The Dissertation Committee for Geoffrey Roger Tumlin

Certifies that this is the approved version of the following dissertation:

Communicating Negative Feedback in Performance Appraisal Interviews:

An Experimental Study

Committee:

_____
Mark L. Knapp, Supervisor

_____
John A. Daly

_____
Madeline M. Maxwell

_____
Howard T. Prince II

_____
Anita L. Vangelisti

Communicating Negative Feedback in Performance Appraisal Interviews:

An Experimental Study

by

Geoffrey Roger Tumlin, B.S., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

In Partial Fulfillment

Of the Requirements

For the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2004

Dedication

To Mark L. Knapp, advisor and friend

Acknowledgements

Thank you to my committee members, my family, and my friends.

Communicating Negative Feedback in Performance Appraisal Interviews:

An Experimental Study

Publication No. _____

Geoffrey Roger Tumlin, Ph.D.
The University of Texas at Austin, 2004

Supervisor: Mark L. Knapp

This study examined the communicative transmission of negative feedback in a performance appraisal context and experimentally tested how subordinate account making, specifically an external excuse and a concession account with an internal excuse, influenced a supervisor's ability to deliver negative feedback and how those same accounts and the corresponding amount of positive feedback interacted with subordinate evaluation ratings. This experiment also charted the implications of feedback specificity on the subordinate's satisfaction with the supervisor's communication during the performance appraisal interview. Results from this study indicated that: (1) supervisors did not mention positive observations about the subordinate significantly more than negative observations, (2) 81.10% of supervisors self-reported that they delivered all of the negative feedback they felt obligated to deliver to the subordinate, (3) 83.00% of supervisors positively distorted their feedback during the appraisal interview when the

subordinate offered an internal excuse account for performance failure; 79.20% of the supervisors positively distorted their feedback when a concession account with external excuse was offered, (4) subordinate evaluation scores were more positive after the appraisal than before the appraisal, regardless of the type of account given, (5) the subordinate was evaluated more positively when an external excuse was given as compared to when a concession account with an internal excuse was offered, (6) supervisors who received the external excuse first in the experimental sequence gave significantly higher evaluation scores across both conditions of the study, and (7) the subordinate's satisfaction with a supervisor's communication during the appraisal interview consistently increased as the supervisor's feedback become more specific.

Table of Contents

List of Tables

Chapter 1

Introduction and Purpose

Behind the adage "don't blame the messenger for the bad news" is the well documented concept that individuals are reluctant to deliver negative feedback, criticism, and bad news to other people (e.g. Baron, 1993; Larson, 1986, 1989; Tesser & Rosen, 1975). In addition to the proverbial fear of "shooting the messenger," other reasons make the delivery of negative feedback problematic: concerns of face violations (Brown & Levinson, 1987; Wagoner & Waldron, 1999) account for some aversion to deliver bad news, potential damage done to interpersonal relationships (Larson, 1984), and a degradation in organizational commitment (Pearce & Porter, 1986) to name but a few.

Furthermore, a string of research done by Tesser, Rosen, and colleagues (e.g. Rosen & Tesser, 1970, 1972; Tesser & Rosen, 1972; Tesser, Rosen, & Batchelor, 1972a, 1972b) details a mum effect—a consistent reluctance to communicate bad news—across numerous situations, communicators, and recipients. The aversion to transmitting bad news is also well researched in the domain of health communication (e.g. Hines, Moss, & McKenzie, 1997; Pomerantz, Fehr, & Ende, 1997; Ptacek & Eberhardt, 1996), for the transparent reason that communicating bad news is a common, if often problematic task for physicians.

In addition to the areas mentioned above, there are innumerable contexts from which to explore the transmission of negative feedback and bad news. This study focuses on performance appraisal. Specifically, the context that is under investigation in this research is a performance appraisal that is (a) formal, (b) scheduled, (c) conducted between the employee and the supervisor (dyadic), and (d) has tangible ramifications on the employee's positional and financial standing in the organization.

In some ways, it is hard to imagine a communicative context between supervisors and subordinates where the stakes are higher than a performance appraisal. For the employee, promotion, raises, or possibly a dismissal loom during a performance appraisal, as every statement is analyzed for potential hidden meanings, nuances, or hints about the employee's future. For the supervisor, the challenge is often to balance accurate, candid evaluation with coaching and support, sustaining exemplary performance when appropriate and refocusing the employee on organizational goals when performance wanes.

It is helpful to be consistent in terminology with respect to the phenomena under investigation in this study. Terms like *bad news, criticism,* and *negative feedback* are used interchangeably in the literature on communicating difficult messages. Given that this experiment focuses on the performance appraisal context, it is most appropriate to use the term *negative feedback* to describe the communicative event under consideration, since performance appraisals are fundamentally vehicles for individual feedback. Furthermore, this study is tailored to explore the feedback that is transmitted or withheld in response to subordinate *performance failure*, another term that will be used throughout this research.

What makes this a study of communication, as opposed to a study of psychology, management, or organizational behavior? First, communication is the avenue through which the performance appraisal occurs. Simply put, there is no performance appraisal without communication; no effective appraisal without effective communication. Numerous empirical studies have documented the primacy of communication in the appraisal process (for a partial listing see Downs, 1990, p.335). And, the relative success or failure of a performance appraisal is frequently judged by its underlying communication; Kikoski and Litterer (1983) argue that

effective communication and ability are the principal components of performance appraisals from a supervisor's vantage point.

Second, this research is grounded in a question that is fundamentally communicative in nature: *Who says what to whom, and of what consequence?* Specifically, this research explores three distinctly communicative acts: (1) how individuals communicate negative feedback, (2) how account making influences (distorts) the transmission of negative performance feedback and employee evaluations within a superior-subordinate dyad, and (3) the interactive influence of perceived specificity of feedback and account making on subordinate satisfaction in performance appraisal. Fundamentally, the entire performance appraisal context is both framed and created through communication. It is this dynamic interaction between source and receiver that provides the framework in which to explore the intricacies of negative feedback transmission in superior-subordinate communication.

Purpose

The purpose of this study is to examine the communicative transmission of negative feedback in performance appraisal, experimentally test how subordinate account making (e.g. external excuses and concessions with internal excuses) influences a supervisor's ability to deliver negative feedback and how those same accounts and the corresponding feedback delivered interacts with subordinate performance ratings, and chart the implications of feedback specificity on subordinate satisfaction with supervisors' communication in the appraisal process.

Understanding the functioning and the interaction of these interpersonal communication dynamics during a performance appraisal is important to a wide variety of organizational stakeholders. For employees, receiving accurate and timely performance feedback during the appraisal interview can warn of potential impediments to their career progression and refocus

them on salient developmental areas in need of attention in the eyes of management. For supervisors, there is an inherent leadership responsibility to deliver accurate performance feedback in hopes of continuously improving subordinates in pursuit of organizational goals. Additionally, performance failures, when identified quickly, have a greater likelihood of being rehabilitated than deficient performance that goes unaddressed.

Voluminous research documents that people have difficulty communicating negatively valenced messages, and a respectable body of knowledge details the performance appraisal process, but far less scholarship focuses exclusively on the interactive, communicative nature of the performance appraisal process in an attempt to determine how negative performance feedback is delivered. This study helps fill that void. It is possible that by better understanding negative feedback transmission in the performance appraisal interview, recommendations can be made to improve the appraisal process and enhance message fidelity within this context. Additionally, discerning how subordinate account making influences the appraisal interview and understanding how employees view different feedback strategies can help both supervisors and organizations to craft appraisals that minimize various structural and cognitive errors and seek to optimize the process of evaluating employees.

Chapter 2

Literature Review and Hypotheses

The Reluctance to Transmit Negative Feedback

Abraham Tesser, Sidney Rosen, and colleagues have coined the term *mum effect* to describe an individual's reluctance to encode and transmit messages perceived as unpleasant for the receiver (Rosen & Tesser, 1970; Tesser & Rosen, 1975). Their research has distilled three categories of reasons why people avoid giving bad news to others: (1) people avoid giving bad news out of their own self-concern (2) people are reluctant to give bad news out of their concern for the other person and (3) people are concerned about the overarching societal norms that govern bad news transmission (Tesser & Rosen, 1977). It is important to note that the mum effect occurs only when the message to be delivered is of specific consequence for the receiver, which makes the mum effect especially salient in performance appraisals. When the bad news is about someone or something that is not directly relevant to the receiver, people are generally *not* reluctant to deliver the bad news, as evidenced by research on rumors. For example, Weenig, Groenenboom, and Wilke (2001) found that people disclose negative rumors without any mum effect when the rumor does not have clearly discernable adverse consequences for the receiver, even though the message is distinctly negative. This is a reasonable point of clarification to the mum effect, which generally assumes that the sender of the bad news makes a determination that there is little to gain and much to lose, interpersonally, from delivering the bad news (Tesser & Rosen, 1975). Rumors present a situation where even when negatively valenced, the rumor's relevance and direct relationship to the receiver is often uncertain. Since negative feedback communicated during a performance appraisal is both highly relevant and certain to the supervisor and to the subordinate, the mum effect is at play within the context of performance

appraisal. And, ironically, a performance appraisal is arguably one of the last places that a mum effect would be desirous, since effective appraisals are based, at least in theory, on premises of honesty, candor, and timeliness. As explained by Tesser and Rosen (1975, p.197):

> The greater reluctance to communicate bad news than good news is restricted to those situations in which the recipient is the person whose fate is altered by the circumstances described in the message. The data have a rather disturbing practical implication. The very individual who needs to know most…may well be the person least likely told about it.

Sender concerned about self.

Self-concern on the part of the sender is one cause that can trigger a mum effect. Tesser and Rosen (1972) conducted a study in which both a subject and a confederate were led to believe that they might be receiving electric shocks as part of an experiment. In cases where the subject was led to believe that there was good news (either the confederate, who reacted strongly to the "test" shock, or both the subject and the confederate would receive no shock), this was communicated with 79% frequency ($p < .05$) as compared with 52% frequency ($p < .05$) when led to believe that the partner would receive a shock. More compellingly, Tesser and Rosen (1972) showed that when the subjects learned that they were to be shocked, but that the confederate was not to be shocked (clearly good news for their confederate partner, who reacted so strongly to the "test" shock), subjects told the confederate of the good news 92% of the time ($p < .01$), compared to 67% of the time ($p < .05$) when both were to receive shocks and 38% ($p < .05$) when the confederate was to be shocked, suggesting that people are reluctant to communicate bad news when they feel some guilt about not sharing in a fate that they perceive as bad news.

Senders can also be concerned about being disliked after giving bad news. In a subsequent experiment, Johnson, Conlee, and Tesser (1974), again found support for the guilt hypothesis as an explanatory variable in the reluctance to communicate bad news. Additionally, they found evidence suggesting that a sender's fear of negative evaluation from telling the bad news also mitigates the relationship.

A series of experiments (Tesser & Conlee, 1973; Tesser, Rosen, & Conlee, 1972) tested the idea that individuals do not like transmitting bad news messages because of mood congruency. The premise behind mood congruency is that people feel that they have to be in a similar mood to deliver valenced news; good news must be delivered with a good mood (positive) affect and bad news with a bad mood (negative) affect. Thus, people will avoid giving bad news when they are in a good mood, since this represents mood incongruence. Findings from these studies, as well as a subsequent experiment (Tesser, Rosen, & Waranch, 1973) found that people in a negative mood communicate both good and bad news more completely and more spontaneously than people in a good mood. Taken together, these results suggest that there is an emotional cost of taking on a bad mood simply to communicate negative news via mood congruency. People are often reluctant to pay this price to transmit bad news; however, research on mood congruency (Tesser & Rosen, 1975; Tesser et al., 1973) shows that people will communicate negative information more readily if they are in a bad mood to begin with.

Sender concerned about receiver.

In addition to concerns about the implications of communicating bad news for the sender, people also worry about transmitting bad news because of the effects on the receiver. Conlee and Tesser (1973) explored the notion that individuals might assume that receivers would not want to hear bad news, and would thus decide to give good news more often than proportionally

occurring bad news. Subjects in their experiment were influenced by what they believed to be the receiver's desire to hear the bad news. The higher the *assumed* reluctance to learn of the bad news, the less frequently the negative news was given.

It is also reasonable to surmise that senders might be reluctant to deliver bad news or negative feedback out of concern for the receiver's reaction to the news (Gibb, 1973). Tesser and Conlee (1973) explored this idea in an experiment in which a subject was to tell a confederate that he or she did not get a desired job (bad news). Subjects were led to believe that the confederate was either emotional, calm, or they were not given any information about the confederate (control). For the calm confederates, news about the job was communicated 79% of the time, compared to a 46% transmission rate for receivers thought to be emotional, and 75% in the control condition ($p < .05$ in all cases).

Lending credence to the idea that transmitting bad news has consequences for the sender, Fisher (1979), in an experiment on the transmission of negative feedback, found that supervisors were evaluated negatively by subordinates after giving a negative feedback message. Earlier, Jones (1966) documented that evaluations tended to be reciprocal, in that positive evaluations *by* a superior resulted in positive evaluations *of* the superior. Similarly, negative feedback resulted in lower evaluations of the superior by the subordinate, supporting the idea that people might be deterred from giving bad news or negative feedback based on the reciprocal impact on their favorability and evaluation by others. In line with this idea, Fisher (1979) also found that superiors who gave negative feedback believed that their subordinates liked them less than supervisors that gave positive feedback. Assuming that a supervisor wants to be liked, this concern with reciprocity also serves to stifle the transmission of negative feedback. More recently, Cloven and Roloff (1993) found that individuals who believed that their partners might

react aggressively to relational complaints were likely not to voice those complaints, resulting in what they called a chilling effect, but could also be characterized as a relational mum effect. Finally, Meyer, Kay, and French (1965) note that the defensiveness that results from critical appraisal frequently produces inferior individual performance, yet another reason for a reluctant manager to avoid transmitting negative feedback.

Sender concerned with norms for sending bad news.

Research on the mum effect also shows that societal norms play a role in the reluctance to communicate bad news. Tesser and Rosen (1975) found that individuals report that they are unwilling to transmit bad news because they believe their role relationship vis-à-vis the receiver is inappropriate for sending the message. Conversely, norms also influence rationales for transmitting bad news, with people reporting that they feel obligated to send the bad news or claiming that they would communicate the negative information simply because there was no compelling reason *not* to communicate the message. In an experiment conducted in five European cities, O'Neal, Levine, and Frank (1979) placed a number of "lost" postcards throughout the cities indicating that a fictitious person had been either under or overcharged $100 on their electric bill. Their results indicate cross-cultural support for a mum effect such that the good news cards ($100 overcharged) were returned more frequently than the bad news cards, $\chi^2 = 4.21$, $p < .05$. These findings lend credence to the idea that the norms for sending good news are more clearly defined than the expectations and norms for communicating negative information. Additionally, Tesser, Rosen, and Batchelor (1972b) found support for the idea that the norms for sending bad news are ambiguous, reporting greater intersubject agreement in both the intention and the perceived obligation to transmit negative as opposed to positive messages.

Positive Distortion

Even when individuals report that they transmitted bad news, closer inspection often reveals that the negative feedback was positively distorted. Distortion has been found in experiments on the mum effect such that individuals will choose to communicate negatively valenced messages by only transmitting the neutral parts of the messages and withholding the explicitly negative portions (Tesser, Rosen et al., 1972a; Tesser, Rosen, & Conlee, 1972). In a performance appraisal, distortion occurs when unfavorable information is withheld and an employee is given a brighter picture of their performance than is otherwise merited (Benedict & Levine, 1988). When tested in a performance appraisal context, research has shown that supervisors will give higher ratings to subordinates when they know that they will be providing those employees with feedback on the ratings, as opposed to times when the evaluations can be given blindly, without employees being able to see the ratings given by the supervisors (Benedict & Levine, 1988; Stockford & Bissell, 1949). Additionally, negative feedback can be positively distorted when accompanied by excessive positive messages designed to soften the blow of the negative feedback.

In light of the mum effect and distortion effects, it is predicted that:

H1a: Supervisors will communicate positive subordinate observations significantly more than negative subordinate observations, regardless of the type of excuse for poor performance offered by the subordinate.

The Interactivity of Receivers and the Influence of Their Account Making

There are a number of things that receivers (subordinates in the case of performance appraisal) do that make the transmission of negative feedback problematic. When commingled with the sender's (supervisor's) potential reluctance to transmit negative feedback, these elements combine to complicate the feedback process.

Self-serving and actor / observer biases.

The tendency for people to overestimate their strengths relative to others is well documented. From driving ability (Svenson, 1981) to managerial skill (Larwood & Whittaker, 1977) to productivity (Cross, 1977), well over half of survey recipients routinely rate themselves in the top half of the population (Babcock & Loewenstein, 1997). This self-serving bias also extends to the groups that people are associated with. In one classic example, Hastrof and Cantril (1954) had students from Dartmouth and Princeton watch a film of a football game between the two schools. Students from Princeton observed three times as many minor penalties and twice as many flagrant penalties committed by the Dartmouth team as compared to the Princeton team. However, Dartmouth students recorded an approximately equal number of penalties between the two teams. Defined from Babcock and Loewenstein (1997, p.110) as the tendency for people "to conflate what is fair with what benefits oneself," the self-serving bias has the potential to contaminate the performance appraisal process, both through self-appraisals, which are routinely conducted as prologue to the performance appraisal itself, and via the interactivity between managers' attribution seeking behavior and subordinates' tendency towards this bias. Research on the self-serving bias shows that people tend to form judgments of what is fair that are slanted towards their own self-interests (Messick & Sentis, 1979), also a potential contaminant in the appraisal process.

Studies of performance appraisal feedback (Ilgen, Peterson, Martin, & Boeschien, 1981) have shown that there are significant differences in the ways that supervisors and subordinates rate their performance, suggesting an actor / observer bias. Murphy and Cleveland (1995) point out that the actor / observer bias is readily visible in performance appraisal systems that use self-appraisal. Even in the (probably) rare case where employees are blisteringly honest in their self-appraisals, the actor / observer bias likely exerts an influence in their ratings. This sets the stage for a potentially contentious scenario in which the performance appraisal interview results in a situation where supervisors are frequently forced to give subordinates a rating below their self-evaluated ratings, which are contaminated by an actor / observer bias and a self-serving tendency.

So strong is this impulse to inflate personal ratings that Ilgen, Peterson, Martin, and Boeschen (1981, p.327) conclude based on their research that "even when the feedback was very straight-forward and presented on a scale with which employees were very familiar, employees still overestimated their own performance." This finding is consistent with other research on appraisal inflation (Borman, 1978; Ilgen, 1971) which documents the influence of self-serving biases and the actor / observer bias on performance appraisals.

Self-serving biases and the tendency to overestimate personal strengths influence both sender and receiver in a performance appraisal interview. Just as subordinates have an inherent desire to bolster their cause during a performance appraisal, so too do supervisors have a bias towards overestimating their managerial effectiveness (Larwood & Whittaker, 1977). Correspondingly, while the mum effect would predict that supervisors would be reluctant to communicate negative observations to subordinates, self-serving biases and the tendency for managers (and others) to overestimate their strengths would indicate that supervisors would

report that they communicated all of the negative messages they needed to transmit to the subordinate. Therefore, it is predicted:

H1b: The majority of supervisors will self-report that they delivered all of the negative messages that they felt obligated to deliver based on their information about the subordinate's performance.

Discounting effect.

The discounting effect refers to the idea that as the number of possible causes for a given event increases, the likelihood of a particular cause being singled out as the primary reason for the event is diminished (Kelley, 1972). Correspondingly, it is expected that as employees present possible causes for performance failure in an appraisal interview, it becomes increasingly difficult for the supervisor to render unequivocal judgment concerning the cause(s) of the performance failure, a concept confirmed by Wood and Mitchell (1981) in their research with supervisory nurses. The implication here is that it becomes more difficult for supervisors to "stick to their guns" and deliver uncontaminated negative feedback as subordinates are given the opportunity to offer multiple causes for their performance failures during the appraisal interview.

Impression management.

Central to understanding the interaction between superior and subordinate in a performance appraisal involving the negative feedback of a performance failure is impression management. There exists a binding relationship between rater and ratee, such that:

> The manager might become informationally dependent upon the subordinate during the diagnosis of a performance failure, placing the subordinate in a position to influence the attributions and disciplinary actions of the manager. In this situation the subordinate will attempt to manage the impressions formed by the supervisor (Wood & Mitchell, 1981, p.358).

In an experiment testing the effects of impression management on the performance appraisal process, Wayne (1991) found that subordinates who employed high impression management techniques received better performance ratings than employees using low impression management techniques. Wayne and Ferris (1990) also found that supervisor-focused impression management tactics in a laboratory setting were positively related to performance ratings, suggesting a positive relationship between impression management and performance evaluations.

Engage in account making.

A crucial characteristic of the performance appraisal is interactivity and mutual attribution formation, modification, and validation. As noted by Landy and Farr (1980, p.100) in their then-landmark review of performance appraisal literature: "It is clear from even a cursory examination of the rating process that all information must ultimately pass through a cognitive filter represented by the rater." Perhaps because the stakes are so high (e.g. promotions, bonuses, and possibly the job itself are on the line), and perhaps because individual performance is such a face sensitive topic, there tends to be bidirectional communication in performance appraisal where the roles of sender and receiver are interchanged with some frequency between employee and supervisor, and performance attributions are influenced by both parties throughout the course of the appraisal (Gioia & Sims, 1986).

When confronted with negative feedback during a performance appraisal, most subordinates will invariably craft a response to the charges. This type of response, called an account, has been studied by Cody and McLaughlin (1990), and Schonbach (1990), among others. McLaughlin, Cody, and Read (1992, p.xvi) argue that:

> An effective accounter needs to assess what the recipient knows and believes because this can place important constraints on the nature of the account that one can construct. In addition, an accounter must pay attention to his or her goals in constructing the account and to whether the account, as constructed, is likely to achieve those goals. Further, in constructing the account, it is useful to try to take the recipient's perspective and try to evaluate whether he or she is likely to find the account to be coherent.

Types of Accounts

Research on accounts, guided by Schonbach's (1990) taxonomy, have distilled four main account types: (1) refusals, (2) justifications, (3) concessions, and (4) excuses (Read, 1992). Although this research experiment focuses on concessions and excuses, a brief explanation of all four account types is helpful to discern the important differences between the various account making strategies.

One type of account from Schonbach's (1990) taxonomy is *refusal*. Refusals usually take two forms. In the first, the accounter denies that the accuser has the right to reproach them. In the second, the accounter claims that the issue under discussion never occurred (McLaughlin, Cody, & O'Hair, 1983). Refusals will not be studied in this experiment since the formal, positional status between supervisor and subordinate affords the right of reproach, thus making the first form of refusal invalid in the context of performance appraisal. In regards to the second point, the perception of poor performance by the supervisor is generally sufficient to legitimize

its existence. While refusals represent valid and sometimes effective account making strategies, their study is outside the scope of this experiment. Here, the concern is with instances of performance failure that both the superior and the subordinate generally agree occurred, although the degrees of acceptance can vary.

In a *justification* account, the accounter admits responsibility for the charge, but then tries to show why the there should be no reprimand or reproach for the transgression. In a performance appraisal, an employee using a justification strategy might admit that their performance during the last rating period was substandard, but they might argue that the organization did not suffer from their low performance since overall industry activity during the last rating period was down sharply across the board. They might continue justifying by claiming that they were simply pausing before the next wave of activity. It is also common in justification strategies for individuals to minimize the severity of the offense by attempting to reframe the failure event in a less serious light (McLaughlin et al., 1983). Justification strategies will not be studied in this experiment because their use generally requires a thorough understanding of a much larger organizational context than is feasible in a laboratory experiment. Asking subjects to evaluate metaorganizational trends and issues in a hypothetical performance appraisal context makes the interaction between subordinate and supervisor speculative and potentially somewhat contrived.

A *concession* account occurs when the accounter admits to the charge that is levied. For example, an employee (the accounter) who is given negative feedback about a period of poor performance during an appraisal interview might make a concession by agreeing with the supervisor that their performance during the last rating period was, indeed, below par. Additionally, concessions are often accompanied by apologies, offers of restitution, and remorse

(Read, 1992). Accordingly, a subordinate admitting to performance failure might promise that they will redouble their efforts during the next appraisal period.

Excuses represent a fourth account making strategy. When using an excuse, people attempt to deny and otherwise minimize culpability and thus avoid blame and punishment (Read, 1992). During a performance appraisal, an employee might offer an excuse for poor performance by claiming that they were unable to concentrate at work during the last evaluation period because they were in the middle of a crisis in their in their personal life; perhaps citing the loss of a parent or another loved one.

Classifying excuses via attribution theory.

In empirical research, excuses have received considerable attention as an account making strategy (e.g. Bies & Sitkin, 1992; Snyder & Higgins, 1988; Weiner, 1992). As noted by Weiner (1992, p.146): "Excuse-giving is a powerful interpersonal tool, effective in its goals of maintaining bonding." Excuses can be more precisely classified on the basis of perceived causality: locus, controllability, and stability (Weiner, 1992). This classification schema is grounded in the ideas of attribution theory, and the most common structural type of excuses have external, uncontrollable, unstable causes.

Based on the idea that people have a fundamental desire to explain the events around them, attribution theory deals with how individuals infer the causes of behavior in their environment (Heider, 1958). After people witness a particular behavior or event, according to attribution theory, they will form explanations about why the behavior occurred through a cognitive process that assigns causes to events—called attribution (McElroy, 1982). Within the context of account making, excuses can be classified based on their causal properties (Coombs,

1995; Weiner, 1992). For clarity, the context of performance failure will be used to provide an example of each structural characteristic of excuses.

First, an excuse can have either an internal of an external locus of control. External excuses project blame and responsibility for the performance in question on events and circumstances other than the actor, such as luck, task difficulty, or a "conspiracy" of fate. Internal excuses focus on internal causes such as intelligence, effort, ability, and character. For example, a subordinate making an external excuse might claim that a performance failure was due to an impossible task. The same subordinate making an internal excuse would claim an organic dearth of intelligence caused the failure (and arguably would illustrate the veracity of the statement simply by rendering it).

Second, excuses can be classified based on controllability. A controllable excuse indicates that the performance failure in the situation under consideration can be directly influenced by the accounter. An example of controllable causality would be claiming that not enough personal effort was expended during the rating period. An uncontrollable causality excuse would highlight that the performance failure was caused by something that the individual could not readily control, such as the relatively stable characteristics of a person's general appearance. Another way of conceptualizing the controllability dimension is based on intentionality, in that an intentional behavior or action is more controllable than an unintentional one (Coombs, 1995). It follows that an individual can control the amount of effort they expend, but that they cannot control their general appearance.

Finally, an excuse can be classified based on stability. Stable excuses are those that indicate that the behavior under consideration is likely to recur while unstable excuses highlight the unlikely nature of a repeat occurrence. A stable excuse for performance failure might be a

subordinate's inability to master fundamental professional tasks, indicating that the chance for future failure remains high. On the other hand, an unstable excuse, like a short-term illness, is likely to indicate that future failure is not likely in the offing.

Diffusion of responsibility.

One previously mentioned quandary of the performance appraisal process is that supervisors are often reliant on subordinates to help them ascertain the causality of the subordinate's performance failures. Since subordinates are unlikely to be overly harsh on themselves because of self-serving biases and an inherent desire for self-preservation, supervisors seeking information from subordinates about the subordinate's performance failures are frequently given information that deflects accountability and responsibility away from the subordinates. This trend is based on the previously discussed discounting effect (Enzle, Hansen, & Lowe, 1975; Kelley, 1972; Wood & Mitchell, 1981). Kelley (1972) postulated that a discounting effect would be the greatest when there is a shift in the loci of control (from external to internal or vice versa), and Wood and Mitchell (1981) designed an experiment to compare the effectiveness of excuses versus concessions (concessions were called apologies in their research).

In two experiments, Wood and Mitchell (1981) found that when subordinates gave excuses as compared to concessions as accounts for performance failure, supervisors: (a) attributed less personal responsibility for the failure, (b) were less personal in their responses to the subordinates, and (c) were more lenient in their responses to subordinates. Wood and Mitchell (1981) also found that although not as effective as excuses in diffusing responsibility and punishment, rendering an apology was more effective than conditions where no apology was given. These findings offer a good reason why Weiner (1992) argues that external,

uncontrollable, and unstable excuses are the most frequently employed in account making: they are very effective in diverting accountability and moderating the consequences of failure events. The tendency to offer excuses as accounts is also likely partially due to the idea that people tend to attribute their successes to internal events, but their failures to external causes (Zuckerman, 1979).

Excuses have been shown to be effective in diffusing attributions when used by accounters to explain performance failure, as have concession accounts. And, the most common types of excuses, potentially because they are so effective at shifting culpability, are those which have external, uncontrollable, and unstable causality (Weiner, 1992). In comparing external, uncontrollable, and unstable excuses to concession accounts, as in the study by Wood and Mitchell (1981), it emerges that concession accounts are frequently accompanied by internal, controllable, unstable excuses. For example, in explaining performance failure, a subordinate might say: *I'm sorry that my performance was substandard* (concession). *I did not put in enough effort on my recent projects* (internal excuse), *but I will redouble my efforts* (controllable causality) *and I know that I can do better in the next rating period* (unstable causality).

One reason why concession accounts are often given in conjunction with internal, controllable, and unstable excuses is that the combination helps to preserve sincerity in the account. Work by Bies and colleagues (Bies, 1987b; Bies & Shapiro, 1988; Bies & Sitkin, 1992) suggests that perceived sincerity is an important factor in legitimizing excuses. In a field study, Bies (1987a) found that both the presence of an excuse and perceived sincerity mitigated perceptions of unfairness. Research by Baron (1988) also determined that perceived sincerity mitigates the severity of responses to the account given. When excuses are combined with concession accounts, the accounter receives two benefits: (1) perceived sincerity by admitting to

and apologizing for the failure and (2) the discounting effect by having an excuse, even if internal, to diffuse attributions concerning the performance failure. For this reason, concession accounts that incorporate internal excuses are compelling to study and compare in their effectiveness to external excuses.

Concession accounts with internal excuses, however effective they may be at diffusing some responsibility, are not predicted to be as effective as external excuses in mitigating negative feedback and diffusing perceived responsibility. Accounts with external causality are predicted to provide the most efficient diffusion of negative feedback, allowing blame to shift from internal to external loci (Weiner, 1992). But, it seems clear from previous research that both concession accounts and external excuses will have some degree of effectiveness in tempering the supervisor's reaction to the subordinate's performance failure—what is in question is the magnitude of those differences.

When politeness becomes problematic.

Additionally, it is expected that there will be a certain amount of positive messages that are transmitted in even the most negatively valenced performance appraisals due to the underlying face concerns that are inherent in any interpersonal interaction, but are heightened in a sensitive situation like a performance appraisal. Theoretically, both politeness theory (Brown & Levinson, 1987) and the mum effect (Tesser & Rosen, 1975) would support the notion that positive messages would be transmitted in even the bleakest of appraisals; politeness theory citing face concerns and the mum effect predicting that supervisors would look to all available communication strategies to avoid giving or dilute the negative feedback.

While a certain amount of positive messages are almost certainly invariable in any performance appraisal, supervisors must walk a fine line to ensure that the positive statements

they might communicate out of face concerns are not misconstrued by their subordinates as positively valenced characterizations of the subordinate's performance when the performance is, in fact, critically substandard. And, this managerial balancing act must be done in an environment where self-serving biases and the overestimation of individual strengths tend to focus the subordinate disproportionately on the positive as opposed to the negative. In the case of performance failure, these biases can result in distorted characterizations of the subordinate's performance.

Further complicating this delicate balance is the fact that there is no shortage of conventional wisdom encouraging the liberal use of positive statements in the process of giving negative feedback. Adages like *build them up before tearing them down*, *two goods for every bad*, and the popular sandwich method of delivering criticism (positive-negative-positive) all seem to indicate that positive feedback should be transmitted with greater frequency than negative feedback, *presumably in situations where the very purpose of the interaction is to deliver negative feedback*. This notion is absurd and foolhardy at worst and confusing to the receiver at the very least.

The critical question is at what point these positive statements accumulate during an appraisal interview and come to positively distort the overall tone of the appraisal. The answer is probably multivariate in nature given that a number of factors like the subordinate's perception of distortion, the relationship between the supervisor and the subordinate, the appraisal interview itself, prior performance information, and other appraisals all influence the characterization of positive distortion. Regardless of the challenge, it is both prudent and important to the larger understanding of negative feedback transmission to consider the issue of positive distortion and examine it empirically.

While the construct of positive distortion is explained in greater detail in the *Hypotheses Testing* section of the Methodology chapter on page 48, a brief explanation is necessary at this point in order to develop testable distortion hypotheses. In seeking to determine the level at which positive messages come to positively distort a communicative interaction, we can choose to divide all communication in an appraisal into three categories: positive, negative, and neutral (a detailed articulation of these categories is in the Methodology chapter). Let us further assume that in a "normal" performance appraisal, each type of message will be communicated in equal amounts. This seems reasonable since a supervisor might approach this hypothetically normal appraisal with an equal number of both positive and negative issues to discuss, as well as messages that are neutral. This would lead to a balance of 33.3% positive, 33.3% negative, and 33.3% neutral messages during the appraisal interview.

Furthermore, let us speculate that in cases where the subordinate's performance is unusually good, we might expect a supervisor to communicate twice as many positive observations as in the balanced example, with the remainder of the statements being either negative or neutral, resulting in 66.6% positive, 16.6% negative, and 16.6% neutral distribution. Correspondingly, in cases where the subordinate's performance is unusually poor, we might expect a distribution of 16.6% positive, 66.6% negative, and 16.6% neutral statements, representing twice as many negative statements as in the balanced appraisal scenario. Based on this reasoning we can establish that in cases where the subordinate's performance has been clearly characterized as a failure that any percentage of positive statements in excess of 16.6% represents positive distortion in the appraisal interview.

The presence of positive distortion, by itself, does not render an appraisal interview ineffective. However, as the positive distortion increases, the likelihood that a subordinate might

get the wrong overall impression about substandard performance likewise increases, thus confounding the fidelity of the supervisor's feedback and possibly impeding subordinate rehabilitation.

In order to explore positive distortion and the interaction between distortion and account types, the following hypotheses are presented for testing:

H2a: When compared to the total number of statements made during the performance appraisal, greater than 16.6% of the supervisor's statements will be positive in nature (positive distortion), both when the subordinate offers an external excuse for performance failure and when the subordinate offers a concession account with an external excuse to explain the performance failure.

H2b: Supervisors will communicate with significantly more positive distortion during the performance appraisal interview when the subordinate offers an external excuse as compared to when the subordinate offers a concession account with an internal excuse.

Performance Evaluation

Research also suggests that in a similar fashion as accounts influence attributions, which in turn moderates the amount of negative feedback given, accounts also impact the evaluations and ratings that supervisors give to subordinates. Within the framework of attribution theory, Kelley (1972) posited the covariation principle, which states that supervisors will try to assign responsibility for performance to the one distinct cause with which the performance seems to covary. According to the principle, when a subordinate's performance is consistent over time (consistency), other tasks are performed to a similar standard (distinctiveness), and other employees typically perform differently (consensus), the supervisor is likely to attribute performance responsibility internally to the subordinate. However, if one or more of these

causes is not met, the supervisor is likely to attribute responsibility for the performance externally.

In crafting a model of the performance appraisal process, Larson (1984) uses both Kelley's idea of covariation (1967; 1972) and a second attribution concept: the differentiation between attributions of effort and attributions of ability.

Drawing from earlier research (Weiner et al., 1972), Larson (1984) argues that ability is relatively stable over time, while effort is subject to much greater variation. Both attributions of effort and attributions of ability assume an internal locus of control according to attribution theory (Gioia & Sims, 1986). When considering effort or ability as explanatory variables on performance valence, those performances that are attributed to effort should be perceived as much more changeable and malleable than performance successes or failures attributed to ability. Attribution research suggests that individuals who exhibit low motivation and effort are thought to be able to work and expend effort when they choose to do so (Tedeschi, 1974). Correspondingly, there has been empirical work showing that performance appraisal evaluations are influenced more by a subordinate's perceived effort than their perceived ability (e.g. Knowlton & Mitchell, 1980; Weiner & Kukla, 1970). For this reason, supervisors are expected to be more stringent in their appraisal ratings of those employees with perceived low effort as compared to those with perceived low ability (Mitchell, Green, & Wood, 1981). Dugan (1989) also showed that supervisors made more punitive decisions when poor subordinate performance was attributed to lack of effort. Similarly, Mitchell and Wood (1980) determined that internal attributions led to more punitive managerial responses to subordinates. In an experiment where students became the managers of a worker who displayed performance failure as a result of either low performance or low ability, Tjosvold (1985) found that managers wanted to work

again with the low ability subordinate, but that they used threats and came to dislike the low effort employee.

Kim and Miller (1990) found that whether a subordinate was perceived to have low effort or low ability made no discernable impact on the selection of negative feedback messages. But, this finding might have been different if the feedback was to an actual person as opposed to respondents indicting what they would hypothetically say on a survey, as was the case in their experiment.

As earlier noted, exploring the interactivity between supervisor and subordinate in the performance appraisal interview is essential to gaining a better understanding of the performance appraisal process. As Gioia and Sims (1985; 1986) posit, there is an attributional shift toward leniency that occurs in the performance appraisal interview itself, suggesting that simply the face to face interaction that occurs during a performance appraisal moderates supervisor responses to performance failure in the favor of the employee.

Based on the ideas and findings above, the following relationships between accounts and performance ratings are predicted:

H3: Supervisors will evaluate a subordinate who makes an external excuse for poor performance higher than a subordinate who makes a concession account with an internal excuse after the appraisal interview.

H4a: Supervisors will evaluate a subordinate who makes an external excuse for poor performance during their appraisal interview higher after the interview than before the interview.

Subordinates who employ concession accounts with internal excuses are likely to experience offsetting reactions from superiors. As suggested by Gioia and Sims (1985; 1986),

leniency occurs as a result of the face to face interaction inherent in the appraisal interview. However, this leniency is likely to be countered by the negative characterizations of the performance failure that occurs when subordinates offer internal excuses in their failure accounts (Dugan, 1989; Mitchell et al., 1981; Mitchell & Wood, 1980). While the combination of a concession account with an internal excuse might be sufficient to engender a slight amount of leniency, the overall leniency inherent in the supervisor's attributions for the performance failure would likely be seriously muted by the internal account of the failure, resulting in insignificant differences in evaluations before and after the appraisal interview. Therefore, it is predicted:

H4b: There will be no significant difference in supervisor's evaluations of a subordinate before and after the performance appraisal when the subordinate offers a concession account with an internal excuse.

Based on the predictions of Hypotheses H3 – H4b, it follows that:

H5a: There will be a significant interaction between preappraisal and postappraisal subordinate evaluations and the types of accounts given by the subordinate to explain performance failure.

H5b: The order of the experimental conditions that the supervisors experience will have no significant effect on the subordinate's evaluation scores.

Other Factors Influencing the Transmission of Negative Feedback

Primacy / recency.

The primacy / recency effect refers to the tendency of people to place an overemphasis on events that occur either first or last in a sequence. Steiner and Rain (1989) found no recency effect on supervisory evaluations of poor performance, but in a second study found a recency effect in that a poor performance viewed last significantly degraded overall evaluations. Steiner

and Rain (1989) also documented an assimilation effect of single inconsistent performances. Kravitz (1992) points out a number of methodological issues in some contrast effect and assimilation studies and proposes a methodology that Steiner and Rain (1989) do indeed use but other key studies do not (Murphy, Balzer, Lockhart, & Eisenman, 1985). In short, findings on primacy / recency are mixed and many studies suffer from methodological flaws.

This current study controls for the potential of contrast effects and primacy / recency by stripping away relational history and by not providing dates that specific performance incidents occurred.

Vagueness.

Eisenberg (1984) points out the potential for individuals to craft purposefully vague messages in order to facilitate relational development. Extending Eisenberg's argument, raters who are purposefully vague during the performance appraisal interview can leave "wiggle room" for subordinates to cast the appraisal in the best possible light. This decision allows for the maintenance of the superior-subordinate relationship and permits the rater to meet the letter of organizational requirements without forcing the candor and honesty that would likely damage the interpersonal relationship between the superior and the subordinate.

The pitfalls of this so-called strategic ambiguity are significant in performance appraisal. First, vagueness deteriorates the notion of trust and candor between supervisor and subordinate, arguably a key element in creating and sustaining a professional relationship. Second, Ilgen, Fisher, and Taylor (1979, p.362) point out that "in the absence of explicit information about performance, the individual often is left to infer what is the desirable behavior," resulting in potentially more confusion about task and other roles after the appraisal than before it.

Rater / ratee affect.

It would be impossible to accurately rate a subordinate without spending some time getting to know them, both on a professional, task-oriented plane and also as individuals. The good news is that superiors and subordinates do seem to spend a great deal of their time interacting with each other. In his review of literature on superior-subordinate communication, Jablin (1979) found that leaders spend one-third to two-thirds of their time interacting with subordinates, with most interaction focusing on task related as opposed to personal issues. Indeed, there is reason to believe that as a relationship develops between supervisors and subordinates, lines of communication will open, better quality and a higher quantity of information will be transmitted, and supervisors will become more prepared to make accurate appraisal evaluations (Beer, 1981). But this idea is fraught with problems. See Zorn (1995) for an excellent discussion of some of the issues and implications of bosses as friends.

Raters have a difficult time accurately evaluating a subordinate that they feel strongly about (Bernardin & Beatty, 1984). When supervisors have a strong affect towards subordinates, a halo effect can inflate and skew ratings. Larson (1984) argues that managers who have a positive affect towards subordinates are reluctant to give negative feedback concerning substandard performance and Decotiis and Petit (1978) point out that close interpersonal relationships between bosses and subordinates have the ability to significantly distort appraisal ratings. Research by Tsui and Barry (1986), among others, demonstrates that raters are more lenient on raters with whom they have a positive affect. Cardy and Dobbins (1986) show that affinity influences ratings bidirectionally, finding that students rate teachers that they like higher in performance ratings. Larson (1984) posits that supervisors are less likely to give negative performance feedback to subordinates that they like and more likely to give those same

subordinates positive feedback because of two reasons. First, the supervisors are concerned about the toll that negative feedback might take in their relationship with the subordinate.

Second, Larson points to empirical research (i.e. Gochman & Smith, 1979; Regan & Totten, 1975) suggesting that a positive affect towards a subordinate biases the attribution process in that the liked subordinate will be seen as personally responsible for success and not accountable for failures (a permutation of the actor / observer bias). Schoorman (1988) also found that supervisors who had a hand in the hiring process of subordinates tended to bias performance appraisal ratings in favor of the subordinate in cases where they approved of the hire and were biased against the subordinate in cases where they opposed the hire.

On the other side, negative affect can lead to erroneous attributions and faulty, excessive judgments against the ratee (Kingstrom & Mainstone, 1985; Tsui & Barry, 1986). Further confounding the causality between affect and evaluation is performance. It is possible that performance is a mediating variable such that good performance leads to higher affect which increases appraisal ratings (DeNisi, 1996). As Green and Mitchell (1979, p. 440) summarize:

> Essentially any factor which makes the leader psychologically closer to the member should increase the tendency for the leader to make self-like attributions regarding the member (Banks, 1976). On the other hand, anything that increases the psychological distance between leader and member will reduce this tendency and increase the likelihood of leader attributions being discrepant from the member's causal explanations of his or her own behavior.

One of the great strengths inherent in the methodology of the present study is that it controls for familiarity by stripping away relational history to focus more narrowly on the communicative transmission of negative feedback. In so doing, a more precise exploration of

negative feedback in performance appraisal can be undertaken, focusing on the communicative dynamics and principles while controlling for relational history.

The Influence of Attributions on Subordinate Satisfaction

The role of attributions are central to a supervisor's decision of how to communicate negative feedback, as discussed previously in the subsection *Classifying excuses via attribution theory* on page 17, as well as being influential in determining the effectiveness of the feedback that supervisors give during the appraisal interview.

Kelley (1967) offered a classification system for causal attributions based on three dimensions most frequently used: (1) consistency, (2) distinctiveness, and (3) consensus. This classification scheme can also be used by supervisors to categorize subordinate performance, as explained in the preceding section, *Performance Evaluation*, on page 24. In other words, the supervisor can classify performance based on this system, and the subordinate can classify the type of feedback received during the appraisal interview in the same manner.

Within the schema, consistency refers to how the behavior in question corresponds to previous instances of the same behavior. Distinctiveness looks at performance on other tasks, and consensus compares the individual's specific performance with the performance of others.

Focusing on how the classification system can be used to describe performance feedback, a supervisor providing negative feedback during an appraisal might give feedback grounded in a *consistency* attribution by pointing out other performance failures on this particular task within the last rating period. Feedback based on *distinctiveness* might point out that the subordinate has failed on a number of other professional duties during in the last rating period. And, *consensus* feedback might tell the subordinate that their performance on the task was significantly worse than the performance of other workers on the same task (Liden & Mitchell, 1985).

The manner in which a supervisor decides to administer feedback is of obvious concern to subordinates. In an experiment that varied the amount of consistency, distinctiveness, and consensus of feedback given to students from teachers, Linden and Mitchell (1985) found that individuals rated specific feedback, which included information relating to the student's past performance (consistency) and the performance of their peers (consensus) higher than nonspecific feedback. More generally, they found that people also rated feedback that incorporated elements of consistency, distinctiveness, and consensus more favorably than feedback with no information on these three dimensions.

This finding is consistent with other research shows that individuals prefer specific to nonspecific feedback (Ilgen, Mitchell, & Fredrickson, 1981; Ilgen, Peterson et al., 1981). Ilgen, Mitchell and Fredrickson (1981) also found that subordinates rated the expertise of their supervisor and the quality of supervision higher to those supervisors who gave specific feedback. Subordinate satisfaction was also determined to be correlated significantly with specificity, timeliness, consideration, and frequency (Ilgen, Peterson et al., 1981). One important caveat on the preference for specific feedback is from Liden and Mitchell (1985) who found that people *did not* prefer specific feedback when the feedback implied that the cause for the feedback was internal. In other words, people seem to prefer specific feedback so long as it does not implicate internal causes for the behavior or performance under scrutiny. Based on these findings, it is predicted that:

H6a: The subordinate will be significantly more satisfied with supervisors who use feedback that incorporates elements of consistency, consensus, and / or distinctiveness (specific feedback) than supervisors who deliver feedback without consistency, consensus, and distinctiveness (nonspecific feedback), when the reason given for performance failure is an external excuse.

H6b: When the reason given for performance failure is a concession account with an internal excuse, the subordinate will be significantly more satisfied with superiors who use feedback that does not incorporate elements of consistency, consensus, and / or distinctiveness (nonspecific feedback) than supervisors who deliver feedback with consistency, consensus, and distinctiveness (specific feedback).

Contrarian Views Concerning the Functionality of Performance Appraisal

There are more critics of performance appraisal than simply those employees who dread their existence and those supervisors who disdain the procedural requirements and potential for confrontation. It might be expected that any system that seeks to operationalize, and frequently incentivize individual performance would not be without its detractors. Scholars such as Eisenberg (1984; Eisenberg & Witten, 1987) argue against the functionality of openness and transparent communication in organizations, which would deal a serious blow to the legitimacy of performance appraisals.

Eisenberg is not alone. Most texts on performance appraisal (e.g. DeNisi, 1996; Murphy & Cleveland, 1995) devote sections to the problematic nature of performance appraisal. At the center of the controversy is the argument that it is extremely difficult, if not impossible, to both evaluate and coach in one interaction (Meyer et al., 1965; Murphy & Cleveland, 1995). Another concern is that there is an incongruence between the performance appraisal interview and its use

in organizations. This incongruence might manifest itself in a company that claims to value the performance appraisal, only to limit its functionality and ability to be used in promotion and salary decisions. Another criticism is with managers who "game" the performance appraisal system. Fried, Tiegs, and Bellamy (1992) note research that indicates that supervisors often inflate, and sometimes even deflate performance ratings for their own political gain. All of these issues make it problematic for subordinates to clearly see the linkage between their behavior and organizational rewards, the mum effect and other communicative issues notwithstanding.

Hobson, Mendel, and Gibson (1981, p.166) argue that "for the appraisal process to function effectively, it is imperative that workers be presented with explicit information linking desired behaviors and organizational rewards." Failure to clearly demonstrate the connections between performance and reward is a violation of the instrumentality principle of expectancy value theory (Porter & Lawler, 1968; Vroom, 1964) and, according to expectancy value, will result in decreased motivation in employees. The linkage from theory to practice is transparent. Some field research also indicates that subordinates who are rated as "satisfactory" in their performance appraisals experience a significant and stable drop in their organizational commitment—to say nothing of those employees rated below the standard (Pearce & Porter, 1986).

With so many pitfalls, why conduct a study on performance appraisal in the first place? Simply put, because performance appraisals are an indispensable tool for leaders, followers, and organizations. And, performance appraisals are unlikely to be replaced or eliminated for the foreseeable future. In addition to providing a formalized vehicle for performance feedback, appraisals also serve the legal function of documenting and quantifying hires, fires, and promotions. Stated differently, because performance appraisals are problematic is precisely why

they are such a compelling topic for scholarship. And, it is because of criticism and scrutiny that performance appraisal systems continue to be refined and improved, benefiting employees, employers, and supervisors alike. As Oberg (1972, p.67) notes:

> Formal systems for appraising performance are neither worthless nor evil, as some critics have implied. Nor are they panaceas, as many managers might wish. A formal appraisal system is, at the very least, a commendable attempt to make visible, and hence improvable, a set of essential organizational activities.

Summary of Hypotheses

H1a: Supervisors will communicate positive subordinate observations significantly more than negative subordinate observations, regardless of the type of excuse for poor performance offered by the subordinate.

H1b: The majority of supervisors will self-report that they delivered all of the negative messages that they felt obligated to deliver based on their information about the subordinate's performance.

H2a: When compared to the total number of statements made during the performance appraisal, greater than 16.6% of the supervisor's statements will be positive in nature (positive distortion), both when the subordinate offers an external excuse for performance failure and when the subordinate offers a concession account with an external excuse to explain the performance failure.

H2b: Supervisors will communicate with significantly more positive distortion during the performance appraisal interview when the subordinate offers an external excuse as compared to when the subordinate offers a concession account with an internal excuse.

H3: Supervisors will evaluate a subordinate who makes an external excuse for poor performance higher than a subordinate who makes a concession account with an internal excuse after the appraisal interview.

H4a: Supervisors will evaluate a subordinate who makes an external excuse for poor performance during their appraisal interview higher after the interview than before the interview.

H4b: There will be no significant difference in supervisor's evaluations of a subordinate before and after the performance appraisal when the subordinate offers a concession account with an internal excuse.

H5a: There will be a significant interaction between preappraisal and postappraisal subordinate evaluations and the types of accounts given by the subordinate to explain performance failure.

H5b: The order of the experimental conditions that the supervisors experience will have no significant effect on the subordinate's evaluation scores.

H6a: The subordinate will be significantly more satisfied with supervisors who use feedback that incorporates elements of consistency, consensus, and / or distinctiveness (specific feedback) than supervisors who deliver feedback without consistency, consensus, and distinctiveness (nonspecific feedback), when the reason given for performance failure is an external excuse.

H6b: When the reason given for performance failure is a concession account with an internal excuse, the subordinate will be significantly more satisfied with superiors who use feedback that does not incorporate elements of consistency, consensus, and / or distinctiveness (nonspecific feedback) than supervisors who deliver feedback with consistency, consensus, and distinctiveness (specific feedback).

Chapter 3

Methodology

Before detailing the methods of the current experiment, it is helpful to clarify the manner in which the supervisors and the subordinate are identified throughout the remainder of this study. The *supervisors* in this experiment are the MBA students who voluntarily agreed to participate in the research. The terms *participants* and *subjects* are also occasionally used as synonyms for supervisors throughout this study. These terms do not include the subordinate. The *subordinate* refers to the one actor who took on the role of the subordinate named Kelley Jarvis for this experiment. The term *actor* is used occasionally as a synonym for subordinate throughout this study.

Experimental Conditions

There were two experimental conditions in this study: (1) a concession account with an internal excuse and (2) an external excuse.

In condition 1: The subordinate combined elements of a *concession* account with an *internal excuse* account to explain the performance failure. In this condition, the subordinate conceded to the supervisor's claims and offered apologies and expressions of remorse for the performance failure (consistent with Read's (1992) description of concession accounts). In this condition, the subordinate also offered an *excuse* for the poor performance that had an *internal* locus of control. Specifically, the subordinate claimed that he was indeed sorry for his poor performance during the last rating period (a concession), pointing out that the performance failure was due to a lack of effort and that more effort would be expended to improve his performance during the next rating period (internal causality).

In Condition 2: The subordinate made an *excuse* for the poor performance that was *external*. In this condition, the subordinate receiving the negative performance feedback offered the most common structural type of excuse, one with external causality (Weiner, 1992). The subordinate communicated to the supervisor that his performance failure was due to non-pervasive, external events that were beyond his control. For all interviews, the subordinate offered a specific external excuse, claiming that his father recently had a stroke (external causality) and that the uncertain recovery prognosis combined with some caretaking duties was the cause for the performance failure.

Protocol

Participants arrived at the research site and were met by either the primary researcher or a research assistant. As noted by Wayne and Kacmar (1991, p.85):

> One might question whether a study on performance appraisal can be conducted in a laboratory setting. However, hundreds of studies on performance appraisal have utilized laboratory settings. In addition to the enhanced control, a major benefit of this design is that previous subordinate performance (i.e. history) and previous supervisor-subordinate interactions do not contaminate the performance ratings.

The research site was an office in the MBA Career Center at a large southwestern university. The office was configured to support an appraisal interview, with a large desk and a chair for the supervisor and a chair for the subordinate. Since the MBA Career Center closed at 5:00 PM each day, it was necessary to conduct the late afternoon and early evening appraisal interviews at a separate location two floors above the MBA Career Center, in a self-contained interview room that could be secured independent of the larger Career Center. A total of 12 subjects out of 53 participated in the experiment in this room. Because of the use of different rooms, a variable

room was inserted as a between subjects variable in all salient hypotheses tests (similar to how experimental *sequence* was treated, as discussed later in this section). As reported in detail the Results chapter starting on page 65, room was not a significant variable in any of the experimental analyses.

Participants filled out a pre-test questionnaire (Appendix A) designed to assess their experience in conducting performance appraisals and to gather baseline demographic information. Subjects were 53 graduate MBA students at a large southwestern university who were screened during recruitment to ensure that they had little or no prior experience in delivering performance appraisals. Correspondingly, 86.80% of participants had never delivered a formal performance appraisal interview. Of the remaining 13.20% of participants, two participants each had delivered one, two, or three prior appraisals and one participant had delivered four prior appraisals.

Participants who had limited experience delivering performance appraisals were desired for this research to mitigate the effects that conducting prior appraisals might present. It is likely that subjects who have significant experience delivering performance appraisals could be significantly influenced by their former experience and their organization's manner of conducting appraisals and they may respond to the experimental conditions in a predetermined way.

MBA students were selected for this experiment because they are a population likely to be involved in performance appraisals in the future and because it is likely that they have some familiarity with the context of the appraisal interview since they are required to have previous work experience to gain entry into the MBA program. By selecting MBA students who have little or no prior experience in delivering performance appraisals, this group will have a

contextual frame for understanding the experiment, but will likely not have preconceived notions about their role in it. It should be noted that MBA students with experience in delivering performance appraisals in a military context were allowed to participate in this experiment. The norms and procedures inherent in the military appraisal system are sufficiently different from the corporate appraisal structure of this experiment to allow for their inclusion in the study without concern that they will be unduly influenced by their prior organization's approach to performance appraisal. A total of five participants in the study had prior military experience.

The experiment was designed so that 14 subjects would participate in the study each day and the experiment would be conducted for four consecutive days, yielding a total capacity of 56 participants. Prior appointments were made for designated time slots in the experiment. Three subjects did not make their appointments, resulting in 53 total subjects in the experiment.

Each participant received both experimental conditions. In order to ensure a similar number of participants received the experimental conditions in the same order, the following sequencing plan was built into the appointment schedule. The two subjects participating on the first day of the experiment were assigned to receive the concession account with internal excuse condition followed by the external excuse condition. The next four subjects were assigned to receive first the external excuse and then the concession account with internal excuse. The next four subjects were assigned to receive first the concession account with internal excuse and then the external excuse condition. The final four subjects on the first day of the experiment were assigned first the external excuse and then the concession account with internal excuse.

On the second day of the experiment, the first two subjects were assigned the external excuse condition followed by the concession account with internal excuse. Then, in groups of four, the sequence of the conditions reversed. Correspondingly, the experimental sequencing on

the third day of data collection was designed to mirror the first day and the fourth day was designed to mirror the second day. This sequencing pattern helped to ensure that a comparable number of subjects received the external condition before the concession account with internal excuse as received the reverse order. For the experiment, 25 participants received the concession account with external excuse condition first and 28 participants received the external excuse condition first. The sequence numbers are not perfectly matched due to the fact that three participants dropped out of the experiment at different times. When this happened, a gap was left in the schedule and the integrity of the preplanned sequence was adhered to. It was important to continue with the preplanned sequence because the pattern helped to safeguard against actor fatigue and confusion by allowing the actor to deliver four similar conditions (types of excuses) in a row after the first two subjects were processed each day. Additionally, the effect of sequence on specific hypothesis was evaluated and is reported in the Results chapter.

After an initial briefing by the primary researcher, subjects read and completed an Institutional Review Board approved human subjects consent form (Appendix B). Subjects then read instructions for conducting the interviews (Appendix C). Subjects were informed that they were participating in a study exploring the performance appraisal interview but they were not told that the overall purpose of the study concerned the delivery of negative feedback in performance appraisal. Subjects were queried to ensure that they understood their role in the experiment after reading the instructions and subjects needing more clarification about their roles in the experiment were given additional instruction.

For each experimental condition, subjects were given a performance evaluation synopsis for the employee (Appendix D), who was an actor trained by the researcher to communicate either an external excuse or a concession account with an internal excuse during the performance

appraisal interview. In all cases, it was clearly evident from the contextual information that the employee's performance was poor, unacceptably below average, and had degraded significantly since the last annual performance appraisal.

Validating employee observations.

In addition to contextual information and instructions, participants were given six behaviorally based observations that were designed to represent "field notes" that a supervisor might make throughout the appraisal period in preparation for the performance appraisal (Appendix E). These field notes will be referred to as *observations* or *note card observations* throughout this study. Three observations were positive in nature and three observations were negative. Because it was important to ensure that supervisors viewed the six behavioral observations about the subordinate as (1) similar in intensity levels and (2) distinct, independent observations from each other, the observations were evaluated in two ways: (1) in exit interviews with participants and (2) in a separate questionnaire study conducted before data collection began.

First, the intensity level and distinctiveness of the observations were validated during the experiment through an exit survey given to subjects at the conclusion of their role in the experiment (Appendix F). For the measure of intensity, 98.10% of the subjects in the experiment agreed with the question on the exit survey assessing intensity: *Do you feel that each observation that you were given on note cards had a similar level of intensity to the other observations?* For distinctiveness, 98.10% of subjects agreed with the statement from the exit survey assessing distinctiveness: *Do you feel that each observation you were given on note cards represented a distinct, different observation about Kelley's behavior and performance?* These two high levels of agreement indicate that the supervisors in the experiment viewed the behaviorally based

observations of the subordinate that were presented on note cards as (1) distinct and independent of each other and (2) similar in levels of intensity.

Additionally, the intensity and distinctiveness of the observations were checked through a survey of the observations given to a pilot sample of subjects (N = 19) before the data collection phase of the experiment began (Appendices G and H). These subjects were not affiliated with the appraisal experiment research in any way. Subjects were asked to rate, on a 1 to 5 scale, the intensity of the subordinate observations (where 1 = trivial and not very intense; 3 = noteworthy and medium intensity; 5 = extremely and severely intense). Table 1 reports the means and standard deviations for the individual intensity scores.

Table 1

Means and Standard Deviations for Observation Intensity Levels

	Neg Obs 1	Neg Obs 2	Neg Obs 3	Pos Obs 1	Pos Obs 2	Pos Obs 3
Mean	3.11	3.63	3.32	3.16	3.47	2.84
SD	.99	.60	.82	.69	.77	.83

The average intensity score for the three negative observations was 3.35 (*SD* = .60) and the average intensity score for the three positive observations was 3.16 (*SD* = .60). The overall intensity score average for all six observations was 3.25 (*SD* = .57). Finally, reliability analysis revealed that the six items had consistent intensity (Cronbach's $\alpha = .82$). This reliability measure, taken in conjunction with the narrow spread in intensity score averages, strongly suggests that the students piloting the six observations evaluated the six observations as similar in intensity.

On the distinctiveness measure, 94.70% of the survey respondents reported that the three negative observations were distinct from each other, 94.70% of the survey respondents reported that the three positive observations were distinct from each other, and 89.50% reported that all

six observations were distinct from each other. With a score of 1 indicating that the observations were distinct from each other and a score of 2 indicating that they were not, the mean scores for the three positive observations was 1.05 (*SD* = .23), for the three negative observations was 1.05 (*SD* = .23), and for all six observations was 1.11 (*SD* = .32). These results strongly suggest that the students piloting the observations viewed each observation as distinct from each other.

When the results from the student pilot are considered in tandem with the results of the experiment exit interview, it is reasonable to conclude that the six behaviorally based observations were considered to be (1) similar to each other in intensity and (2) describing distinct, independent observations about the subordinate.

Subjects were also asked to evaluate the subordinate using a multi-item instrument (Appendix I) prior to conducting the performance appraisal. Subjects were given time to prepare for the interview and, once they indicated that they were ready, the interview began. All interviews were captured on video tape. Wiemann (1981) has shown that videotaping has no effect on conversational behavior. The interviews were also audio taped to ensure that there was a redundant system in place for data collection process.

At the conclusion of the appraisal interview, the confederate and the subject separated and both filled out a postinterview survey (Appendices J and K for the supervisor and Appendix L for the subordinate). Then, the participants were given an opportunity to prepare for the second experimental condition and the process repeated itself with interview preparation, preappraisal rating, appraisal interview, and postinterview survey. To reduce unintended variation, each experimental condition occurred within the same organizational context and all work histories were identical (i.e. the actor was portraying the same employee in the same hypothetical organization and the subject was the "same" supervisor for every interview; all

performance observations were identical). Subjects were asked to treat each interview as an independent event, with one appraisal interview having no bearing on the other.

The employee actor was given training before conducting appraisal interviews, including participation in two pilot studies. The actor was given clear instructions for his interactions with supervisors, as well as being given specific excuses to deliver. As mentioned previously, excuses were identical within each of the two experimental conditions (e.g. the external excuse condition blamed his father's recent stroke for the performance failure; the concession account with internal excuse condition gave an apology, cited an underestimation of the effort required, and promised to improve). Appendix M details the guidance to the actor and also describes the excuses in greater detail.

Validating actor behavior.

The same actor was used in both conditions throughout the experiment to reduce the unintended variation that using multiple actors would cause. Even though only one actor was used in the experiment to reduce unintended variation, it is still prudent to assess the consistency of the actor's behavior over the course of the experiment. The actor's reliability was tested in the following manner: A random sample of six different interview video tapes was selected from all interviews conducted. This sample returned the interviews of subjects M08, M16, M26, M36, M39, and M50. This sample included at least one interview from each of the four days of data collection and no more than two interviews from any one day. Each tape was queued to a random position in the appraisal interview.

Seven survey participants who had never met the actor completed a survey designed to establish the actor's consistency throughout the appraisal interviews (Appendix N). The participants watched the same six interview segments described above, with each segment

lasting for approximately 90 seconds. After viewing each interview segment, the participants answered five questions about the actor's communication, behavior, and appearance: *(1) What is your impression about the subordinate's formality of dress (clothing) (2) What is your impression about the subordinate's level of fatigue (3) What is your impression about the subordinate's level of attentiveness (4) What is your impression about the subordinate's level of dynamism (5) What is your impression about the subordinate's level of composure.* Reliability scores (Cronbach's α) for these five questions are reported in Table 2.

Table 2

Reliability Analysis for Actor's Performance Across Interviews

	Question 1	Question 2	Question 3	Question 4	Question 5
Reliability (Cronbach's α)	.96	.69	.93	.86	.90

All of the questions with the exception of question two, which asks for the participant's impression about the subordinate's level of fatigue (where 1 = not fatigued and 5 = very fatigued), have reliability levels of .86 or higher. Given that the generally accepted minimum standard for reliability is .70 (Bobko, 2001; Nunnaly, 1978), question two falls just short at the level of .69. To explore question two further, the means and standard deviations for question two from each interview segment are reported in Table 3.

Table 3

Means for Survey Question Assessing Actor Fatigue

	Video Segment 1	Video Segment 2	Video Segment 3	Video Segment 4	Video Segment 5	Video Segment 6
Mean	2.86	2.43	2.29	2.43	2.57	2.57
SD	.69	.53	.49	.53	.79	.53

The overall mean for all responses to question two (across all six segments) is 2.52 (*SD* = .20). Given that there is a relatively small spread between the average scores for question two, it is reasonable to conclude that the survey participants generally viewed the actor as similarly fatigued in each interview segment.

Additionally, a reason that might contribute to understanding why the reliability score of .69 is slightly below the desired benchmark of .70 could be that it was more difficult for the survey participants to evaluate fatigue than the other items that they were asked to evaluate. For example, the item with the highest reliability of .96 queried about the actor's formality of dress. This is a relatively straightforward item to answer, especially given that the actor wore the same type of clothing during each day of the performance appraisal interviews. It is most likely easier for people watching a video of the appraisal to answer questions about things that they can easily *see* in the video, such as clothing, attentiveness, dynamism, and composure than for the same people to evaluate things that are harder to visually detect, such as fatigue.

Additionally, the overall reliability of all the individual items that evaluate the actor's consistency is .92. When all of the above factors are taken into account, the evidence strongly supports the actor's reliability across the appraisal interviews.

At the conclusion of the first experimental protocol, a second appraisal interview was conducted, the subject was out briefed, completed a short exit survey (Appendix F as described earlier), and was dismissed.

Hypotheses Testing

H1a: Supervisors will communicate positive subordinate observations significantly more than negative subordinate observations, regardless of the type of excuse for poor performance offered by the subordinate.

H1b: The majority of supervisors will self-report that they delivered all of the negative messages that they felt obligated to deliver based on their information about the subordinate's performance.

In order to test hypotheses H1a, a content analysis was conducted of all appraisal interviews. One coder was selected to code the interviews according to a content analysis form (Appendix O) in order to control for variation among multiple coders. To ensure that the coder had sufficient training, the primary researcher randomly selected 10% of the interviews and both the primary researcher and the coder completed the content analysis form for those interviews.

Once the primary researcher and the coder had independently completed the content analysis form for the selected interviews, both parties met to assess agreement on the coding. Where there were areas of nonagreement, the audio tapes were played and the coder and the primary researcher discussed the nonagreement and revised the characterization of certain messages if necessary and if possible.

Two of the questions on the content analysis form were dichotomous in nature, questions one and three. Question one asked if the supervisor:

(A) Compared Kelley's performance to the performance of others?

(B) Indicated that the subordinate had failed on a number of other tasks?

(C) Compared Kelley's current performance with Kelley's past performance?

Question three asked if: *In the concession account with internal excuse ONLY, did the supervisor ask the subordinate if there was some other external reason (i.e. something going on in their personal life) to explain the performance failure AFTER the subordinate had already admitted responsibility for the poor performance?*

The inter-rater reliability, which was calculated using Cohen's kappa for all dichotomous questions and an intraclass correlation coefficient for nondichotomous questions, between the coder and the primary researcher in response to these two questions was $\kappa = 1.00$, $p < .01$ for both the external excuse and the concession account with internal excuse conditions.

Question two on the content analysis form asked how many times the supervisor cited the following performance observations (which were from the note card observations that supervisors were given for the performance appraisal) during the appraisal interview:

(A) Observation A (Kelley late to meeting)

(B) Observation B (Kelley argument with client)

(C) Observation C (Kelley unprepared and gave poor presentation to boss)

(D) Observation D (Kelley stayed late helping a salesperson learn about a product)

(E) Observation E (Kelley solved challenging client integration issue)

(F) Observation F (Kelley good job working with a team on a client proposal)

The intraclass correlation coefficient for each of these items was $\alpha = 1.00$, $p < .01$ for both the concession account with internal excuse and the external excuse conditions.

It is reasonable to expect perfect agreement on items such as questions one through three on the content analysis form. These items are straightforward to code and easily verifiable (by both the coder and the primary researcher) by listening to the audiotapes of the appraisal interviews. The coding methodology is explained below.

Items four through six on the content analysis form ask the coder to provide count data of negative, positive, and neutral statements from the appraisals. After meeting with the primary researcher and discussing areas of nonagreement, acceptable intraclass correlation coefficients were attained. These intraclass correlation coefficients are reported in Table 4.

Table 4

Intra-class Correlation Coefficients for Content Analysis Form

	Concession with Internal Excuse			External Excuse		
	Negative	Positive	Neutral	Negative	Positive	Neutral
α	.9963	1.0	.9863	1.0	.9820	.9852

Note. $p < .01$ in all cases.

Based on the intra-class correlation coefficients from each item on the content analysis form, there is a high degree of agreement between the primary researcher and the coder on the items comprising the content analysis form. Because of this level of agreement, the remainder of the interviews were coded by the single coder.

Hypothesis H1a was tested by using items from the content analysis form (Appendix O). Appraisal interview tapes were evaluated to determine the number of times that a subject cited each particular performance observation (given to participants from note cards in their appraisal preparation) during the performance appraisal interview. Observations could be counted more than once if the supervisor discussed the observation at one point in the appraisal, moved on to another topic, and later revisited or mentioned the previous observation. All discussions of a particular observation that were contained with one conversational turn were counted as one observation. A paired sample *t*-test was used to test Hypothesis H1a.

Hypothesis H1b was tested by determining the overall response, across both conditions, to item 11 on the supervisor feedback form (Appendix K) which asks: *Do you feel that you gave Kelley all of the negative feedback that you needed to deliver based on Kelley's performance?* Fifty percent and higher constituted a majority of supervisors.

H2a: When compared to the total number of statements made during the performance appraisal, greater than 16.6% of the supervisor's statements will be positive in nature

(positive distortion), both when the subordinate offers an external excuse for performance failure and when the subordinate offers a concession with an internal excuse to explain the performance failure.

H2b: Supervisors will communicate with significantly more positive distortion during the performance appraisal interview when the subordinate offers an external excuse as compared to when the subordinate offers a concession account with an internal excuse.

In order to test hypotheses H2a and H2b, it is necessary to provide an operationalized definition of the dependent variable positive distortion. Positive distortion refers to the propensity for communicators to counter the unpleasant effects, both for the sender and for the receiver, of giving negative feedback by either omitting negative parts of messages or by only communicating neutral message elements. Communicators can also positively distort feedback by inflating the appraisal interview with positive, supportive comments which serve to offset the negative feedback.

In this experiment, positive distortion was measured through a content analysis of the interview audio tapes. The number of positive messages, the number of negative messages, and the number of neutral messages were evaluated for each appraisal. The sum of all three message types (positive, negative, and neutral) yielded the total number of messages in the performance appraisal.

Positive messages are messages that reflect positive, optimistic, or other generally favorable comments about the employee or the employee's performance. Positive messages might include comments such as: (a) *You are a valuable member of this team* (b) *I know that you are a good employee* (c) *Your technical skills are outstanding*. Conversely, negative messages are messages that reflect negative, pessimistic, or generally unfavorable comments about the

employee or the employee's performance. Some negative messages might include comments like: (a) *This performance is unacceptable* (b) *You are really falling down on the job* (c) *I have never seen such poor work from an employee*.

Some comments during the performance appraisal will be neutral in nature. These comments will not be counted as either positive or negative but constitute their own category of neutral messages. Examples could include: (a) *Today we are going to discuss last year's performance* (b) *The company has some important projects upcoming* (c) *Please tell me your opinion of your performance*.

It is foreseeable that positive and negative comments might appear within a single sentence, particularly if the supervisor hedges the negative feedback. In this case, both a positive and a negative message will be counted. Examples include: (a) *You are a great employee, but your performance this period has been horrible* (b) *I know you can do so much better than this dismal work indicates* (c) *You are very talented and important to this team, but you simply have been letting all of us down*.

Questions that are leading (non-neutral) will be assessed as positive or negative depending on their valence. For example, *What happened to the proactive employee that I know you can be* would be classified as positive and *How do you explain this miserable performance* would be classified as negative.

Statements that have no bearing on the subordinate's performance or on the subordinate's professional career will not be counted as positive, negative, or neutral. These statements will also not count in the grand total of messages tallied during the performance appraisal. Pilot study interviews revealed that the majority of these statements occurred at the beginning or at the ending of the interviews. Examples include: *Hello Kelley, How's the family, How about those*

Astros, So long, and *See you later*. In most cases, these statements represented one or perhaps two conversational turns at the beginning or the end of the conversation. Once the supervisor communicated the first appraisal-related comment, these peripheral comments usually did not reappear until the end of the appraisal interview as the supervisor prepared to leave the interview room.

Finally, noncommunicative speech such as *ummm* and *ahhh* were not coded by themselves as positive, negative, or neutral, consistent with Stiles' (1992) taxonomy of verbal response modes. When this noncommunicative speech occurred in the middle of other messages and not as a distinct part of conversation, the speech was simply grouped with the larger message categorization unit.

Distortion is the percentage derived from the number of positive messages compared to the total number of messages in the appraisal; such that an interview with 15 positive messages and 55 total messages will have a 27.3% (15 / 55) positive distortion score. Evaluations of positive distortion are based on the percentage of positive messages delivered during the performance appraisal. Consistent with the discussion of positive distortion and the context of the performance appraisal in this experiment as explained in the Chapter 2 subsection on page 21 *When politeness becomes problematic*, Table 5 delineates the characterizations of positive distortion used in this experiment.

Table 5

Positive Distortion Characterizations

Distortion	Characterization
0.0 – 16.6%	No positive distortion
16.7 - 30.9%	Moderate positive distortion
31.0 – 44.9%	High positive distortion
45.0% and above	Extremely high positive distortion

The decision to select 16.6% as the level that indicates positive distortion is based on a consideration of experiment protocol in conjunction with the literature reviewed earlier on the reluctance to communicate negative messages and issues of face concern (Brown & Levinson, 1987) and is explained in the *When politeness becomes problematic* subsection starting on page 21 of Chapter 2. This literature indicates that there will be a certain amount of positive messages communicated out of concerns for both the sender and the receiver, even in decidedly negative circumstances such as in the current experiment which clearly states that the subordinate's performance is compellingly poor. As documented in the *When politeness becomes problematic subsection*, the level of 16.6% positive messages was established as the threshold at which an appraisal interview was characterized as being positively distorted. Any quantity 16.6% or less is considered to be small enough that the interaction cannot be characterized as positively distorted because the effects might simply be due to the frictional reluctance to communicate negative messages and face concerns. At levels above 16.6%, however, the incidence of positive

messages is prevalent enough to characterize the interaction, in this context, as positively distorted.

Hypothesis H2a was evaluated by assessing the overall distortion in both experimental conditions and analyzing the overall percentage means of positive statements for the two experimental conditions.

Hypothesis H2b was tested by comparing the positive distortion in each experimental condition. Hypothesis H2b was statistically evaluated by using a repeated measures ANOVA with the dependent variable being positive distortion and the independent variable as the type of account given by the subordinate. It should be noted that all repeated measures ANOVAs used in this experiment were run with a Bonferroni correction to conservatively adjust the observed significance level because of multiple comparisons with the dataset. Doing this provided a safeguard against the possible cumulative effect of familywise error.

H3: Supervisors will evaluate a subordinate who makes an external excuse for poor performance higher than a subordinate who makes a concession account with an internal excuse after the appraisal interview.

H4a: Supervisors will evaluate a subordinate who makes an external excuse for poor performance during their appraisal interview higher after the interview than before the interview.

H4b: There will be no significant difference in supervisor's evaluations of a subordinate before and after the performance appraisal when the subordinate offers a concession account with an internal excuse.

Based on the predictions of Hypotheses H3 – H4b, it follows that:

H5a: There will be a significant interaction between preappraisal and postappraisal subordinate evaluations and the types of accounts given by the subordinate to explain performance failure.

H5b: The order of the experimental conditions that the supervisors experience will have no significant effect on the subordinate's evaluation scores.

Evaluating the relationship between positive distortion and subordinate evaluation.

In order to properly test Hypotheses H3 - H5b, it first must be determined if there is a consistent relationship between the subordinate's evaluations and the supervisors' distortion (the latter of which was determined from analyses of H2a and H2b). Pearson's correlations were required to be analyzed between evaluation scores and distortion scores to assess the relationship between the two variables. If a correlation existed at the r = .30 level or greater (suggesting that these variables might need to be analyzed in tandem), a repeated measures MANOVA would be run to assess the significance between evaluation scores and distortion scores. Then, if the relationship was significant at the $\alpha = .05$ level, the repeated measures MANOVA would be used to test both Hypotheses H3 – H5a (which concern evaluation scores) and Hypotheses H2a – H2b (which concern distortion scores).

Pearson's correlations between distortion scores and evaluation scores did not reveal a consistent relationship between the distortion scores and subordinate evaluation scores. The correlation between postappraisal scores in the external excuse condition and positive distortion in the external excuse condition was not significant, $r = .16$, $p = .25$. The correlation between postappraisal scores in the concession account with internal excuse condition and positive distortion in the concession account with internal excuse condition was only significant if

rounded at two decimal places but was not at the minimum threshold established prior to data collection of $r = .30$ to warrant additional consideration, $r = .27$, $p = .049$.

Since the relationship between evaluation scores and distortion scores was not significant in the external excuse condition and not of a sufficient magnitude in the concession account with internal excuse condition, Hypotheses H2a - H5b were evaluated by using a repeated measure ANOVA, which allowed for the consideration of three effects: (1) differences over time, (2) differences between groups, and (3) the interaction between groups and time. Analysis of these three effects satisfies the testing requirements for Hypotheses H2a - H5b.

Scale development for subordinate evaluation measure.

In order to evaluate hypotheses H3 - H5b which deal with the supervisor's evaluation of the subordinate, a factor analysis and a reliability validation were conducted on the items (from Appendix I for the preappraisal measure and Appendix K for the postappraisal measure) that were initially selected to constitute the composite score of the supervisor's satisfaction with the subordinate. A factor analysis revealed that all of the initial items selected to comprise the composite score of supervisor satisfaction (1-6, 8, and 10 from Appendix I) did not load on one factor. Specifically, item number one, *I feel that Kelley is a valuable employee,* and item number ten, *Kelley is a model employee,* did not group with the other items from the supervisor questionnaire. This was not entirely surprising since during the data collection process, a few subjects sought clarification on the frame of reference with which they were to compare Kelley, indicating that they might find it problematic to answer item ten about Kelley being a model employee. In other words, these participants indicated that they did not have enough contextual information about the employee or the company to rate the question *Kelley is a model employee*.

Item one which asks about Kelley being a valuable employee was also problematic to include in the composite measure of subordinate evaluation potentially because the supervisors could assess the statement *I feel that Kelley is a valuable employee* as an indictment of his current work performance or as a more general, macro statement of Kelley's worth as an employee. Those supervisors with the former referent would likely rate one way while those supervisors who read the statement *I feel that Kelley is a valuable employee* and wanted to convey that every subordinate is valuable would likely rate in the other direction.

Finally, item nine which states: *It will be fine if Kelley's performance does not change* was included in the composite measure of subordinate evaluation to increase the validity of the measure. Given that item eight states *Kelley's current level of performance is acceptable to me*, item nine was a reasonable item to include in the composite. After these changes, the composite measure for supervisor's evaluation of a subordinate was comprised of the following items:

2. I am pleased with Kelley's quality of work

3. Kelley is successfully performing all work duties

4. The quality of Kelley's work is consistently good

5. Kelley's work performance is adequate

6. Kelley is demonstrating sufficient personal effort on work-related tasks

8. Kelley's current level of performance is acceptable to me

9. It will be fine if Kelley's performance does not change

These items comprising subordinate satisfaction loaded on to one factor. Maximum likelihood method of extraction was used for the factor analyses. No rotation was necessary. Both the concession account with internal excuse and the external excuse preappraisal questionnaires were validated by factor analysis. Table 6 shows how the items loaded onto the

single factor in the concession account with internal excuse and Table 7 does the same for the external excuse condition.

Table 6

Items Comprising Supervisor's Evaluation of the Subordinate in the Concession Account with Internal Excuse Condition

	Factor Loadings
Item 5	.79
Item 3	.77
Item 2	.68
Item 8	.63
Item 4	.60
Item 9	.51
Item 6	.50

Table 7

Items Comprising Supervisor's Evaluations of Subordinates in the External Excuse Condition

	Factor Loadings
Item 8	.90
Item 4	.76
Item 9	.71
Item 5	.71
Item 3	.67
Item 2	.60
Item 6	.58

The second step in validating the items comprising the supervisor's evaluation of the subordinates' composite measure was to check the reliability of the items. Reliability scores (Cronbach's α) for the items in the concession account with internal excuse were $\alpha = .81$. For the external excuse condition, the reliability score was $\alpha = .85$, indicating that the supervisor's scoring of the items comprising the composite measure have acceptable levels of internal

consistency. Having established that the items used in the composite measure loaded on a single factor and that there was an acceptable level of reliability within the scores, hypotheses H3 – H5b were evaluated using these items.

H6a: The subordinate will be significantly more satisfied with supervisors who use feedback that incorporates elements of consistency, consensus, and / or distinctiveness (specific feedback) than supervisors who deliver feedback without consistency, consensus, and distinctiveness (nonspecific feedback), when the reason given for performance failure is an external excuse.

H6b: When the reason for performance failure is a concession account with an internal excuse, the subordinate will be significantly more satisfied with superiors who use feedback that does not incorporate elements of consistency, consensus, and / or distinctiveness (nonspecific feedback) than supervisors who deliver feedback with consistency, consensus, and distinctiveness (specific feedback).

Hypotheses 6a and 6b were tested by conducting two univariate ANOVAs between perceived specificity of feedback from subordinates and the account type made during the appraisal. Subordinate satisfaction will be derived from a composite score of items 1 – 4, 9, and 10 on the subordinate feedback form (Appendix L).

Scale development for subordinate satisfaction measure.

In order to properly evaluate Hypotheses 6a and 6b, it was necessary to conduct a factor analysis and a reliability check of the items comprising the composite scale for the subordinate's satisfaction with the supervisor's communication during the performance appraisal interview. Items 1 - 4, 9, and 10 comprised the composite scale. The items were:

1. I was satisfied with the supervisor's communication during the performance appraisal interview.

2. I was satisfied with the way that the supervisor provided feedback to me during the appraisal.

3. The supervisor was too vague in delivering the feedback (this item was reverse coded).

4. The supervisor made it clear to me how they felt about my performance.

9. I felt like the supervisor was very effective in conducting the performance appraisal.

10. I felt like the supervisor was very skilled at delivering performance feedback.

A factor analysis was conducted on the items using principal component analysis. No rotations were necessary. In both the concession account with internal excuse and the external excuse conditions, the items measuring the subordinate's satisfaction with the supervisors' communication during the appraisal interview loaded on to one factor and this factor explains 73.27% of the overall variance. These loadings are in Table 8 for the concession account with internal excuse and Table 9 for the external excuse condition.

Reliability analysis of the items comprising the composite scale reveals a reliability score of $\alpha = .91$ for the items in the concession account with internal excuse. The reliability score for the items in the external excuse condition is $\alpha = .91$. Based on the factor analysis and reliability

analysis, the items 1 - 4, 9, and 10 from the subordinate were used to measure subordinate satisfaction with the supervisor's communication during the performance appraisal interview.

Table 8

Items Comprising Subordinate Satisfaction in Concession Account with Internal Excuse Condition

Item	Factor Loadings
Item 1	.90
Item 2	.94
Item 3	.71
Item 4	.68
Item 9	.93
Item 10	.93

Table 9

Items Comprising Subordinate Satisfaction in the External Excuse Condition

Item	Factor Loadings
Item 1	.92
Item 2	.96
Item 3	.73
Item 4	.61
Item 9	.89
Item 10	.92

The three characteristics of specific feedback are captured by items 11 (consensus), 12 (distinctiveness), and 13 (consistency) on the subordinate's feedback form: *(11) The supervisor compared my performance to the performance of others during the appraisal, (12) The supervisor indicated that I had failed on a number of other tasks during the appraisal, (13) The supervisor compared my current performance to my past performance during the appraisal.*

Specificity of feedback will be determined by evaluating the self-reported responses to items 11-13 from the subordinate's feedback form. These individual items have been shown to comprise the construct of specificity of feedback in prior empirical and theoretical studies (Kelley, 1967; Larson, 1984; Liden & Mitchell, 1985).

Chapter 4

Results

Communication of Negative Feedback

Hypothesis H1a.

H1a: Supervisors will communicate positive subordinate observations significantly more than negative subordinate observations, regardless of the type of excuse for poor performance offered by the subordinate.

Hypothesis H1a is not supported. A paired samples *t*-test between the total number of times that a supervisor communicated the positive observations (from the note cards) of the subordinate and the total number of times that a supervisor mentioned the negative observations revealed that, on average, the supervisors communicated negative observations 2.26 times more frequently than positive observations, $t(52) = 6.32$, $p < .01$. When counting observations made in both appraisal interviews that a participant conducted, negative observations were mentioned an average of 7.30 total times (in both interviews) and positive observations were mentioned 5.04 total times. Evaluating the number of times that the positive and negative observations were mentioned by subordinate excuse type also revealed that the positive observations were mentioned with significantly less frequency than the negative observations.

In the concession account with internal excuse condition, negative observations, on average, were communicated 4.23 times as compared to 2.66 times for the positive observations, for a difference of 1.57 times, $t(52) = 6.32$, $p < .01$.

In the external excuse condition, negative observations were communicated 3.08 times as compared to 2.38 for the positive observations, for a difference of .70 times, $t(52) = 3.42$, $p < .01$.

Hypothesis H1b.

H1b: The majority of supervisors will self-report that they delivered all of the negative messages that they felt obligated to deliver based on their information about the subordinate's performance.

Hypothesis H1b is supported. In the concession account with internal excuse condition, 81.10% of the supervisors in the experiment reported that they delivered all of the negative feedback that they felt obligated to deliver based on Kelley's performance. In the external excuse condition, the same percentage, 81.10%, of the supervisors reported that they delivered the negative feedback they wanted to during the performance appraisal. Mean scores (where 1 = yes, delivered all feedback desired and 2 = no, did not deliver all feedback desired) for both the concession account with internal excuse and the external excuse condition were 1.19 (*SD* = .40).

Positive Distortion

Hypothesis H2a.

H2a: When compared to the total number of statements made during the performance appraisal, greater than 16.6% of the supervisor's statements will be positive in nature (positive distortion), both when the subordinate offers an external excuse for performance failure and when the subordinate offers a concession with internal excuse to explain the performance failure.

Hypothesis H2a is supported. In both the concession account with internal excuse and in the external excuse condition, over 16.6% of supervisors' statements were positive in nature. In fact, as shown in Table 10, less than 25% of the supervisors did not positively distort their messages during the performance appraisal interview regardless of the subordinate's account making. In the concession account with internal excuse condition, 83.00% of the supervisors

positively distorted the interview; in the external excuse condition 79.20% of the supervisors positively distorted the interview.

Table 10

Positive Distortion Scores

		Concession account with internal excuse	External excuse condition
Mean Positive Distortion Score		.24	.22
Standard Deviation		.08	.07
Percentiles	25^{th}	.19	.19
	50^{th}	.25	.22
	75^{th}	.29	.27

Note. All scores > .166 indicate positive distortion. Percentiles reflect mean positive distortion scores for the indicated percentile.

Hypothesis H2b.

H2b: Supervisors will communicate with significantly more positive distortion during the performance appraisal interview when the subordinate offers an external excuse as compared to when the subordinate offers a concession account with an internal excuse.

Hypothesis H2b is not supported. A repeated measures ANOVA revealed nonsignificant results for the influence of excuse type on distortion scores. A repeated measures ANOVA was first conducted with positive distortion scores in each experimental condition as the main effects and sequence and room as between subjects effects. The effect of room was not significant in the model, $F(1,49) = 2.10$, $p = .15$, and was subsequently eliminated from the analysis. The resulting repeated measures ANOVA evaluated distortion in each experimental condition with a between subjects effect of sequence. All observed interactions were nonsignificant.

The difference between the supervisors receiving the concession account with internal excuse as compared with the internal excuse condition and the influence on positive distortion

was not significant, $F(1,51) = 2.60$, $p = .11$. The effect of sequence also was not significant, $F(1,51) = 1.72$, $p = .20$. Based on these results, H2b is not supported.

Subordinate Evaluation by Supervisors

Hypotheses H3 – H5b were evaluated using a repeated measures ANOVA. The dependent variables were the supervisor's evaluation of the subordinate, measured using the same series of questions (that where validated through scale construction as explained in the Methods chapter) before and after the appraisal interview. The independent variables are the types of accounts (concession with internal or external) and the sequence in which the supervisors are exposed to the accounts (either concession with internal excuse or external excuse first, then the other). Additionally, room was included as a between subjects variable to ensure that the two different interview rooms used had no significant influence on the observed variables. When the particular room (either primary or alternate) was included in the first repeated measures ANOVA that was conducted as an independent variable, results showed that room had no effect on evaluation scores, $F(1, 49) = 1.30$, $p = .26$. Subsequently, room was eliminated from the repeated measures ANOVA and the model was conducted with evaluation scores as the dependent variables and type of account and sequence as the independent variables.

Hypothesis H3.

H3: Supervisors will evaluate a subordinate who makes an external excuse for poor performance higher than a subordinate who makes a concession account with an internal excuse after the appraisal interview.

Hypothesis H3 is supported. Comparisons resulting from a repeated measures ANOVA reveal that the supervisors evaluate the subordinate making external excuses significantly higher than the subordinate offering a concession account with an internal excuse, $F(1,51) = 18.98$,

$p < .01$. There is a mean difference of 3.66 in subordinate evaluations between internal and external conditions. This is based on a composite score for the subordinate's evaluation that could range from a low (unfavorable evaluation) of 7 to a high (favorable evaluation) of 49.

Hypothesis H4a.

H4a: Supervisors will evaluate a subordinate who makes an external excuse for poor performance during their appraisal interview higher after the interview than before the interview.

Hypothesis H4a is supported. Results from the repeated measures ANOVA reveal that supervisors rate the subordinate who offers an external excuse for the performance failure higher after the appraisal interview than before the appraisal interview, $F(1,51) = 55.92$, $p < .01$. Pairwise comparisons from the repeated measures ANOVA show a mean difference of 5.73 in subordinate evaluation scores between the preappraisal and the postappraisal evaluation scores in the external condition.

Hypothesis H4b.

H4b: There will be no significant difference in supervisor's evaluations of a subordinate before and after the performance appraisal when the subordinate offers a concession account with an internal excuse.

Hypothesis H4b is not supported. Results from the repeated measures ANOVA reveal that supervisors rate the subordinate who offers a concession account with an internal excuse for the performance failure higher after the appraisal interview than before the appraisal interview, $F(1,51) = 14.47$, $p < .01$. Pairwise comparisons from the repeated measures ANOVA show a mean difference of 2.34 in subordinate evaluation scores between the preappraisal and the postappraisal evaluation when a concession account with internal excuse is offered.

Hypothesis H5a.

H5a: There will be a significant interaction between preappraisal and postappraisal subordinate evaluation and the types of accounts given by the subordinate to explain performance failure.

Hypothesis H5a is supported. The interaction between the preappraisal and the postappraisal subordinate evaluation scores and the types of accounts given by the subordinate is significant, $F(1,51) = 14.96$, $p < .01$. The salient results of this interaction are reported above in Hypotheses H3 – H4b.

Hypothesis H5b.

H5b: The order of the experimental conditions that the supervisors experience will have no significant effect on the subordinate's evaluation scores.

Hypothesis H5b is partially supported. The three-way omnibus interaction between sequence, condition, and preappraisal and postappraisal evaluation scores is not significant, $F(1, 51) = .95$, $p = .34$. Nor is there a between subjects effect found based on sequence $F(1,51) = .19$, $p = .66$. This indicates that there is no main effect due to sequence. However, there is an interaction effect between sequence and condition $F(1,51) = 5.97$, $p = .018$. Evaluating which groups were responsible for this effect, multiple comparisons using a Bonferroni correction showed that when supervisors are given the external excuse condition before the concession account with internal excuse, there is a mean difference in evaluation scores of 3.45, $p < .01$. There is no significant effect when participants first experience the concession account with internal excuse condition; the mean difference is .48, $p = .59$. Table 11 illustrates the effect on the subordinate's evaluation scores when supervisors receive the external condition first in the sequence.

Table 11

Preappraisal and Postappraisal Evaluation Scores by Sequence

Sequence	Eval Score	Mean	Std. Error
Internal 1st	Preappraisal	18.82	1.09
	Postappraisal	19.30	1.15
External 1st	Preappraisal	16.71	1.03
	Postappraisal	20.16	1.08

Subordinate Satisfaction with Supervisor's Communication

Hypothesis H6a.

H6a: The subordinate will be significantly more satisfied with supervisors who use feedback that incorporates elements of consistency, consensus, and / or distinctiveness (specific feedback) than supervisors who deliver feedback without consistency, consensus, and distinctiveness (nonspecific feedback), when the reason given for performance failure is an external excuse.

Hypothesis H6a is supported. A univariate ANOVA with the amount of specificity as the independent variable and the subordinate's satisfaction as the dependent variable revealed a significant interaction between the two variables, $F(3,49) = 7.40$, $p < .01$.

The amount of specificity that a supervisor used in the performance appraisal could vary from zero to three depending on whether the content analysis revealed that the supervisor: (1) compared the subordinate's performance to the performance of others (2) indicated that the subordinate had failed on a number of other tasks (3) compared the subordinate's past performance with the subordinate's future performance. Supervisors mentioning none of these three items received a specificity score of zero, supervisors mentioning one received a score of one, supervisors mentioning two received a score of two, and a score of three was given for mentioning all three elements of specific feedback. Table 12 shows that the subordinate's

satisfaction with the supervisor's communication during the appraisal interview increased as the supervisor's level of specificity increased in delivering the performance feedback.

Table 12

Satisfaction Scores with Supervisor's Communication by Specificity Level

Specificity	Mean Satisfaction	*SD*	n
0	18.00	10.58	3
1	25.85	6.27	13
2	31.12	6.05	25
3	32.50	5.68	12
All levels	29.40	7.17	53

Hypothesis H6b.

H6b: When the reason for performance failure is a concession account with an internal excuses, the subordinate will be significantly more satisfied with superiors who use feedback that does not incorporate elements of consistency, consensus, and / or distinctiveness (nonspecific feedback) than supervisors who deliver feedback with consistency, consensus, and distinctiveness (specific feedback).

Hypothesis 6b is not supported. A univariate ANOVA with the amount of specificity as the independent variable and the subordinate's satisfaction as the dependent variable revealed a significant interaction between the two variables, $F(3,49) = 6.29$, $p < .01$. In fact, the effect of specificity on subordinate satisfaction when the reason given for performance failure is a concession account with an internal excuse is significant in the opposite direction as predicted in the hypothesis. Table 13 shows that as the degree of specificity increases, the subordinate's satisfaction with the supervisor's communication also increases.

Table 13

Satisfaction Scores with Supervisor's Communication by Specificity Level

Specificity	Mean Satisfaction	*SD*	n
0	11.50	4.95	2
1	27.11	9.47	9
2	30.64	5.39	36
3	34.50	6.50	6
All levels	29.75	7.42	53

Chapter 5

Discussion

Summary of Findings

Results from the current experiment indicate that: (1) supervisors do *not* mention the three positive note card observations significantly more than the three negative note card observations, (2) 81.10% of supervisors self-report that they delivered all of the negative feedback they felt obligated to deliver to the subordinate, (3) 83.00% of supervisors positively distort their feedback when the subordinate offers an concession with internal excuse account during the appraisal interview and 79.20% of supervisors positively distort their feedback when the subordinate offers an external excuse account, (4) subordinate evaluation scores are more positive after the appraisal than before the appraisal, regardless of the type of account given, (5) the subordinate is evaluated more positively when an external excuse is given as compared to when a concession account with an internal excuse is given, (6) supervisors who received the external excuse first in the experimental sequence gave significantly higher evaluation scores across both conditions, (7) the subordinate's satisfaction with the supervisor's communication during the appraisal interview increases as the supervisor's feedback becomes more specific.

Communication of Negative Feedback

A novel finding from this research is that when given an equal number of positive and negative behavioral observations (from the note cards) about the subordinate, supervisors choose to mention the negative observations significantly more often than the positive observations. This runs counter to the initial hypothesis that supervisors would be more reluctant to mention the negative observations and use the positive observations as an expedient crutch during the appraisal interview to soften their feedback.

The initial hypothesis that supervisors would be more reluctant to mention the negative observations than the positive observations was theoretically grounded in research findings from Tesser, Rosen, and colleagues on the mum effect (e.g. Rosen & Tesser, 1970, 1972; Tesser & Rosen, 1972; Tesser & Rosen, 1975; Tesser, Rosen et al., 1972b). This body of work which documents the relatively pervasive nature of communicators' reluctance to transmit bad news would seem to indicate that supervisors, in the context of the current experiment, would be less likely to articulate negative as opposed to positive observations of the subordinate's performance.

The finding from this current research that negative observations were reported more frequently might suggest that the mum effect is not triggered when supervisors are reporting observable, verifiable behaviors. A supervisor's reluctance to communicate negative feedback might be mitigated when the message to be transmitted is a relatively straightforward reporting of an actual behavioral event as opposed to a more subjective observation about the subordinate's performance, reflecting the idea that it is less face threatening to communicate facts as opposed to judgments. There does not seem to be any mum effect present with respect to the communication of the negative observations from the note cards that were provided for the supervisors.

As discussed in the Chapter 2 section *The Reluctance to Transmit Negative Feedback* on page 5*,* Tesser and Rosen (1977) highlight three general reasons why people are reluctant to communicate negatively valenced messages: (1) out of concern for self, (2) out of concern for others, and (3) out of concern for the societal norms for transmitting negative feedback. Unquestionably, there can be individual differences between communicators concerning the degree to which they are influenced by these three issues. These individual differences can be

easily illustrated by considering the negative observations that supervisors mentioned in this experiment. Across both appraisal interviews, supervisors communicated negative observations an average of 7.30 times (*SD* = 2.55), and 66% of the supervisors mentioned negative observations 5, 6, 7, or 8 times. Frequencies revealed individual differences: one supervisor at the low end of the range communicated only one negative observation in both appraisal interviews, and at the high end of the range, one supervisor communicated 14 negative observations, indicating that some supervisors communicated almost none of the negative observations to the subordinate while other supervisors repeatedly mentioned the negative observations in both of the two interviews.

If three main reasons for the mum effect are concern for self, concern for others, and concern for societal norms, then some individual differences from this experiment might be explained by supervisors' different levels of empathy for the other person (concern for others). Some of the supervisors who communicated few of the negative observations might simply have placed such a primary on the subordinate's feelings that they eschewed communicating the negative observations. Given the finding that most supervisors did *not* avoid transmitting the negative feedback and, in fact, communicated more negative observations than positive observations, it is interesting to speculate about the generalizability of this finding to other groups.

In the current experiment, the sample of supervisors was drawn from a population of graduate MBA students at a large university. What if the sample had been taken instead from a population of nursing, social work, or seminary graduate students? It is possible that graduate students in these fields might place greater emphasis on how the other person (subordinate)

would feel after receiving the negative feedback, perhaps due to the respective norms and culture of these fields.

Just as occupational norms and culture can influence individuals, so can larger societal norms and culture influence people. For this reason it is important to note that the majority of Tesser, Rosen, and colleagues' research on the mum effect was conducted in the 1970s. It is reasonable to assume that societal norms have changed since the 1970s and, since concern for societal norms is given by Tesser and Rosen (1977) as one of three main reasons why people are reluctant to communicate negative feedback, as societal norms change over time, the mum effect would correspondingly be impacted by those changes. It could be that negative feedback that was considered socially unacceptable and impolite in the 1970s might be viewed as acceptable to transmit today. These differences in societal norms could help explain why the mum effect was found in multiple experiments in the 1970s but no mum effect was readily apparent in the current experiment.

Another possible reason why the mum effect was not found in the current experiment is that supervisors might have determined that the possible future benefits that could spring from giving the negative feedback outweighed the immediate discomfort that transmitting the negative feedback entailed. As noted by Bass (1990, p.340): "Rewarding when positive, feedback can be highly punitive when negative. However, negative feedback may quickly come to be interpreted as rewarding if it is seen as intended to be helpful and if it actually results in improved performance."

In many of the mum effect experiments conducted by Tesser, Rosen, and others, there was no reason for the participants to believe that their communicating a negative message would result in any possible future good for the recipient. For example, consider two of the mum effect

experiments discussed in the Chapter 2 section starting on page 5, *The Reluctance to Transmit Negative Feedback*. In the study by Tesser and Rosen (1972) where a subject was faced with deciding whether or not to tell the confederate that the confederate would receive an electric shock, there was no future benefit that might have made communicating the negative message more palatable, such as being able to tell the confederate that they would receive an electric shock but that the shock would prevent them from having mental illness in the future. As a second example, in the experiment by O'Neal, Levine, and Frank (1979) where a number of "lost" postcards were placed throughout European cities indicating that a fictitious person had been either under or overcharged $100 on their electric bill there was a mum effect reported because people mailed the positive postcard more frequently than the negative postcard. But again, there was no indication that sending the bad news postcard might benefit the receiver in some way. It is possible that by modifying the postcard to indicate that the receiver owed $100 but noting that prompt payment would prevent them from being sued and having their credit being ruined might have mitigated the mum effect. Both of these examples highlight a critical distinction between many mum effect studies and the current study: In the current study there is a presumably clear future benefit that comes from delivering the negative feedback, namely that the subordinate's performance improves and the supervisor's bosses know that the supervisor rehabilitated the performance failure. The presence of this future benefit might be sufficient to trump the mum effect, which happened in this experiment.

It is also reasonable to speculate that Tesser, Rosen, and colleagues' mood congruency hypothesis (e.g. Tesser & Conlee, 1973; Tesser, Rosen, & Conlee, 1972) might be influenced by the presence of compelling future benefits that might derive from the negative feedback. Although mood congruency predicts that senders would prefer to be in a negative mood before

transmitting negative messages, a sender who believed that it was in the long term best interests of the receiver to get the negative feedback is likely to send the message regardless of their current mood. Additionally, it could be argued that supervisors who know that they have to deliver a negative performance appraisal might already be in a bad mood because they would rather not deliver the feedback, thus ironically making it more likely that they would, in fact, deliver the negative messages. It should be mentioned that the mood congruency hypothesis might be somewhat problematic to extend to a performance appraisal situation simply because the appraisal interview is a circumstance where negative feedback is often expected and does not seem out of place, presumably setting up an environment where mood is peripheral to candor and accuracy.

Finally, it is important to note that Tesser and Rosen (1975, p.207) were aware that the mum effect might not apply in certain situations and they called for empirical testing, such as the current experiment, to explore these situations:

> Although a communicator may fear negative evaluation when communicating bad news, how does he think a recipient will evaluate him if he fails to communicate the news? He might anticipate being characterized as irresponsible, cowardly, and deceitful. Hardly a positive evaluation….There are circumstances under which *failure* [emphasis in original] to communicate bad news may lead to a negative evaluation. Further theoretical and empirical work will have to spell out the conditions under which this might occur.

It appears that the current experiment has uncovered such a condition.

In further exploring why the mum effect was not triggered in the current study, the framework of conversational constraint theory provides some clues. Conversational constraint theory is based on the premise that five conversational constraints influence a communicator's

selection of conversational strategies. The five constraints are: (1) concern for hurting the receiver's feelings, (2) concern for minimizing imposition, (3) concern for avoiding negative evaluation by the receiver, (4) concern for clarity, and (5) concern for effectiveness (M. S. Kim, 1994).

These five constraints that act on communication strategy selection highlight the tension between *social appropriateness* (constraints 1-3) and *efficiency* (constraints 4-5). As noted by Kellerman (1992, p.289): "social appropriateness and efficiency are two constraints that communicative behavior is responsive to and regulated by. Communication is selected, fashioned, edited, enacted, and evaluated on these grounds." The first three constraints are all reasons given by Tesser and Rosen (1975) to explain the mum effect, suggesting some overlap between the two theories. Conversational constraint theory builds on the mum effect by exploring the efficiency constraints and assessing their influence in relationship to social appropriateness.

In a study that considered the influence of situational urgency on social appropriateness and efficiency, Kellerman and Park (2001) found that when people perceive situational urgency, the preferred level of efficiency increases. Situational urgency, however, did not influence the preferred level of social appropriateness. Additionally, Kellerman and Park (2001) found that the socially appropriate but less efficient tactics that are acceptable in nonurgent situations are not acceptable in urgent situations.

Kellerman and Park's (2001) findings about situational urgency reinforce tenets from Brown and Levinson's (1987) politeness theory. Politeness theory evolved in part from Brown and Levinson's desire to explain deviation from such models of rational communication as Grice's conversational maxims (1975). In developing their theory of politeness, Brown and

Levinson point out that there are a number of instances where concern for another person's face is trumped by issues of efficiency. As Brown and Levinson (1987, p.95) note:

> The prime reason for bald-on-record usage may be stated simply: in general, whenever S [speaker] wants to do the FTA [face threatening act] with maximum efficiency more than he wants to satisfy H's [hearer's] face, even to any degree, he will choose the bald-on-record strategy.

These principles and findings from politeness theory and conversational constraint theory shed some light on possible reasons why supervisors were not reluctant to communicate the negative observations about the subordinate's performance in the current research. Specifically, there were a number of components embedded within the experimental protocol that might lead a supervisor to believe that there was a premium placed on efficiency over social appropriateness, and possibly even lead the participants to believe that there was an element of urgency in the performance appraisal.

For example, the following is an excerpt from the contextual information about the subordinate given to the supervisors in preparation for the appraisal interview (Appendix D):

> Your supervisors believe in the adage "an employee is a reflection of the boss" and you know that you need to honestly address Kelley's performance decline accurately and swiftly or you run the risk of having your supervisors step-in and handle the problem. And that would not he good for you or for Kelley. In short, you need to address this performance problem right now.

Additionally, the instructions to participants (Appendix C) stated:

> Although the contextual instructions will indicate that the appraisal interview is scheduled for 15 minutes, you may spend either more or less time conducting the

interview as you deem it appropriate. Simply end the appraisal when you feel that the appraisal has reached its end.

Telling participants that they had tentatively scheduled 15 minutes for the appraisal interview was done because pilot study appraisals revealed large variations between supervisors' interview lengths when no time referent was given. A subsequent pilot study revealed that telling participants in the written instructions that they had reserved 15 minutes on their calendar was helpful in preventing these large time ranges.

However, this time referent, combined with the instructions to the participants indicating that their supervisors wanted the performance failures addressed immediately, might well have caused participants to feel like there was a sense of urgency in the performance appraisal. There were probably a number of the supervisors who inferred that it was in their best professional interest to quickly attend to the subordinate's performance failures, lest they draw the ire of their supervisors. And, consistent with the general tenets of conversational constraint theory and the specific findings from Kellerman and Kim (2001), this sense of urgency would trump social appropriateness in favor of efficiency as a communication strategy in the appraisal interview, muting the influence of the mum effect. Although there were an equal number of positive and negative observations given on note cards to the supervisors (Appendix E), every sentence of the contextual information given to supervisors (Appendix D) made it abundantly clear that the subordinate's performance was decidedly negative in nature, and it would seem to be much more efficient for supervisors to select and comment on the negative performance observations. This would help explain why supervisors chose to communicate the negative observations more frequently than the positive observations during the performance appraisal interview.

It is interesting to note that paired sample *t*-tests show that the excuse given by the subordinate has a significant effect on the amount of times that the negative observations are mentioned. Giving an external excuse attenuates the amount of negative observations mentioned by an average of 1.15 times, $t(52) = 4.89$, $p < .01$. In the external excuse condition, negative observations are mentioned an average of 3.08 times compared to 4.23 times when a concession account with an internal excuse is given. Additionally, there is no significant effect of subordinate excuse on the positive observations, $t(52) = .144$, $p = .16$. These findings suggest that although supervisors in this experiment were more inclined to mention negative as opposed to positive observations during the appraisal interview, the amount of times that the negative observations were discussed could be reduced when an external excuse was offered. It appears that supervisors are less likely to dwell on negative observations when the subordinate offers an external excuse for the performance failure.

Although self-serving biases combined with the tendency for managers to overestimate their strengths would posit that supervisors would claim that they communicated all the negative messages that they felt obligated to deliver no matter if salient information was not transmitted, the fact that supervisors chose to articulate negative observations more than positive ones leads to the sensible finding (from H1b) that the supervisors would report they had delivered all of the negative messages that they felt obligated to deliver to the subordinate. The finding (from H1a) that supervisors were not reluctant to communicate the negative observations takes away some of the explanatory power from the finding that the supervisors self-reported they had delivered all of the negative messages that they felt obligated to deliver based on their information about the subordinate's performance. Had H1a been supported and had supervisors mentioned positive observations more than negative observations, it could have been speculated that self-serving

biases and managerial skill overestimation was at play. As it stands, it is not possible to separate these effects from the possibility that the supervisors genuinely, and accurately, believed that they delivered all necessary negative feedback.

It is interesting to note that a rather large percentage of supervisors, 81.10%, indicated that they had delivered all of the negative messages that they felt obligated to deliver based on their information about the subordinate because that large percentage is counter to certain observations from the subordinate. To explain, consider a one paragraph section of the contextual information (from Appendix D) that all supervisors read prior to conducting the appraisal interview:

> It is clear to you that other employees are having much more success than Kelley in performing the same professional duties, so it does not appear that you are asking too much of Kelley. Furthermore, Kelley seems to be failing on a wide variety of tasks—everything from interacting with clients to working with colleagues to meeting deadlines—Kelley's performance is pervasively poor. Finally, you have noticed that this performance degradation is slowly getting worse. In prior years, Kelley's performance was significantly better than this year's performance.

This paragraph was initially constructed to ensure that the supervisors had enough information to be able to communicate the three elements of specific feedback: distinctiveness, consistency, and consensus, which will be discussed further in the section on *Subordinate Satisfaction with Supervisor's Feedback* on page 94. There were three yes / no questions that the subordinate answered after each appraisal interview as part of the subordinate postinterview questionnaire (Appendix L): *(11) The supervisor compared my performance to the performance of others during the appraisal (12) The supervisor indicated that I had failed at a number of*

other tasks during the appraisal and (13) The supervisor compared my current performance with my past performance. Table 14 captures the frequency of the subordinate's responses to these questions. *Internal* indicates that the responses are from appraisal interviews where the subordinate gave a concession account with internal excuse and *external* indicates that the responses are from appraisals where the subordinate gave an external excuse.

Table 14

Subordinate Frequency of Responses

	Internal Q11	Internal Q12	Internal Q13	External Q11	External Q12	External Q13
Frequency Yes	11.30	9.40	84.60	13.20	7.50	77.40
Frequency No	88.70	90.60	15.10	86.80	92.50	22.60

As Table 14 indicates, the subordinate clearly did not believe that the supervisors compared his performance to the performance of others during the appraisal (question 11) nor did the subordinate believe that the supervisor indicated that he failed on a number of other tasks during the appraisal (question 12). In fact, across both conditions, the subordinate believed that his performance *was not* compared to the performance of others in 87.75% of the appraisals and the subordinate believed that the supervisor *did not* indicate that he failed on a number of other tasks in 91.55% of the interviews.

Since the contextual information specifically gave the supervisors information about the subordinate's performance in relation to others as well as information detailing other instances of performance failure, the fact that 81.10% of the supervisors self-reported that they had given all of the negative information that they felt obligated to deliver seems in line with the tenets of self-serving biases and managerial overestimation of strengths. Namely, while supervisors might feel like they delivered all of the negative information that they were obligated to, the subordinate

clearly indicated that the supervisors failed to mention two key areas that were highlighted as negative issues in the contextual information for the supervisors.

Positive Distortion

The current findings about positive distortion highlight an underlying communication tendency, inherent not only in performance appraisal but apparent more broadly in interpersonal communication, that it is the exception as opposed to the rule when negative communication is delivered without some type of positive buffer. This tendency is explored in theories from Brown and Levinson's (1987) politeness theory to Kim and colleagues' (M. S. Kim, 1994; M. S. Kim, Sharkey, & Singelis, 1994) conversational constraint theory to Tesser and Rosen's (e.g. 1975) mum effect to Grice's (1975) conversational maxims.

The finding from this research that there is positive distortion in the performance appraisal interviews both when the subordinate offers a concession account with an internal excuse and when the subordinate offers an external excuse highlight the universal inclination explored by Brown and Levinson (1987) of underlying politeness in most interpersonal communications. As Brown and Levinson (1987, p.68) point out: "In the context of the mutual vulnerability of face, any rational agent will seek to avoid these face-threatening acts, or will employ certain strategies to minimize the threat." The performance appraisal context created for this experiment does not seem to be entirely immune from this effect. The fact that supervisors were not reluctant to transmit the negative observations did not prevent the vast majority of them from positively distorting their appraisal interviews.

At issue here is not whether there will or will not be some amount of positivity expressed by the supervisor in the performance appraisal but, at what point this inherent politeness serves to dilute or even misrepresent the true nature of the feedback intended to be transmitted during

the performance appraisal. Given the context of this experiment, where the information to supervisors made it clear that the performance was negative to the point of being career threatening, the level at which an appraisal was characterized as positively distorted was 16.6%, indicating that 16.6% of the total messages communicated about the subordinate and the subordinate's performance were positive in nature. As explained earlier in the *Hypotheses Testing* section starting on page 48, this was a reasonable level in light of the subordinate's work performance as described to supervisors. Using this 16.6% threshold, a remarkable 83.00% of appraisals in the concession account with internal excuse condition were positively distorted and 79.20% were positively distorted in the external excuse condition.

Positive distortion and inflation during performance appraisal interviews are hardly new issues. Longenecker, Sims, and Gioia (1987, p.189) present findings from executives detailing a host of different reasons why managers inflate and distort during performance appraisals and Laird and Clampitt (1985) found that inflation was one of four problems reported by *all* supervisors in their sample of 60 managers who routinely delivered performance appraisals. Yet it is precisely these recurring issues of positive distortion and inflation that are at the heart of this research effort.

It is hard to overlook the fact that the mean distortion score across both experimental conditions was .23 (SD = .06), indicating that an average of 23% of the total number of messages communicated during the performance appraisal interviews were positive in nature. And, if in fact some supervisors felt that there was a sense of urgency in the performance appraisal as communicated through the contextual instructions, it might be expected that this urgency would have prevented the distortion from going even higher. Additionally, the fifteen minute time referent might also have caused those supervisors who felt some sense of urgency to truncate

their positive communication in favor of more efficient and salient (in accordance with the context) negative communication. Had there been no sense of urgency and no indicated time referent, the positive distortion scores, already high, might have gone even higher.

There are many possible explanations for the significant amount of positive distortion that appeared in the appraisal interviews, and a number of these reasons have been explored elsewhere in this study (e.g. mum effect, politeness). One additional explanation that deserves to be highlighted concerns the structural nature of the performance appraisal interview.

There are a number of factors inherent in the performance appraisal interview that cause a poorly performing subordinate to receive more lenient treatment than may be warranted. The first reason is that there is evidence to suggest that a leniency effect is triggered simply by the face to face interaction with another human in a performance appraisal (Gioia & Sims, 1985, 1986). Within this omnibus effect of leniency there are, of course, individual variations, with some people being more predisposed to leniency than others (for more thorough coverage of these idiosyncratic differences, see Bernardin and Orban (1990); for discussion and validation of an instrument for assessing rater leniency see Villanova, Bernardin, Dahmus, and Sims (1993)). Regardless of the amount of individual variation that may exist between supervisors, the latent leniency is likely to tilt the appraisal process away from candor and in favor of the poorly performing subordinate.

Secondly, this leniency is exacerbated by a great deal of conventional wisdom—scholarly and popular—that places a primacy on actively involving the subordinate in the performance appraisal. While some degree of subordinate interaction in the appraisal is absolutely essential, much well-meaning appraisal advice places a supervisor's ability to accurately and candidly assess subordinate performance in jeopardy and potentially taints the entire appraisal interview in

favor of leniency toward the subordinate. Consider the following advice from Beer (1981, p.34), which is step six of a seven-step process recommended for conducting a performance appraisal interview:

> 6. Starting the discussion. Give the initiative to the subordinate in the discussion that follows the opening statement. Specifically start the discussion by asking, "How do you feel things are going on the job? What's going well and what problems are you experiencing? How do you see your performance?" Such general questions will stimulate the subordinate to take the initiative in the problem identification and solving discussion. To facilitate this, a subordinate may be asked to appraise his or her own performance. If the manager starts by expressing views about the employee's performance, the interview inevitably becomes directive.

Based on the findings from this experiment, managers should heed the above advice at their own risk.

Giving the initiative to the subordinate at the beginning of the appraisal interview is fraught with problems, the most salient of which were previously discussed in the subsection *Self-serving and actor / observer biases* on page 11. The natural tendency for an employee to engage in a self-serving bias when questioned about performance is likely to color the framework of the appraisal interview. And, given the findings from this experiment about the effectiveness of excuses, allowing subordinates to take the initiative in appraisal interviews and asking them through open ended questions to characterize their performance is more likely than not giving subordinates a golden opportunity to take control of the tenor of the interaction and frame the interview in their favor, at the expense of accuracy, remediation, and candor in the appraisal interview.

A final structural component of the appraisal interview that can bias the odds in favor of a poorly performing subordinate occurs when supervisors follow the above mentioned conventional wisdom and advice to allow subordinates to begin the appraisal interview with a characterization of their current performance and end the appraisal interview with a summary of the discussion. When this occurs, a number of biases are triggered that make it likely that poor performance will be underemphasized and under-reported. First, the subordinates are allowed to seize valuable parts of the appraisal interview, namely the beginning and the ending, important in light of the primacy / recency effect. Second, subordinates will almost certainly present self-serving information during their turns in the conversation. Third, managers are prone to response biases that cause them to assess the probability of outcomes they desire as greater than the evidence warrants. Similarly, managers tend to view as less probable the outcomes that they do not desire (Wagner, 2002). When this bias toward wishful thinking on the part of supervisors is coupled with the fact that subordinates are often given such wide latitude in characterizing their performance during the appraisal, it becomes quite likely that the perceived magnitude of the performance failure will be minimized. All of these factors and biases lend a structural advantage to a poorly performing subordinate and increase the odds that performance failures will be de-emphasized and subordinates will receive more leniency for failures during the appraisal interview.

Subordinate Evaluation by Supervisors

Before conducting the current experiment, it was theorized that there might be a positive correlation between positive distortion and subordinate evaluation such that as positive distortion increases in an appraisal, so would a subordinate's evaluation tend to inflate. This seems like a reasonable supposition since we might expect that supervisors likely to positively distort might

also be likely to inflate ratings. However, this relationship was not borne out by the data in this experiment, as reported in the earlier subsection *Evaluating the relationship between positive distortion and subordinate evaluation* on page 57*,* which suggested that positive distortion and rating inflation, in this instance, are independent effects.

Although there was no conclusive evidence of a linkage between subordinate evaluation and positive distortion, this experiment did return important findings about the influence of account type on subordinate evaluation rankings. Perhaps the most important single finding from this portion of the experiment is that the external excuse used in this experiment was significantly more effective than the concession account with internal excuse in positively altering the way a supervisor perceives the subordinate. So significant is the external excuse in modifying supervisor's judgments of the subordinate that a sequencing effect was detected when supervisors in the experiment first received the external excuse condition. In other words, when the first excuse given by the subordinate was the external excuse, supervisors in this sequence tended to significantly increase their ratings in both the first (external excuse) and in the second (concession with internal excuse) condition, suggesting that the supervisors carry over some of the external condition excuses in their minds to the second experimental condition. Receiving the concession account with internal excuse first in the sequence did not significantly influence evaluation scores in the second condition (external excuse).

It was predicted that the external excuse would be more effective at diffusing supervisor discontent with the subordinate's performance failure than concession accounts with internal excuses. Studies such as Weiner, Amirkhan, Folkes, and Verette (1987) have found that external causality excuses compared to internal causality excuses (1) produce greater tendencies to demonstrate positive behaviors toward the excuse-giver (2) produce more positive emotional

reactions and (3) increase perceptions that the excuse-giver has positive traits. So, it is not entirely surprising that the external excuses were more effective than the concession accounts in enhancing attributions of the subordinate. What is surprising is the magnitude of this difference and the overarching influence of the external excuse on *both* of the experimental conditions.

Further evidence of the primacy of external excuses in this experiment was an unexpected finding discovered during pilot studies of the experimental protocol. During the pilots, many supervisors in the concession account with internal excuse condition were asking the subordinate *after hearing the internal excuse* if there was some external reason that could possibly explain the performance failures. In other words, even after the subordinate took the blame for the performance failures and said that the failures were "his fault", many supervisors continued to seek an external reason to explain the failures and thus diffuse responsibility from the subordinate. Due to the high incidence of this occurrence in the pilot studies, a question was added to the content analysis form before the data collection process. The question asked *In the concession account with internal excuse only, did the supervisor ask the subordinate if there was some other external reason (i.e. something going on in their personal life) to explain the performance failure after the subordinate had already admitted responsibility for the poor performance*.

Results from this question revealed that 43.4% of the supervisors prodded the subordinate for an external excuse even after the subordinate accepted responsibility internally for the performance failures. The fact that 43.4% of supervisors went looking for an external excuse in the internal excuse condition probably speaks to not only the centrality of external excuses in this experiment, but also to the popularity of external excuses in everyday life. The provocative question from this finding is why a subordinate taking responsibility for his failures, as was the

case here, was not a sufficient explanation for the supervisors. And, since an external excuse as opposed to a concession account with an internal excuse evoked a higher subordinate evaluation, there seems to be little evidence from this study to support the functionality of ascribing internal causality to performance failure.

It should be noted, in defense of the idea of accepting responsibility for failures, that there was indeed a significant, positive increase in the subordinate's evaluations after a concession account with internal excuse was offered. However, it is arguable how much of this difference was due to the account alone and how much might be attributed to a leniency effect being triggered simply on the basis of the face to face interaction between supervisor and subordinate. Although the same argument might be proffered for the external excuse condition as well, the magnitude of the differences between the external excuse and the concession account with internal excuse evaluations, plus the fact that receipt of the external condition first in sequence influences overall evaluations renders the argument extremely problematic in the case of the external excuse. However, concession accounts with internal excuses should not be dismissed out of hand because they were not as effective as an external excuse in this experiment as there are certainly situations wherein an account maker would be well advised to accept responsibility in the present to prevent a greater punitive response from the supervisor in the future.

In sum, the findings concerning the influence of account type on subordinate evaluations strongly support the primacy of external excuses as having a stronger, significantly more positive effect on supervisor's attributions of the subordinate. Additionally, the external excuse exerted an unexpected influence on the overall experimental sequencing and impacted evaluations of those supervisors even in the concession account with internal excuse condition. Also interesting

is the high incidence of supervisors looking for external causality even after the subordinate admitted responsibility for the performance failures. Taken together, it is clear that the external causality accounts exert a stronger influence on supervisors than the internal causality accounts from this research.

Subordinate Satisfaction with Supervisor's Communication

Evaluation of the specificity of feedback and its influence on the subordinate's satisfaction with the supervisor's communication during the performance appraisal revealed a consistent positive relationship between level of specificity and subordinate satisfaction. In the concession account with internal excuse and in the external excuse condition, the subordinate's satisfaction with the supervisor's communication consistently increased as the supervisor's feedback became more specific.

A note of caution should be made before this finding is generalized outside of the experimental context and recommended for application by appraisal practitioners. The initial hypothesis that people would not prefer specific feedback when the cause for their performance failure was internal was derived from the finding by Liden and Mitchell (1985) who discovered that people *did not* prefer specific feedback when the feedback implied that the cause for the feedback was internal. Results from this experiment are counter to the Linden and Mitchell (1985) finding. Although both studies queried participants who were not actually commenting on feedback that was from their own personal experience, the Linden and Mitchell (1985) experiment used multiple participants (N = 298) to assess the effects of feedback specificity.

This difference in subjects makes it prudent to cautiously interpret the finding from the current experiment in that the one subordinate preferred specific feedback even when the reason for performance failure had internal causality. It is possible that the subject in this experiment

felt differently than the norm from Linden and Mitchell's (1985) study. It could also be possible that the subordinate was inaccurate or did not listen carefully for the tenets of feedback specificity, although it is just as reasonable to assume that the subordinate was more accurate and consistent since he was exposed to a large number (106) of appraisal interviews and he trained himself to readily and effectively discern these feedback tenets over time. Additionally, it is reasonable to expect that when the feedback given is about an individual's real life, as opposed to a hypothetical role play (as in the case of this experiment) or student responses to a hypothetical example (as in the case of Linden and Mitchell's (1985) study) that the individuals receiving the feedback might indeed prefer nonspecific feedback over specific feedback out of an inherent desire to protect their face.

Although the above are reasons to cautiously interpret the feedback specificity scores, that alone does not diminish the consistent relationship exhibited in this experiment between specificity of feedback and the subordinate's satisfaction with the supervisor's communication during the appraisal interview. At least within this experimental context, the subordinate demonstrated a distinct preference for specific feedback during the performance appraisal interview and a linear relationship indicated that the greater the specificity, the more satisfied the subordinate was with the supervisor's communication during the appraisal interview.

Summary of Discussion

To summarize, the most salient learning points to come from the current experiment are (1) the mum effect was not in play during the performance appraisal, possibly because supervisors determined that social appropriateness was secondary to message efficiency, (2) the external excuse account was more effective in raising subordinate evaluations than the concession account with internal excuse, (3) overestimating managerial effectiveness likely

caused supervisors report that they transmitted all the negative feedback they were required to, but the subordinate reported that some critical feedback was missing, (4) over three-fourths of the appraisal interviews were positively distorted in both experimental conditions, (5) the subordinate consistently preferred specific feedback from the supervisors regardless of the account type offered to explain the performance failure.

Chapter 6

Limitations and Future Directions

Limitations

This laboratory experiment, like all others, trades a degree of generalizability for the precision that can be achieved in a controlled environment. In constructing the experiment, a number of decisions were made that influenced the research. These decisions also influence the interpretation of some of the experimental results.

Relational history.

The decision to control for relational history in this experiment by having no previous history between the supervisor and the subordinate is both a strength and a limitation inherent in this research. On the positive side, stripping out relational history prevents the supervisors from having any preconceived attributions about the subordinate prior to their participation in the experiment. This allowed each supervisor to be given the same baseline of information about the subordinate and thus ensured that attributions would be formed based on similar data and observations (given to the supervisors) before the appraisal interview. During the interview, this lack of relational history and the standardized performance synopsis enabled the research to focus on the effects of the accounts given by the subordinate. In addition, the lack of relational history also mitigated the influence of affinity or antipathy from entering into the performance appraisal and contaminating the interview and supervisors' attributions, which adds strength to the findings.

Relational history can be expected to positively influence the appraisal process (in terms of the subordinate's vantage point) when the relational history is good between the supervisor and the subordinate and negatively influence the appraisal process when the history is poor.

And, the effects of relational history might be so pronounced as to thoroughly mask the influence of subordinate account giving, thus clouding the thrust of this experiment. Said differently, it would certainly be problematic, if not impossible, to study the phenomenon under investigation in a field setting where any conclusions drawn would have to consider the relational history between the supervisor and the subordinate.

The lack of relational history in the experiment also makes caution prudent when interpreting the research findings. In performance appraisals conducted outside a laboratory, relational history is the norm. This is not to say that situations do not arise, from time to time, when a subordinate might receive a performance appraisal from a supervisor who shares little relational history with the employee. Examples of this might include a supervisor and a subordinate working in a highly decentralized and geographically dispersed organization, a supervisor–subordinate relationship where one of the two recently began the position, a supervisor with appraisal responsibilities for a large number of subordinates, an organizational or personal change event that causes an appraisal to be conducted earlier than originally scheduled, or a situation where a supervisor is disconnected and aloof from subordinates. These examples—and others—notwithstanding, most supervisor-subordinate relationships bring a degree of relational history to the table (or desk) during the performance appraisal. And this history directly influences the attributions formed by both parties before, during, and after the appraisal interview.

As an example of the effects of relational history, consider the finding from this research which indicates that supervisors evaluate subordinates higher when an external excuse is offered than when a concession account with an internal excuse is given. It is likely that the efficacy of this external excuse would be highly influenced by relational history. We might expect a

subordinate with a history of offering external excuses for performance failures to have much less luck convincing a supervisor of the external causality of the performance failures than a subordinate who rarely offers external excuses. This suggests that while external excuses were quite effective the first time they were proffered, as exemplified by this research, their effectiveness might markedly diminish over time. This is, of course, the challenge of giving any account: What works *this time* might not work next time, as both senders and receivers dynamically influence each others' attributions with each interpersonal interaction.

In a similar vein, while the merits of a concession account with an internal excuse paled in comparison to the external excuse account in this experiment, it is not unreasonable to surmise that in the case where there is a supervisor and a subordinate with a good relational history, a concession account with an internal excuse might be extremely effective. However, this scenario is speculation only and is not supported by results from this study.

What this experiment does shed light on is what happens in a performance appraisal when the supervisor brings in no preconceived attributions of the subordinate before the appraisal. To be sure, attributions about the subordinate in this experiment are formed prior to the appraisal through the information given to the supervisors that sets up the scenario, and these initial attributions are captured via survey before and after the appraisal in an attempt to discern attributional shifts due to subordinate account giving. What this experiment does is capture a simulated performance appraisal environment between two individuals who share no prior relational history. This enables us more accurately to comment on how people are predisposed to communicate with one another as opposed to shining light on pervasive communication patterns that occur during performance appraisal interviews. This does not mean that this experiment is salient to interpersonal communication but irrelevant with respect to performance

appraisal. To the contrary, providing a perspective into a performance appraisal without any relational history offers a more transparent look at the fundamental interpersonal communication dynamics involved in account making, attributions, and feedback in this particular context.

Supervisor experience with appraisals.

Another factor that helped to make the supervisor-subordinate communication more transparent was the decision to choose participants for the experiment who had limited experience conducting performance appraisals. In a similar way that the decision to eliminate the effects of relational history removed the supervisors' preconceived attributions about the subordinate, choosing supervisors with limited appraisal experience limited their preconceived notions about the way that performance appraisals should be conducted.

The limitation from this choice is that many supervisors are influenced, to one degree or another, by the unique appraisal protocol endorsed by their employer or organization. Additionally, other factors such as organizational culture and norms, as well as socialization and assimilation processes in organizations serve to influence the manner in which people conduct—and conduct themselves in—performance appraisals. In addition to influencing the way supervisors and subordinates approach appraisals, factors such as culture and norms can influence the attributions formed during the appraisals, as well as also impact postappraisal attributions. For example, an organization that highlights individual accountability and responsibility through its organizational norms and culture might influence subordinates to eschew external excuses during appraisal interviews and also might influence supervisors to be especially critical of excuses offering external causality since externalizing performance failures goes against the core values of the organization. As a result, generalizing the results from this experiment, especially those findings concerning the primacy of external excuses, must also

consider the specific context into which the results are seeking to be generalized. Failure to do so could result in misapplication of the tenets uncovered in this research.

Note card observations.

Offering observations of the subordinate on note cards in an attempt to reproduce observations that a supervisor might make throughout the year was done to provide a reasonable insight into the differences between both positive and negative feedback. Because half of the observations were positive and half of the observations were negative, this research was able to look at differences between supervisors as they determined which observations to comment on and how often they choose to discuss the positive observations in comparison to the negative observations. And, pilot studies conducted without the observations revealed that supervisors would make up the observations in their absence, most likely out of a desire to quantify their feedback with specific examples of employee performance. However, creating these impromptu examples was problematic in the experiment because (1) it introduced an unanticipated level of variation into the appraisals, (2) it was an added stress for the supervisor to come up with the examples, and (3) it caused an awkward and somewhat contrived moment during the appraisal when both the subordinate and the supervisor had to acknowledge and assimilate a new, unexpected, and often extremely salient piece of spontaneous information into the conversation. It is quite likely that failure to offer these note card observations would have led to such a wide range of impromptu "evidence" being contrived that some of the appraisals with the most outlying evidence would have to be excluded from the study.

For all the benefits of using the observations, there was an unintended consequence in that the participants, in some cases, seemed to feel compelled to articulate all of the observations on the note cards during the appraisal interview, regardless of the natural direction of the

conversation. Because of this, there were times when the note card observations appeared more like a burden and a constraint to the supervisors than the helpful contextual aids that they were designed to be. This was most noticeable in the external excuse condition. In a few instances, supervisors would scarcely hesitate after learning that the subordinate's father recently had a stroke (and that was causing the performance failure) before continuing to discuss the negative observations, almost as if the supervisors felt like they were being evaluated exclusively on their ability to cover every observation printed on the note cards. In future experiments using a similar protocol, ensuring that supervisors clearly understand that they can pick and choose which observations to discuss in the performance appraisal would go a long way to ensuring that the observations are seen more as aids for the supervisors than as burdensome during the appraisal interview.

Future Directions

While this experiment reports on a number of interesting and stimulating findings, perhaps equally as exciting are the future possibilities for this line of inquiry. Extending the ideas from above, one area that would be compelling to investigate is an exploration of the amount of positive and negative observations that supervisors might mention if they were given no note card observations or any other specific instances of the subordinate's performance. Analyzing this would involve studying the observations that the supervisors created on the spot to back up their feedback and assessing the valence, frequency, and intensity of those observations.

Perhaps the most natural progression for this line of research is to conduct future studies that explore performance appraisal and account making on the basis of (1) gender and (2) culture. Communication, psychology, and leadership literatures are filled with gender difference

studies, indicating that there might very well be considerable interest, both scholarly and applied, in analyzing the differences and similarities between men and women in performance appraisal contexts.

Conducting this research with an eye for culture differences and similarities would be interesting and potentially quite informative, especially as organizations become more diverse and the likelihood of supervisors and subordinates being from different cultures increases. A logical theoretical start point for developing this research would be Kim and colleagues' (M. S. Kim, 1994; M. S. Kim et al., 1994) work on conversational constraint theory as it applies to different cultures. Her findings (1994) that (a) the perceived importance of clarity is higher in individualistic cultures and (2) the perceived importance of not hurting the receiver's feelings and of minimizing impositions is higher in collectivistic cultures lend themselves to replication and further investigation in a performance appraisal context.

In addition to exploring demographic differences to possibly explain why some people are reluctant to give negative feedback and others are not, other individual differences that might be fruitful to isolate in future research are (1) discomfort with the appraisal process and (2) affinity-seeking.

Research by Villanova, Bernardin, Dahmus, and Sims (1993) developed a scale to measure individual discomfort with performance appraisal situations and they subsequently used the scale to show a positive and relatively stable correlation between individual discomfort with appraisals and leniency. Although the scale was tested by undergraduate students and not actual supervisors, this would be a worthwhile instrument to use in future research exploring possible linkages between the reluctance to communicate negative feedback, discomfort with the appraisal process, and supervisor leniency.

Another individual difference for future study is affinity-seeking. Drawing from the research of Bell and Daly (1984), affinity-seeking refers to the manner in which individuals go about getting others to like them. It would be interesting to explore how affinity-seeking might influence the transmission of negative feedback, speculating that people with a high need to be liked would be less likely to transmit negative feedback to others. Bell, Tremblay, and Buerkel-Rothfuss (1987) devised an instrument which measures the ability to say and do what will engender interpersonal attraction from others as well as the ability to take on roles to be liked by others. Future research might also seek to determine the relationship between these affinity-seeking skills and transmission of negative feedback and subordinate satisfaction with the supervisor's communication during the appraisal interview.

Future studies could also seek to invert the current experimental protocol and have the participants serve as subordinates instead of supervisors and respond to the feedback of a single actor serving as the supervisor. Doing this would allow for a more comprehensive perspective on the subordinate's reactions to the negative feedback as well as the specificity of the feedback and satisfaction with the appraisal interview. Such a role reversal would enable the examination of a vital constituent in the performance appraisal process—the subordinate. Since one of the most important facets of the appraisal interview is to sustain positive performance and remediate performance shortfalls, understanding the impact of the appraisal on the subordinate is critical towards this end. Although this research focused more on the supervisor's communication during the appraisal, future studies employing multiple subordinates and a single supervisor could provide more robust data on subordinates during the appraisal process.

There is reason to believe that a small number of performance appraisal interviews in this experiment were handled very poorly when some interviews on the extreme ends of the spectrum

are considered. For example, such instances included an interview that was conducted in less than five minutes, an interview citing only one negative observation, and an interview citing no positive observations. Given the contextual information provided in this experiment, these interviews are most likely markers of appraisals that were conducted very poorly. In an applied context, conducting these interviews in such a slovenly manner can have devastating personal and organizational consequences. As documented repeatedly throughout this research, giving negative feedback is challenging even in the best of circumstances. When the supervisor is careless or, even worse, malevolent during the appraisal interview, the communication can be tremendously damaging to the relationship between the supervisor and the subordinate, as well as personally damaging to the subordinate.

While it is unknown how many of these extreme cases would disappear if the appraisals were conducted in a field setting, it is prudent to err on the side of caution and point out the distinct possibility that a handful of the participants in this research were unequipped, at the time of the experiment, to professionally conduct a performance appraisal interview. This lack of skill, if not corrected, can leave incredible, irreparable damage and hard feelings in its wake.

It is important to highlight the applied nature of the current experiment because performance appraisal is truly a context where rigorous empirical research can be of great utility to practitioners. The responsibility falls to scholars to ensure that what is studied in the lab and in the field is salient and useful to practitioners. Performance appraisal scholars should keep in the back of their mind the timeless tongue-in-cheek warning from appraisal researcher Robert Wherry (1957, p.1): "We don't know what we are doing, but we are doing it very carefully and hope you are pleased with our unintelligent diligence." His caution is fair warning for appraisal researchers to keep their scholarship practical and applied.

Understanding how individuals communicate to one another during a performance appraisal interview, exploring the influence of account making on supervisors, and determining how positive and negative feedback is given during an appraisal interview are all salient, relevant, and important issues for all parties in the performance appraisal process. Separating good, accurate scholarship from junk science is beneficial to everyone involved in the planning, design, and conduct of performance appraisal interviews. When done correctly, a performance appraisal interview has the potential of starting a virtuous cycle, where communicating candid, honest, and accurate feedback serves to improve another's potential and in doing so creates benefit and positive momentum for the supervisor, the subordinate, and the organization itself.

The mere possibility of helping to create such a virtuous cycle is reason enough to continue this line of applied and relevant research well into the future.

Appendix A

Supervisor Pre-Interview Questionnaire

1. Have you ever conducted, as a supervisor, a formal performance appraisal interview? (NOTE: Military appraisal experience is not applicable in this case)

YES NO

2. If yes, how many formal performance appraisals have you conducted as a supervisor?

_____

(NOTE: If this number is 5 or greater, please see researcher before participating in experiment)

3. Please fill in the following demographic information. This information will remain confidential.

Name:

Gender:

Age:

4. Are you an international student?

YES NO

If yes, what is your nationality? _____

5. Are you prior service or active duty military?

YES NO

If yes, what was / is your branch of service? _____

Appendix B

Institutional Review Board Approved Human Subjects Consent Form

IRB#: 2003-11-0054

Informed Consent to Participate in Research

The University of Texas at Austin

You are being asked to participate in a research study. This form provides you with information about the study. The Principal Investigator (the person in charge of this research) or his/her representative will also describe this study to you and answer all of your questions. Please read the information below and ask questions about anything you don't understand before deciding whether or not to take part. Your participation is entirely voluntary and you can refuse to participate without penalty or loss of benefits to which you are otherwise entitled.

Title of Research Study: Communicating negative feedback in performance appraisal: Issues, impediments, interference, and interventions

Principal Investigator, UT affiliation, and Telephone Number(s):
Principal Investigator: Geoffrey R. Tumlin, Department of Communication Studies, 512 471-5251
Faculty Advisor: Mark L. Knapp, Department of Communication Studies, 512 471-3787

Funding source: There is no external funding source for this research. Compensation for participation in the study is provided by the principal researcher.

What is the purpose of this study? The purpose of this study is to explore communication variables that are at play during the course of a standard performance appraisal interview between a supervisor and a subordinate. Approximately one hundred people will participate in this study.

What will be done if you take part in this research study? If you take part in this study, you will be asked to conduct two performance appraisal interviews. You will be asked to assume the role of a supervisor and conduct a performance appraisal with another study participant who will assume the role of a subordinate. You will be asked to complete two short surveys, one before and one after the performance appraisal interview. You will have the option of conducting as many as two performance appraisal interviews (the subordinate will be the same person each time). At the completion of your last performance appraisal interview, you will be out briefed and your participation in the study will be complete.

What are the possible discomforts and risks? The only foreseeable discomfort or risk during this experiment is the internal discomfort that you might experience while conducting the performance appraisal. You might equate this to the discomfort that you might feel if you were asked to comment on a teammate's contribution to a group project.

In the unlikely event that you have a mental or emotional reaction during the conduct of this experiment, treatment will not be provided. Should you feel the need to talk to someone following the experiment, you can reach a trained University of Texas nurse at 512 475-6877 or a counselor at the National Graduate Student Crisis Line at 877-GRAD-HLP.

What are the possible benefits to you or to others? It is possible that you will gain beneficial experience by conducting these performance appraisals. Additionally, it is hoped that other people can increase their knowledge of the performance appraisal process based on this research.

If you choose to take part in this study, will it cost you anything? There is no cost to participate in this study.

Will you receive compensation for your participation in this study? You will receive $100 for conducting two performance appraisals.

What if you are injured because of the study? There is no physical risk associated with this study.

If you do not want to take part in this study, what other options are available to you? Participation in this study is entirely voluntary. You are free to refuse to be in the study, and your refusal will not influence current or future relationships with The University of Texas at Austin

How can you withdraw from this research study and who should I call if I have questions?

If you wish to stop your participation in this research study for any reason, you should contact: Geoffrey R. Tumlin at (512) 471-5251. You are free to withdraw your consent and stop participation in this research study at any time without penalty or loss of benefits for which you may be entitled. Throughout the study, the researchers will notify you of new information that may become available and that might affect your decision to remain in the study.

In addition, if you have questions about your rights as a research participant, please contact Clarke A. Burnham, Ph.D., Chair, The University of Texas at Austin Institutional Review Board for the Protection of Human Subjects, 512 232-4383.

How will your privacy and the confidentiality of your research records be protected?

Authorized persons from The University of Texas at Austin and the Institutional Review Board have the legal right to review your research records and will protect the confidentiality of those records to the extent permitted by law. If the research project is sponsored then the sponsor will also have the legal right to review your research records. Otherwise, your research records will not be released without your consent unless required by law or a court order.

If the results of this research are published or presented at scientific meetings, your identity will not be disclosed.
The performance appraisal interviews in this study will be videotaped and audiotaped. The cassettes will be coded so that no personally identifying information is visible on them and they will be kept in a secure location. These tapes will be heard or viewed only for research purposes by the investigator and his associates. These tapes will be retained for possible future analyses by the researcher due to the requirements of researcher's academic discipline.

The researcher and his associates may wish to present some of the tapes from this study at scientific conventions or as demonstrations in classrooms. Please sign below if you are willing to allow us to do so with the tape of your performance.

Will the researchers benefit from your participation in this *study?* Beyond publishing or presenting the results, the researchers will not benefit from your participation in this study.

Signatures:

As a representative of this study, I have explained the purpose, the procedures, the benefits, and the risks that are involved in this research study:

_____

Signature and printed name of person obtaining consent **Date**

You have been informed about this study's purpose, procedures, possible benefits and risks, and you have received a copy of this Form. You have been given the opportunity to ask questions before you sign, and you have been told that you can ask other questions at any time. You voluntarily agree to participate in this study. By signing this form, you are not waiving any of your legal rights.

_____

Printed Name of Subject **Date**

_____

Signature of Subject **Date**

_____

Signature of Principal Investigator **Date**

I hereby give permission for the videotape and audiotape made for this research study to be also used for educational purposes.

_____

Printed Name of Subject **Date**

_____

Signature of Subject **Date**

_____

Signature of Principal Investigator **Date**

Appendix C

Instructions for Supervisors

You are being asked to assume the role of a supervisor and conduct a performance appraisal interview. A performance appraisal interview is a formal interaction between a supervisor and a subordinate where the supervisor assesses the subordinate's performance during the last evaluation period, usually one year in length, and provides performance related feedback to the subordinate. Although a written performance appraisal can often complement an appraisal interview, in this case we are only concerned with the oral performance appraisal interview. At the conclusion of these instructions, you will be given a scenario describing the subordinate that you are to conduct the performance appraisal with. All scenarios are based on actual cases from for-profit organizations.

These instructions will be detailed enough so that you will be able to accurately provide performance feedback to the subordinate, but, as with any type of feedback, there is plenty of room for you to exercise your own judgment in conducting the performance appraisal. While you will be role playing, we ask that you do not *act*. You are role playing yourself! Be yourself and conduct yourself as you most likely would in the given situation, and remember that you will be interacting with another "real live" human being.

After you have read the background information about the subordinate that you are appraising, you will be given the time that you need to prepare for the performance appraisal interview. When you are ready, the subordinate will report to you and the performance appraisal will begin. Although the contextual instructions will indicate that the appraisal interview is scheduled for 15 minutes, you may spend either more or less time conducting the interview as

you deem it appropriate. Simply end the appraisal when you feel that the appraisal has reached its end.

Finally, please treat each of the two appraisals that you will be conducting as independent of each other. Although the contextual information will be the same in both cases and although you will be interviewing the same subordinate, please consider each interaction as completely unrelated to the other interview.

At this point, if you have any questions please ask the researcher before conducting the appraisal interviews.

Appendix D

Performance Appraisal Context

You are the senior project manager for a group of hardware engineers at Texicon Innovations, an industry leader in the development and production of cutting-edge technology products. The hardware engineers that you supervise have three primary responsibilities: (1) helping non-technical Texicon employees (e.g. sales and marketing) understand Texicon's products, which sometimes involves your employees accompanying sales representatives for offsite sales meetings (2) troubleshooting and solving product issues with customers, and (3) working in teams with other hardware engineers and Texicon employees to solve large scale integration and implementation problems involving Texicon products.

It is time for you to conduct the annual performance appraisal interview of your subordinates. Of the ten employees that you rate, one of your subordinates, Kelley Jarvis, has had clearly substandard performance during this evaluation period. Kelley's work in the last year has markedly declined in quality, and you are concerned about this performance degradation. You are not exactly sure what the cause of Kelley's performance decline is, but you are certain that you must honestly address the situation during the performance appraisal interview and provide accurate feedback to Kelley. You are concerned about the long-term career implications of Kelley's continued poor performance and you know that failure on Kelley's part to improve performance could result in probation or termination in the future.

It is clear to you that other employees are having much more success than Kelley in performing the same professional duties, so it does not appear that you are asking too much of Kelley. Furthermore, Kelley seems to be failing on a wide variety of tasks—everything from interacting with clients to working with colleagues to meeting deadlines—Kelley's performance is pervasively poor. Finally, you have noticed that this performance degradation is slowly getting worse. In prior years, Kelley's performance was significantly better than this year's performance.

As a general manager, you have some latitude in conducting these performance appraisals. You enjoy considerable support and autonomy from your supervisors to interact with your subordinates as you see fit. Furthermore, you have the authority to take corrective action if you decide that is necessary. Your supervisors believe in the adage "an employee is a reflection of the boss" and you know that you need to honestly address Kelley's performance decline accurately and swiftly or you run the risk of having your supervisors step-in and handle the problem. And that would not he good for you or for Kelley. In short, you need to address this performance problem right now.

Throughout the last year, you have jotted down specific observations about Kelley's performance on note cards—as you do for all of your employees—so that you will be properly prepared for the annual performance appraisal. Please review those observations NOW.

Appendix E

Observations for Kelley Jarvis

Observation 1: Kelley was late to a client meeting. The meeting had been scheduled for over a month but he still arrived 20 minutes late, even though the meeting was in the same building as Kelley's office. His tardiness was embarrassing for the sales team which was trying to win a contract from the client.

Observation 2: Kelley had an argument with a client. The client believed that Texicon sent his company the wrong piece of hardware. Kelley argued that the client had misstated the technical specifications of their server which led to the wrong part being sent. The argument apparently became slightly heated and the client notified you (Kelley's boss) via email about what he perceived as "discourteous" treatment.

Observation 3: Kelley delivered a poor presentation. Kelley was scheduled to give a technical briefing to you (Kelley's boss) and to your boss on the operating system for a new hardware product. Kelley was unprepared for the presentation and he could not answer the majority of your questions about the product and he seemed distracted throughout the briefing.

Observation 4: Kelley stayed late helping a salesperson learn about a product. You (Kelley's boss) observed Kelley staying considerably past the normal time that everyone goes home to help a salesperson understand a new product's features. There was no looming deadline which would have required their staying late to learn about the product; Kelley was simply sharing his expertise with a salesperson who wanted to learn about a new product.

Observation 5: Kelley solved a challenging integration issue. A client was having trouble networking two pieces of hardware together. Another hardware engineer had attempted to fix the problem without success. Kelley went to the client's location and figured out a unique way to solve the problem.

Observation 6: Kelley worked with a team in developing a client proposal. The sales department put together an ad hoc team to develop a client sales pitch and Kelley was the designated hardware engineer. Following the sales proposal to the client, the sales manager in charge of the ad hoc team sent you an email that was complimentary to Kelley and his role on the team.

Appendix F

Supervisor Exit Survey

1. Do you feel that each observation you were given on note cards represented a distinct, different observation about Kelley's behavior and performance?

(please circle one)

YES NO

If no, please explain:

2. Do you feel that each observation that you were given on note cards had a similar level of intensity to the other observations?

(please circle one)

YES NO

If no, please explain:

Appendix G

Instrument for Assessing Intensity Level and Distinctiveness of Subordinate Observations (Negative)

Below you will find 3 observations about the work performance of an employee named Kelley who works at a company called Texicon. These 3 observations are part of a larger research project and your opinion is being sought to validate two questions about these observations: (1) Are the 3 observations similar in their intensity level? (2) Do the 3 observations report different incidents?

To answer the questions about intensity level, think of an intensity scale for these observations between 1 and 5 in your head where 1 is trivial and not very noteworthy and 5 is extremely and severely intense. For example, forgetting to send a thank you note might be a 1 (trivial; not very noteworthy; low intensity) but killing two people in car wreck would be a 10 (extremely intense).

To answer the questions about each observation reporting a different incident, you need to ask yourself if any of the 3 observations reports a similar incident to another. For example, if one observation reported that Kelley was late to a meeting and another observation reported that Kelley was late to a sales presentation, both of these observations would be highlighting a similar incident (Kelley late to something). As another example, if Kelley says something unkind to a coworker in one observation and Kelley says something unkind to a client in another, this would be reporting a similar incident (Kelley saying something unkind to someone). As a final example, if Kelley was late to a meeting in one observation and said something unkind to a client in another observation, this would be NOT be reporting a similar incident but rather it would be reporting a different incident.

Please note that all three of these observations report negative incidents. THIS FACT ALONE DOES NOT MAKE THEM SIMILAR TO EACH OTHER. Rather, you are being asked to decide if the incidents being reported are different from each other, not if there is an underlying theme of poor performance in the observations.

With the above explanations of intensity and differences in your mind, please read the 3 observations at this time and answer the questions at the end of the each observation.

Observation 1: Kelley was late to a client meeting. The meeting had been scheduled for over a month but he still arrived 20 minutes late, even though the meeting was in the same building as Kelley's office. His tardiness was embarrassing for the sales team which was trying to win a contract from the client.

Use the scale below to rate the intensity of Observation 1:

1 being "trivial and not very intense" [example: forgetting to send a thank you note]
2 being "somewhat noteworthy and low intensity"
3 being "noteworthy and medium intensity"
4 being "high intensity"
5 being "extremely and severely intense" [example: killing two people in a car wreck]

Your rating for Observation 1: _____

Observation 2: Kelley had an argument with a client. The client believed that Texicon sent his company the wrong piece of hardware. Kelley argued that the client had misstated the technical specifications of their server which led to the wrong part being sent. The argument apparently became slightly heated and the client notified you (Kelley's boss) via email about what he perceived as "discourteous" treatment.

Use the scale below to rate the intensity of Observation 2:

1 being “trivial and not very intense” [example: forgetting to send thank you note]
2 being “somewhat noteworthy and low intensity”
3 being “noteworthy and medium intensity”
4 being “high intensity”
5 being "extremely and severely intense” [example: killing two people in a car wreck]

Your rating for Observation 2: _____

Observation 3: Kelley delivered a poor presentation. Kelley was scheduled to give a technical briefing to you (Kelley's boss) and to your boss on the operating system for a new hardware product. Kelley was unprepared for the presentation and he could not answer the majority of your questions about the product and he seemed distracted throughout the briefing.

Use the scale below to rate the intensity of Observation 3:

1 being “trivial and not very intense” [example: forgetting to send thank you note]
2 being “somewhat noteworthy and low intensity”
3 being “noteworthy and medium intensity”
4 being “high intensity”
5 being "extremely and severely intense” [example: killing two people in a car wreck]

Your rating for Observation 3: _____

2. If you feel that any of these observations have very different levels of intensity (more than one point difference between any two), please indicate which observation(s) are different in their levels of intensity and why you feel this way (leave blank if not applicable):

3. Do these three observations each report a different, distinct observation? (Note: recall from the instructions above that the fact that the three observations are negative does not by itself indicate similarity. This question is seeking to determine if each observation is reporting a different, distinct incident.)

Please indicate "YES, DIFFERENT" or "NO, SAME" here: _____

4. If you answered no, please explain which observations are reporting indistinct incidents in the space below (leave blank if you answered yes):

Appendix H

Instrument for Assessing Intensity Level and Distinctiveness of Subordinate Observations (Positive)

Below you will find 3 additional observations about the work performance of an employee named Kelley who works at a company called Texicon. While the first three observations (from the previous email) were negative in nature, these three observations are positive in nature. Your opinion is being sought to validate two questions about these observations: (1) Are the 3 observations similar in their intensity level? (2) Do the 3 observations report different incidents?

To answer the questions about intensity level, think of an intensity scale for these observations between 1 and 5 in your head where 1 is trivial and not very noteworthy and 5 is extremely and severely intense. For example, remembering to send a thank you note might be a 1 (trivial; not very noteworthy; low intensity) but personally saving the lives of two people involved in car wreck would be a 10 (extremely intense).

To answer the questions about each observation reporting a different incident, you need to ask yourself if any of the 3 observations reports a similar incident to another. For example, if one observation reported that Kelley was early to a meeting and another observation reported that Kelley was early to a sales presentation, both of these observations would be highlighting a similar incident (Kelley early to something). As another example, if Kelley says something kind to a coworker in one observation and Kelley says something kind to a client in another, this would be reporting a similar incident (Kelley saying something kind to someone). As a final example, if Kelley was early to a meeting in one observation and said something kind to a client in another observation, this would be NOT be reporting a similar incident but rather it would be reporting a different incident.

Please note that all three of these observations report positive incidents. THIS FACT ALONE DOES NOT MAKE THEM SIMILAR TO EACH OTHER. Rather, you are being asked to decide if the incidents being reported are different from each other, not if there is an underlying theme of good performance in the observations.

With the above explanations of intensity and differences in your mind, please read the 3 observations at this time and answer the questions at the end of the each observation.

Observation 1: Kelley stayed late helping a salesperson learn about a product. You (Kelley's boss) observed Kelley staying considerably past the normal time that everyone goes home to help a salesperson understand a new product's features. There was no looming deadline which would have required their staying late to learn about the product; Kelley was simply sharing his expertise with a salesperson who wanted to learn about a new product.

Use the scale below to rate the intensity of Observation 1:

1 being "trivial and not very intense" [example: remembering to send a thank you note]
2 being "somewhat noteworthy and low intensity"
3 being "noteworthy and medium intensity"
4 being "high intensity"
5 being "extremely and severely intense" [example: saving two people's lives in a car wreck]

Your rating for Observation 1: _____

Observation 2: Kelley solved a challenging integration issue. A client was having trouble networking two pieces of hardware together. Another hardware engineer had attempted to fix the problem without success. Kelley went to the client's location and figured out a unique way to solve the problem.

Use the scale below to rate the intensity of Observation 2:

1 being “trivial and not very intense” [example: remembering to send a thank you note]
2 being “somewhat noteworthy and low intensity”
3 being “noteworthy and medium intensity”
4 being “high intensity”
5 being "extremely and severely intense” [example: saving two people’s lives in a car wreck]

Your rating for Observation 2: _____

Observation 3: Kelley worked with a team in developing a client proposal. The sales department put together an ad hoc team to develop a client sales pitch and Kelley was the designated hardware engineer. Following the sales proposal to the client, the sales manager in charge of the ad hoc team sent you an email that was complimentary to Kelley and his role on the team.

Use the scale below to rate the intensity of Observation 3:

1 being “trivial and not very intense” [example: remembering to send a thank you note]
2 being “somewhat noteworthy and low intensity”
3 being “noteworthy and medium intensity”
4 being “high intensity”
5 being "extremely and severely intense” [example: saving two people’s lives in a car wreck]

Your rating for Observation 3: _____

2. If you feel that any of these observations have very different levels of intensity (more than one point difference between any two), please indicate which observation(s) are different in their levels of intensity and why you feel this way (leave blank if not applicable):

3. Do these three observations each report a different, distinct observation? (Note: recall from the instructions above that the fact that the three observations are positive does not by itself indicate similarity. This question is seeking to determine if each observation is reporting a different, distinct incident.)

Please indicate "YES, DIFFERENT" or "NO, SAME" here: _____

4. If you answered no, please explain which observations are reporting indistinct (same) incidents in the space below (leave blank if you answered yes):

5. Finally, think back to the previous 3 observations from your earlier questionnaire. Do all 6 observations (previous 3 plus the 3 on this document) report different, distinct observations?

Please indicate "YES, DIFFERENT" or "NO, SAME" here: _____

6. If you answered no, please explain which observations are reporting indistinct (same) incidents in the space below (leave blank if you answered yes):

Appendix I

Supervisor Preappraisal Interview Form

After reviewing the observations on the note cards but BEFORE conducting the performance appraisal, please answer the questions below:

Based on the above information and the observations on the note cards, CIRCLE YOUR ASSESSMENT of Kelley's performance during this past evaluation period on the following scale:

1- Strongly Disagree
2- Disagree
3- Somewhat Disagree
4- Neutral
5- Somewhat Agree
6- Agree
7- Strongly Agree

1. I feel that Kelley is a valuable employee.
   1 2 3 4 5 6 7
2. I am pleased with Kelley's quality of work.
   1 2 3 4 5 6 7
3. Kelley is successfully performing all work duties.
   1 2 3 4 5 6 7
4. The quality of Kelley's work is consistently good.
   1 2 3 4 5 6 7
5. Kelley's work performance is adequate.
   1 2 3 4 5 6 7
6. Kelley's is demonstrating sufficient personal effort on work-related tasks.
   1 2 3 4 5 6 7
7. Kelley's performance decline is likely due to lack of ability.
   1 2 3 4 5 6 7
8. Kelley's current level of performance is acceptable to me.
   1 2 3 4 5 6 7
9. It will be fine if Kelley's performance does not change.
   1 2 3 4 5 6 7
10. Kelley is a model employee.
    1 2 3 4 5 6 7

Appendix J

Supervisor Postappraisal Interview Form Part 1

AFTER you have conducted the performance appraisal interview, please circle the number that best describes how you feel in response to the statement below.

1- Strongly Disagree
2- Disagree
3- Somewhat Disagree
4- Neutral
5- Somewhat Agree
6- Agree
7- Strongly Agree

1. I was satisfied with my communication during performance appraisal interview.
   1 2 3 4 5 6 7
2. I was satisfied with the way that I provided feedback to the subordinate during the appraisal.
   1 2 3 4 5 6 7
3. I was too vague in delivering the feedback.
   1 2 3 4 5 6 7
4. I made it clear to the subordinate how I felt about their performance.
   1 2 3 4 5 6 7
5. I felt like I was being overly cautious and sensitive while delivering the feedback.
   1 2 3 4 5 6 7
6. I was not specific enough in delivering the feedback.
   1 2 3 4 5 6 7
7. I felt like I was too harsh and critical in delivering feedback during the appraisal.
   1 2 3 4 5 6 7
8. I felt like I was too lenient in delivering feedback during the performance appraisal.
   1 2 3 4 5 6 7

Appendix K

Supervisor Postappraisal Interview Form Part 2

Based on the information that you learned DURING the performance appraisal interview, CIRCLE YOUR ASSESSMENT of Kelley's performance during this past evaluation period on the following scale:

1- Strongly Disagree
2- Disagree
3- Somewhat Disagree
4- Neutral
5- Somewhat Agree
6- Agree
7- Strongly Agree

1. I feel that Kelley is a valuable employee.
   1 2 3 4 5 6 7
2. I am pleased with Kelley's quality of work.
   1 2 3 4 5 6 7
3. Kelley is successfully performing all work duties.
   1 2 3 4 5 6 7
4. The quality of Kelley's work is consistently good.
   1 2 3 4 5 6 7
5. Kelley's work performance is adequate.
   1 2 3 4 5 6 7
6. Kelley is demonstrating sufficient personal effort on work-related tasks.
   1 2 3 4 5 6 7
7. Kelley's performance decline is likely due to lack of ability.
   1 2 3 4 5 6 7
8. Kelley's current level of performance is acceptable to me.
   1 2 3 4 5 6 7
9. It will be fine if Kelley's performance does not change.
   1 2 3 4 5 6 7
10. Kelley is a model employee.
    1 2 3 4 5 6 7
11. Do you feel that you gave Kelley all of the negative feedback that you needed to deliver based on Kelley's performance?
    YES NO
12. I compared the subordinate's performance to the performance of others during the appraisal.
    YES NO
13. I indicated that the subordinate had failed on a number of other tasks during the appraisal.
    YES NO
14. I compared the subordinate's current performance to their past performance during the appraisal.
    YES NO

Please use the space below to comment on any other aspect of your communication during the appraisal interview:

Appendix L

Subordinate Feedback From

Instructions: Please circle the number that best describes how you feel in response to the statement.

1- Strongly Disagree
2- Disagree
3- Somewhat Disagree
4- Neutral
5- Somewhat Agree
6- Agree
7- Strongly Agree

1. I was satisfied with the supervisor's communication during performance appraisal interview.
1 2 3 4 5 6 7
2. I was satisfied with the way that the supervisor provided feedback to me during the appraisal.
1 2 3 4 5 6 7
3. The supervisor was too vague in delivering the feedback.
1 2 3 4 5 6 7
4. The supervisor made it clear to me how they felt about my performance.
1 2 3 4 5 6 7
5. I felt like the supervisor was being overly cautious and sensitive while delivering the feedback.
1 2 3 4 5 6 7
6. The supervisor was not specific enough in delivering the feedback.
1 2 3 4 5 6 7
7. I felt like the supervisor was too harsh and critical in delivering feedback during the appraisal.
1 2 3 4 5 6 7
8. I felt like the supervisor was too lenient in delivering feedback during the performance appraisal.
1 2 3 4 5 6 7
9. I felt like the supervisor was very effective in conducting the performance appraisal.
1 2 3 4 5 6 7
10. The supervisor was very skilled at delivering performance feedback.
1 2 3 4 5 6 7
11. The supervisor compared my performance to the performance of others during the appraisal.
YES NO
12. The supervisor indicated that I had failed on a number of other tasks during the appraisal.
YES NO
13. The supervisor compared my current performance to my past performance during the appraisal.
YES NO

Appendix M

Instructions for the Subordinate (Actor)

The actor will be trained by the researcher prior to their participation in the study. During the training, subordinates will receive detailed instructions about the way they are to make accounts and communicate during the performance appraisal. This document serves as a synopsis of the instructions that will be given to the actor and this document will be made available to the actor for reference.

General instructions to the actor:

Thank you for participating in this research experiment. This document serves as a written reminder of the training that we conducted in preparation for the experiment. To reiterate a point from our training session: Your interaction with the supervisor during the performance appraisal is essential to the success of this experiment. Failure to follow the protocol of the experiment will result in unusable data from your performance appraisal interview. The principles explained below clearly spell out your role within this experiment so that every appraisal conducted will result in the collection of usable data.

As explained in the training session, there are only two types of subordinates: (1) subordinates offering external excuses and (2) subordinates offering internal excuses with concession accounts. You will only be role playing one of these types, but it is essential that you are familiar with both types of excuses so you do NOT offer external excuses if you are assigned to the internal excuse type and vice versa. Remember, consistency is the key. You will be role playing the same subordinate in every performance appraisal that you conduct. Below is a list of do's and don'ts that applies in all conditions.

DO

1. Act the same in every performance appraisal. Be as consistent as possible from one appraisal to the next. Wide variation in your behavior from one interview to the next will result in unusable data.
2. Follow the conversational lead of the supervisor.
3. Allow the supervisor to begin and end the interview.
4. Be as natural as possible. Although you are role playing, try to convey authenticity and sincerity in your responses.
5. Remain attentive and actively engaged from the beginning to the end of the appraisal.

DO NOT

1. Do not initiate conversation or make small talk with the supervisor. If the supervisor makes small talk, respond as you normally might but do not seek to extend the conversation.
2. Do not try to draw out conversation from the supervisor. Some supervisors will talk more than others, and that's fine.
3. Do not show signs of boredom or disinterest in the appraisal. Remember that performance appraisals have enormous consequences for subordinates and act accordingly.

Instructions for the external excuse:

The central reason that you will use to explain your performance failure is: *You father recently had a serious stroke and is bedridden. Your father's recovery prognosis is uncertain;*

anything ranging from a full recovery to continued paralysis is possible. You do not know how your father's recovery will progress. It is not likely that you will be probed extensively about this personal issue. However, if a supervisor asks for more information, you can simply repeat the nature of your father's debilitating stroke. It is critical that you DO NOT apologize for your poor work performance nor do you admit that your performance failure was due to any lack of effort or ability on your part. Also, DO NOT tell the supervisor that you will do better during the next performance period (after all, you cannot promise this because you are not certain about your father's recovery).

Instructions for the concession with internal excuse:

When questioned about your performance failure, you will reply that: Y*ou are very sorry for the performance failure and you will work harder, starting immediately, to improve your performance. You simply did not properly calculate the amount of effort that you needed to expend to be successful during the last evaluation period.* If the supervisor presses you for a reason why you did not expend enough effort, simply reply that you misjudged how much effort would be necessary to have performance success. In other words, the crux of your account for the performance failure is that you will (1) admit and apologize for the performance failure (2) promise better performance starting immediately and (3) tell the supervisor that you simply did not realize that you were not expending enough effort to be successful.

Key to your role playing will be in not "overselling" the apology. Do not excessively apologize for the performance failure and do not express extreme grief and remorse over the performance failure. Doing this might make it excessively difficult for the supervisor to offer

you feedback of any kind, since overemphasis on guilt or remorse will likely influence how the supervisor communicates with you.

Guidance to the actor for possible appraisal circumstances:

1. *What if the supervisor repeatedly asks me why my performance was so poor?* Reiterate the answer that you have given before. If you are role playing the external excuse, repeat your comments about your father's stroke. If you are role playing the internal excuse with concession, tell the supervisor that you simply did not understand how much effort would be required to perform to standard.

2. For the internal excuse: *But I've been performing solidly for the past few years so how can I say that I simply didn't understand how much effort I needed to expend?* If the supervisor brings this up, explain that you are involved in different projects this year. Without being too defensive, point out to the supervisor that, just like in every job, responsibilities change from year to year. This year you misjudged how much effort you needed to expend to be successful.

3. For the internal excuse: *What if the supervisor asks me personal question, such as asking me if everything's fine at home?* It is critical that if a supervisor goes "looking" for other reasons to explain your performance failure than the ones that you give that you direct the supervisor back to the reasons that you have been communicating. Therefore, answer all probing questions so that the supervisor does not come to believe that there are other causes for your performance failure.

4. *What if the supervisor gets angry with me and starts yelling?* Although this is not likely, it is possible. If the supervisor gets upset, stay in character. If you are offering an

external excuse, tactfully repeat the stroke story. If you are giving an internal excuse with concession, offer a statement of remorse to diffuse the anger.

Appendix N

Subordinate (Actor) Reliability Instrument

Instructions for participants: For the next few minutes, you will be watching short video clips (approximately one minute each) of six performance appraisal interviews. In all six interview segments, the subordinate is being role played by the same actor. He appears on the right side of your screen in each interview. For this questionnaire, you are only concerned with the subordinate. Please focus your attention primarily on this person as you watch each interview segment.

You will be asked to answer the five questions **ABOUT THE SUBORDINATE** after you watch each interview segment. So, the sequence will be (a) watch interview segment #1 (b) answer questions (c) watch interview segment #2 (d) answer questions. This sequence will repeat until you have watched all 6 interview segments.

Please start the tape and watch interview segment #1 NOW.

After watching interview segment #1, please answer the following questions ABOUT THE SUBORDINATE'S communication, behavior, and appearance in interview segment #1. In each case, circle the appropriate number that indicates your response.

What is your impression about the subordinate's formality of dress (clothing)?

(not formal) 1 2 3 4 5 (very formal)

What is your impression about the subordinate's level of fatigue?

(not fatigued) 1 2 3 4 5 (very fatigued)

What is your impression about the subordinate's level of attentiveness?

(not attentive) 1 2 3 4 5 (very attentive)

What is your impression about the subordinate's level of dynamism?

(not dynamic) 1 2 3 4 5 (very dynamic)

What is your impression about the subordinate's level of composure?

(not composed)1 2 3 4 5 (very composed)

Thank you for answering the questions. **Please watch interview segment #2 NOW.**

After watching interview segment #2, please answer the following questions ABOUT THE SUBORDINATE'S communication, behavior, and appearance in interview segment #2. In each case, circle the appropriate number that indicates your response.

What is your impression about the subordinate's formality of dress (clothing)?

(not formal) 1 2 3 4 5 (very formal)

What is your impression about the subordinate's level of fatigue?

(not fatigued) 1 2 3 4 5 (very fatigued)

What is your impression about the subordinate's level of attentiveness?

(not attentive) 1 2 3 4 5 (very attentive)

What is your impression about the subordinate's level of dynamism?

(not dynamic) 1 2 3 4 5 (very dynamic)

What is your impression about the subordinate's level of composure?

(not composed)1 2 3 4 5 (very composed)

Thank you for answering the questions. **Please watch interview segment #3 NOW.**

After watching interview segment #3, please answer the following questions ABOUT THE SUBORDINATE'S communication, behavior, and appearance in interview segment #3. In each case, circle the appropriate number that indicates your response.

What is your impression about the subordinate's formality of dress (clothing)?

(not formal) 1 2 3 4 5 (very formal)

What is your impression about the subordinate's level of fatigue?

(not fatigued) 1 2 3 4 5 (very fatigued)

What is your impression about the subordinate's level of attentiveness?

(not attentive) 1 2 3 4 5 (very attentive)

What is your impression about the subordinate's level of dynamism?

(not dynamic) 1 2 3 4 5 (very dynamic)

What is your impression about the subordinate's level of composure?

(not composed)1 2 3 4 5 (very composed)

Thank you for answering the questions. **Please watch interview segment #4 NOW.**

After watching interview segment #4, please answer the following questions ABOUT THE SUBORDINATE'S communication, behavior, and appearance in interview segment #4. In each case, circle the appropriate number that indicated your response.

What is your impression about the subordinate's formality of dress (clothing)?

(not formal) 1 2 3 4 5 (very formal)

What is your impression about the subordinate's level of fatigue?

(not fatigued) 1 2 3 4 5 (very fatigued)

What is your impression about the subordinate's level of attentiveness?

(not attentive) 1 2 3 4 5 (very attentive)

What is your impression about the subordinate's level of dynamism?

(not dynamic) 1 2 3 4 5 (very dynamic)

What is your impression about the subordinate's level of composure?

(not composed)1 2 3 4 5 (very composed)

Thank you for answering the questions. **Please watch interview segment #5 NOW.**

After watching interview segment #5, please answer the following questions ABOUT THE SUBORDINATE'S communication, behavior, and appearance in interview segment #5. In each case, circle the appropriate number that indicates your response.

What is your impression about the subordinate's formality of dress (clothing)?

(not formal) 1 2 3 4 5 (very formal)

What is your impression about the subordinate's level of fatigue?

(not fatigued) 1 2 3 4 5 (very fatigued)

What is your impression about the subordinate's level of attentiveness?

(not attentive) 1 2 3 4 5 (very attentive)

What is your impression about the subordinate's level of dynamism?

(not dynamic) 1 2 3 4 5 (very dynamic)

What is your impression about the subordinate's level of composure?

(not composed)1 2 3 4 5 (very composed)

Thank you for answering the questions. **Please watch interview segment #6 NOW.**

After watching interview segment #6, please answer the following questions ABOUT THE SUBORDINATE'S communication, behavior, and appearance in interview segment #6. In each case, circle the appropriate number that indicated your response.

What is your impression about the subordinate's formality of dress (clothing)?

(not formal) 1 2 3 4 5 (very formal)

What is your impression about the subordinate's level of fatigue?

(not fatigued) 1 2 3 4 5 (very fatigued)

What is your impression about the subordinate's level of attentiveness?

(not attentive) 1 2 3 4 5 (very attentive)

What is your impression about the subordinate's level of dynamism?

(not dynamic) 1 2 3 4 5 (very dynamic)

What is your impression about the subordinate's level of composure?

(not composed)1 2 3 4 5 (very composed)

Thank you for participating in this survey.

Appendix O

Content Analysis Form

1. Did the supervisor:
(A) Compare Kelley's performance to the performance of others
YES NO
(B) Indicate that the subordinate had failed on a number of other tasks
YES NO
(C) Compare Kelley's current performance with Kelley's past performance
YES NO

2. How many times did the supervisor cite the following performance observations:

(A) Observation A (Kelley late to meeting)
(B) Observation B (Kelley argument with client)
(C) Observation C (Kelley unprepared and gave poor presentation to boss)
(D) Observation D (Kelley stayed late helping a salesperson learn about a product)
(E) Observation E (Kelley solved challenging client integration issue)
(F) Observation F (Kelley good job working with a team on a client proposal)

3. In the concession account with internal excuse ONLY, did the supervisor ask the subordinate if there was some other external reason (i.e. something going on in their personal life) to explain the performance failure AFTER the subordinate had already admitted responsibility for the poor performance?

YES NO Not applicable external condition

4. Record total number of negative statements made

5. Record total number of positive statements made

6. Record total number of neutral statements made

7. Record grand total of all statements made

References

Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives, 11*(1), 109-126.

Banks, W. C. (1976). The effects of perceived similarity upon the use of reward and punishment. *Journal of Experimental Social Psychology, 12*, 131-138.

Baron, R. A. (1988). Attributions and organizational conflict: The mediating role of apparent sincerity. *Organizational Behavior and Human Decision Processes, 41*, 111-127.

Baron, R. A. (1993). Criticism (informal negative feedback) as a source of perceived unfairness in organizations: Effects, mechanisms, and countermeasures. In R. Cropanzano (Ed.), *Justice in the workplace: Approaching fairness in human resource management* (pp. 155-170). Hillside, NJ: Erlbaum Associates.

Bass, B. M. (1990). *Bass and Stogdill's handbook of leadership* (Third ed.). New York: Free Press.

Beer, M. (1981). Performance appraisal: Dilemmas and possibilities. *Organizational Dynamics, 9*(3), 24-36.

Bell, R. A., & Daly, J. A. (1984). The affinity-seeking function of communication. *Communication Monographs, 51*, 91-115.

Bell, R. A., Tremblay, S. W., & Buerkel-Rothfuss, N. L. (1987). Interpersonal attraction as a communication accomplishment: Development of a measure of affinity-seeking competence. *Western Journal of Speech Communication, 51*, 1-18.

Benedict, M. E., & Levine, E. L. (1988). Delay and distortion: Tacit influences on performance appraisal effectiveness. *Journal of Applied Psychology, 73*(3), 507-514.

Bernardin, H. J., & Beatty, R. W. (1984). *Performance appraisal: Assessing human behavior at work*. Boston: Kent Publishing.

Bernardin, H. J., & Orban, J. (1990). Leniency effect as a function of rating format, purpose for appraisal, and rater individual differences. *Journal of Business and Psychology, 5*, 197-211.

Bies, R. J. (1987a). Beyond "voice": The influence of decision-maker justification and sincerity on procedural fairness judgments. *Representative Research in Social Psychology, 17*, 3-17.

Bies, R. J. (1987b). The predicament of injustice: The management of moral outrage. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 9, pp. 289-319). Greenwich, CT: JAI Press.

Bies, R. J., & Shapiro, D. L. (1988). Voice and justification: Their influence on procedural fairness judgments. *Academy of Management Journal, 31*, 676-685.

Bies, R. J., & Sitkin, S. B. (1992). Explanation as legitimation: Excuse-making in organizations. In M. L. McLaughlin, M. J. Cody & S. J. Read (Eds.), *Explaining one's self to others: Reason-giving in a social context* (pp. 183-198). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bobko, P. (2001). *Correlation and regression: Applications for industrial organizational psychology and management* (2nd ed.). Thousand Oaks, CA: Sage.

Borman, W. (1978). Exploring the upper limits of reliability and validity in job performance ratings. *Journal of Applied Psychology, 63*, 135-144.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge, England: Cambridge University Press.

Cardy, R. L., & Dobbins, G. H. (1986). Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance. *Journal of Applied Psychology, 71*(4), 672-678.

Cloven, D. H., & Roloff, M. E. (1993). The chilling effect of aggressive potential on the expression of complaints in intimate relationships. *Communication Monographs, 60*, 199-219.

Cody, M. J., & McLaughlin, M. L. (1990). Interpersonal accounting. In H. Giles & P. Robinson (Eds.), *Handbook of language and social psychology* (pp. 227-255). London: Wiley.

Conlee, M. C., & Tesser, A. (1973). The effects of recipient desire to hear on news transmission. *Sociometry, 36*, 588-599.

Coombs, W. T. (1995). Choosing the right words: The development of guidelines for the selection of the appropriate crisis-response strategies. *Management Communication Quarterly, 8*(4), 447-476.

Cross, K. P. (1977). Not can, but will college teaching be improved. *New Directions for Higher Education, 17*, 1-15.

DeCotiis, T., & Petit, A. (1978). The performance appraisal process: A model and some testable propositions. *Academy of Management Review, 3*, 635-646.

DeNisi, A. S. (1996). *Cognitive approach to performance appraisal: A program of research*. New York: Routledge.

Downs, T. M. (1990). Predictors of communication satisfaction during performance appraisal interviews. *Management Communication Quarterly, 3*(3), 334-354.

Dugan, K. W. (1989). Ability and effort attributions: Do they affect how managers communicate performance feedback information? *Academy of Management Journal, 32*(1), 87-134.

Eisenberg, E. M. (1984). Ambiguity as strategy in organizational communication. *Communication Monographs, 51*, 227-242.

Eisenberg, E. M., & Witten, M. G. (1987). Reconsidering openness in organizational communication. *Academy of Management Review, 12*(3), 418-426.

Enzle, M. D., Hansen, R. D., & Lowe, C. A. (1975). Causal attribution in the mixed-motive game. *Journal of Personality and Social Psychology, 31*(1), 50-54.

Fisher, C. D. (1979). Transmission of positive and negative feedback to subordinates: A laboratory investigation. *Journal of Applied Psychology, 64*(5), 533-540.

Fried, Y., Tiegs, R. B., & Bellamy, A. R. (1992). Personal and interpersonal predictors of supervisors' avoidance of evaluating subordinates. *Journal of Applied Psychology, 77*(4), 462-468.

Gibb, J. R. (1973). Defensive communication. In H. J. Leavitt & L. R. Pondy (Eds.), *Readings in managerial psychology* (Second ed.). Chicago: University of Chicago Press.

Gioia, D. A., & Sims, H. P., Jr. (1985). Superior-subordinate attributional and conversational processes. *Academy of Management Proceedings*, 230-235.

Gioia, D. A., & Sims, H. P., Jr. (1986). Cognition-behavior connections: Attribution and verbal behavior in leader-subordinate interactions. *Organizational Behavior and Human Decision Processes, 37*, 197-229.

Gochman, I. R., & Smith, R. E. (1979). Attraction and behavioral outcome as determinants of observer attributions. *Representative Research in Social Psychology, 9*, 114-121.

Green, S. G., & Mitchell, T. R. (1979). Attributional processes of leaders in leader-member interactions. *Organizational Behavior and Human Performance, 23*, 429-458.

Grice, H. P. (1975). Logic and conversation. In P. Cole & M. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.

Hastrof, A., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49*(1), 129-134.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Hines, S. C., Moss, A. H., & McKenzie, J. (1997). Prolonging life or prolonging death: Communication's role in difficult dialysis decisions. *Health Communication, 9*(4), 369-388.

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

Ilgen, D. R. (1971). Satisfaction with performance as a function of the initial level of expected performance and the deviation from expectation. *Organizational Behavior and Human Performance, 6*, 345-361.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology, 64*, 349-371.

Ilgen, D. R., Mitchell, T. R., & Fredrickson, J. W. (1981). Poor performers: Supervisors' and subordinates' responses. *Organizational Behavior and Human Performance, 27*, 386-410.

Ilgen, D. R., Peterson, R. B., Martin, B. A., & Boeschien, D. A. (1981). Supervisor and subordinate reactions to performance appraisal sessions. *Organizational Behavior and Human Performance, 28*, 311-330.

Jablin, F. M. (1979). Superior-subordinate communication: The state of the art. *Psychological Bulletin, 86*, 1201-1222.

Johnson, R. E., Conlee, M. C., & Tesser, A. (1974). Effects of similarity of fate on bad news transmission: A reexamination. *Journal of Personality and Social Psychology, 29*, 644-648.

Jones, S. C. (1966). Some determinates of interpersonal evaluating behavior. *Journal of Personality and Social Psychology, 3*, 397-403.

Kellerman, K. (1992). Communication: Inherently strategic and primarily automatic. *Communication Monographs, 59*, 288-300.

Kellerman, K., & Park, S. P. (2001). Situational urgency and conversational retreat. *Communication Research, 28*(1), 3-47.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.

Kelley, H. H. (1972). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1-26). Morristown, NJ: General Learning Press.

Kikoski, J. F., & Litterer, J. A. (1983). Effective communication in the performance appraisal interview. *Public Personnel Management, 12*, 33-42.

Kim, M. S. (1994). Cross-cultural comparisons of the perceived importance of conversational constraints. *Human Communication Research, 21*(1), 128-151.

Kim, M. S., Sharkey, W. F., & Singelis, T. M. (1994). The relationship between individuals' self-construals and perceived importance of interactive constraints. *International Journal of Intercultural Relations, 18*, 117-140.

Kim, Y. Y., & Miller, K. I. (1990). The effects of attributions and feedback goals on the generation of supervisory feedback message strategies. *Management Communication Quarterly, 4*(1), 6-29.

Kingstrom, P. O., & Mainstone, L. E. (1985). An investigation of the rater-ratee acquaintance and rater bias. *Academy of Management Journal, 28*(3), 841-853.

Knowlton, W. A., & Mitchell, T. R. (1980). Effects of causal attributions on a supervisor's evaluation of subordinate performance. *Journal of Applied Psychology, 65*, 459-466.

Kravitz, D. A., & Balzer, W. K. (1992). Context effects in performance appraisal: A methodological critique and empirical study. *Journal of Applied Psychology, 77*(1), 24-31.

Laird, A., & Clampitt, P. G. (1985). Effective performance appraisal: Viewpoints from managers. *The Journal of Business Communication, 22*(3), 49-57.

Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87*(1), 72-107.

Larson, J. R. (1984). The performance feedback process: A preliminary model. *Organizational Behavior and Human Performance, 33*, 42-76.

Larson, J. R. (1986). Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. *Organizational Behavior and Human Decision Processes, 37*, 391-408.

Larson, J. R. (1989). The dynamic interplay between employees feedback-seeking strategies and supervisors' delivery of performance feedback. *Academy of Management Review, 14*, 408-422.

Larwood, L., & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology, 62*, 194-198.

Liden, R. C., & Mitchell, T. R. (1985). Reactions to feedback: The role of attributions. *Academy of Management Journal, 28*(2), 291-308.

Longenecker, C. O., Sims, H. P., Jr., & Gioia, D. A. (1987). Behind the mask: The politics of employee appraisal. *The Academy of Management Executive, 1*(3), 183-193.

McElroy, J. C. (1982). A typology of attribution leadership research. *Journal of Management Review, 7*(3), 413-417.

McLaughlin, M. L., Cody, M. J., & O'Hair, H. D. (1983). The management of failure events: Some contextual determinants of accounting behavior. *Human Communication Research, 9*(3), 208-224.

McLaughlin, M. L., Cody, M. J., & Read, S. J. (Eds.). (1992). *Explaining one's self to others: Reason-giving in social context*. Hillside, NJ: Lawrence Erlbaum Associates.

Messick, D., & Sentis, K. (1979). Fairness and preference. *Journal of Experimental Social Psychology, 15*(4), 418-434.

Meyer, H. H., Emanuel, K., & French, J. R. P., Jr. (1965). Split roles in performance appraisal. *Harvard Business Review, 43*(1), 123-129.

Mitchell, T. R., Green, S. C., & Wood, R. E. (1981). An attributional model of leadership and the poor performing subordinate. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 3, pp. 197-234). Greenwich, CT: JAI Press.

Mitchell, T. R., & Wood, R. E. (1980). Supervisor's responses to subordinate poor performance: A test of an attributional model. *Organizational Behavior and Human Performance, 25*, 123-138.

Murphy, K. R., Balzer, W. K., Lockhart, M. C., & Eisenman, E. J. (1985). Effects of previous performance on evaluations of present performance. *Journal of Applied Psychology, 70*, 72-84.

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

Nunnaly, J. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Oberg, W. (1972). Make performance appraisal relevant. *Harvard Business Review, 50*(1), 61-67.

O'Neal, E. C., Levine, D. W., & Frank, J. F. (1979). Reluctance to transmit bad news when the recipient is unknown: Experiments in five nations. *Social Behavior and Personality, 7*(1), 39-47.

Pearce, J. L., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology, 71*(2), 211-218.

Pomerantz, A., Fehr, B. J., & Ende, J. (1997). When supervising physicians see patients: Strategies used in difficult situations. *Human Communication Research, 23*(4), 589-615.

Porter, L. W., & Lawler, E. E., III. (1968). *Managerial attitudes and performance*. Homewood, IL: Dorsey.

Ptacek, J. T., & Eberhardt, T. L. (1996). Breaking bad news: A review of the literature. *The Journal of the American Medical Association, 276*(6), 496-503.

Read, S. J. (1992). Constructing accounts: The role of explanatory coherence. In M. L. McLaughlin, M. J. Cody & S. J. Read (Eds.), *Explaining one's self to others: Reason-giving in a social context* (pp. 3-20). Hillsdale, NJ: Lawrence Erlbaum Associates.

Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology, 32*, 850-856.

Rosen, S., & Tesser, A. (1970). On the reluctance to communicate undesirable information: The MUM effect. *Sociometry, 33*, 253-263.

Rosen, S., & Tesser, A. (1972). Fear of negative evaluation and the reluctance to transmit bad news. *Journal of Communication, 22*, 124-141.

Schonbach, P. (1990). *Account episodes: The management or escalation of conflict*. Cambridge, England: Cambridge University Press.

Schoorman, F. D. (1988). Escalation bias in performance appraisals: An unintended consequence of supervisor participation in hiring decisions. *Journal of Applied Psychology, 73*, 58-62.

Snyder, C. R., & Higgins, R. L. (1988). Excuses: Their effective role in the negotiation of reality. *Psychological Bulletin, 104*, 23-35.

Steiner, D. D., & Rain, J. S. (1989). Immediate and delayed primacy and recency effects in performance evaluation. *Journal of Applied Psychology, 74*(1), 136-142.

Stiles, W. B. (1992). *Describing talk* (Vol. 12). Newbury Park, CA: Sage.

Stockford, L., & Bissell, H. W. (1949). Factors involved in establishing a merit rating scale. *Personnel, 26*, 94-116.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica, 94*, 143-148.

Tedeschi, J. T. (1974). Attributions, liking, and power. In T. L. Huston (Ed.), *Foundations of interpersonal attraction* (pp. 183-193). New York: Academic Press.

Tesser, A., & Conlee, M. C. (1973). *Recipient emotionality as a determinant of the transmission of bad news.* Paper presented at the 81st Annual Convention of the APA (Division 8).

Tesser, A., & Rosen, S. (1972). On understanding the reluctance to transmit negative information (the MUM effect): The effects of similarity of objective fate. *Journal of Personality and Social Psychology, 23*, 46-54.

Tesser, A., & Rosen, S. (1975). The reluctance to transmit bad news. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 9, pp. 193-232). New York: Academic Press.

Tesser, A., & Rosen, S. (1977). Why subjects say they would or would not communicate affectively toned messages. *Psychological Reports, 40*, 231-240.

Tesser, A., Rosen, S., & Batchelor, T. (1972a). On the reluctance to communicate bad news (the MUM effect): A role play extension. *Journal of Personality, 40*, 88-103.

Tesser, A., Rosen, S., & Batchelor, T. (1972b). Some message variables and the MUM effect. *Journal of Communication, 22*, 239-256.

Tesser, A., Rosen, S., & Conlee, M. C. (1972). News valence and available recipient as determinants of news transmission. *Sociometry, 35*, 619-628.

Tesser, A., Rosen, S., & Waranch, E. (1973). Communicator mood and the reluctance to transmit undesirable messages. *Journal of Communication, 23*, 266-283.

Tjosvold, D. (1985). The effects of attribution and social context on superiors' influence and interaction with low performing subordinates. *Personnel Psychology, 38*, 361-376.

Tsui, A. S., & Barry, B. (1986). Interpersonal affect and rating errors. *Academy of Management Journal, 29*(3), 586-599.

Villanova, P., Bernardin, H. J., Dahmus, S. A., & Sims, R. L. (1993). Rater leniency and performance appraisal discomfort. *Educational and psychological measurement, 53*, 789-799.

Vroom, V. H. (1964). *Work and motivation*. New York: John Wiley and Sons.

Wagner, R. K. (2002). Smart people doing dumb things: The case of managerial incompetence. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 42-63). New Haven, CT: Yale University Press.

Wagoner, R., & Waldron, V. R. (1999). How supervisors convey routine bad news: Facework at UPS. *Southern Communication Journal, 64*(3), 193-210.

Wayne, S. J., & Ferris, G. R. (1990). Influence tactics, affect, and exchange quality in supervisor-subordinate interactions: A laboratory experiment and a field study. *Journal of Applied Psychology, 75*(5), 487-499.

Wayne, S. J., & Kacmar, K. M. (1991). The effects of impression management on the performance appraisal process. *Organizational Behavior and Human Decision Processes, 48*, 70-88.

Weenig, M. W. H., Groenenboom, A. C. W. J., & Wilke, H. A. M. (2001). Bad news transmission as a function of the definitiveness of consequences and the relationship between communicator and recipient. *Journal of Personality and Social Psychology, 80*(3), 449-461.

Weiner, B. (1992). Excuses in everyday interaction. In M. L. McLaughlin, M. J. Cody & S. J. Read (Eds.), *Explaining one's self to others: Reason giving in a social context* (pp. 131-146). Hillsdale, NJ: Lawrence Erlbaum Associates.

Weiner, B., Amirkhan, J., Folkes, V. S., & Verette, J. A. (1987). An attributional analysis of excuse giving: Studies of a native theory of emotion. *Journal of Personality and Social Psychology, 52*, 316-324.

Weiner, B., Frieze, L., Kukla, A., Reed, L., Rest, S., & Rosenbaum, R. M. (1972). Perceiving the causes of success and failure. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 95-120). Morristown, NJ: General Learning Press.

Weiner, B., & Kukla, A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology, 15*, 1-20.

Wherry, R. J. (1957). The past and future of criterion evaluation. *Personnel Psychology, 10*, 1-5.

Wiemann, J. M. (1981). Effects of laboratory videotaping procedures on selected conversation behaviors. *Human Communication Research, 7*(4), 302-311.

Wood, R. E., & Mitchell, T. R. (1981). Manager behavior in a social context: The impact of impression management on attributions and disciplinary actions. *Organizational Behavior and Human Performance, 28*, 356-378.

Zorn, T. E. (1995). Bosses and buddies: Constructing and performing simultaneously hierarchical and close friendship relationships. In J. T. Wood & S. Duck (Eds.), *Under-studied relationships: Off the beaten path* (Vol. 6, pp. 122-147). Thousand Oaks, CA: Sage.

Zuckerman, M. (1979). Attributions of success and failure revisited, or: The motivational bias is alive and well in attribution theory. *Journal of Personality, 47*, 245-287.

VITA

Geoffrey Roger Tumlin was born in Houston, Texas on November 15, 1971, the son of Daphne Belbin Tumlin and Ronald Wayne Tumlin. After completing his work at Eisenhower High School, Houston, Texas, in 1990, he entered the United States Military Academy at West Point, New York. Upon graduation from West Point in 1994, he was commissioned as an infantry officer in the United States Army, serving in Georgia and in Hawaii. Following his military service, he entered the Graduate School of The University of Texas where he earned a Master of Arts in Communication Studies in December 2000.

Permanent Address: 12445 Alameda Trace Circle #236, Austin, Texas 78727-6478

This dissertation was typed by the author.