undergraduate courses during the semester.	5.22	0.28	4.22	0.43
The teaching assistant who receives the lowest rating will be sent to remedial training and prevented from teaching until the training is complete. The teaching assistant will be replaced by another graduate student for the following academic year.	4.44	0.6	3.56	0.56

Appendix J

Importance and Likelihood Ratings of Group Performance Appraisal Outcomes

Appendix J

Importance and Likelihood Ratings of Group Performance Appraisal Outcomes

	Importance		Likelihood	
Statement	*M*	*SD*	*M*	*SD*
All teaching assistants who receive a below average rating will NOT have their teaching contracts renewed for the next semester and will be replaced by other graduate students	4.44	0.47	3.44	0.44
All teaching assistants who receive a below average rating will NOT have their teaching contracts renewed for the next semester. The teaching assistant will not be replaced, which will result in a reduction in the number of courses offered by the department.	5	0.37	3.89	0.48
All teaching assistants who receive a below average rating will NOT have their teaching contracts renewed for the next semester. Additionally, the teaching assistant position will be eliminated, which will result in the elimination of the graduate program for that department. The elimination of the graduate program will force the department to offer fewer undergraduate courses during the semester.	5.33	0.29	4.22	0.43
All teaching assistants who receive a below average rating will be sent to remedial training and prevented from teaching until the training is complete. Additionally, the teaching assistant will be replaced by another graduate student for the following academic year.	4.78	0.49	3.78	0.6