ASSESSING THE IMPACT OF AN OUTCOME BIAS AND AN ORDER EFFECTS BIAS ON PERFORMANCE EVALUATIONS OF FIELD SALES PERSONNEL BY SALES MANAGERS

By

GREGORY W. MARSHALL

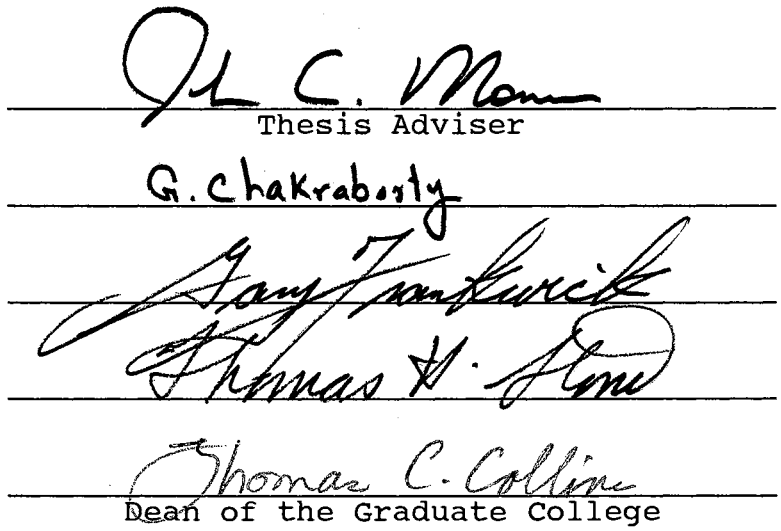
Bachelor of Science
in Business Administration
The University of Tulsa
Tulsa, Oklahoma
1978

Master of Business Administration
The University of Tulsa
Tulsa, Oklahoma
1983

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 1993

ASSESSING THE IMPACT OF AN OUTCOME BIAS
AND AN ORDER EFFECTS BIAS ON
PERFORMANCE EVALUATIONS OF
FIELD SALES PERSONNEL
BY SALES MANAGERS

Thesis Approved:



Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGEMENTS

Several people deserve special recognition for their role in this dissertation and in my doctoral program at Oklahoma State University. First, I would like to express my thanks to Dr. John C. Mowen, my program and dissertation committee chairperson, for his guidance throughout the process. Dr. Mowen is a mentor in the truest sense of the word, and I aspire to perform that role as well myself some day.

Second, my appreciation goes to the other members of my dissertation committee, Dr. Goutam Chakraborty, Dr. Gary L. Frankwick, and Dr. Thomas H. Stone, for all their valuable assistance with the dissertation. Their comments and suggestions have resulted in a document that is far better than what it would have been without their contributions.

Third, thanks goes to Dr. Steven J. Miller and Dr. Joshua L. Wiener, both of whom taught me a great deal about research and about marketing. Likewise, Dr. Jerry R. Goolsby deserves much credit for his commitment that the marketing doctoral students at Oklahoma State get the best training possible. And, I am grateful to Dr. Clifford E. Young, my first faculty contact at Oklahoma State, for championing my admission into the doctoral program.

Fourth, to Mr. Keith A. Buchanan, Ms. Linda Gamble, Mr.

Joseph W. Masse, and Ms. Rosemary E. Masse, all of the national consumer products sales company that provided subjects for the research, you have my sincerest gratitude. I hope this research provides you and others in sales management with helpful insights.

Finally, I want to recognize my wife Patti for all her patience and support throughout the doctoral program. We were married two weeks prior to the start of my first semester, and the fact that after four years of a doctoral program we are still the best of friends is a testament to her commitment and understanding.

To all who played a role in this project and in my doctoral program, I will always be grateful.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

> *We know very little about the behavior and characteristics of salespeople and even less about the people who manage them.*
>
> Richard P. Bagozzi
> Sloan Management Review
> (Winter 1980, p. 15)

In the years since Bagozzi's (1980a) comment, the sales manager has presided over some radical changes in selling, and has been under a microscope professionally as never before (O'Connell and Keenan 1990). Changes in technology are affecting personal selling, such as just-in-time production and purchasing, sophisticated computer-based materials planning, availability of scanner data by channel partners, and the growing concentration of buying points for companies selling in many industries. Other changes are arising out of the way goods and services are sold (e.g., new distribution options, telemarketing, and computer-aided target marketing). These changes put the sales manager in a prime position to directly affect their company's bottom line based upon the way the sales team is managed.

This influence by sales managers on the bottom line is accentuated by the fact that selling costs in general have spiraled in recent years to an estimated median cost range

of $86 - $228 (depending upon industry type) for a single sales call, according to the publication Sales and Marketing Management's "1990 Survey of Selling Costs" (February 26, 1990, p. 8). Typical costs to recruit, hire, and train a salesperson to production status can range from $10 thousand to $50 thousand, with average compensation for an experienced salesperson usually exceeding $45 thousand a year (Ingram and LaForge 1992, p. 22). Clearly, the rapidly changing environment and rising costs of the personal selling effort signal a need for an increased emphasis on research designed to improve salesforce performance.

In a keynote address to members of the 1992 American Marketing Association Faculty Consortium on Personal Selling and Sales Management, Professor David Cravens challenged a national audience of sales researchers to respond to the rapidly increasing need for an expansion of the knowledge base of sales management. The suggested topical research priorities he listed for the next decade included the following:

1) management control systems - especially behavioral versus outcome control systems, as exemplified by the Anderson and Oliver (1987) model.
2) role of the field sales manager in the success of sales organizations.
3) salesforce performance systems.

Another important theme emerged at the Faculty Consortium: a call for utilizing a wider variety of

research methodologies to investigate salesforce performance. In the past, virtually all studies have employed a survey design, and most have relied exclusively upon correlational analyses. Such a unidimensional methodological approach points out a key weakness in the stream of research on sales management. This weakness was verified in the literature review undertaken within this dissertation, which indicated that out of the many published studies in salesforce performance, less than a dozen have employed an experimental design.

As a follow-up to the Faculty Consortium, in its Fall 1992 issue the Journal of Personal Selling and Sales Management published an agenda for improving personal selling and sales management research that included the following consensus conclusions of Consortium attendees:

1) Move beyond large descriptive models to more micro-models, especially in sub-functional areas such as evaluation of performance.
2) Focus more on programmatic research.
3) Draw from theoretical perspectives of other disciplines to supplement sales and sales management research.
4) Give more consideration to the sales situation, especially salesperson/situational context interaction.
5) Be more open to different methods of gathering and analyzing data.
6) Incorporate experimental design into future studies.
7) Pay more attention to issues of internal validity.

Research Questions

This dissertation specifically seeks to meet the challenge for sales management research in the 1990's as outlined above. It employs an experimental design to investigate the impact of two judgmental biases drawn from the field of behavioral decision theory that may influence salesforce performance evaluations: an outcome bias and an order effects bias. The research questions addressed are as follows:

1) Do field sales managers exhibit an outcome bias when evaluating sales personnel?
2) Does the order of presentation of performance information affect managerial ratings of sales personnel?

Two separate experiments were conducted in a field setting, one to investigate each type of bias.

Purposes of the Dissertation

The dissertation has four purposes: 1) the development of a comprehensive literature review on salesforce performance, utilizing the Walker, Churchill, and Ford (1977) model as a pedagogical framework for the review; 2) an empirical test for an outcome bias in managerial ratings of sales personnel; 3) an empirical test for an order effects bias in managerial ratings of sales personnel; and 4) an answer to the call for alternative research methods in investigating salesforce performance through the use of two

experiments. Each of these purposes will now be discussed in more detail.

Literature Review on Salesforce Performance

One purpose of the dissertation is to provide a comprehensive literature review utilizing an adaptation of the Walker, Churchill, and Ford (1977) salesforce performance model to organize the literature. The adapted Walker, Churchill, and Ford model is presented later in Chapter II as Figure 1.

This comprehensive review of the salesforce performance literature is accomplished for two reasons:

1) It is important to show where and how the present research in the area of judgmental biases on salesforce performance evaluation fits into the most widely accepted paradigm for salesforce performance research. By reviewing the overall literature in salesforce performance we may gain a greater appreciation of the gaps in knowledge as well as the potential contributions of the present research.

2) No such systematic explication of the salesforce performance research subsequent to the development of the Walker, Churchill, and Ford (1977) model has been offered in the literature. The author believes there is inherent value in a thorough review of the salesforce performance literature using the model to organize, categorize, and present the previous research

findings. Thus, such a literature review is a key purpose of this dissertation.

Outcome Bias

A second purpose of this dissertation is to examine the issue of managerial utilization of outcome versus decision appropriateness information in evaluating sales personnel. From a review of the extant literature related to the differential usage of outcome versus other measures of salesforce performance (c.f., Behrman and Perreault 1982; Jackson, Keith, and Schlacter 1983; Anderson and Oliver 1987; and Morris et al. 1991), as well as from my own experience as a consumer products sales manager, an empirical question emerges as to how sales managers differentially use information about sales results in conjunction with information about the quality of the decision making by the salesperson in achieving those results when managers rate sales personnel. Literature from the field of behavioral decision theory describes the potential for an "outcome bias" to occur in which raters take into account outcomes in a manner that is irrelevant to the appropriateness of the decisions made by a ratee (c.f., Baron and Hershey 1988; Hawkins and Hastie 1990).

The susceptibility of salesforce performance ratings to an overreliance on outcomes has also been noted in the sales management literature. Based upon a meta-analysis of 116 articles on salesforce performance, Churchill et al. (1985)

identified the tendency to emphasize outcomes rather than behavioral performance measures as a particularly pervasive problem worthy of further investigation. The implication is that sales organizations end up judging sales personnel based largely upon results for which these salespeople may or may not be able to directly control, a condition labeled by a leading writer in human resources management as the "Achilles heel of our profession" (Heneman 1975, p. 91.)

Thus, the purpose of the first study within this dissertation is to empirically test the impact of an outcome bias on salesforce performance evaluations.

Order Effects Bias

A third purpose of this dissertation is to provide the first empirical test in the salesforce performance literature of a second judgmental bias, an order effects bias. Order effects biases occur when the mere order in which information is presented to an evaluator affects the opinions of the evaluator about the individual or object being rated (Hogarth 1987, p.55). Normatively, one would assume information content would be utilized in judgment without regard to order. However, a rich literature exists based upon information integration theory (c.f., Anderson 1981) indicating that, in a string of information, sometimes earlier informational inputs dominate evaluators' ratings (a primacy effect), and sometimes later informational inputs dominate (a recency effect).

A new theory has been developed by Hogarth and Einhorn (1992) called the "belief-adjustment model" to explain and predict order effects in evaluation. A second study within this dissertation utilized predictions of the belief-adjustment model to test the impact of order of receipt of performance information on managerial evaluations of sales personnel. There are no published studies related to this phenomenon in the salesforce performance literature.

(The reader should note that, while there are dozens of articles in the human resources and organizational behavior literatures about primacy and recency effects, this dissertation makes no attempt to summarize that literature. This is because such a synthesis was exactly the purpose of the Hogarth and Einhorn (1992) integrating paper. The model and accompanying predictions they provide are based upon their own exhaustive review and analysis of that somewhat contradictory prior literature, in an attempt to define the parameters of primacy-based and recency-based predictions. The resulting model is quite parsimonious, and its predictions are the subject of the second study within this dissertation.)

Alternative Research Methods

A fourth purpose of the dissertation is to answer the call for the use of alternative research methods (i.e., methods other than survey research) to investigate salesforce performance. This issue is addressed by

employing two experiments conducted in a field setting with practicing sales managers as subjects. As such, the dissertation meets the dual objectives of contributing to the literature through empirical research that tests specific hypotheses while at the same time contributing through the use of a particular research design that has been underutilized within the domain of salesforce performance.

Organization of the Dissertation

This dissertation consists of six chapters. Chapter I has provided an introduction to the research questions and the purposes of the study. Chapter II reviews the relevant literature in salesforce performance, with particular emphasis given to salesforce performance evaluation. Chapter III provides a theoretical synthesis and a framework for the study by examining in detail the prior literature on the outcome bias and the Hogarth and Einhorn (1992) belief-adjustment model for order effects. The research hypotheses are derived from this theory base, and are presented in Chapter III. Chapter IV details the methodology utilized in the dissertation. For each of the two studies (i.e., outcome bias and order effects bias), the chapter provides a description of the design, stimuli, subjects, procedure, measurements, and methods of data analysis. Chapter V presents the findings of the data analysis. Finally, Chapter VI presents a discussion of the results, limitations

of the research, and implications of the findings for sales management practice.

CHAPTER II

LITERATURE REVIEW

Introduction

The primary goal of this chapter is to demonstrate that while our knowledge of the dynamics of salesforce performance in general has greatly increased since the research stream began in the early 1970's, many opportunities exist for future research within this domain. A broad survey of the literature is presented in order to fully appreciate where the present study fits within the context of existing knowledge, as well as to highlight the contribution made by the present research. Because the focus of the present research is at the evaluation stage of sales management, relatively more detail is presented of the literature relevant to that stage.

In reviewing the various studies on salesforce performance it became apparent that the vast majority of the empirical work in the area has employed a single type of research methodology (field survey), utilizing primarily correlational analyses. The critical need for alternate research methods in salesforce performance as pointed out in Chapter I is accentuated when the existing literature is thoroughly examined.

A caveat is necessary at this point. The present study is cast within the general nomological net of salesforce performance management as originally conceptualized by Churchill, Ford, and Walker in the 1970's. Another important stream of research exists within the marketing literature related to the efficacy of salespeople within the sales encounter (or selling task) itself. Representative of this second literature stream is the work by Saxe and Weitz (1982) on sales-oriented versus customer-oriented selling and research by Weitz (1981), Weitz, Sujan, and Sujan (1986), and Spiro and Weitz (1990) on adaptivity in personal selling effectiveness. While the significant contribution of the sales efficacy literature is explicitly recognized, the present study focuses instead on issues of salesforce performance evaluation. (See Rhoads 1988 for an extensive review of the literature on personal selling efficacy.)

The Walker, Churchill, and Ford (1977) Salesforce Performance Model

A paradigm for studying salesforce performance as a separate field of marketing research was defined by Walker, Churchill, and Ford (1977) in the classic article, "Motivation and Performance in Industrial Selling: Present Knowledge and Needed Research." In the article, the authors presented an adaptation of the Vroomian expectancy model (Vroom 1964), accompanied by a series of research propositions which have served to drive the majority of

scholarly effort in salesforce performance research for the last fifteen years. An adaptation of the Walker, Churchill, and Ford salesforce performance model is presented in Figure 1. (Note: Future references to the model in this dissertation will refer to the "WCF" model.)

The WCF model is grounded in expectancy theory, which is characterized by the view that behavior by an individual is purposeful, based upon conscious intention, and goal-directed. The theory views behavior as a function of an individual's anticipations for the future. In general, expectancy theory postulates that individuals receive input into their decision-making process, the effect of the input on the individual's anticipation of future events is cognitively determined, and motivation is subsequently increased, decreased, or unaffected (Evans, Margheim, and Schlacter 1982).

Putting Vroomian expectancy theory into the terminology of salesforce performance, the expectancy process is posited to operate as follows: 1) The salesperson's motivation to expend effort on a particular task is impacted by his/her level of expectation that such effort will result in improved performance against some performance dimension; 2) subsequently, this achieved level of performance will lead to increased attainment of a particular reward that is desirable and valuable to the salesperson; 3) the implication is that organizations can train salespeople in key activities that clearly lead to desired outcomes; 4)

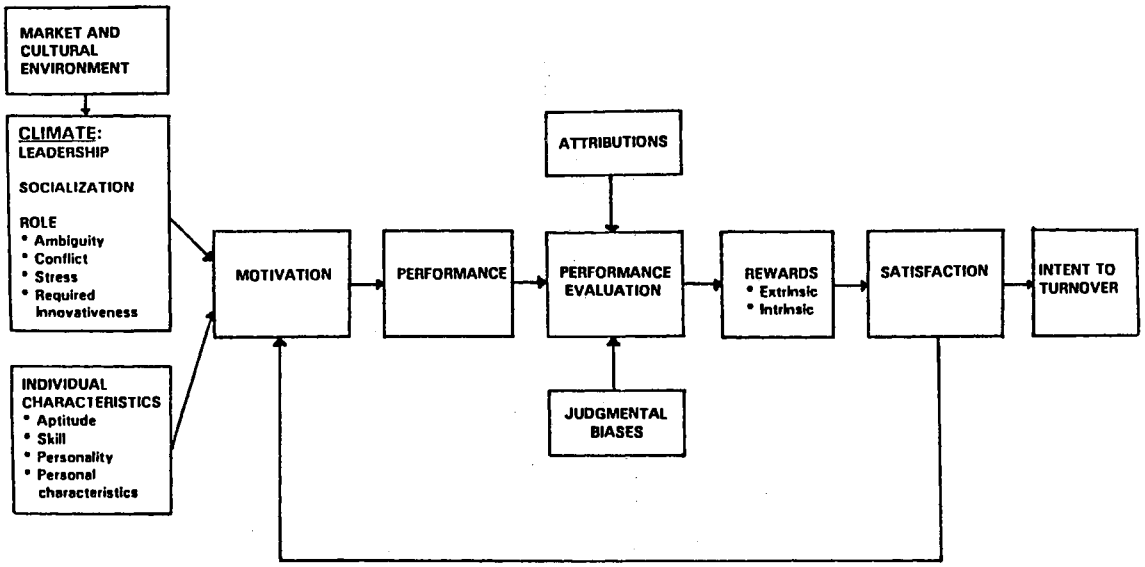


Figure 1. Adaptation of the Walker, Churchill, and Ford (1977) Salesforce Performance Model

then, employees are rewarded for such outcomes (Marshall and Miller 1991). The WCF adaptation of the Vroomian expectancy equation is presented in Figure 2.

Because the overall paradigm for salesforce performance research over the past fifteen years has been an expectancy approach utilizing the WCF model, the review of the key findings across the model will be presented in a format that utilizes the order of the variables in the adapted model presented in Figure 1 as a method of organizing the literature review. It should be noted that to-date the WCF model has never been successfully tested as a <u>whole process</u> within the sales domain, although various authors have called for such a comprehensive empirical test (c.f., Oliver 1974; Teas 1981). Thus, it seems appropriate to break the individual elements in the model apart for examination of the empirical findings related to that element. (Note: Throughout the sections that follow the reader may wish to refer back to Figure 1 as needed.)

<u>Market Environment</u>

Observe in Figure 1 that the external environment initiates the process of salesforce performance. However, in general external environmental factors have received little attention with regard to their impact on salesforce performance. An exception is the work by Ryans and Weinberg (1979) and LaForge and Cravens (1981) utilizing market response models for sales management decision making.

$$M_i = \Sigma E_{ij} \times (\Sigma I_{jk} \times V_k)$$

M_i = motivation: the motivation to expend effort on any task (i).

E_{ij} = expectancy: the estimate of the probability that expending a given amount of effort on task (i) will lead to an improved level of performance on some performance dimension (j).

I_{jk} = instrumentality: the estimate of the probability that achieving an improved level of performance on performance dimension (j) will lead to increased attainment of a particular reward (k).

V_k = valence for rewards: the perceptions of the desirability of receiving increased amounts of each of a variety of rewards that might be attained as a result of improved performance.

Figure 2. Walker, Churchill, and Ford (1977) Adaptation of Vroom's (1964) Expectancy Equation

Taking an approach to market response models drawn from marketing strategy (c.f., Abell and Hammond 1979), such models seek to bring order to complex decision processes involving interactive effects of environmental factors (such as control unit attractiveness and business position), organizational factors (such as marketing effort and sales management effort of the firm), and salesperson factors as defined within the WCF model (i.e., aptitude, skill, personality, personal characteristics, and motivation).

The attraction of market response modelling is its ability to depict the general relationships between factors that sales management cannot control and those that can generally be controlled. As the uncontrollable factors change, sales management must respond by adapting the levels of the controllable factors. As such, market response models provide a framework for analyzing these changes and evaluating alternative adaptations. Unfortunately, no work has been done in the last decade to further the knowledge of the use of such models in assessing the impact of environmental variables within the domain of salesforce performance.

Cultural Environment

As indicated in Figure 1, internal environmental forces also play a role in salesforce performance in the form of organizational culture. Organizational culture may be defined as patterns of shared values and beliefs that

provide norms and direct behavior within an organization (Schein 1984; Ouchi 1981). Despite the call by Deshpande and Webster (1989) for a systematic research program on organizational culture issues within the domain of marketing, no published work exists in the marketing literature addressing the potential impact of organizational culture factors on salesforce performance. Williams (1992) developed and tested a scale for organizational culture within a sales context that holds promise for future application. Evidence that organizational culture research has the potential to benefit sales management practice comes from findings in a human resources management study by Pozner, Kouzes, and Schmidt (1985) indicating that clarity of company values and strong work group norms significantly lowered worker turnover.

Organizational Climate

Figure 1 depicts the construct of organizational climate as differing from organizational culture. Deshpande and Webster (1989) defined organizational climate as the operationalization of themes that pervade everyday organizational behavior. Climate includes the routines that are rewarded, supported, and expected by the organization. Thus, while culture has traditionally been viewed as an overall "gestalt" of influence within a firm, climate has been operationalized as the "things" that actually take place within firms that impact the workplace and the people

in it.

As defined above, the impact of climate on salesforce performance has been a popular research topic. Disagreement exists, however, as to what specific variables actually comprise the construct of organizational climate. In the organizational behavior literature, Campbell et al. (1970) posited that four factors are common to most measures of perceived organizational climate: autonomy, structure of the workplace and position, reward orientation, and the nature of interpersonal relationships. In their seminal article on the subject, Churchill, Ford, and Walker (1976) adapted the Campbell et al. (1970) framework into the sales domain, recasting the original four climate variables into the following seven: three supervisory variables (closeness of supervision, the amount of influence the salesperson has in determining supervisory standards, and the frequency of communication between the sales manager and salesperson), authority structure, innovativeness demanded of the salesperson, and two interpersonal variables (role conflict and role ambiguity). The Churchill, Ford, and Walker (1976) study represented the first empirical test of the impact of organizational climate on job satisfaction in sales, and found that more than 40 percent of the variation in total job satisfaction among sales personnel was explained by salespersons' perceptions of the seven climate variables (the result was found to be moderated by the amount of time a salesperson held his/her position).

In a subsequent study, Tyagi (1985a) broke organizational climate into the following three components: 1) job characteristics - including challenge, variety, and autonomy; 2) leadership characteristics - supervisory styles; and 3) the extent of organizational identification (job involvement) by the salesperson. Tyagi's dependent variable was not job satisfaction, however, but preference for extrinsic versus intrinsic rewards. The major conclusion was that a number of organizational climate dimensions significantly influence salespersons' perceived desirability of extrinsic and intrinsic rewards, with the effect being more significant on the desirability of the extrinsic reward types.

Leadership. One sub-set of studies within the general context of organizational climate has addressed issues of leadership in selling organizations. In a comprehensive review of the leadership literature, Yukl (1989) proposed that the overall approach to leadership research is moving away from trait-based approaches toward a more behaviorally based, integrative view. Such a behavioral approach is consistent with the focus of studies that have examined the role of leadership in salesforce performance. In the first study in this area, Churchill, Ford, and Walker (1976) found that job satisfaction by salespeople was positively impacted by closer supervision by their managers, more frequent communication, and influence by the salesperson in determining performance standards against which he/she will

ultimately be evaluated.

Tyagi (1985b) investigated the differential impact of leadership style versus job dimensions on motivation (intrinsic versus extrinsic) of salespeople. Using an expectancy approach, Tyagi found that selection of leadership style by the sales manager was more important for extrinsic motivational value, while dimensions of the salespersons job were more important for intrinsic motivation. This finding has the major impact of implying that when intrinsic motivation is the issue, exhorting sales managers to practice adaptive leadership or any other particular approach to leadership may be of little motivational value.

A recent study has the potential to make a similarly important impact on sales management practice. Jaworski and Kohli (1991) developed a fourfold typology of supervisor feedback to salespeople based upon locus of feedback (outcome versus behavior) and valence of feedback (positive versus negative). Results of their study indicate that positive performance feedback focusing on salespeople's behaviors tends to increase salesperson job satisfaction to a greater extent than either positive or negative feedback on outcomes. On the other hand, positive feedback on outcomes had a more significant impact on job performance (a logical result, since the outcome was positive regardless of whether the behaviors that went into attaining the outcome were favorable or unfavorable).

Though Jaworski and Kohli (1991) did not explicitly examine reasons for the differential impact of the two types of feedback on job satisfaction in their study, salespeople are likely to be more desirous of behavioral feedback in general because it pertains to aspects that they can more readily control (e.g., decision making or "process" factors), as opposed to outcomes that may depend upon a number of extraneous factors over which the salesperson has little or no direct control.

Salesforce Socialization. A second sub-set of studies under the umbrella of organizational climate includes work related to salesforce socialization. In an organizational behavior context, socialization may be defined as a process by which an individual acquires the social knowledge and skills necessary to assume an organizational role (Van Maanan and Schein 1979). Socialization is typically accomplished in sales organizations by such means as training, education, apprenticeship, debasement experiences, and cooperation (Van Maanan 1976).

The general thrust of research on salesforce socialization has been in the context of the following quote from Churchill et al.'s meta-analysis on salesforce performance (1985, p. 117): "...from a manager's point of view, whom one recruits is important, but probably not as important as what one does with them---and to them---after they have been hired." Two articles have addressed socialization issues in the sales literature. Dubinsky et

al. (1986) drew upon Feldman's (1976) original organizational socialization model to introduce the concept into the sales literature, establishing a goal of facilitating long-term "exchange" relationships among salespeople and managers (as representatives of "the company"). In an empirical test of the Feldman model, the authors' results indicated that salesforce socialization significantly impacts performance, job satisfaction, and job commitment.

Taking a somewhat different approach to socialization, Lagace (1990) used a leader-member exchange model (Graen and Schiemann 1978) to study the importance of developing "cadres" - a nucleus of trained people capable of socializing new salespeople, instead of "hired hands" - individuals whose stakehold in the sales organization is minimal. The development of such cadres is predicated upon importance being placed by the organization on long-term salesperson/sales manager relationships, and clearly increases the capability to more effectively transfer the knowledge and skills required for socialization of new organization members.

Role. Probably more research has been done examining the impact of various role conditions on salesforce performance than any other set of variables in the WCF model. Before discussing the general findings in this area, some definitions are in order:

1) role ambiguity - the degree of uncertainty experienced

by the individual with regard to relevant dimensions of the job role (Bagozzi 1978).

2) role conflict - expectations or demands of two or more role partners are incompatible and cannot be served effectively at the same time (Walker, Churchill, and Ford 1972).

3) boundary role person - organization members who occupy positions in the firm that require them to interact with role partners beyond the formal boundaries of their own organization (Adams 1976).

3) role stress - a potential deleterious result of the boundary-spanning nature of the sales job. May include role ambiguity and role conflict (Behrman and Perreault 1984).

5) required innovativeness - the degree to which the salesperson must produce innovative solutions to nonroutine problems (Churchill, Ford, and Walker 1976).

Most of the early studies by Churchill, Ford, and Walker (i.e., prior to the 1977 WCF synthesis model) were targeted toward determining the impact of the above role variables on salesforce performance (c.f., Walker, Churchill, and Ford 1972; Ford, Walker, and Churchill 1975; Walker, Churchill, and Ford 1975; Ford, Walker, and Churchill 1976; and Churchill, Ford, and Walker 1976). The general findings of these early studies are as follows:

1) Perceptions of role conflict tend to have a significant negative impact on a salesperson's satisfaction with

role partners (e.g., supervisors and customers), but do not impact satisfaction with the nature of the sales job itself.

2) Perceptions of role ambiguity tend to have a negative impact both on the level of satisfaction with role partners and on satisfaction with the job itself.

3) Ambiguity appears to be relatively more readily actionable by sales management through better communication systems, closer supervision, more accurate job descriptions and goal-setting, and other similar methods. Conflict, on the other hand, is much less actionable by management than ambiguity because to a great extent the potential for conflict is between the salesperson and external role partners over which sales management has no control.

4) Increased requirements for innovativeness within the sales role tend to result in feelings by salespeople that they are receiving inadequate support from their companies and supervisors. Consequently, salespeople may experience lowered job satisfaction specific to this lack of support. However, required innovativeness does not appear to significantly negatively impact overall general satisfaction with the sales job.

Regarding role stress as a by-product of the conflict inherent to the boundary-spanning nature of sales positions, Behrman and Perreault (1984) found a differential effect for conflict on job satisfaction versus performance. In their

study, stress-producing role conflict negatively impacted satisfaction but positively impacted performance, leading the authors to conclude that some aspects of role conflict may be basic to performance of the sales job---even if those aspects potentially reduce the salesperson's job satisfaction in the process. In response to the recognition that previous attempts at understanding the impact of role stress on performance and satisfaction have been conceptually over-simplified, Goolsby (1992) developed a theoretical model for role stress that takes into account a number of other variables beyond boundary-role conditions that are hypothesized to impact organizational and personal outcomes. The addition of these other variables, categorized as extrinsic supports (social support and organization strategies) and intrinsic supports (coping skills and individual resources), results in a theory of role stress that is contextually much richer than previous conceptualizations. However, the model has yet to be empirically tested.

Individual Characteristics

The next major box in Figure 1 is labeled "Individual Characteristics." Individual characteristics include aptitude (i.e., inherent ability), skill (i.e., learned proficiency), personality variables, and personal characteristics (e.g., height, weight, etc.). Some of the earliest research into sales performance effectiveness

centered around such characteristics or traits that were hypothesized to be predictive of sales success (c.f., Dunnette and Kirchner 1960; Miner 1962; Mosel 1952; Rich 1966.) Typically, this genre of studies was not theory-driven, and even if hypothesized relationships were prespecified, an explanation of how the salesperson's characteristics affected performance were not considered (Weitz 1979). The explanatory power of differing combinations of characteristics on performance tended to be weak---for example, Ghiselli (1973) demonstrated that even the most commonly used type of selection test for salespeople based upon such criteria typically accounted for only ten percent of variance in performance.

In one well-regarded study, Lamont and Lundstrom (1977) attempted to break down the broad context of salesperson characteristics into related factors in order to better facilitate research. The authors proposed two overall categories of characteristics: personality variables and personal characteristics. Personality variables included dominance, endurance, social recognition, empathy, and ego strength. Personal characteristics assessed were age, height, weight, formal education, outside activities, and civic and professional memberships.

Unlike much of the earlier work in which a few "off-the-shelf" ability variables were related to salesperson satisfaction and performance, Lamont and Lundstrom conducted an analysis of the job in order to identify *a priori* the

characteristics underlying sales performance. The authors then searched for theoretical frameworks that included variables defined as potentially having an impact on performance so that existing scales might be employed to measure the desired characteristics in a reliable and valid manner. As a result, in comparison with much of the previous research in sales management involving personality and personal characteristics as predictor variables, the Lamont and Lundstrom models have considerably greater explanatory power---34 percent of variance in overall ratings by managers. A resulting "Profile of the Successful Salesman" emerged that described a person who is (among other traits) tall, exhibits perseverance, has a broad range of interests, and is adaptable and flexible.

Churchill, Ford, and Walker (1979) examined the impact of personal characteristics only (not personality variables) on salespersons' preferences (valences) for different reward types. Lower order rewards were operationalized as pay and job security; higher order rewards were recognition, promotion, and liking and respect. Among the findings, older salespeople tended to value higher order rewards more, salespeople with large families tended to value lower order rewards more, and overall education level tended to positively impact the valence for higher order rewards.

Meta-analysis. A culmination of the work on individual characteristics was Churchill et al.'s (1985) meta-analysis of the determinants of salesperson performance. The massive

project reviewed 116 published and unpublished studies yielding 1653 reported associations between performance and determinants of that performance. Five broad categories of personal, psychological, and situational impactors on salesperson performance were examined: 1) aptitude, 2) skill, 3) motivation, 4) role perception, and 5) personal, organizational, and environmental factors. None of the factors accounted for a great amount of the variance in performance---less than ten percent on average---though the variance explained could be much higher in any single study. (It should be noted that a particularly troublesome finding was that some of the factors were actually positively related to performance in one context and negatively related to performance in some other context.)

A number of potentially useful implications may be drawn from this meta-analysis:

1) Enduring characteristics were less important to performance than influenceable characteristics, thus accentuating the importance of training and performance feedback.
2) Multi-dimensional models of salesperson performance are superior to unidimensional approaches.
3) Inherent benefits exist of hiring a salesperson who is already trained and familiar with the nature of the sales role.
4) The overall impact of the predictor variables on performance was moderated by type of product being

sold. Therefore, job-specific and company-specific studies on salesperson performance are more apt to yield usable results than global performance studies, despite obvious concerns with generalizability of results gathered via such methodologies.

In a follow-up meta-analysis that was more narrowly focused, Ford et al. (1987) included only empirical studies that used biographical or psychological variables to predict or explain variance in performance across sales jobs. As with the Churchill et al. (1985) meta-analysis, no single set of such factors was found to be a consistently robust predictor of sales performance.

Motivation

Figure 1 portrays each of the previously mentioned variables as ultimately impacting "motivation." The motivation construct in the WCF model is analogous to the "black box" concept in consumer behavior in that it is not directly observable. Rather, motivation manifests itself via behaviors (in this case, performance). In Vroomian expectancy theory, the motivation component is defined as the amount of effort expended on a particular task (Walker, Churchill, and Ford 1977). In essence, it represents the criterion variable in the expectancy equation as outlined previously in Figure 2.

A variety of studies have produced empirical support for the Vroomian model with respect to non-selling employees

(e.g., Lawler 1968). Also, two studies have been reported in the salesforce performance literature that support the robust nature of expectancy theory in predicting salesforce motivation (Oliver 1974; Teas 1981). (For a thorough review of expectancy theory research in selling, see Evans, Margheim, and Schlacter 1982).

One criticism of the salesforce performance literature in general has been its almost exclusive reliance on expectancy theory as a theoretical base. In the organizational behavior literature, Scholl (1981) has questioned the ability of expectancy theory to consistently explain employee behavior. His skepticism is based upon a number of studies demonstrating that many employees whose expectations were not being met still continued to work for their respective organizations (c.f., Vroom and Deci 1971).

Recently, two alternatives to the expectancy paradigm of salesforce performance management have been proposed in the literature. Scott et al. (1986) presented an organizational behavior modification (OBM) approach that relies on the application of operant conditioning principles to the selling environment, complete with a taxonomy for utilizing the standard feedback strategies of positive and negative reinforcement, punishment, and extinction. A second alternative to expectancy is behavioral self management (BSM), proposed by Sauers, Hunt, and Bass (1990). Whereas OBM requires that the behavioral consequences be imposed by others, BSM requires that such consequences by

self-imposed. A framework was presented for implementing BSM principles. BSM relies heavily on a strong commitment by employees to specified behavioral goals and on empowerment of employees by management to allow freedom of action to achieve the goals. To date, the efficacy of OBM and BSM has not been tested empirically.

Performance and Satisfaction

The two most commonly used criterion variables across the entire body of sales management research are performance and satisfaction. Because the variables have been used so extensively in this manner, they will be considered together here. Research in industrial and organizational psychology has focused on the order of the causal relationship of job satisfaction and performance, with mixed results (Locke 1970; Porter and Lawler 1968). Although marketers have not explicitly examined the direction of causality issue, it is clear that there is a relationship between job satisfaction and dimensions of the work itself (Bagozzi 1980b).

Reflective of traditional Vroomian expectancy theory, the WCF model orders satisfaction after performance, but with a feedback loop to motivation. Performance evaluation and rewards are portrayed as mediators (refer to Figure 1).

Salesforce Performance Evaluation

This section of the review of the salesforce performance literature is given special attention because

the focus of the present study is on performance evaluations of field salespeople by their sales managers. Compared to other variables in the adapted WCF model (Figure 1), little research has been done to investigate the performance evaluation phase. In fact, it has only recently been proposed that the role of causal attributions (Teas and McElroy 1986) and judgmental biases (Gentry, Mowen, and Tasaki 1991) be included in salesforce performance research. This evaluation phase of the adapted WCF model is reintroduced in Figure 3.

Hiring and motivating qualified salespeople is a high priority in most firms. One indispensable motivational tool available to the sales manager is the performance appraisal. An accurately and fairly executed performance appraisal provides input for decisions regarding salary and promotions, as well as essential two-way communication between the salesperson and the sales manager for purposes of goal-setting, training, and performance feedback (Dubinsky, Skinner, and Whittler 1989).

Focus of Recent Appraisal Literature. Because performance appraisals play a critical role in organizations, an important research goal in the human resources and organizational behavior literatures has been devising ways to improve the accuracy of ratings (DeNisi, Cafferty, and Meglino 1984). One approach to improving the performance appraisal process has been based upon attempts to eliminate rating errors by analyzing scale construction

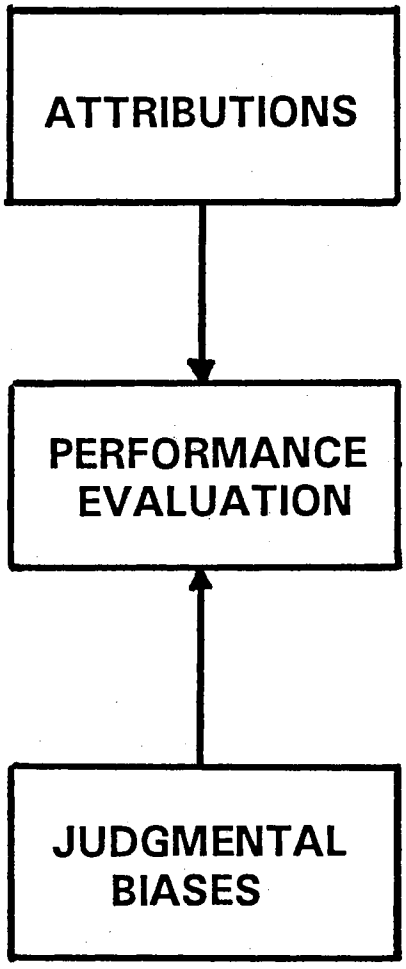


Figure 3. The Evaluation Stage of the Adapted WCF (1977) Model

and administrative techniques. However, as noted by Ilgen and Feldman (1983), future advances resulting from such an approach will probably be very slow and of limited magnitude. Therefore, if performance appraisal is to advance, the total appraisal process must be better understood. To do this, we must focus upon the appraiser, the nature of the appraisal setting (contextual variables), and the motives and desires of those being appraised---i.e., the appraisee (Ilgen and Feldman 1983). Thus, much of the recent literature on performance evaluation has focused on cognitive processes and on integration and application of theories instead of development of procedures for appraisal administration.

In a review article on the performance appraisal literature from 1986-1993, Latham et al.(1993) concluded that much of the research on performance appraisal during that time span focused on the identification and removal of rater biases. This continues a trend evident during the prior five year period as well (Latham 1986). Specifically, the research has focused on rating error, and on both rater and ratee characteristics that contribute to rating error.

In particular, halo error has continued to be the subject of prolific research. Much of the work on halo has been designed to effectively separate true halo from halo error (c.f., Murphy and Jako 1989; Nathan and Tippins 1990; and Pulakos, Schmitt and Ostroff 1986.) Both Mount and Thompson (1987) and Kozlowski and Kirsch (1987) conducted

studies to identify cognitive processes that contribute to halo error. Other rater biases recently investigated within the human resources domain include contrast/assimilation effects (Steiner and Rain 1989; Smither, Reilly, and Buda 1988; and Maurer and Alexander 1991), and primacy/recency effects under conditions of immediate and delayed performance appraisal (Steiner and Rain 1989). In sum, the recent human resources management and organizational behavior literatures are rich with studies designed to extend knowledge of appraisers, appraisal settings, and the motives and desires of appraisees.

In contrast to the proliferation of research on performance appraisal processes in the human resources and organizational behavior literatures, until recently comparatively little work on performance appraisal has appeared in publications specific to the domain of salesforce performance management (recent exceptions include Avila, Fern, and Mann 1988; Dubinsky, Skinner, and Whittler 1989; Jaworski and Kohli 1991; Marshall, Mowen, and Fabes 1992; McKay et al. 1991; and Morris et al. 1991). Three distinct research streams are evident within the salesforce performance appraisal literature: 1) identification of appropriate evaluative criteria to utilize in making ratings; 2) investigation of the role of causal attributions in sales manager ratings of sales personnel; and 3) the impact of human judgment (especially judgmental biases) on salesforce performance ratings. Each of these literatures

will now be reviewed.

Determining Evaluative Criteria. In the introduction to a 1956 Conference Board report titled "Measuring Salesmen's Performance" (Umemura 1956, p. 3), G. Clark Thompson, Director of the Division of Business Practices for the Conference Board, made the following observation:

> Sales volume alone is no longer considered a good measure of a salesman's worth. For volume without profit is generally valueless, and volume achieved at the expense of future sales can prove to be a liability. There are so many uncontrollable factors that may affect the amount of sales booked by individual salesmen that the sole use of volume as a measurement of performance may be very misleading. The size and richness of the territory, the distance from the factory, the intensity of the competition in the area, the amount of promotional and supervisional support are just a few of the variables which can greatly alter sales results even when salesmen are evenly matched.
>
> ...It follows that a good system of measuring sales performance is a prerequisite to an intelligent training program and the key to personal counseling. It is also a safeguard against the loss of salesmen who have a good potential but have not yet realized it.

Fifteen years later, Cotham and Cravens (1969, p.79)

noted that, "The exclusive use of actual (raw) performance data to measure salesmen's contributions to the firm can be misleading. It should be limited to comparisons of selling under similar circumstances and when a single performance measure is used."

Such early recognition that selecting appropriate performance criteria for sales personnel is critical to the motivational value of the entire performance management system set the tone for a research agenda which, though dormant during the 1970's, has seen a resurgence of interest in the last ten years. The first article in the recent stream was by Behrman and Perreault (1982). A self-report performance scale was developed and evaluated based upon the responses of 200 salespeople and 42 managers from five major industrial firms. Factor analysis of the job performance variables revealed five general categories of evaluative criteria for use in salesperson evaluation: sales objectives (e.g., market share, profit, sales volume), technical knowledge, providing information (e.g., paperwork handling, maintaining company records), controlling expenses, and sales presentations (e.g., overall selling skills).

Jackson, Keith, and Schlacter (1983) took a different approach, by asking 213 sales managers from a variety of companies to indicate what performance measures they actually used to evaluate sales personnel. These measures were grouped into output bases (e.g., sales, market share,

accounts lost/gained, profit, and orders), input bases (e.g., calls, expenses, and ancillary activities such as reports, training meetings conducted, and letters/calls to prospects), and qualitative bases (e.g., product knowledge, selling skills, planning ability, and judgment). In a follow-up study, Morris et al. (1990) used a similar methodology in which senior sales managers from 200 industrial product and service companies were asked to rate the importance of a variety of performance measures. In this study the performance measures were categorized as quantitative and qualitative in nature.

Anderson and Oliver (1987) reviewed and contrasted two salesforce control systems: outcome-based and behavior-based. Outcome-based control approximates a market contracting arrangement wherein salespeople are left alone to achieve results in their own way using their own strategies. Salespeople are held accountable for their results (outcomes) but not for how they achieve the results (the behavior). On the other hand, behavior-based control systems require active managers, backed by a significant management information-gathering staff, who vigorously monitor and direct the operations of the salesforce. Managers typically have a well-defined idea of what they want salespeople to do and work to ensure the salesforce behaves accordingly. Sales results are presumed to follow, often in the long term.

Anderson and Oliver recognized that these polar

extremes are stereotypical and that most salesforce control systems involve a mixture of the two evaluative criteria, albeit tending to lean in one direction to another. However, one finding of the Churchill et al. (1985) meta-analysis previously mentioned was the overwhelming tendency for sales managers (and the performance appraisal systems they use) to emphasize outcomes rather than process, particularly in determining compensation. A major reason is the ready availability of simple, seemingly equitable measures of sales volume or dollars (Anderson and Oliver 1983). As such, a set of outcome criteria may become a surrogate for the process undertaken by the salesperson to achieve the outcome, with a tendency to believe that the end is reflective of the means.

Causal Attributions in Salesperson Evaluation. As previously mentioned, Teas and McElroy (1986) proposed the inclusion of causal attribution research within the domain of salesforce performance. These authors used concepts of attribution theory developed by Heider (1958), Kelley (1967), and Weiner (1972) to examine the role played by attributions within the expectancy based formulation of salesperson motivation, with particular emphasis on attribution theory's potential usefulness in explaining some of the linkages among perceptions of salesforce performance and expectancy perceptions.

Up to this point in the present literature review, every salesforce performance study cited has employed some

form of survey methodology, and most have relied on a variation of correlational analyses for data analysis. One stream of research in the salesforce performance literature to utilize an alternative research methodology is a series of studies investigating the role of causal attributions on evaluations of salespeople by their managers. Six studies on this topic have been reported in the literature to-date, each of which employed an experimental design.

Four of the studies on causal attributions in salesforce evaluation were authored by Mowen and his colleagues (Mowen, Brown, and Jackson 1981; Mowen et al. 1985; Mowen, Fabes, and LaForge 1986; and Marshall, Mowen, and Fabes 1992). These studies utilized Heider's (1958) attributional framework to make predictions concerning sales manager evaluations of employees. Heider developed the cornerstone concept that individuals tend to operate as "naive psychologists" when they observe and analyze the behavior of others. He classified variables used by individuals to interpret the actions of others into three categories: 1) a performance variable (i.e., task success); 2) environmental variables (task difficulty and luck); and 3) person variables (ability and effort). Heider proposed that evaluators assess performance based upon relationships among these factors. Results of the studies by Mowen and his colleagues have provided general support for Heider's model.

One particular focus of this research stream has been

the effect of territory difficulty on performance ratings. Research on territory difficulty is important because in order to administer accurate and fair performance appraisals, sales managers must adjust ratings by taking into account the differences in territory difficulty among the salespeople they supervise. However, a phenomenon known as the "fundamental attribution error" (Ross 1977) predicts just the opposite: that contextual or background information (such as territory difficulty) will be systematically ignored by raters, and instead evaluations will be based upon "person" factors such as perceived ability and effort. The two earlier studies (Mowen, Brown, and Jackson 1981; Mowen et al. 1985) found support for the fundamental attribution error. The two later studies (Mowen, Fabes, and LaForge 1986; Marshall, Mowen, and Fabes 1992) found that sales managers utilized territory difficulty in making their ratings. The authors suggested that an improved, perhaps more vivid format for presenting performance information in the later studies (i.e., a spreadsheet format as opposed to a scenario format in the earlier studies) may have contributed to the increased saliency of territory difficulty information in those later studies. An implication for sales managers is the importance of arranging performance information in a highly useable format prior to making ratings.

Two other studies utilizing experimental designs have contributed to the knowledge of the role of attributions in

salesforce performance evaluation. Using an attributional model of leadership formulated by Green and Mitchell (1979) and Mitchell, Green, and Wood (1981), Dubinsky, Skinner, and Whittler (1989) examined the effects of two different levels of work history (good/poor) and two different levels of task difficulty (high/low) on sales managers' attributions and responses to a salesperson's failure to obtain a sale. The results suggested that both internal information (work history) and external/situational information (task difficulty) about the salesperson affected causal attributions sales managers made concerning a salesperson's performance, thus confirming the findings of the two later studies by Mowen and his colleagues.

McKay et al. (1991) used a scenario format with an experimental design that varied level of salesperson effort and ability in order to examine the relationship between managers' perceptions of salespersons' effort and ability to perform and the types of corrective actions or rewards offered to salespeople. The perceptions of effort and ability were found to differentially impact both corrective actions selected for poor performers and reward actions for good performers.

Judgmental Biases in Salesperson Evaluation. Two articles have specifically examined the role of human judgment in managerial evaluations of sales personnel. Patton and King (1985) developed a model to predict the use of five choice models (i.e., simple linear compensatory,

weighted linear compensatory, lexicographic, disjunctive, and conjunctive) in salesperson evaluation. Findings indicated that evaluative decisions, attribute importance, and the judgment model utilized may vary across and within decision type. The authors concluded that knowledge of the manner in which these various choice models operate, the attribute make-up and predictive ability of each model, and the type of decision each model predicts most accurately should be major objectives of a sales manager who is contemplating improving the evaluation aspect of the job.

Gentry, Mowen, and Tasaki (1991) provided a broad overview of cognitive biases that may affect the performance evaluation of sales personnel. They proposed the inclusion of the potential for judgmental biases into the evaluation stage of salesforce performance research. Their goal was to develop an approach to improving rater accuracy based upon principles of behavioral decision theory.

For over twenty years, behavioral decision researchers have investigated human judgment and choice, focusing on gaining a better understanding of the factors that cause decisions to be "suboptimal" (e.g., Nisbett and Ross 1980). Findings in behavioral decision theory indicate that managers take decision making short-cuts by applying judgmental heuristics, or rules of thumb, that may lead to poor decisions (for a review see Bazerman 1990).

The use of judgmental heuristics introduces systematic biases due to a shortcoming in information processing by the

manager. Examples of such biases include the fundamental attribution error (Ross 1977), the availability heuristic (Tversky and Kahneman 1974), decision framing (Tversky and Kahneman 1981), the preference-reversal phenomenon (Mowen and Gentry 1980), the hindsight bias (Fischhoff 1975), the outcome bias (Baron and Hershey 1988), and the order effects bias (Einhorn and Hogarth 1992). Each of these biases has the potential to cause "suboptimal" decisions.

A number of factors may contribute to the susceptibility of sales managers to judgmental biases in the evaluation of sales personnel. First, the day-to-day activities of many practicing salespeople tend to take place apart from the direct observation of the sales manager. Second, most sales positions typically contain a fairly large set of performance dimensions with varying degrees of priority attached. In fact, the relative priority of any given task can change quickly as customer needs, company demands, or perceived rewards dictate. Third, by the very nature of personal selling, a tendency exists for managers to encourage entrepreneurial behavior from their sales personnel. Churchill, Ford, and Walker (1976) described such behavior in terms of the "innovative role" of the salesperson in which he/she is required to produce innovative solutions to nonroutine problems. The expectation is that effective salespeople can make appropriate decisions, independent of the supervisor, that will lead to sales success. To do this, salespeople must be

empowered to make these decisions and to follow the courses of action they believe are best. Unfortunately, such choices may be inconsistent with those of the sales manager. The importance of effective decision making skills by sales personnel has been well-documented in the literature (c.f., Behrman and Perreault 1984; Churchill, Ford, and Walker 1976; Lamont and Lundstrom 1977; Weitz 1981; and Weitz, Sujan, and Sujan 1986).

A fourth factor that may contribute to rater biases is the boundary-spanning role held by salespersons between their companies and constituents. As previously mentioned, salespeople are susceptible to a number of well-documented role problems (e.g., conflict, ambiguity). When such role problems occur loyalties may be compromised and jobs incompletely understood, with specific duties inadequately described or totally unspecified (Feldman 1981). Finally, salespeople are among the first in any organization to be directly impacted by externalities, such as the environment, the economy, competition, and the overall market (Adkins 1979). Such externalities may differentially influence the decision processes of salespeople and their sales managers. In sum, these domain-specific factors build a persuasive case that salesforce performance appraisal is a quite difficult process, potentially resulting in haphazard, unsystematic approaches to evaluation by managers (Dubinsky and Barry 1982; Jackson, Keith, and Schlacter 1983).

In response to these difficulties inherent in

evaluating sales personnel, a tendency exists for sales managers to focus on readily observable outcomes of selling efforts when developing performance ratings. Classic examples include sales volume, sales-to-quota, number of new customers, and other similar outcome measures (Behrman and Perreault 1982). In practice, truly optimal selling processes are quite difficult to specify. Because of uncertainty about how various factors interact to result in a particular sales outcome, managers may rely solely on those outcomes as a surrogate measure of the quality of the corresponding sales effort. However, it has been frequently asserted that overreliance on such outcomes as evaluative criteria can sometimes be misleading, since outcome measures are highly impacted by the very role difficulties and externalities over which the salesperson has little control and that make salesforce performance appraisal so difficult in the first place (Morris et al. 1991).

In the context of salesforce performance appraisal, the biases introduced by the use of judgmental heuristics by sales managers could negatively impact the accuracy and fairness of ratings. As a result, rewards may be misappropriated. Ultimately, the motivational value of the entire performance appraisal process may break down. Gentry, Mowen, and Tasaki (1991) suggested several approaches to debiasing the performance evaluation process in sales organizations, including an emphasis on training managers to avoid falling victim to judgmental biases when

rating salespeople.

Rewards

The next-to-last box represented in Figure 1 is labeled "rewards." Surprisingly, few studies have been undertaken in the literature specifically designed to examine the role of rewards within the domain of salesforce performance. One such study by Churchill, Ford, and Walker (1979) was reviewed earlier in this chapter within the context of personal characteristics of salespeople and reward preferences.

Two types of rewards are available, as defined below by Tyagi (1985a):

1) intrinsic rewards - come directly from the performance itself. The salesperson bestows these upon him-/herself. Examples are feelings of accomplishment, self-worth, and developing one's skills and abilities.
2) extrinsic rewards - rewards that are bestowed upon the salesperson by someone else. Examples are monetary income, promotion, and recognition/respect received from a supervisor.

Two studies conducted within the sales domain examined the relative impact of various organizational climate variables on internal versus external motivation (Tyagi 1982) and internal versus external reward desirability (Tyagi 1985a). These studies produced mixed results, prompting Tyagi to suggest that more work in the area is

needed, especially regarding how situation-specific climate dimensions can impact the relative desirability of internal versus external rewards.

Intent to Turnover

The last element in Figure 1 is "intent to turnover." Much empirical support has been generated that establishes a linkage between job dissatisfaction and intent to turnover (c.f., Fern, Avila, and Grewal 1989; Futrell 1984; Lucas et al. 1987). Organizational commitment, a construct that might be considered an antithesis to intent to turnover, has also recently begun to receive attention in the sales management literature (c.f., Chonko 1986; Ingram et al. 1989; Sager and Johnston 1989).

On a broader level, organizational commitment may be viewed as another alternative to the expectancy paradigm. Scholl (1981) has suggested that organizational commitment is a stabilizing force that acts to maintain a behavioral direction when expectancy conditions are not met and do not operate. Briefly, an individual internalizes expectations of others concerning his/her behavior. According to commitment theorists, when an individual's behavior is directed by these internal normalized pressures, behavior no longer depends upon relationships with outcomes and rewards (as posited by expectancy theory). Viewed this way, a salesperson who has a strong commitment to the organization is likely to behave according to internalized norms rather

than performance-reward relationships (Chonko 1986).

Whereas expectancy theory assumes that work effort is the result of the interaction between the individual and aspects of the work situation (e.g., rewards), an organizational commitment approach suggests that certain individuals may exhibit behavioral tendencies (e.g., commitment) that may vary between individuals but are relatively constant across work situations. Viewed in this way, commitment is a value-based, normative evaluation of alternative behaviors leading to performance. As such, an organizational commitment approach to motivating salespeople addresses the problems of task definition, observability, performance measurement, and role complexity associated with appraising and rewarding sales personnel. High commitment among employees has been found to lead to lower turnover and, thus, to higher organizational performance (Mowday, Steers, and Porter (1979), as well as to higher levels of satisfaction on the job (Hunt, Chonko, and Wood 1985).

CHAPTER III

THEORETICAL SYNTHESIS AND RESEARCH HYPOTHESES

Introduction

Based upon the review of the salesforce performance literature presented in Chapter II, a gap in knowledge has been identified based upon the fact that, despite attention in the literature to different types of evaluative criteria (e.g., outcome versus behavioral or process criteria), little empirical work has been done to explain the mechanisms by which these types of criteria are utilized by sales managers in rating their salespeople.

This chapter is developed in two major sections. The first section provides a background and theoretical framework for a study to examine the effects of an outcome bias in salesforce performance evaluations (designated "Study One"). The section closes by advancing several hypotheses related to an outcome bias in such evaluations. The second section provides a background and theoretical framework for a study to examine the effects of an order effect bias in salesforce performance evaluations (designated "Study Two"). The section also closes by advancing several hypotheses related to this order effects

bias.

Outcome Bias

In order to address the knowledge gap identified in the introductory section of this chapter, a theoretical basis will be drawn from the behavioral decision theory literature on the outcome bias. In this dissertation, Early et al.'s (1990) definitions of the two forms of performance feedback are adopted: outcome, meaning information concerning performance outcomes; and process, meaning information concerning the manner in which an individual implements a work strategy.

As previously mentioned, the tendency in salesforce performance evaluation has been to focus on evaluating the outcomes of performance (Churchill et al. 1985). The presumption across the previous studies in the sales domain on outcome versus other bases for evaluation has been that outcome information will tend to be overutilized and behavioral performance or process information underutilized when salespeople are evaluated by their managers (c.f., Jackson, Keith, and Schlacter 1983; Anderson and Oliver 1987; and Morris et al. 1991.) As pointed out by Anderson and Oliver (1987), outcome measures are obvious and readily observable.

An alternative approach is to focus on the decisions made by the salesperson (i.e., the process). Thus, a process-based approach focuses on the quality of the

salesperson's decisions by analyzing the appropriateness of decision making, given the circumstances encountered by the salesperson in implementing a selling strategy. It has often been assumed that the more subjective nature of process-based evaluation permits managers to introduce preconceptions or personal biases into the evaluation (for a review of judgmental measures of work performance see Landy and Farr 1983, pp. 57-90). Instead, it may actually be the omission of process-based criteria from evaluations, rather than the inclusion of such criteria, that introduces the greater potential for appraisal bias. This systematic overweighting of outcomes and underweighting of process is the essence of the outcome bias (Baron and Hershey 1988; Hawkins and Hastie 1990). Thus, when a decision results in a positive (negative) outcome, evaluators tend to rate the quality of the decision and the competence of the decision maker positively (negatively) regardless of the actual appropriateness of the decision itself. Despite evidence suggesting an overreliance on outcome measures in salesforce performance appraisal, no empirical work has been done to demonstrate the mechanisms by which an outcome bias operates within the context of sales/sales management decision making.

Jaworski and Kohli's (1991) finding of a differential impact on job satisfaction of outcome versus behavioral information feedback (as outlined in the preceding salesforce performance literature review) is central to the

question addressed within the outcome bias portion of this dissertation because if it can be empirically demonstrated that an outcome bias exists within the domain of salesforce performance evaluation, the potential for such a bias to impact job satisfaction (and, according to the WCF model to ultimately impact performance and turnover) is clear.

Prior Research on the Outcome Bias

Early work relevant to the outcome bias was conducted by Fischhoff (1975) on a closely related judgmental bias, the hindsight bias. As described by Fischhoff (1982), hindsight refers to the tendency of individuals to:

> ...consistently exaggerate what could have been anticipated in foresight. They not only tend to view what has happened as having been inevitable but also to view it as having appeared "relatively inevitable" before it happened.

Because of the hindsight bias, people tend to believe that others should have been able to anticipate events much better than was actually the case, and people even misremember their own predictions, resulting in the exaggeration in hindsight of what they actually knew in foresight (Fischhoff and Beyth 1975).

Five articles on the outcome bias are known to exist. These are reviewed in the following sections, and serve as a conceptual basis for the hypotheses for Study One.

Mitchell and Kalb (1981). In the earliest of these studies, Mitchell and Kalb (1981) investigated an outcome bias on supervisors' evaluations of subordinates in a health

care delivery setting. Their work revealed that those subjects with outcome knowledge, particularly in the case of a negative outcome, rated the outcome as more probable, saw the subordinate as more responsible for the behavior, and made more internal attributions for the behavior than did subjects with no outcome knowledge.

Baron and Hershey (1988). In a series of five studies, Baron and Hershey (1988) gave undergraduate student subjects a set of twelve to sixteen medical and gambling decisions to evaluate and the outcome of each of the decisions. The results were highly consistent across the five experiments, with the outcome of the decision (good or bad) systematically influencing subjects' evaluations of the quality of the decision. In addition, in their Study 4 the bias was shown to extend beyond evaluations of mere decision quality to the evaluation of the individual who made the decision, prompting the authors to claim evidence of an outcome bias in predictions of future competence of the decision maker.

Both the Mitchell and Kalb (1981) and Baron and Hershey (1988) studies held decision appropriateness constant while manipulating outcome. Good evidence of an outcome bias was generated when the decision was bad (i.e., Mitchell and Kalb 1981) or when the quality of the decision itself was ambiguous (i.e., Baron and Hershey 1988). But in order to truly investigate the phenomenon, decision outcome must be manipulated independently of decision appropriateness. The

question is, will the outcome bias still be apparent even when the appropriateness of the decision is varied?

Lipschitz (1989). Such an approach was taken in an experiment conducted by Lipschitz (1989) in which both decision outcome (success versus failure) and what he termed "decision appropriateness" (appropriate versus inappropriate decision, given the circumstances) were manipulated. Decision appropriateness was operationalized based upon whether an action would be expected to be normatively correct. Results revealed an interaction between outcome and decision appropriateness. When a decision was perceived to be inappropriate, outcome information had a strong impact on evaluations. When a decision was considered to be appropriate, the decision maker was evaluated relatively positively, regardless of the outcome of the decision.

Unfortunately, the study had methodological shortcomings. The context of the study was decision making by actual Israeli military officers who were evaluating a fellow officer's decision. What was labeled as the normatively correct (appropriate) decision actually involved violating orders. Indeed, on the dependent variable of "discipline," higher ratings were given to those taking the "inappropriate action" (i.e., those who followed orders) than those taking the "appropriate action" (i.e., those who violated orders). Thus, the manipulation of the appropriateness of the decision was ambiguous.

Mowen and Stone (1992). A fourth study on the outcome bias was the first to investigate the phenomenon within a marketing setting (Mowen and Stone 1992). In the study, adult subjects role-played consumers of services offered by the Corps of Engineers. The scenario was modeled upon actual events that took place in a Midwestern state. Specifically, the subjects' houses were threatened by the potential release of flood waters below a Corps of Engineers dam. Subjects were given information on the appropriateness (either inappropriate or appropriate) of the decision of the Corps of Engineers official to hold excess water behind the dam for as long as possible in order to avoid causing certain minor flooding. In addition, subjects received information on the outcome of the decision in which either major flooding occurred or no flooding occurred depending upon whether new torrential rains fell. As hypothesized, an interaction occurred between outcome and decision appropriateness information. When the decision was appropriate, ratings of decision quality did not differ according to the outcome. In contrast, when the decision was inappropriate, ratings of decision quality were significantly worse when the outcomes were bad than when they were good.

A critical question raised by Mowen and Stone (1992) concerns whether the display of an outcome bias really represents suboptimal decision making. Behavioral decision theory researchers have consistently argued that outcomes

should be used carefully, if at all, as a basis for evaluations. In what has become known as Edwards' dictum, Edwards (1984) stated that because all decisions are made under uncertainty, "A decision is therefore a bet, and evaluating it as good or not must depend on the stakes and the odds, not the outcome" (Edwards 1984, p. 7). As noted by Mowen and Stone (1992), however, a question exists as to whether Edward's dictum applies to evaluations of public policy or marketing decision makers. Particularly when the evaluator has multiple occasions on which to observe behavior, employing outcome as well as decision appropriateness information may be advisable. As a result, Mowen and Stone proposed a "weak form" of Edward's dictum, which states that evaluators may use information on outcome as well as the stakes and the odds (i.e., decision appropriateness) when assessing decision quality. The results of their study supported the hypothesis that evaluators would follow the weak form, rather than the strong form, of Edward's dictum.

Marshall and Mowen (1993). The most recent study to test for an outcome bias was conducted by Marshall and Mowen (1992). Following the general approach utilized by Lipschitz (1989) and Mowen and Stone (1992), decision appropriateness and outcome were varied independently. The context of the research involved a salesperson's decision to pursue one of two possible companies from whom a large sales order might be obtained. Decision appropriateness was

manipulated by varying the likelihood of the salesperson successfully achieving orders from the two firms. Outcome was manipulated based upon whether or not the salesperson in fact achieved the order from the firm chosen. Two dependent measures were employed: decision quality - subjects' perceptions of the correctness/competence of the salesperson's decision; and a more personal-related measure based upon subjects' attributions related to general evaluations of the salesperson and his/her performance.

The results of the study strongly supported the hypotheses. The evaluations of the salesperson revealed the expected interaction between the appropriateness and the outcome of the decision on the dependent variable that assessed the quality of the choice. When the decision was appropriate, no differences in ratings occurred. In contrast, when the decision was inappropriate, outcome strongly impacted ratings. That is, when the salesperson chose to target the low probability customer, ratings were significantly higher when the sale was made than when the sale was not made. On the other hand, when the salesperson chose to target the high probability customer, outcome had no measurable impact on ratings of decision quality. These results are consistent with the weak form of Edwards' dictum, that evaluators will use both outcome and decision appropriateness information when rating the quality of decisions.

Mowen and Stone (1992) and Marshall and Mowen (1993)

also sought to provide evidence for an information processing explanation for the interaction between decision appropriateness and outcome in evaluating the quality of the salesperson's decision. Both studies found that more cognitive responses occurred in the inappropriate decision condition than in the appropriate decision condition. Thus, when a decision was inappropriate, expectations were violated, which caused evaluators to engage in greater amounts of cognitive processing (as revealed by the increased number of cognitive responses). The result was a magnification of available outcome information such that when the decision was inappropriate, a greater amount of information processing lead the evaluator to include outcome information (if available) in the evaluation, acting to magnify (i.e., drive apart) the ratings in the "bad" versus "good" outcome conditions. In contrast, when the decision was appropriate, less information processing occurred because expectations were not violated, resulting in minimal impact of outcome information on the evaluation.

In contrast to the ratings of decision quality, in the Marshall and Mowen (1993) study the general attributional ratings of the salesperson showed a different pattern. The results revealed that on the index of attribution-based items and general evaluations of the salesperson, only a main effect for outcome occurred. Thus, while decision appropriateness information interacted with outcome information to affect ratings of decision quality, only the

outcome impacted the more personal evaluation of the salesperson.

Drawn from Heider's (1958) attributional model, the personal evaluation index consisted of scales that assessed the salesperson's ability, effort, job performance, promotion potential, and skill level. The findings of only a main effect for outcome on this dependent measure were disturbing to the authors because, consistent with Baron and Hershey (1988), the result shows the overwhelming impact of outcome information on personal evaluations. That is, when the focus of the evaluators' attention moved away from the direct evaluation of decision quality to more general assessments of the salesperson, the influence of outcome information overwhelmed the effects of decision appropriateness information. Thus, on the personal evaluation index, the results failed to support the weak form of Edwards' dictum. Rather, these results indicated that decision appropriateness information was completely ignored.

Outcome Bias in Multiple Evaluations

One criticism of the outcome bias stream of research has been the use of single decision scenarios rather than scenarios incorporating multiple decisions. In the "real world," evaluations take place across a variety of decisions over time. An important empirical question that is thus-far unanswered is whether an outcome bias will occur over

multiple time periods.

One way to conceptualize the possible effect on the outcome bias of making multiple evaluations over time is through Kelley's (1967) model of covariance (frequently called the "Kelley Cube." Kelley displayed three criteria utilized by an evaluator in a three-dimensional cube with the following axes: distinctiveness, consistency over time and modality, and consensus (Mizerski, Golden, and Kernan 1979). Kelley proposed that these criteria are used by the evaluator to ascertain whether the impression reflects the inherent properties of the entity rather than some environmental influences. Mizerski, Golden, and Kernan (1979) provide the following description of the dimensions of the Kelley Cube:

1) Distinctiveness - the effect is attributed to the entity if it uniquely occurs when the entity is present and does not occur in its absence.
2) Consistency over time - each time the entity is present, the individual's reaction must be the same, or nearly so.
3) Consistency over modality - the reaction must be consistent even though the mode of interaction with the entity varies.
4) Consensus - actions of their effects are perceived the same way by all observers.

Specifically, Kelley (1967, p. 197) proposed that:

To the degree that a person's attributions fulfill

> these criteria, he feels confident that he has a true picture of his external world. He makes judgments quickly and with subjective confidence...When his attributions do not satisfy the criteria, he is uncertain in his views and hesitant in action.

Within the context of the outcome bias in salesforce performance evaluation, repeated ratings of a salesperson over time would be expected to be differentially impacted depending upon the level of distinctiveness, consistency, and consensus across levels of outcome and decision appropriateness and across time. In their study, Mitchell and Kalb (1981) indicated suspicion that major determinants of both internal versus external attributions and responsibility judgments regarding outcome and behavior by raters are as follows: 1) whether the subordinate has done the task before (consistency); 2) whether the subordinate makes mistakes on other tasks (distinctiveness); and 3) whether other subordinates make the error in their scenario frequently (consensus).

Kelley's model of covariation provides one theoretical underpinning for a hypothesis presented later in this chapter regarding the role of multiple evaluations over time on the outcome bias. Another theory of attribution developed by Kelley (1973), the discounting and augmentation principles, provides additional theoretical rational for predicting the potential impact of more than one rating on the outcome bias. The discounting principle comes into play

when an attributor perceives multiple causes for a given effect. Some causes are picked as representing the subject's "real causes, while other causes are discarded as fictitious (discounted). Kelley (1973) posits that the role of a given cause in producing a given effect is discounted if other plausible causes are also present. That is, if external (environmental) factors propel the action, the attributor will tend to believe causality is attributable to the environment, as opposed to making internal attributions to the subject.

Augmentation is a reverse version of the discounting principle. Here, the subject is taking action contrary to the pressures of the environment. Since the action is surprising and unexpected given externalities, the observer will tend to believe that the action is attributable to internal motivating factors. In the case of a salesperson making a series of inappropriate decisions (i.e., decisions counter to some norm of appropriateness based upon company training or corporate culture), the augmentation principle would suggest that the sales manager will likely attribute the cause to internal factors the individual's ability and effort, not to territory difficulty, luck, or other external factors.

Study One: Outcome Bias

Study One of the dissertation draws upon the outcome bias literature from the field of behavioral decision theory

as discussed above to test whether sales managers systematically overrely on outcome information when evaluating sales personnel. This prior research suggests the following are the primary dependent variables of interest:

1) decision quality - the sales manager's perceptions of the correctness/competence of the salesperson's decisions. Such competence previously has been assessed in terms of the quality of the decision in light of stated probabilities provided within a scenario.
2) performance evaluation - the sales managers' attributions related to general evaluations of the salesperson and his/her performance.
3) the number of cognitive responses elicited from the subjects as an assessment of the level of information processing.
4) internal versus external attributions of the manager related to the salesperson's performance.

As noted previously, the outcome bias has already been empirically tested across several work domains (c.f., Mitchell and Kalb 1981; Baron and Hershey 1988; Lipschitz 1989; Mowen and Stone 1992; and Marshall and Mowen 1993). These studies either held decision appropriateness constant while manipulating outcome (resulting in only a partial test of the true dynamics of the outcome bias), or manipulated both decision appropriateness and outcome via scenarios.

The outcome manipulation (e.g., bad/good) has been consistently strong across studies. Unfortunately, the decision appropriateness manipulation (appropriate/ inappropriate) has been somewhat problematic.

A key criticism of previous decision appropriateness manipulations is the issue of just what a "normatively appropriate" decision is for a given situation. For example, Lipschitz' (1989) manipulation of decision appropriateness actually resulted in a scenario in which military officers were forced to violate orders in order to make the normatively correct decision. Marshall and Mowen (1993) utilized a sales planning department's recommendation in their scenario in order to establish a normatively correct decision, yet upon debriefing numerous respondents indicated that it could be appropriate for the salesperson to ignore the advice of the sales planning department if the salesperson had "insider" information him-/herself. Another criticism of the decision appropriateness manipulation in previous studies is that respondents were only required to make one very narrowly focused decision.

Thus, it was critical that the present study provide a better manipulation of decision appropriateness, and as a result allow for a better understanding of the dynamics of the outcome bias. To this end, meetings were held with personnel from the participating sales organization in order to ascertain what sorts of decisions are typically made by their salespeople and what might constitute appropriate and

inappropriate decisions. Ultimately, scenarios were developed to maximize the level of realism for the respondents.

Hypotheses. Based upon the previous discussion of the literature and prior research findings, the following hypotheses were developed related to the outcome bias.

H1 A two-way interaction will occur between decision appropriateness and outcome when the sales manager rates the quality of the salesperson's decision. This hypothesized two-way interaction is depicted in Figure 4.

H1a In the inappropriate decision condition, the sales manager's evaluations of the quality of the salesperson's decisions will be significantly less favorable when the outcome is bad than when the outcome is good.

H1b In the appropriate decision condition, no differences in the sales manager's evaluations of the quality of the salesperson's decisions will occur across levels of outcome.

H2 When rating the general performance of the salesperson rather than the quality of the salesperson's decisions, a main effect will occur for outcome such that the sales manager will only take into account outcome information when making ratings, without regard to decision appropriateness information.

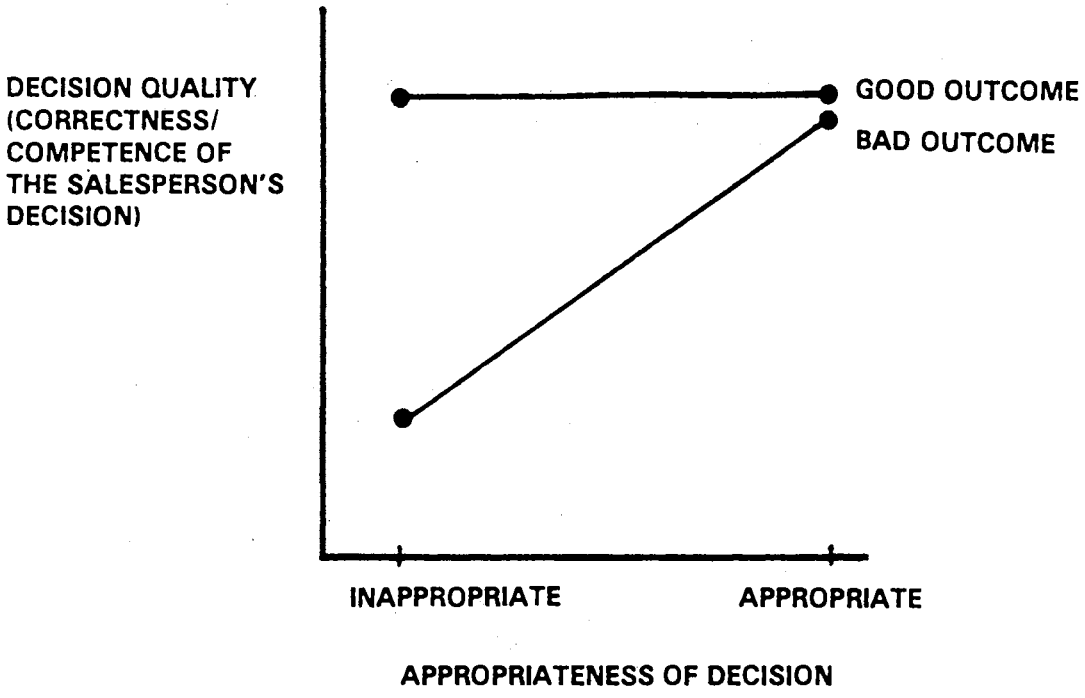


Figure 4. Portrayal of Interaction Described in H1

H3 When rating the general performance of the salesperson, a three-way interaction will occur among decision appropriateness, outcome, and rating period that may be described in the following manner:

1) When decision appropriateness and outcome information are both consistently good or both consistently bad across the three rating periods over time, general performance ratings of the salesperson by the sales manager will be stable at a relatively high and low level respectively across the three time periods.
2) When an appropriate decision is made but the outcome is bad, general performance ratings of the salesperson will decline over the three time periods.
3) When an inappropriate decision is made but the outcome is good, general performance ratings of the salesperson will improve over the three time periods.

This three-way interaction is depicted in Figure 5.

H4 Sales managers will exhibit significantly greater levels of information processing in the inappropriate decision condition than in the appropriate decision condition.

H5 Sales managers who are provided outcome information will make significantly more internal attributions regarding the performance of the salesperson than sales managers who are not provided outcome information

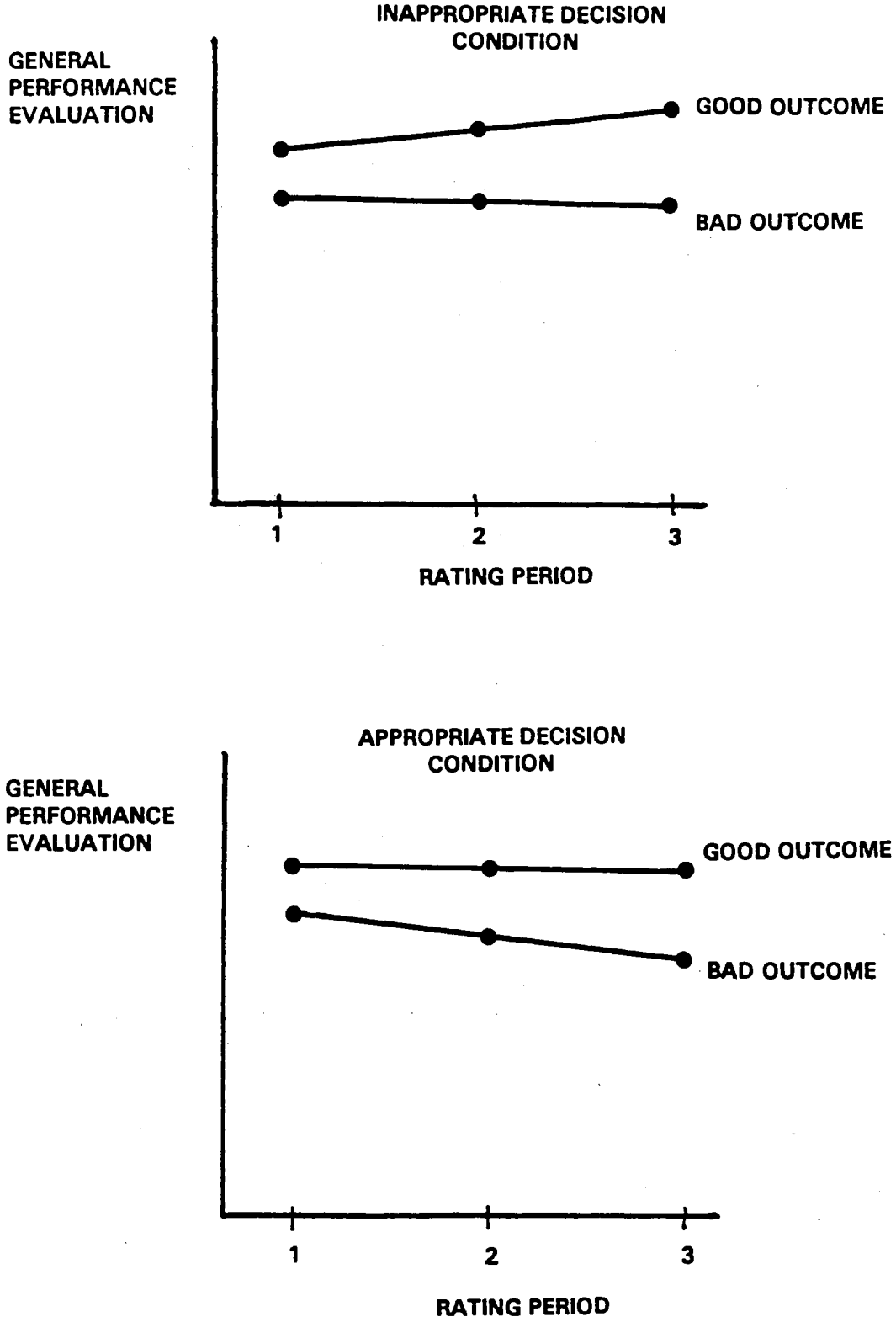


Figure 5. Portrayal of Interaction Described in H3

Order Effects Bias

A new theory of order (i.e., primacy/recency) effects seeks to explain and predict differential effects on evaluative judgment based upon the order in which information is presented. This belief-adjustment model (Hogarth and Einhorn 1992) shows promise based on a few initial studies, but the applicability of its predictions across different domains of decision making needs to be empirically tested. Because sales managers receive performance information on their salespeople that may be good or bad, because this information may be received in differential order, and because the resulting evaluations based upon this information may be performed incrementally or at the end of a rating period, the domain of salesforce performance evaluation regimen seems an ideal fit for an empirical test of the predictions of the belief-adjustment model.

Prior Research on Order Effects

An order effects bias concerns how the manner in which order of presentation of information can affect its salience and thus its importance as a predictive cue (Hogarth 1987). For example, many studies have shown that the order in which information is presented can produce so-called "primacy" or "recency" effects. (As discussed in Chapter I, Hogarth and Einhorn (1992) synthesized the results of years of studies on primacy and recency effects into their model.)

In general, the order effects bias operates as follows: When presented with a sequence of information inputs, sometimes the earlier items dominate the individual's final opinion (a primacy effect) and sometimes the latter (a recency effect). From a normative viewpoint, however, the order of presentation should not affect one's final opinion. Moreover, it appears that primacy and recency effects can be manipulated to some extent by task characteristics (Hogarth 1987). (Note: The order effects bias is part of a huge body of literature related to information integration. For a detailed review beyond what is presented in this chapter please see Anderson 1981.)

Anderson (1971) has explained primacy and recency effects found in various experimental conditions in impression-formation studies by an attention hypothesis. When only a final judgment is required, primacy effects result from the decreased attention paid to information presented later to the evaluator. The attention hypothesis explains recency effects when repeated judgments are required by proposing that the additional response requirements force an increase in attention to the later information.

Unfortunately, the empirical evidence for primacy versus recency across various conditions has been mixed and there is disagreement concerning what kinds of order effects are most prevalent. For example, Nisbett and Ross (1980) have stated that several decades of psychological research

have shown that primacy effects are overwhelmingly more probable, a conclusion contradicted by Davis (1984) in a review of studies of decision making by juries that indicated greater prevalence of recency effects. The work by Anderson and his colleagues have demonstrated both primacy and recency (Anderson 1981). Einhorn and Hogarth (1992) argued that much of the discrepancy in the literature is attributable to the fact that there are at least four paradigms in which order effects have been investigated, and each paradigm uses different operational definitions. Thus, in addition to the inherent complexity of the phenomenon, it should not be surprising that investigators from different traditions disagree about the relative likelihood of attaining primacy or recency effects.

Belief-Adjustment Model. In an effort to achieve some "order" out of the chaos of previous order effects research, Einhorn and Hogarth (1992) have proposed a general model of order effects called the "belief-adjustment model." Their conceptualization is grounded in two important factors: First, although the object of the belief updating task can cover a wide range of types of beliefs (e.g., causal hypotheses, attitudes, estimates of quantities, and so on) it must be well-specified. The operational restriction is that opinion can be represented on a predetermined scale---it is the location on the scale that changes when beliefs are revised. Second, order effects are estimated by comparing the final judgments of subjects that have

processed the same items of information but in different orders. Typically, order effects studies have involved comparisons between two or more groups of subjects. However, on occasion within-subject analysis is also possible (c.f., Shanteau 1970).

The belief-adjustment model proposed by Einhorn and Hogarth (1992) is based upon their review of over 40 order effects studies reported in the literature. The authors sought a parsimonious classification scheme for the effects based upon several conditions inherent across those studies. These conditions are defined below:

1) type (consistency) of evidence - is the evidence presented to subjects consistent or mixed? That is, is the evidence being evaluated over the stages all good, all bad, or a combination of the two?
2) order of evidence - when the type of evidence is mixed, in what order does the positive/negative evidence fall?
3) response mode - the manner in which subjects' judgments are elicited. Two response modes have commonly been employed in the order effects literature: a) a step-by-step (abbreviated SbS) procedure in which subjects are asked to express their beliefs after integrating each piece of evidence in a given sequence; and b) an end-of-sequence procedure (abbreviated EoS), where subjects only report their opinions after all the information has been presented. (Note: Throughout the remainder of this dissertation the abbreviations of

these alternative response modes will be used.)

4) task complexity - the amount of information to which the subject must respond and the subject's familiarity with the task involved. Thus, task complexity is viewed as an increasing function of the amount of information and lack of familiarity with the task.

5) number of stages of evidence items to be evaluated - two kinds of effects would be predicted as the number of stages increases. First, subjects could become fatigued if asked to process many items of information concerning the same topic. Second, as more information accumulates, one would expect beliefs to become less sensitive to the impact of new information (i.e., for someone very knowledgeable about a topic, an incremental bit of information will represent a small part of the total relevant evidence already processed such that beliefs are more resistant to change). Both of these effects imply a force toward primacy over time.

The predictions of the belief-adjustment model are depicted in Table I.

Empirical Support for the Belief-Adjustment Model. Three empirical tests of the belief-adjustment model have been reported in the literature. Hogarth and Einhorn (1992) conducted a series of five studies employing an experimental design with student subjects. Experiments 1 and 2 tested for order effects in the updating of beliefs based upon

TABLE I

SUMMARY OF ORDER EFFECTS PREDICTIONS BASED UPON THE HOGARTH AND EINHORN (1992) BELIEF-ADJUSTMENT MODEL

		Type of evidence			
		Mixed		Consistent	
		Response mode		Response mode	
		End-of-Sequence	Step-by-Step	End-of-Sequence	Step-by-Step
Short series					
	Simple	Primacy	Recency	Primacy	No effect
Complexity					
	Complex	Recency	Recency	No effect	No effect
Long series					
	Simple	Force toward primacy	Force toward primacy	Primacy	Primacy
Complexity					
	Complex	Force toward primacy	Force toward primacy	Primacy	Primacy

consistent positive and negative evidence, respectively. Subjects evaluated four scenarios covering a gamut of content issues. Students received an initial description (the stem) and made an initial rating. Then, two additional pieces of information were presented in separate paragraphs. Strong and weak forms of evidence were presented in either a strong-weak or weak-strong order. Those in the SbS condition made two additional ratings; those in the EoS condition made only one additional rating after receiving all the information. As predicted, no order effects were exhibited when the information was consistently positive or negative.

Experiments 3, 4, and 5 utilized mixed evidence (positive-negative/negative-positive) and predicted a recency effect. Experiment 3 involved two pieces of evidence, while Experiment 4 involved four pieces of evidence. Experiment 5 differed from Experiment 4 in that instead of having subjects rate their initial beliefs after reading the stem of the scenario, they were told to imagine that their initial beliefs were a particular value (based on the averages for the stems of the scenarios obtained in the earlier experiments). Strong support for a recency effect was attained in all cases. However, in addition to a significant main effect for response mode, a significant interaction occurred between response mode and order of information such that in the SbS condition, judgments were significantly impacted by recency of information but in the

EoS condition, no differences in judgments occurred based upon order of information. This finding was consistent with Einhorn and Hogarth's assertion that EoS responding tends to reduce the recency effect.

The two other similar experimental tests of the belief-adjustment model reported in the accounting literature have been administered to practicing auditors. Both Ashton and Ashton (1988) and Tubbs, Messier, and Knechel (1989) reported results consistent with the predictions of the belief-adjustment model.

To summarize, the Hogarth and Einhorn (1992) belief-adjustment model predicts a recency effect overall, but also a significant interaction between order of evidence and response mode such that order of evidence is significant only in the SbS condition (as demonstrated by a simple effects test). Thus, EoS responding tends to reduce the recency effect.

Study Two: Order Effects Bias

As previously indicated, order effects (i.e., primacy versus recency) have not been examined within the domain of salesforce performance evaluation. The present study utilized predictions from Hogarth and Einhorn's (1992) belief-adjustment model to test for order effects when sales managers evaluate sales personnel. The dependent variable of interest is the sales manager's general attributional evaluation of the salesperson's performance. This is the

same index utilized by Marshall and Mowen (1993) and in Study One of this dissertation.

As outlined in the previous section, the belief-adjustment model requires that to predict order effects, five questions must be addressed. For purposes of the present study, this section presents those questions placed into the context of salesperson evaluation. The answers to the questions are then given to provide the design framework for the study.

1) Is the information used by the sales managers to rate sales personnel all positive, all negative, or a combination of positive and negative (type of evidence - consistent or mixed)?
2) When evidence in mixed, is the positive/negative evidence about the salesperson first or last in a series of evaluations (order of evidence)?
3) After an initial evaluation, is another evaluation completed only after all available information is reviewed (an EoS response mode), or are interim evaluations completed after each piece of new information is received (an SbS response mode)?
4) Is the task of evaluating a salesperson simple (involving few words) and relatively familiar to the sales manager, or complex (involving many words) and relatively unfamiliar to the sales manager?
5) Is the number of stages in which evidence is presented short (2-12 iterations), long (greater than 20

iterations), or somewhere in between?

In a salesperson evaluation situation, answers to the first three questions would likely be situationally dependent, while the fourth and fifth questions should be quite stable across most occurrences. That is, one would expect sales managers to encounter both positive and negative information about their ratees, that this information might appear in a variety of sequential orders, and that evaluations might be completed after each round of evidence is received and/or after all evidence has been marshalled. In order to address the first three questions effectively within the domain of salesforce evaluation, both order of evidence and response mode must be manipulated in the present study. Because the belief-adjustment model predicts no order effect on evaluations when evidence is consistent, the present study considered only a mixed evidence situation.

Because the task of evaluating a salesperson is clearly complex, the present study kept task complexity constant at a complex level. Likewise, it is highly unlikely that more than 12 iterations of performance evidence would be utilized by the sales manager during a given evaluation period, therefore the number of stages in which evidence is presented in the present study was kept constant at three (including the initial information, or stem). This allowed for sufficient iterations to test for order effects, yet was few enough to avoid subject burn-out.

Hypotheses. Based upon the previous discussion of the predictions of the adjustment model and prior research findings regarding those predictions, the following hypotheses were developed related to the order effects bias.

H6: Sales managers will overrely on more recent information when evaluating the salesperson, thus resulting in a significant main effect for order of information on general performance ratings of the salesperson.

H7: A two-way interaction will occur between order of information and response mode.

H7a: In the step by step (SbS) response mode condition, sales manager evaluations of the salesperson will be significantly impacted by recency of information.

H7b: In the end of sequence (EoS) condition, no differences in sales manager evaluations of the salesperson will occur based upon order of information.

Figure 6 depicts the hypothetical "fishtail" pattern showing recency effects for mixed evidence. Table II presents a summary of all the research hypotheses tested in this dissertation, by study.

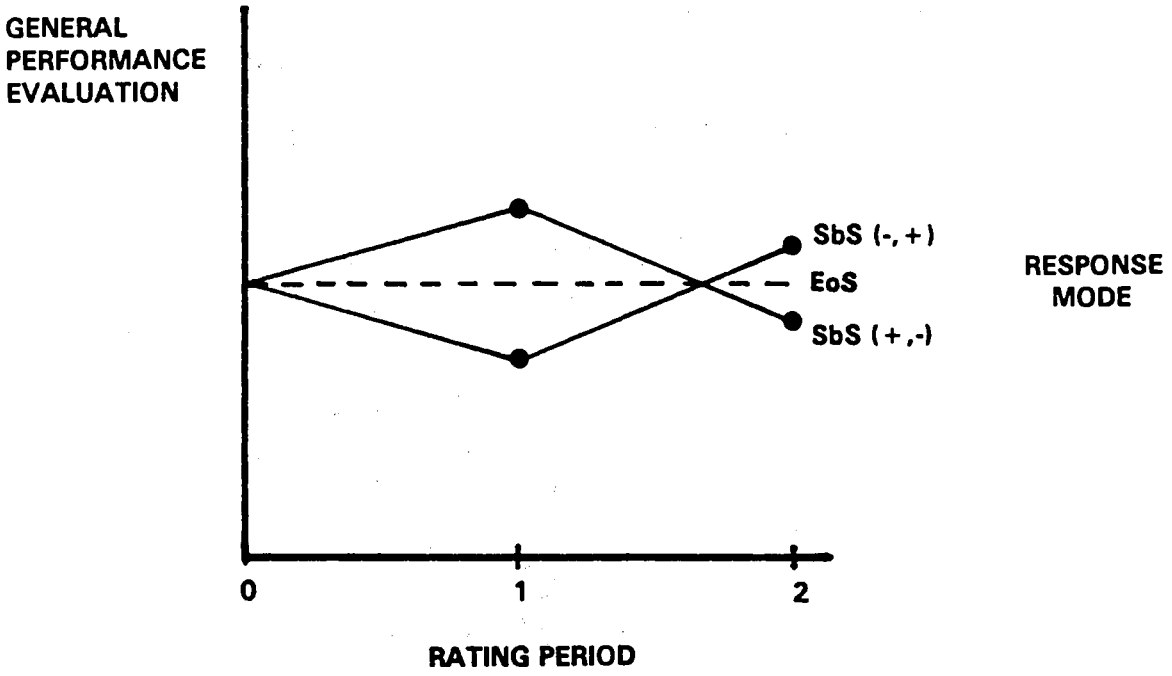


Figure 6. Hypothetical "Fishtail" Pattern Showing Recency Effects for Mixed Evidence

TABLE II

SUMMARY OF RESEARCH HYPOTHESES

STUDY ONE: OUTCOME BIAS

H1 A two-way interaction will occur between decision appropriateness and outcome when the sales manager rates the quality of the salesperson's decision. This hypothesized two-way interaction is depicted in Figure 4.

H1a In the inappropriate decision condition, the sales manager's evaluations of the quality of the salesperson's decisions will be significantly less favorable when the outcome is bad than when the outcome is good.

H1b In the appropriate decision condition, no differences in the sales manager's evaluations of the quality of the salesperson's decisions will occur across levels of outcome.

H2 When rating the general performance of the salesperson rather than the quality of the salesperson's decisions, a main effect will occur for outcome such that the sales manager will only take into account outcome information when making ratings, without regard to decision appropriateness information.

H3 When rating the general performance of the salesperson, a three-way interaction will occur among decision appropriateness, outcome, and rating period that may be described in the following manner:

1) When decision appropriateness and outcome information are both consistently good or both consistently bad across the three rating periods over time, general performance ratings of the salesperson by the sales manager will be stable at a relatively high and low level respectively across the three time periods.
2) When an appropriate decision is made but the outcome is bad, general performance ratings of the salesperson will decline over the three time periods.
3) When an inappropriate decision is made but the outcome is good, general performance ratings of the salesperson will improve over the three time periods.

This three-way interaction is depicted in Figure 5.

H4 Sales managers will exhibit significantly greater levels of information processing in the inappropriate decision condition than in the appropriate decision condition.

H5 Sales managers who are provided outcome information will make significantly more internal attributions regarding the performance of the salesperson than sales managers who are not provided outcome information

STUDY TWO: ORDER EFFECTS BIAS

H6: Sales managers will overrely on more recent information when evaluating the salesperson, thus resulting in a significant main effect for order of information on general performance ratings of the salesperson.

H7: A two-way interaction will occur between order of information and response mode.

H7a: In the step by step (SbS) response mode condition, sales manager evaluations of the salesperson will be significantly impacted by recency of information.

H7b: In the end of sequence (EoS) condition, no differences in sales manager evaluations of the salesperson will occur based upon order of information.

CHAPTER IV

RESEARCH METHODOLOGY

Introduction

This dissertation tested theory-based hypotheses about the causal relationships between an outcome bias and sales manager evaluations of field sales personnel and an order effects bias and sales manager evaluations of field sales personnel. In order to establish such causal relationships, two experimental studies were performed, one related to each of the two biases. The experiments took place in a field setting. Stimulus materials were mailed to a total of 300 sales managers from two divisions of a major U.S. consumer products company.

This chapter consists of six sections for each of the two studies: 1) an overview of the design; 2) the stimulus materials employed in the study; 3) the subjects utilized; 4) the procedure; 5) the measurement of the variables of interest; and 6) the analytical methods used to test the research hypotheses. The results of the data analyses are presented in the Chapter V.

Study One: Outcome Bias

Design Overview

The outcome bias study utilized a 2 X 3 X 3 mixed factorial design with two between subjects factors and a repeated measures factor. The variables manipulated in the between subjects design were information about the appropriateness of a salesperson's decision (bad/good) and information about an outcome achieved by the salesperson (bad/good/no outcome-control group). The repeated measures are three rating periods for the same salesperson. The group that received no outcome information served as a control group for conducting a manipulation check for decision appropriateness. Additionally, the no outcome group was necessary to test H5. Information was presented in a scenario format. The basic design of Study One is depicted in Figure 7.

Stimuli

Development of Stimuli. The specific performance factors (both the decisions and the outcomes) utilized in the scenarios, as well as the particular sales situations depicted in the scenarios, were developed based upon a two-step procedure. First, the author worked in the field with a veteran sales manager from the participating organization. Recent performance appraisal forms, job descriptions, and standards of performance were gathered during the trip.

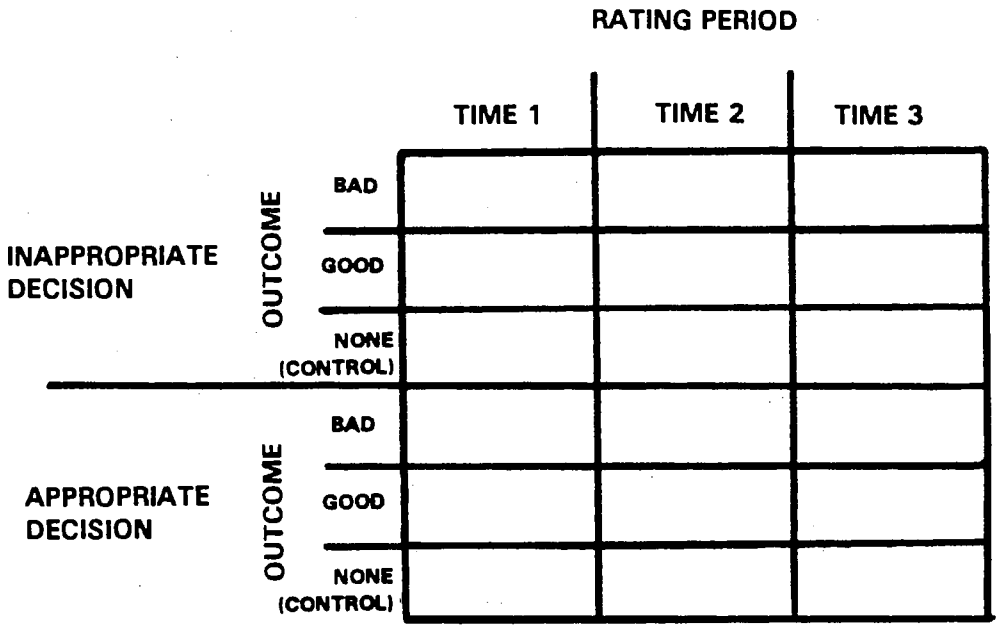


Figure 7. Outcome Bias Experimental Design

Second, the author met with members of upper level sales management at the company's home office to gain specific suggestions on scenario content. Drafts of the scenarios were then sent to the home office staff to check for realism.

Pilot Study. The pilot study for the dissertation is available for review in the published literature (see Marshall and Mowen 1993). Additionally, the three scenarios developed for the present study were tested on a sample of business students for readability and realism. Comments elicited via cognitive responses and discussions with students after administration of the materials indicated that the subjects understood the task they were asked to perform, took the task seriously, and viewed it as realistic.

Format of Stimuli. Subjects received a cover letter from the appropriate divisional vice president of sales and a packet of information. The first page provided general instructions to the respondent and described the task at hand---a series of evaluations of the performance of a salesperson named "Smith." Emphasis was placed on the importance of the subjects' tracking through the booklet in sequential order.

The preliminary information about Smith was constant at a "satisfactory" level across all experimental conditions. The next few pages of the materials provided a

series of three role-play scenarios with information varied about the decision appropriateness and outcome of recent sales opportunities. Subjects made their ratings after each scenario. Cognitive responses, biographical information, and answers to several other questions were gathered at the end of the questionnaire. Packets were mailed from and returned to the University Center at Tulsa.

Upper management at the participating company expressed great concern about two issues: 1) protection of anonymity of respondents, and 2) length of the packets. Consequently, only one mailing was possible and the number of ancillary exploratory questions was limited. The entire packet of materials for Study One is presented in Appendix A.

Subjects

Subjects for Study One were 180 sales managers from two divisions of a major United States consumer products marketer. Based upon the design, this allowed for 15 subjects per cell even with only a 50 percent response rate. The actual number of usable responses to Study One was 143, a 79 percent response rate.

Procedure

Subjects were assigned to the treatment conditions on a random basis, ensuring that each cell had the appropriate ratio of subjects from each of the two sales divisions of the company. The three scenarios were counterbalanced

across all conditions. As previously mentioned, the packet was sent out by mail with a cover letter from the appropriate divisional vice president of sales soliciting full support. Added appeal for response was generated by providing each subject with a gift of an Oklahoma State University ballpoint pen. Postage stamps were used in place of metering, and a stamped, pre-addressed return envelope was provided with each packet. Subjects were told that the study had the potential to substantially improve sales management practice.

Measurement

Manipulation Checks. In order to assess the effectiveness of the manipulation of decision appropriateness (inappropriate/appropriate), subjects in the no outcome/control condition were utilized and ratings on decision quality scale items were compared for significant differences between those assigned to the inappropriate versus appropriate decision conditions. As a manipulation check for outcome, subjects in the bad or good outcome conditions were asked to describe the immediate outcome (result) of salesperson Smith's decision utilizing a single item 7-point Likert scale ranging from "very unfavorable" to "very favorable."

Dependent Measures. As outlined in Chapter IV, the dependent variables of interest in Study One were as follows: 1) decision quality (i.e., the correctness/

competence of the salesperson's decision); 2) a general attributional performance evaluation of the salesperson; 3) global measures of internal/external attributions made by the sales manager; and, 4) the level of information processing by the sales manager as indicated by the number of cognitive responses[1]. In none of these cases was any new scale development necessary. Scales employed as dependent measures and their sources are summarized in Table III.

Exploratory Covariates. At the end of the questionnaire, three measures were taken to be utilized as potential covariates in the study.

1) A single-item measure of how much like oneself the subject considered the salesperson in the scenario to be. The theoretical basis for taking this measure was the so-called "similar to me" effect (Byrne 1961). Rand and Wexley (1975, p. 536) described the effect, "similar to me," as follows: "to the extent that a person offers consensual validation by demonstrating similarity to us in some way, an interaction with that individual will be perceived as being rewarding and will lead to positive feelings toward that individual."

---

[1] The format for soliciting cognitive responses in this study varied slightly from some classic definitions of cognitive responses (c.f., Bettman 1979, p. 113). The modified approach was necessitated by the nature of the evaluation subjects were asked to perform. A discussion of the development of the measure employed here is provided in Chapter V.

TABLE III

DEPENDENT MEASURES

NOTE: See Appendix A for actual questionnaire and more detail of scales.

A. Decision Quality (DECQUAL): Correctness/competence of the salesperson's approach to his/her business. Five items, each utilizing a 7-point Likert scale with 7 representing a more favorable evaluation or a higher level of the construct. Source: Marshall and Mowen 1992. Cronbach alpha in pilot study: .89. Items:

* Smith made an excellent decision under the circumstances.
* I consider Smith to be a very poor decision maker. (Reverse scored)
* I view Smith to be highly competent as a decision maker.
* Smith made the wrong decision. (Reverse scored)
* Given the circumstances, the decision made by Smith was correct.

B. Performance Evaluation (PEREVAL): Attribution-based general performance evaluation of the salesperson. Seven items, each utilizing a 7-point Likert scale with 7 representing a more favorable evaluation or a higher level of the construct. Source: Marshall and Mowen 1992. Cronbach alpha in pilot study: .90. Items:

* Please rate Smith's sales ability.
* How would you rate Smith's overall level of effort in obtaining sales?
* Rate Smith's overall job performance.
* How would you categorize Smith? (Not likeable to very likeable)
* Do you think Smith deserves a promotion or bonus?
* How would you describe Smith as a person? (Bad to good)
* Rate Smith's skill level as a salesperson.

C. Global Internal/External Attributions (INATTRIB and EXATTRIB): Two items, each utilizing a 7-point Likert scale with 1 indicating "extremely unimportant" and 7 indicating "extremely important." Source: Adapted from Mitchell and Kalb (1981).

* How important do you feel circumstances of the situation were in contributing to Smith's decisions?
* How important do you feel Smith's personal characteristics were in contributing to Smith's decisions?

D. Information Processing (INFOPROC): After the three scenarios and evaluations, subjects were asked to write down what factors they considered in making their ratings. These were summed and compared across subjects assigned to the inappropriate versus appropriate decisions conditions in order to test H4.

2) A two-item index to assess perceived overall level of realism of the events described in the scenario by the respondents. This measure was created for this study, and was included as a check for ecological validity.
3) A five-item index for propensity toward risk taking on the job by respondents, created for this study.

These exploratory covariate measures are presented in Table IV. It should be noted that these questions were asked after the main experimental questions. This was because only one mailing was possible, thus all questions were answered in one sitting. If contamination were to take place by exposure to stimuli or prior measures it was deemed preferable that the covariate measures experience the contamination rather than the key dependent measures.

Exploratory Dependent Measure. A global rating of salesperson Smith was taken utilizing a 0 - 100 scale anchored from "Unacceptable" to "Far Exceeds Expectations." These descriptors were drawn from the actual performance appraisal document of the participating company. This item is also presented in Table IV.

Other Exploratory Measures. Several diverse measures were taken on the last two pages of the questionnaire that are not directly related to the dissertation but may prove useful later in further analysis. Additionally, all questionnaires were coded so that the responses from the two different sales divisions may later be analyzed separately

TABLE IV

EXPLORATORY MEASURES

NOTE: See Appendix A for actual questionnaire and more detail of scales.

A. Similar to Me (LIKEME): One item, utilizing a 7-point Likert scale with 7 representing "like me" and 1 representing "not like me." Developed specifically for this study, based upon the work by Byrne (1961) and Rand and Wexley (1975) on the effect, "similar to me."

   * How much like yourself do you consider Smith to be?

B. Ecological Validity (ECOLVAL): Level of perceived realism of the scenarios to respondents. Two items, each utilizing a 7-point Likert scale with 7 representing "very likely" and 1 representing "very unlikely." Developed specifically for this study. Items:

   * Rate the degree to which the types of decisions Smith faced could happen to you or your people on the job.
   * Rate the extent that the problems you've read in the previous scenarios could actually happen.

C. Propensity for Individual Risk Taking on the Job (RISKTKNG): Five items, each utilizing a 7-point Likert scale with 7 representing "strongly agree" and 1 representing "strongly disagree." Items:

   * As a manager, I am willing to risk a small loss in order to achieve a large gain.
   * I find that on the job I'm a highly conservative manager. (Reverse scored)
   * As a manager, I am willing to take stands my boss may disapprove of.
   * I find that most of my decisions on the job are made from my gut.
   * I enjoy the thrill of taking chances in my decision making on the job.

D. Global Rating of salesperson Smith (GLOBAL): A single item measure scaled 0 - 100, with anchor words drawn from the participating company's performance appraisal document. Item:

   * On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

and compared for purposes of reporting results back to the company. Again, the reader is referred to Appendix A for the complete set of materials from Study One.

Data Analysis

Please refer back to Table II for the specific hypotheses. First, principal component analysis, Cronbach alphas, and item-total correlations were performed on the indices to assess structure and reliability.

In terms of the research hypotheses, H1 was tested via analysis of variance (ANOVA) using the decision quality scale (DECQUAL) as the dependent measure. H2 was tested via ANOVA using the attribution-based general performance scale (PEREVAL) as the dependent measure. H3 was tested via repeated measures ANOVA. Polynomial contrast tests for trend analysis were conducted to assess linearity of ratings across time. The dependent measure was the attribution-based general performance scale (PEREVAL). H4 was tested via ANOVA utilizing the summed cognitive response scores as a dependent measure of information processing (INFOPROC). Finally, H5 was tested via ANOVA using the internal attribution scale item (INATTRIB) as a dependent measure.

Additionally, where appropriate several exploratory analyses of covariance (ANCOVAS) were performed utilizing the three covariates described previously---similar to me (LIKEME), ecological validity (ECOLVAL), and risk taking (RISKTKNG). Also, the global rating (GLOBAL) was employed

in several analyses as an exploratory dependent measure.

Study Two: Order Effects Bias

As previously noted, Study Two ran concurrently with Study One, but with different subjects from the samepopulation. The two studies have numerous procedural similarities. Thus, to avoid redundancy this section highlights only the relevant differences in Study Two versus Study One.

Design Overview

The order effects study employed a 2 X 2 between subjects full factorial design in which response mode (SbS)/(EoS) and order of presentation of information (bad/good - good/bad) were varied. No control group was utilized in Study Two. The basic design of Study Two is presented in Figure 8.

Subjects first responded to an initial set of information representing a salesperson named Smith's performance. Across all subjects, the initial information showed performance that "meets expectations." (Note: this performance descriptor was drawn from the participating company's documents and indicates an acceptable level of performance.) The managers then received two more sets of performance information related to two decisions made by Smith regarding targeting orders from two customers. Subjects assigned to the SbS condition made two more

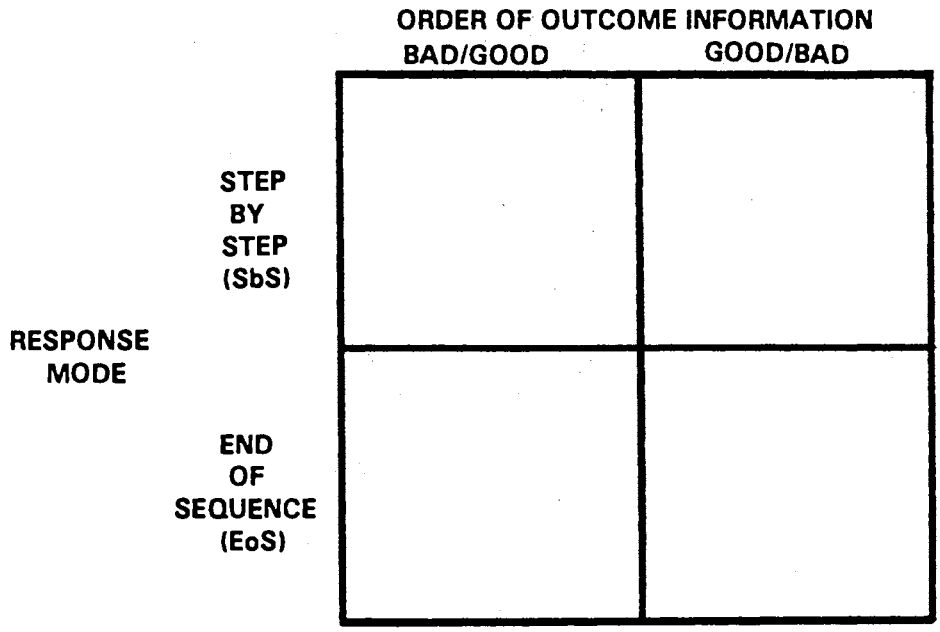


Figure 8. Order Effects Bias Experimental Design

ratings, one after each of the two scenarios. Subjects assigned to the EoS condition also received two additional scenarios after the initial baseline information, however they made only one more rating after both additional pieces of information had been received.

Stimuli

The specific scenarios utilized after the stem in Study Two were exactly the same as two of the three utilized in Study One. The two selected for use in Study Two involved the more similar decisions to be made by salesperson Smith. The entire packet of materials for the SbS condition in Study Two is presented in Appendix B, and for the EoS condition in Appendix C.

In order to maximize the saliency and effectiveness of the bad/good performance information conditions, scenarios in the bad condition presented subjects with the inappropriate decision/bad outcome combination of information that was used in Study One. Similarly, the good condition scenarios presented subjects with the corresponding appropriate decision/good outcome combination of information that was used in Study One. It was believed that such combinations would create the strongest possible manipulation of valence of performance information. Because of the exploratory nature of Study Two, a strong manipulation was highly desirable.

Pretest

The only difference in the materials utilized in Study Two versus Study One was the initial performance information (stem) that was kept at a constant "meets expectations" level. (Note that this was done to replicate previous studies utilizing the belief-adjustment model.) Therefore, in order to ensure that this initial information was readable and easily comprehensible, undergraduate business students were given this stimuli and a rating sheet. A review of the responses indicated that subjects had no trouble comprehending the task, therefore the initial information was incorporated into the Study Two materials. (This "stem" information appears in Appendix B.)

Subjects

Subjects for the study were 120 sales managers from the same company utilized in Study One. It should be made clear that this was a different set of subjects drawn from the same population as those in Study One. As before, this sample allowed for 15 subjects per cell even with only a 50 percent response rate. The actual number of usable responses to Study Two was 90, a 75 percent response rate. Thus, the overall response rate for the two studies combined was 233 out of 300 mailed, or 78 percent.

Procedure

One key difference existed in the procedure for Study

Two versus Study One. This procedural change was necessitated by the response mode manipulation (either SbS or EoS, as previously described). Those subjects assigned to the EoS condition rated the salesperson only twice---once in the beginning after the stem and once at the end, but were provided one set of interim information for review without rating. In contrast, those subjects in the SbS condition actually performed a rating on the salesperson after receiving the interim information. Thus, SbS subjects made three iterations of ratings instead of just two. (Again this is consistent with the previous studies utilizing the belief-adjustment model.)

Measurement

A manipulation check for perception of good versus bad scenarios was taken utilizing the manipulation check question for outcome and the decision quality scale (DECQUAL) from Study One.

The key dependent variable of interest in Study Two was the general attributional performance rating for the salesperson by the manager. Items comprising this scale are the same as those described in Study One, and will not be reviewed here. Other measures taken were identical to those in Study One and likewise will not be recapitulated. Please refer to Tables III and IV for these measures.

Data Analysis

As in Study One, principal component analysis, Cronbach alphas, and item-total correlations were performed on the indices to assess structure and reliability.

Analysis of Variance (ANOVA) was performed to test H6 and H7 (please refer to Table II for these hypotheses).

CHAPTER V

RESEARCH RESULTS

Introduction

This chapter presents the findings of the two studies. Study One (Outcome Bias) and Study Two (Order Effects Bias) results are presented separately. Presentation of the results for each study is organized by the following sections: 1) issues of structure and reliability of measures; 2) description of the sample; 3) tests of hypotheses; and 4) exploratory tests. In the interest of brevity, throughout this chapter for both studies the names of the indices described in Chapter IV will be abbreviated as follows (please refer to Tables III and IV for details of the measures):

1) DECQUAL - the five-item scale measuring the sales managers' perceptions of the correctness/competence of salesperson Smith's decisions.
2) PEREVAL - the seven-item scale of attribution-based general performance evaluation of salesperson Smith.
3) INFOPROC - the measure of level of information processing by the subject, created by summing cognitive responses taken at the end of the scenarios.
4) EXATTRIB - the single-item measure of subjects' level

of external attributions regarding salesperson Smith.

5) INATTRIB - the single-item measure of subjects' level of internal attributions regarding salesperson Smith.

6) GLOBAL - the exploratory dependent measure of salesperson Smith utilizing a 0 - 100 scale.

7) LIKEME - the single-item measure of the effect, "similar to me," designed to be utilized as an exploratory covariate.

8) ECOLVAL - the two-item measure of ecological validity, designed to be utilized both as an exploratory covariate and as an overall check on the perceived realism of the scenarios.

9) RISKTKNG - the five-item measure of individual risk taking, designed to be utilized as an exploratory covariate.

Descriptive statistics for both studies are provided in Appendix D.

Study One: Outcome Bias

Structure and Reliability of Indices

The key indices of DECQUAL and PEREVAL have been utilized in previous research. Therefore, a comparison exists for checking the stability of their structure in the present application versus past applications, as well as the reliability of the measures across applications. Because these scales were administered three times during the present experiment (after exposure to each scenario), it was

critical that similar structure and reliability be exhibited across these three current applications. Thus, results of the analysis of scale structure and reliability are reported here for each of the three administrations of the scales.

Structure. Principal component analysis was employed with a Promax (non-orthogonal) rotation to assess underlying structure. Promax rotation was utilized due to anticipated interfactor correlations. In fact, the interfactor correlations were 0.506 for rating period one, 0.577 for rating period two, and 0.535 for rating period three.

The principal component analysis yielded two distinct factors, structured exactly as prespecified in Chapter IV. The structure was identical in each of the three applications of the indices. A summary of the factor loadings is presented in Table V.

Reliability. Cronbach alphas and item-total correlations were calculated for the DECQUAL and PEREVAL scales for each of the three rating periods. Alphas for the scales ranged from .87 to .94 across the three rating periods. (Note that in the Marshall and Mowen 1993 pilot study, alphas were .89 and .90 for the DECQUAL and PEREVAL scales, respectively.) Item-total correlations were generally high---78 percent of the correlations reported across the three applications of the two scales were greater than r=.70. Details of the reliability analysis for the DECQUAL and PEREVAL scales are presented in Tables VI and VII.

TABLE V

FACTOR LOADINGS ACROSS
THREE RATING PERIODS
STUDY ONE

	RATING PERIOD ONE		RATING PERIOD TWO		RATING PERIOD THREE	
	FACTOR1 DECQUAL	FACTOR2 PEREVAL	FACTOR1 DECQUAL	FACTOR2 PEREVAL	FACTOR1 DECQUAL	FACTOR2 PEREVAL
Q2	.88	.06	.84	.09	.88	.04
Q4	.82	.13	.83	.17	.87	.10
Q7	.88	-.05	.89	-.11	.82	.00
Q9	.98	-.13	.91	-.03	.92	-.02
Q11	.92	.02	.86	.10	.89	.05
Q1	.00	.87	.01	.90	-.04	.94
Q3	.20	.74	.04	.90	-.06	.95
Q5	.20	.79	.15	.82	.17	.84
Q6	.04	.65	.10	.66	.08	.75
Q8	.00	.46	-.08	.70	.16	.60
Q10	-.20	.72	.07	.64	.04	.69
Q12	.01	.88	-.01	.93	-.02	.95

NOTE: Please refer to the questionnaire in Appendix A for specific items.

TABLE VI

RELIABILITY ANALYSIS OF DECQUAL SCALE
USING STANDARDIZED VARIABLES
STUDY ONE

	ITEM-TOTAL CORRELATIONS		
	RATING PERIOD 1	RATING PERIOD 2	RATING PERIOD 3
Q2	.87	.85	.85
Q4	.84	.89	.88
Q7	.77	.72	.74
Q9	.86	.84	.86
Q11	.90	.88	.84
Cronbach α	.94	.94	.94

NOTE: Please refer to the questionnaire in Appendix A for specific items.

TABLE VII

RELIABILITY ANALYSIS OF PEREVAL SCALE
USING STANDARDIZED VARIABLES
STUDY ONE

	ITEM-TOTAL CORRELATIONS		
	RATING PERIOD 1	RATING PERIOD 2	RATING PERIOD 3
Q1	.78	.85	.87
Q3	.76	.88	.87
Q5	.84	.86	.90
Q6	.56	.65	.73
Q8	.36	.54	.61
Q10	.47	.60	.63
Q12	.81	.87	.91
Cronbach α	.87	.92	.93

NOTE: Please refer to the questionnaire in Appendix A for specific items.

Based upon the results of this analysis, the two scales were deemed sufficiently internally reliable for usein the present study.

Unfortunately, the five-item scale for RISKTAKING targeted for use as an exploratory covariate measure performed poorly in the reliability analysis. Item-total correlations ranged from 0.06 to 0.17, with a Cronbach alpha of .26. As a result, in all subsequent analyses only a single-item measure of RISKTKNG was employed: "I enjoy the thrill of taking chances in my decision making on the job." This item employed a 7-point Likert scale with 1 representing "strongly disagree" and 7 representing "strongly agree."

The two-item ECOLVAL scale was assessed for internal reliability to ensure its viability as a covariate measure. Item-total correlation for each item was .74, with a Cronbach alpha of .85. Thus, this measure was deemed satisfactory for exploratory research purposes.

No other multiple item measures were employed in Study One.

Development of INFOPROC Measure

As previously stated, H4 predicted a higher level of information processing in the inappropriate versus the appropriate decision conditions. The standard approach for assessing the level of information processing is to take cognitive responses from the subjects, which was

accomplished at the end of the questionnaire. Subjects were asked to write down what factors they considered in making their previous ratings of salesperson Smith. Two judges, one of which is the author, independently summed the number of inferences drawn by each subject in the cognitive response section of the questionnaires. Each judge was blind to which condition each subject was in during the assessment. The authors then compared their results (initial r=.78). The majority of disagreements involved a minor difference in the assessment of the number of inferences for a given subject. For final coding purposes, any disagreements were rectified by judges' reviewing the specific questionnaire jointly for resolution. The resulting total cognitive response score was then coded for each subject and utilized as the dependent measure INFOPROC to test H4.

Description of Sample

As reported in Chapter IV, the number of usable responses in Study One was 143, a 79 percent response rate. Median age of respondents was 38, with an age range of 26 to 56 years old. The male/female split was 77 percent/23 percent respectively.

As would be expected from a national sample of practicing sales managers for a major consumer packaged goods company, the level of experience was very high. Median years of sales management experience was 7 years,

with a range of 1 to 38 years. In terms of sales experience, subjects exhibited a median of 13 years with a range of 2 to 40 years. Median years of total work experience was 19.5 years, with a range of 5 to 40 years.

Subjects were asked to give their best estimate of the total number of people for whom they had completed formal performance appraisals during their career. The median number of appraisals given was 20. However, comments by several of the respondents indicated confusion about whether they were to provide the number of appraisals given or the number of different people appraised. Subjects also were asked to estimate the number of times they personally had been evaluated by superiors during a formal performance appraisal process during their career. This question yielded a median of 14. However, once again evidence of confusion surfaced, as some respondents were quoting quarterly evaluations and others yearly evaluations. Due to this confusion, the overall accuracy of the responses to the two questions about performance appraisals is suspect.

Still, the overall profile of the respondents is a group of seasoned veterans of sales and sales management who are comfortable with giving and receiving performance feedback.

Manipulation Checks

As outlined in Chapter IV, both decision appropriateness and outcome were manipulated. The results

of the manipulation checks described in Chapter IV are presented separately below. Also, the result of a check on ecological validity is reported.

Decision Appropriateness. In order to determine the effectiveness of the decision appropriateness manipulation, ANOVAs were performed utilizing only those subjects assigned to the two control conditions (n=48) in which no outcome information was given. A significant difference in subjects' perceptions of DECQUAL based upon whether the subject was assigned to the inappropriate or appropriate decision condition would yield evidence of a successful manipulation of decision appropriateness. As expected, in each of the three rating periods subjects in the inappropriate decision condition rated DECQUAL significantly lower than subjects in the appropriate decision condition (Rating Period 1 - F: 53.66, $p<.0001$, MEANS = 3.10/5.49; Rating Period 2 - F: 103.50, $p<.0001$, MEANS = 2.64/5.55; Rating Period 3 - F: 127.95, $p<.0001$, MEANS = 2.31/5.55). Thus, the manipulation of decision appropriateness was deemed successful.

Outcome. As previously described, response sheets after each of the scenarios included a question asking subjects to describe the immediate outcome (result) of Smith's decision on a 7-point Likert scale from very unfavorable to very favorable. In an ANOVA, a significant main effect for outcome on the above outcome manipulation

check item would provide evidence that the manipulation of outcome was successful.

ANOVAs yielded the expected main effect for the outcome condition on the outcome manipulation check item for each of the three rating periods. In each rating period, subjects clearly perceived the bad outcome condition as being significantly worse than the good outcome condition (Rating Period 1 - F: 84.53, $p<.0001$, MEANS = 2.37/5.05; Rating Period 2 - F: 149.95, $p<.0001$, MEANS = 2.08/5.21; Rating Period 3 - F: 130.80, $p<.0001$, MEANS = 1.90/5.07). The results indicated a successful manipulation of outcome in the present study. (It should be noted that comparable results for both manipulation checks were attained for each of the individual rating periods.)

Ecological Validity. As previously mentioned, a two-item measure was taken of perceived likelihood that the problems and decisions faced in the scenarios could happen to the manager or his/her salespeople (ECOLVAL). It should be noted here that the mean score on this 7-point Likert scale in Study One was 4.82 (n=141, S.D.=1.80), indicating an overall perception that the scenarios described realistic events. This result is not surprising in light of the involvement of the host company in the development of the stimulus materials.

Tests of Hypotheses

Hypotheses 1, 1a, and 1b. These hypotheses predicted an interaction between decision appropriateness and outcome on DECQUAL. In the appropriate decision condition it was predicted that ratings would be unaffected by outcome, but in the inappropriate decision conditions it was predicted that good outcomes would lead to significantly higher ratings than bad outcomes. Analysis of Variance (ANOVA) was performed to test H1, H1a, and H1b. ANOVA results for the dependent measure DECQUAL are presented in Table VIII. The means in each of the conditions are presented in Table IX.

A significant main effect was obtained for decision appropriateness (F: 305.14, $p<.0001$, $\omega^2=60.3\%$). Subjects rated the DECQUAL of salesperson Smith significantly lower (MEAN = 2.70) when the decision was inappropriate versus when the decision was appropriate (MEAN = 5.44). Also, a significant main effect was obtained for outcome (F: 13.60, $p<.0001$, $\omega^2=5.1\%$). Subjects in the bad outcome condition rated the salesperson's DECQUAL significantly lower (MEAN = 3.59) than did those in the good outcome condition (MEAN = 4.74).

As predicted by H1, these main effects were superseded by a two-way interaction of decision appropriateness by outcome on DECQUAL (F: 3.27, $p<.05$, $\omega^2=0.9\%$). This result is also consistent with the predicted magnification effect in which more negative responses occur when a bad outcome resulted from an inappropriate decision. The interaction is

TABLE VIII

ANOVA FOR DECQUAL
STUDY ONE

Independent Variable	DF	Type III SS	F-value	PR > F	ω^2
Decision	1	250.48	305.14	.0001	60.4%
Outcome	2	22.33	13.60	.0001	5.1%
Decision * Outcome	2	5.37	3.27	.0409	.9%

r^2 for model = .72

TABLE IX

MEANS FOR DECQUAL BY CONDITION
STUDY ONE

	OUTCOME		
DECISION	BAD	GOOD	NONE
INAPPROPRIATE DECISION	2.11 SD=.85 n=26	3.54 SD=1.03 n=19	2.68 SD=.95 n=26
APPROPRIATE DECISION	5.14 SD=.87 n=25	5.65 SD=.94 n=25	5.53 SD=.80 n=22

depicted graphically in Figure 9.

Supporting H1a, when the decision was inappropriate, ratings of the salesperson's DECQUAL were significantly lower when the outcome was bad than when the outcome was good. This conclusion is supported by an *a priori* F test showing a significant difference in the mean ratings of DECQUAL by subjects assigned to the inappropriate decision condition depending upon whether subjects received bad or good outcome information (F: 27.34, $p<.001$, MEAN in the bad outcome condition = 2.11, MEAN in the good outcome condition = 3.54).

Hypothesis 1b was not supported. An *a priori* F test revealed a significant difference in the mean ratings of DECQUAL by subjects assigned to the appropriate decision condition based upon receipt of bad versus good outcome information (F: 3.96, $p<.05$, MEAN in the bad outcome condition = 5.14, MEAN in the good outcome condition = 5.65). Thus, subjects exposed to unfavorable outcomes rated the correctness/competence of salesperson Smith's decisions significantly lower than subjects exposed to favorable outcomes, even when Smith's decisions were appropriate.

A *post hoc* test revealed that the means in the good outcome condition were significantly different (F: 58.58, $P<.0001$, MEAN in the inappropriate decision condition 3.54, MEAN in the appropriate decision condition 5.65). This pattern closely matched that found in the pilot study, but was different from the finding by Mowen and Stone (1992) of

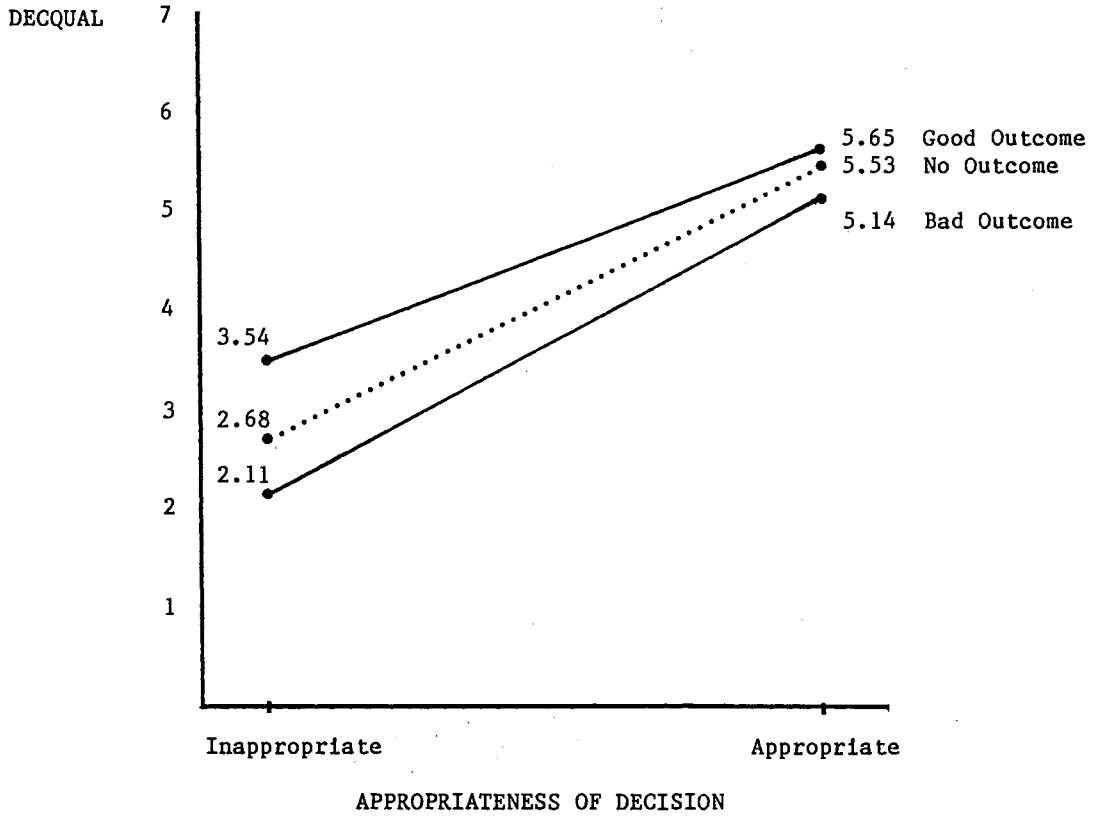


Figure 9. Two-way Interaction of Decision Appropriateness and Outcome on DECQUAL in Study One

no effect for decision appropriateness in the good outcome condition (which was the basis for H1b---see Figure 4 for a pictorial representation of the hypothesized pattern).

The pattern of the mean ratings of DECQUAL by subjects in the no outcome/control condition also deserves mention. When the decision was appropriate, *post hoc* tests showed no significant differences in mean ratings of DECQUAL between subjects in either the no outcome (MEAN = 5.53) versus bad outcome (MEAN = 5.14, F: 2.16) conditions or subjects in the no outcome (MEAN = 5.53) versus good outcome (MEAN = 5.65, F: 0.21) conditions. *Post hoc* tests indicated that when the decision was inappropriate, the mean rating of DECQUAL by subjects in the no outcome/control condition (MEAN = 2.68) fell between the mean ratings of DECQUAL for those in the bad outcome condition (MEAN = 2.11, F: 5.14, $p<.01$) and good outcome condition (MEAN = 3.54, F: 9.89, $p<.001$). Thus, in the inappropriate decision condition the mean rating of DECQUAL when no outcome information was available may be viewed as an anchor point from which ratings of DECQUAL may become more or less favorable depending upon whether information about a bad or good outcome was included in the evaluation.

To summarize, evidence was found for the hypothesized interaction of decision appropriateness and outcome on perceptions of the competence/correctness of salesperson Smith's decisions (DECQUAL). Overall, when the decisions made were inappropriate, ratings of DECQUAL were more highly

impacted by outcome than when appropriate decisions were made.

Hypothesis 2. H2 predicted a main effect for outcome on ratings of PEREVAL. To test H2, ANOVA was performed utilizing the PEREVAL index as the dependent measure. ANOVA results are summarized in Table X. The means for PEREVAL in each of the conditions are presented in Table XI.

The analysis yielded not only a main effect for outcome on PEREVAL as predicted (F: 58.51, $p<.0001$, $\omega^2=41.4\%$, MEAN in the bad outcome condition = 3.07, MEAN in the good outcome condition 4.75), but also a significant main effect for decision appropriateness (F: 17.16, $p<.0001$, $\omega^2=5.8\%$) and a significant two-way interaction between decision appropriateness and outcome (F: 4.46, $p<.02$, $\omega^2=2.5\%$). Because the predicted main effect for outcome was superseded by the interaction, the interaction merits further discussion. The interaction is depicted graphically in Figure 10.

The pattern of the means for this interaction on the dependent measure PEREVAL exhibited one key difference versus the interaction found for DECQUAL. Mean ratings of PEREVAL were equal across the decision appropriate conditions (4.75/4.76), as depicted by the horizontal line for the good outcome condition in Figure 10. As such, in the good outcome condition performance ratings of salesperson Smith were the same regardless of whether the decisions made were appropriate or inappropriate, indicating

TABLE X

ANOVA FOR PEREVAL
STUDY ONE

Independent Variable	DF	Type III SS	F-value	PR > F	ω^2
Decision	1	9.42	17.16	.0001	5.8%
Outcome	2	64.25	58.51	.0001	41.4%
Decision * Outcome	2	4.90	4.46	.0133	2.5%

r^2 for model = .53

TABLE XI

MEANS FOR PEREVAL BY CONDITION
STUDY ONE

	OUTCOME		
DECISION	BAD	GOOD	NONE
INAPPROPRIATE DECISION	2.74 SD=.76 n=26	4.76 SD=.75 n=18	3.33 SD=.81 n=25
APPROPRIATE DECISION	3.41 SD=.72 n=25	4.75 SD=.75 n=24	4.26 SD=.62 n=19

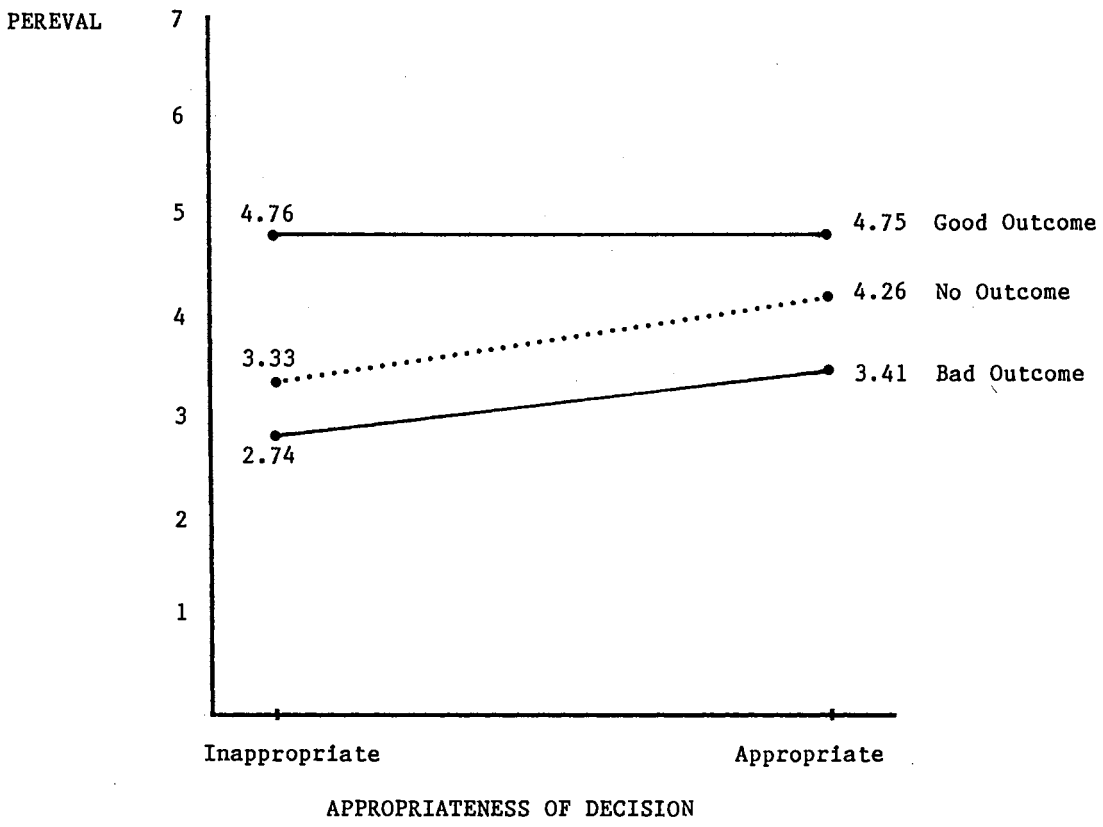


Figure 10. Two-way Interaction of Decision Appropriateness and Outcome on PEREVAL in Study One

decision appropriateness had no impact whatsoever on ratings when the outcomes of the decisions were good.

A priori F tests revealed that, similar to the findings for H1 on ratings of DECQUAL, PEREVAL ratings were significantly different both in the inappropriate and appropriate decision conditions, depending upon outcome (in the inappropriate decision condition, F: 78.92, $p<.0001$, MEAN for bad outcome = 2.74, MEAN for good outcome = 4.76; in the appropriate decision condition, F: 39.99, $p<.0001$, MEAN for bad outcome = 3.41, MEAN for good outcome = 4.75). However, as was also true for DECQUAL, the difference in the means for PEREVAL was greater when the decisions were inappropriate than when the decisions were appropriate, thus yielding the magnitude interaction portrayed in Figure 10.

Hypothesis 3. H3 predicted a three-way interaction among decision appropriateness, outcome, and rating period. H3 was tested by means of a repeated measures analysis of variance with polynomial contrast tests for linear trend. A polynomial contrast test was performed to provide evidence of the predicted linearity of the pattern of the means for PEREVAL across rating periods (please refer to Figure 5 for a graphical representation of this predicted three-way interaction).

In SAS, the "repeated" statement automatically places the analysis into a multivariate mode. After the multivariate statistics are provided, univariate ANOVAs are then produced. However, due to the potential for violation

of the homogeneity assumption in within-subjects analysis, use of conventional univariate repeated measures ANOVA is not generally recommended in cases such as the present analysis without attention paid first to the MANOVA results (LaTour and Miniard 1983).

At the multivariate level, the analysis for PEREVAL provided evidence of a main effect for rating period (Wilks' $\Lambda = 0.88$, F: 8.50, $p<.001$) and two-way interactions for decision appropriateness * rating period (Wilks' $\Lambda = 0.90$, F: 7.59, $p<.001$) and outcome * rating period (Wilks' $\Lambda = 0.77$, F: 9.08, $p<.0001$). These two-way interactions are pictured in Figures 11 and 12, respectively. The decision appropriateness * outcome * rating period interaction predicted in H3 approached significance (Wilks' $\Lambda = 0.94$, F: 2.11, $p<.08$). A summary of the multivariate test statistics for the three-way interaction is presented in Table XII.

One way to overcome the potential for bias due to understated Type I errors (due to the homogeneity problem) when reporting univariate statistics in repeated measures is by employing Greenhouse and Geisser's (1959) conservative test (LaTour and Miniard 1983). This approach modifies the degrees of freedom to be used in establishing the critical value of F for rejection of the null hypotheses. The univariate repeated measures ANOVA for PEREVAL is presented in Table XIII. The table presents Pr>F values both with and without the Greenhouse-Geisser adjustment. In all cases, the results mirror those at the multivariate level.

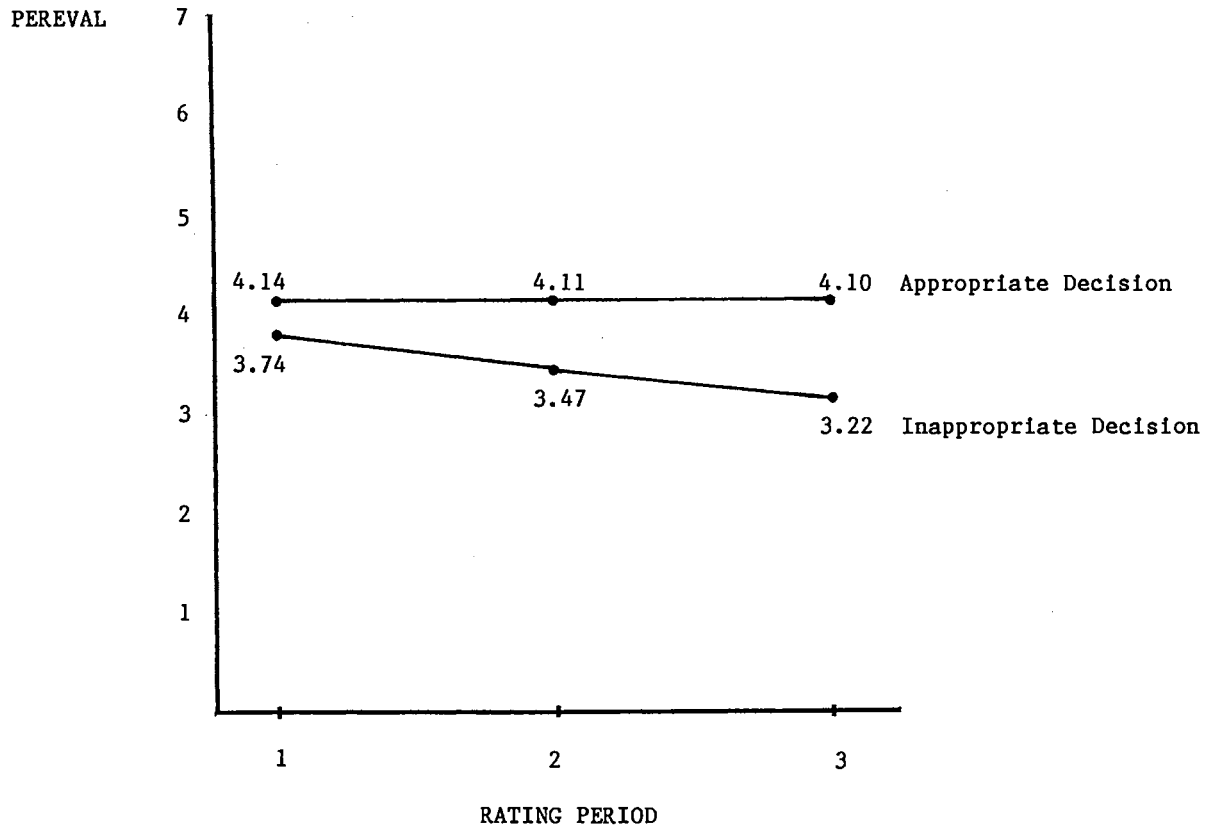


Figure 11. Two-way Interaction of Decision Appropriateness and Rating Period on PEREVAL in Study One

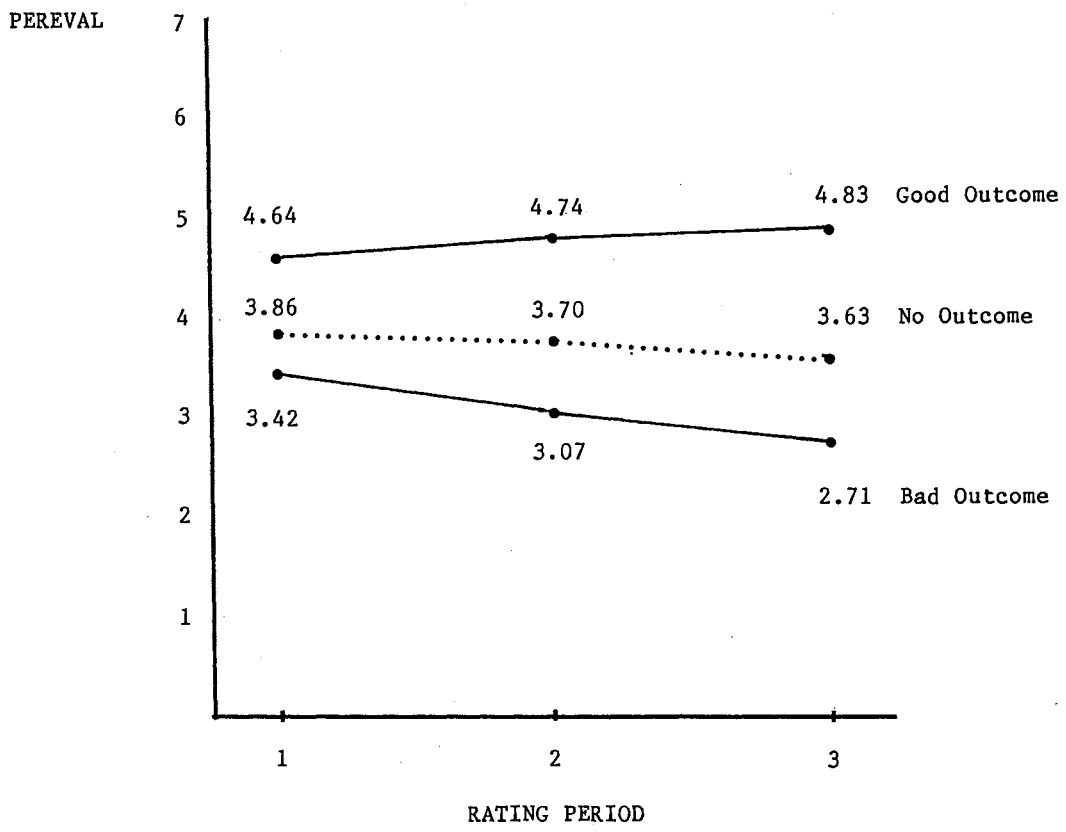


Figure 12. Two-way Interaction of Outcome and Rating Period on PEREVAL in Study One

TABLE XII

SUMMARY OF MULTIVARIATE TEST STATISTICS, REPEATED MEASURES MANOVA, FOR THREE-WAY INTERACTION OF DECISION APPROPRIATENESS, OUTCOME, AND TIME ON PEREVAL IN STUDY ONE

Statistic	Value	F-value	Num DF	Den DF	Pr>F
Wilks' Lambda	.94	2.11	4	260	.0801
Pillai's Trace	.06	2.12	4	262	.0785
Hotelling-Lawley Trace	.06	2.10	4	258	.0818
Roy's Greatest Root	.04	2.83	2	131	.0629

TABLE XIII

UNIVARIATE REPEATED MEASURES ANOVA
FOR WITHIN-SUBJECTS EFFECTS ON
PEREVAL - STUDY ONE

Independent Variable	DF	Type III SS	F-value	PR > F	Pr>F G-G
Time	2	3.84	11.38	.0001	.0001
Time * Decision	2	3.41	10.11	.0001	.0001
Time * Outcome	4	8.81	13.06	.0001	.0001
Time*Decision*Outcome	4	1.60	2.37	.0534	.0621

Greenhouse-Geisser Epsilon = .8753

NOTE: The relevant between subjects ANOVA for PEREVAL was presented in Table X.

The next question becomes, do the patterns of the means for PEREVAL follow those predicted by H3. Three-way interactions are inherently unwieldy, and are frequently best assessed visually. To this end, Figure 13, presents a plot of the resulting three-way interaction for decision appropriateness * outcome * rating period, showing the means for PEREVAL across all three rating periods for all conditions. Figure 13 should be compared to the hypothesized patterns of means depicted in Figure 5.

The predicted three-way interaction may be described as follows. When the decision was appropriate and the outcome was good, and when the decision was inappropriate and the outcome was bad, ratings of PEREVAL were hypothesized to be consistent across rating periods at relatively high or low levels, respectively. However, when the decision was appropriate with a bad outcome, ratings were predicted to decline across rating periods. Likewise, when the decision was inappropriate with a good outcome, ratings were predicted to increase across rating periods. As discussed in Chapter IV, the theoretical bases for these predictions was the Kelley Covariation Model (Kelley 1967) and discounting/augmentation principles (Kelley 1973).

Partial support was generated for the above predicted patterns of means. *A priori* F tests supported the hypothesized pattern of means in the appropriate decision condition. That is, ratings of PEREVAL did not change across rating periods when the outcomes were good,

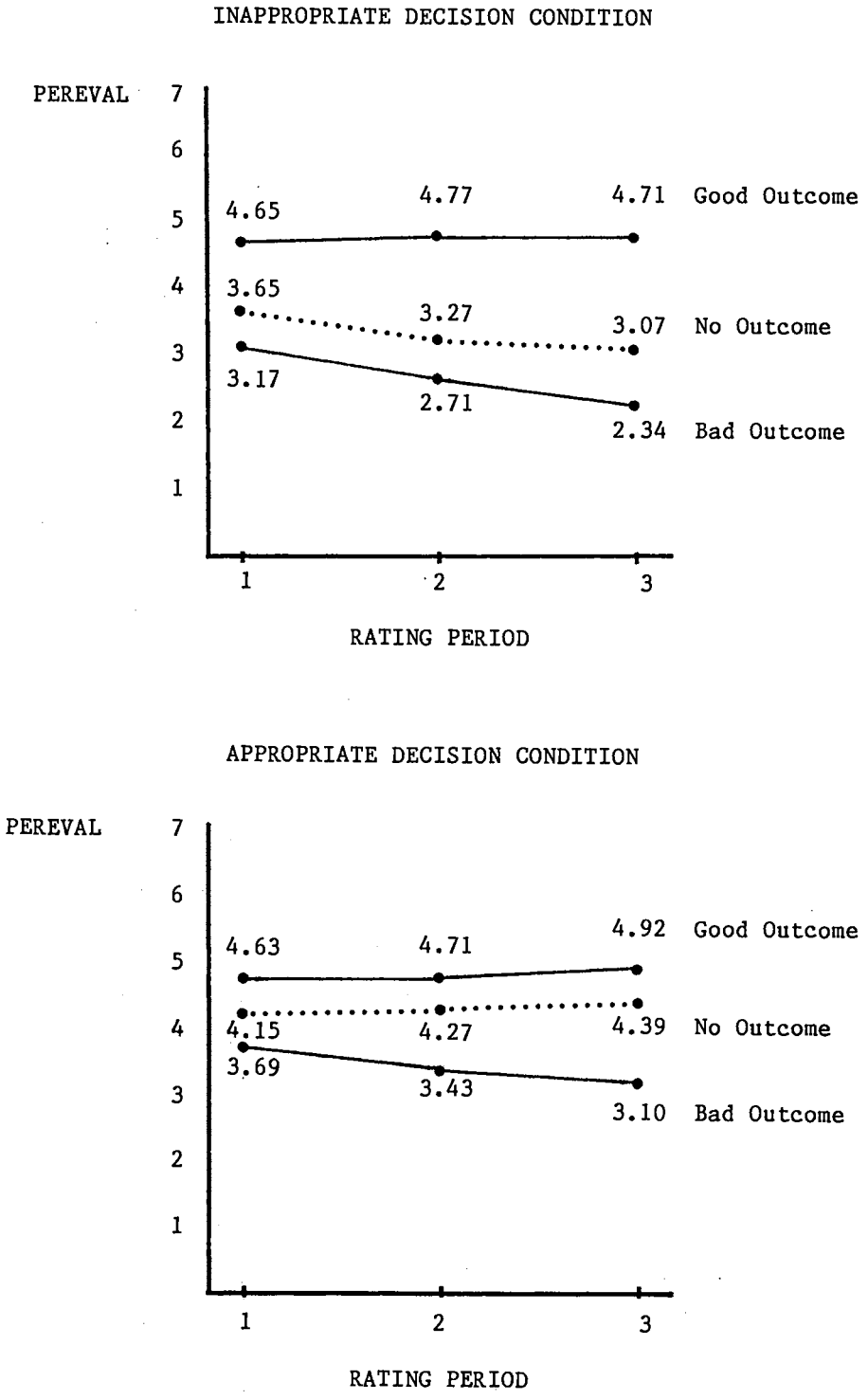


Figure 13. Three-way Interaction of Decision Appropriateness, Outcome, and Rating Period on PEREVAL in Study One

(F: 1.83, MEAN in rating period 1 = 4.63, MEAN in rating period 3 = 4.92), but PEREVAL ratings did decline over rating periods when the outcomes were bad, (F: 7.93, $p<.01$, MEAN in rating period 1 = 3.69, MEAN in rating period 3 = 3.10).

However, *a priori* F tests did not support the hypothesized pattern of means in the inappropriate decision condition. In fact, the results indicate a pattern opposite of that predicted, and similar to the pattern exhibited in the appropriate decision condition. Specifically, ratings of PEREVAL did not change across rating periods when the outcomes were good, (F: 0.06, MEAN in rating period 1 = 4.65, MEAN in rating period 3 = 4.71), but PEREVAL ratings again declined across rating periods when the outcomes were bad (F: 15.90, $p<.001$, MEAN in rating period 1 = 3.17, MEAN in rating period 3 = 2.35).

As expected, and consistent with the findings in H1 as well as the pilot study, mean ratings for PEREVAL for subjects in the no outcome condition fell between the ratings in the bad versus good outcome conditions (refer to Figure 13). This pattern was true in both decision conditions and across all three rating periods. In the appropriate decision/no outcome condition, no difference in mean ratings for PEREVAL was exhibited across rating periods (F: 1.02, MEAN in rating period 1 = 4.15, MEAN in rating period 3 = 4.39). This pattern of means in the appropriate decision/no outcome condition closely matched the pattern

previously reported for subjects in the appropriate decision/good outcome condition.

In contrast, in the inappropriate decision/no outcome condition, mean ratings of PEREVAL declined across rating periods (F: 7.75, $p<.01$, MEAN in rating period 1 = 3.65, MEAN in rating period 3 = 3.07). This pattern of means in the inappropriate decision/no outcome condition closely matched the pattern previously reported for subjects in the inappropriate decision/bad outcome condition. This similarity of results for the control and good outcome groups (in the appropriate decision condition) and the control and bad outcome groups (in the inappropriate decision condition) would appear to be partly responsible for the three-way interaction.

To summarize, a three-way interaction was demonstrated. However, contrary to H3, regardless of the decisions, good outcomes yielded consistent and relatively higher ratings across rating periods, and bad outcomes yielded declining and relatively lower ratings across rating periods.

In order to test for linearity of the PEREVAL ratings trend between rating periods one and three, a polynomial contrast test was employed. Results yielded evidence for linearity of the ratings trend. Contrasting PEREVAL ratings between rating period 1 and rating period 3 yielded a main effect for outcome between the two rating periods, signifying linearity of PEREVAL ratings in the outcome conditions between those periods (F: 19.08, $p<.0001$). The

direction of the linear trends in the outcome conditions is portrayed graphically in Figure 13.

Hypothesis 4. As discussed in Chapter III, this hypothesis was tested to provide further evidence of an information processing explanation for the magnitude interaction between decision appropriateness and outcome on DECQUAL. An analysis of the cognitive responses failed to support H4. An ANOVA for INFOPROC produced no significant effects (F: 0.02, MEAN in the inappropriate decision condition = 2.07, MEAN in the appropriate decision condition = 2.08).

Hypothesis 5. H5 predicted more internal attributions when outcome information was available versus the no outcome condition. To test this hypothesis, ANOVAS were conducted on INATTRIB in two ways. First, the two outcome conditions (bad and good) were collapsed into one, and an ANOVA was run for INATTRIB. No effect was found for outcome on INATTRIB (F: 0.05, MEAN when outcome information provided = 4.10, MEAN when outcome information not provided = 4.17.) In order to further examine this result, a second ANOVA was performed with all three outcome groups (bad, good, and none) represented separately. Again, no effect was found for outcome on INATTRIB (F: 2.10, MEAN with bad outcome = 3.74, MEAN with good outcome = 4.51, MEAN with no outcome = 4.17.) Similarly, no effects were found for outcome on the measure of external attributions for salesperson Smith's

decisions (EXATTRIB). Thus, H5 was not supported.

Tests With Exploratory Variables

As previously mentioned, several ANCOVA analyses were conducted utilizing three exploratory variables as covariates (LIKEME, ECOLVAL, and RISKTKNG). It should be noted that prior to utilizing these covariates, each was individually employed as a blocking variable in ANOVAS with decision appropriateness and outcome to check for interactions with the independent variables. Because no interactions were found, it was deemed appropriate to proceed with the ANCOVAS. Also, the GLOBAL rating of salesperson Smith's performance was utilized as a dependent measure in several tests.

Universally, the only finding in these exploratory tests that added explanatory power to the overall analyses of the hypotheses in Study One was the strength of the covariate LIKEME in accounting for a large amount of variance across the analyses. Table XIV presents the ANCOVA results for DECQUAL and Table XV presents the ANCOVA results for PEREVAL. These results should be compared with the comparable ANOVA results previously discussed. In both cases, the LIKEME variable had a large impact (F: 23.82, $p<.0001$, $\omega^2=4.0\%$ for DECQUAL; F: 36.71, $p<.0001$, $\omega^2=9.7\%$ for PEREVAL). The need for future research into the impact of this variable on salesforce performance evaluation is clear.

In no case did the addition of the covariates to the

TABLE XIV

ANCOVA FOR DECQUAL
STUDY ONE

Independent Variable	DF	Type III SS	F-value	PR > F	ω^2
Decision	1	126.93	186.89	.0001	32.2%
Outcome	2	8.58	6.32	.0024	1.8%
Decision * Outcome	2	7.11	5.24	.0065	1.5%
LIKEME	1	16.17	23.82	.0001	4.0%
ECOLVAL	1	.37	.55	.4600	
RISKTKNG	1	.85	1.26	.2642	

r^2 for model = .77

TABLE XV

ANCOVA FOR PEREVAL
STUDY ONE

Independent Variable	DF	Type III SS	F-value	PR > F	ω^2
Decision	1	.86	2.14	.1464	
Outcome	2	27.21	33.61	.0001	17.7%
Decision * Outcome	2	4.59	5.66	.0044	2.5%
LIKEME	1	14.86	36.71	.0001	9.7%
ECOLVAL	1	1.01	2.50	.1165	
RISKTKNG	1	.03	.07	.7948	

r^2 for model = .66

models change the basic results. ECOLVAL and RISKTKNG provided no additional explanatory power. Likewise, analyses utilizing the exploratory dependent variable GLOBAL failed to provide insight. It is conceivable that requesting such a broad, overall single-number rating of salesperson Smith such as GLOBAL simple was not realistic in the context of a snapshot review of a few of Smith's decisions by the rater.

Study Two: Order Effects Bias

Study Two sought to provide evidence of an order effects bias in evaluations of field sales personnel by their managers. Because the Hogarth and Einhorn's (1992) belief-adjustment model of primacy/recency effects is so new and its propositions have just begun to be empirically tested, Study Two was viewed as more exploratory in nature than Study One.

Structure and Reliability of Indices

The key dependent measure in Study Two was PEREVAL. Subjects read an initial description (called, the stem) that remained constant across all conditions. Next, they rated salesperson Smith on the PEREVAL scale. After receiving updating information on Smith's performance, subjects completed the PEREVAL scale either once or twice more, depending upon whether they were assigned to the EoS or SbS evaluation condition.

Structure. Initial principal component analysis of the DECQUAL scale at the three rating periods yielded a two-factor solution for periods one and three and a single factor solution for period two. This was unacceptable, since the multiple rating requirement necessitated similar factor structure across rating periods. The two problem items in periods one and three were the two 7-point Likert questions related to salesperson Smith as an individual: 1) "How would you categorize Smith?" (not likeable to very likeable); and 2) "How would you describe Smith as a person?" (bad to good). It may have been more difficult for respondents to make such personal ratings of Smith in this study for two reasons: 1) the stem information was not a full scenario (as was the case in Study One), and 2) the performance information about Smith (i.e., decision appropriateness and outcomes) was mixed over the two rating periods (bad/good or good/bad), instead of similar across periods (as in Study One). When the two problem items were removed from the set of variables for principal component analysis, a unitary factor solution resulted in all three rating periods. Therefore, for purposes of Study Two it was necessary to reduce the PEREVAL index to a 5-item measure, which will be labelled R-PEREVAL for "reduced PEREVAL." A summary of the factor loadings is presented in Table XVI.

Reliability. Cronbach alphas and item-total correlations were calculated for the R-PEREVAL scale for each of the three rating periods. Alphas for the scales

TABLE XVI

FACTOR LOADINGS ACROSS
THREE RATING PERIODS
STUDY TWO

		RATING PERIOD 1	RATING PERIOD 2	RATING PERIOD 3
		FACTOR 1 (R-PEREVAL)	FACTOR 1 (R-PEREVAL)	FACTOR 1 (R-PEREVAL)
(Q1)	Q1	.79	.88	.86
(Q2)	Q3	.78	.86	.83
(Q3)	Q5	.86	.85	.88
(Q5)	Q8	.62	.51	.64
(Q7)	Q12	.77	.92	.85

NOTE: Question numbers in parentheses are from the initial rating period, and were numbered differently from the other two rating periods. Here, the initial questions are matched to the corresponding questions from the other rating periods. Please refer to the questionnaire in Appendix B for specific items.

were slightly lower than those reported in Study One, ranging from .82 to .87. However, due to the more exploratory nature of Study Two, the scale was deemed acceptable. Results of the reliability analysis for the R-PEREVAL scale are presented in Table XVII.

The DECQUAL measure had no bearing on the tests of hypotheses in Study Two, but was taken (along with the single-item outcome manipulation check employed in Study One) in the second and third rating periods as a manipulation check on subjects' perceptions of the performance information on salesperson Smith presented in the two updating scenarios (i.e., bad or good). Cronbach alphas and item-total correlations were calculated for DECQUAL for rating periods two and three. Alphas were .96 and .88 for periods two and three, respectively, and item-total correlations were acceptable (results of the reliability analysis for the DECQUAL scale are presented in Table XVIII). Thus, it was deemed appropriate to utilize the DECQUAL scale as intended for a manipulation check on the valence of the performance information.

As in Study One, reliability analysis on the five-item scale for RISKTAKING designed to be used as an exploratory covariate measure yielded poor results. Item-total correlations ranged from 0.09 to 0.38, with a Cronbach alpha of .45. As before, in all subsequent analyses RISKTKNG was converted to the same single-item measure as described in Study One.

TABLE XVII

RELIABILITY ANALYSIS OF R-PEREVAL SCALE
USING STANDARDIZED VARIABLES
STUDY TWO

		ITEM-TOTAL CORRELATIONS		
		RATING PERIOD 1	RATING PERIOD 2	RATING PERIOD 3
(Q1)	Q1	.64	.78	.75
(Q2)	Q3	.64	.77	.72
(Q3)	Q5	.74	.72	.79
(Q5)	Q8	.46	.39	.50
(Q7)	Q12	.62	.84	.74
Cronbach α		.82	.87	.87

NOTE: Question numbers in parentheses are from the initial rating period, and were numbered differently from the other two rating periods. Here, the initial questions are matched to the corresponding questions from the other rating periods. Please refer to the questionnaire in Appendix B for specific items.

TABLE XVIII

RELIABILITY ANALYSIS OF DECQUAL SCALE
USING STANDARDIZED VARIABLES
STUDY TWO

	ITEM-TOTAL CORRELATIONS	
	RATING PERIOD 2	RATING PERIOD 3
Q2	.92	.81
Q4	.89	.82
Q7	.78	.41
Q9	.92	.75
Q11	.96	.82
Cronbach α	.96	.88

NOTE: Please refer to the questionnaire in Appendix B for specific items.

Reliability of the two-item ECOLVAL scale was highly similar to that reported in Study One, with an item-total correlation for each item of .75 and a Cronbach alpha of .86. Thus, as before this measure was deemed satisfactory for exploratory research purposes.

No other multiple item measures were employed in Study Two.

Description of Sample

As reported in Chapter IV, the number of usable responses in Study Two was 90, a 75 percent response rate. Characteristics of the subjects were highly similar to those of the subjects in Study One. Median age of respondents was 37, with an age range of 27 to 53 years old. The male/female split was 78 percent/22 percent respectively.

Again, level of experience was very high. Median years of sales management experience was 7 years, with a range of 1 to 28 years. In terms of sales experience, subjects exhibited a median of 13 years with a range of 13 years with a range of 1 to 35 years. Median years of total work experience was 18 years, with a range of 4 to 35 years. As noted in Study One, some respondents indicated confusion about the questions regarding number of performance appraisals completed and received. Therefore, the results of those questions will not be discussed with regard to Study Two.

Manipulation Check

For Study Two, because subjects received performance information about salesperson Smith in varied order of valence (bad/good or good/bad), a manipulation check was required to ensure that subjects perceived the valence of the performance information on salesperson Smith as expected. To assess this perception of the relative badness/goodness of the performance information, separate ANOVAS were performed for rating periods two and three (note that the information in the stem was not varied). A significant difference in the mean ratings of DECQUAL at both rating points, coupled with a significant different in the means of the outcome manipulation check item at both rating points, would indicate a successful manipulation of performance information valence. Specifically, decision quality should be rated lower and outcome of the decision worse when information about Smith's performance was bad versus when it was good.

ANOVA yielded the predicted results at both rating points on both dependent measures. In all cases, the means were significantly different in the predicted directions. These results are summarized in Table XIX. Based upon the clear differences in perception of the bad versus good information about salesperson Smith's performance, the manipulation of valence of performance information was deemed successful.

The check for ecological validity with the two-item

TABLE XIX

SUMMARY OF MANIPULATION CHECK FOR VALENCE
OF PERFORMANCE INFORMATION
STUDY TWO

	Rating Period Two	
Performance Information	Mean Perceived Decision Quality	Mean Perceived Outcome
Bad Performance Information	2.67	1.86
Good Performance Information	5.92	5.91
	DF=1 Type III SS=115.86 F: 116.93 p<.0001 n=44	DF=1 Type III SS=180.02 F: 207.67 p<.0001 n=44
	Rating Period Three	
Performance Information	Mean Perceived Decision Quality	Mean Perceived Outcome
Bad Performance Information	3.07	2.56
Good Performance Information	4.81	4.87
	DF=1 Type III SS=65.41 F: 60.49 p<.0001 n=86	DF=1 Type III SS=120.26 F: 60.88 p<.0001 n=90

ECOLVAL scale yielded a mean score of 5.30 (n=89, S.D.=1.57) on the 7-point Likert scale, where 7 indicated subjects' perceived the scenarios as highly realistic. This ECOLVAL score for subjects in Study Two was slightly higher than that for subjects in Study One (MEAN = 4.82, n=141, S.D.=1.80), perhaps because in Study Two the valence of the performance information for salesperson Smith was varied (either bad or good) from scenario to scenario, while in Study One the valence of the performance information was the same across the scenarios.

Tests of Hypotheses

Because the two hypotheses in Study Two are closely related, they will be considered together in this section.

Hypotheses 6 and 7. H6 and H7 predicted a main effect for order of performance information (bad/good versus good/bad), superseded by an order of information * response mode (SbS/EoS) interaction. More specifically, a recency effect was predicted overall, but with an interaction in which recency impacted R-PEREVAL ratings in the SbS response mode but not in the EoS response mode.

To test H6 and H7, an ANOVA was performed utilizing ratings of the reduced index of attribution-based general performance measures (R-PEREVAL). Following the procedure utilized by Tubbs, Messier, and Knechel (1990) and Hogarth and Einhorn (1992), initial ratings (i.e., ratings after the stem information) were subtracted from final ratings (i.e.,

ratings after exposure to all updating information), and the differences in these ratings were employed as a dependent variable, which will be labeled DIFF[1].

The ANOVA for DIFF is presented in Table XX. The analysis failed to reveal the hypothesized main effect for order of information (F: 2.12, MEAN DIFF in the bad/good order condition = 0.17, MEAN DIFF in the good/bad order condition = -0.05). However, an order of information * response mode interaction approached significance (F: 2.78, $p<.10$, $\omega^2=2.0$). Mean DIFF by cell is presented in Table XXI. The interaction is portrayed graphically in Figure 14.

The graphic portrayal of the interaction shows no change in DIFF based upon order of performance information in the EoS response mode, but a dramatic change in DIFF in the SbS response mode. This rating differential in the SbS response mode depended upon whether the good performance information was received last in the sequence of updating information (MEAN DIFF = +0.45, indicating that the final rating was more favorable than the initial rating) or first in the sequence of updating information (MEAN DIFF = -0.07, indicating that the final rating was less favorable than the

---

[1] In a recent article, Peter, Churchill, and Brown (1993) argued that the use of difference scores may often be problematic. However, their arguments referred to the subtraction of one measure from another to create a measure of a distinct construct---for example, subtracting expectations from perceptions to create a measure of perceived service quality. In contrast, in the present study DIFF represented a change in a single measure (the rating of salesperson Smith's performance) based upon subjects' receipt of updated performance information.

TABLE XX

ANOVA FOR DIFF
STUDY TWO

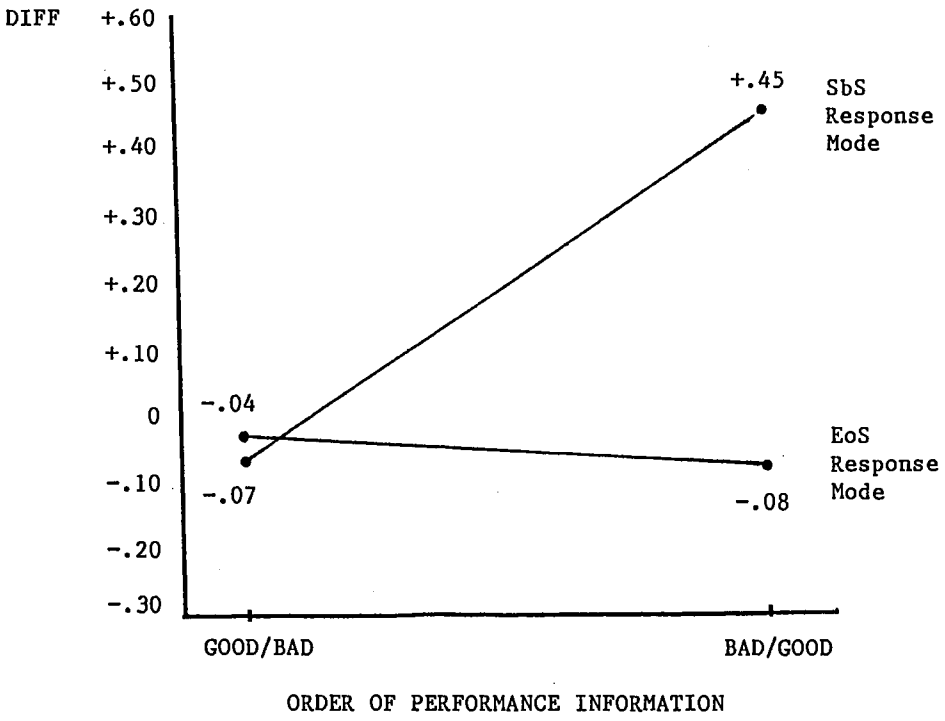
Independent Variable	DF	Type III SS	F-value	PR > F	ω^2
Response Mode	1	1.37	2.27	.1356	
Order	1	1.28	2.12	.1488	
Response Mode*Order	1	1.67	2.78	.0995	2.0%

r^2 for model = .08

TABLE XXI

MEANS FOR DIFF BY CONDITION
STUDY TWO

	ORDER	
RESPONSE MODE	BAD/GOOD	GOOD/BAD
SbS	.45 SD=.82 n=21	-.07 SD=.80 n=21
EoS	-.08 SD=.78 n=24	-.04 SD=.69 n=20



NOTE: DIFF represents the difference when the initial rating of R-PEREVAL is subtracted from the final R-PEREVAL rating.

Figure 14. Two-way Interaction of Order of Information and Response Mode on DIFF in Study Two

initial rating).

Based upon the predictions in H7, an *a priori* F test was performed to check for a significant difference in the means in the SbS response mode. The test revealed that the means described above were, in fact, significantly different (F: 18.84, $p<.001$). An *a priori* F test on the means in the EoS condition yielded no effect for order of performance information on DIFF (F: 0.03, MEAN DIFF in the bad/good order condition = -0.08, MEAN DIFF in the good/bad order condition = -0.04). Thus, the overall ANOVA had not revealed the effect.

The results provided evidence of a recency effect in the SbS response mode but no order effect (i.e., an attenuation of the recency effect) in the EoS response mode. This finding confirms H7, and also mirrors the findings of Hogarth and Einhorn (1992) and Tubbs, Messier, and Knechel (1990). A graphic portrayal of the overall pattern of the mean ratings for R-PEREVAL in both response modes across the three rating periods produces the "fishtail" pattern predicted by Hogarth and Einhorn (1992). This pictorial representation of the R-PEREVAL means is presented in Figure 15.

Tests With Exploratory Variables

As in Study One, an ANCOVA was conducted (here, on R-PEREVAL only) utilizing the three covariates (LIKEME, ECOLVAL, and RISKTKNG). Again, each variable was first

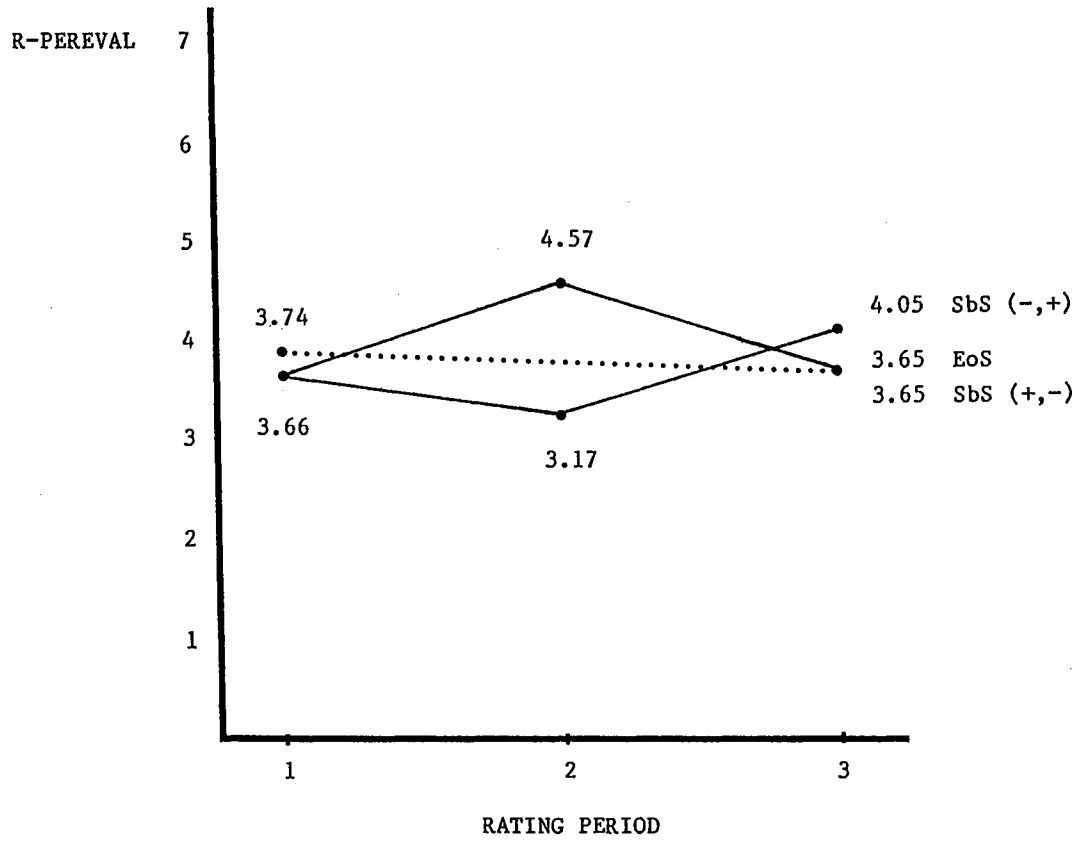


Figure 15. "Fishtail" Pattern of Means for R-Pereval Showing Recency Effects in SbS Response Mode in Study Two

individually employed as a blocking variable in ANOVAS with response mode and order of information to check for interactions with the independent variables. No such interactions were observed. Also, several exploratory analyses were conducted using the GLOBAL rating of salesperson Smith's performance as a dependent measure.

Similar to Study One, the only finding in these exploratory tests that added explanatory power to the overall analysis of the hypotheses in Study Two was the strength of the covariate LIKEME in accounting for a large amount of variance in the analysis. Table XXII presents the ANCOVA for DIFF. It should especially be noted that in the ANCOVA the hypothesized order of information * response mode interaction was significant ($F: 4.36$, $p<.05$, $\omega^2=3.6\%$). The pattern of the means and graphic portrayal of the interaction were highly similar to those reported above in the discussion of the ANOVA for DIFF. The same concerns expressed in Study One about the GLOBAL measure hold true for Study Two as well.

TABLE XXII

ANCOVA FOR DIFF
STUDY TWO

Independent Variable	DF	Type III SS	F-value	PR > F	ω^2
Response Mode	1	.59	.99	.3238	
Order	1	1.60	2.78	.0994	1.9%
Response Mode*Order	1	2.51	4.36	.0400	3.6%
LIKEME	1	4.07	7.07	.0095	6.5%
ECOLVAL	1	.74	1.20	.2606	
RISKTKNG	1	.07	.12	.7270	

r^2 for model = .17

CHAPTER VI

SUMMARY AND CONCLUSIONS

This chapter is divided into four sections. The first section is a discussion section containing an analysis of the specific results, including potential avenues for future research; section two presents a general discussion of implications; the third section addresses limitations of the research; and the fourth section offers a set of specific steps that sales organizations are encouraged to take to maximize the effectiveness of the salesforce performance evaluation process.

Discussion

Overview

It is important to reiterate the key purposes of the dissertation. First, a comprehensive literature review on salesforce performance was developed utilizing the most widely accepted paradigm in the domain, the Walker, Churchill, and Ford (1977) model, as a means of organizing the various studies. Second, the dissertation highlighted an area of marketing management decision making that is underdeveloped in the literature, judgmental heuristics and biases, by developing and empirically testing hypotheses

concerning an outcome bias and an order effects bias on performance ratings of field sales personnel by sales managers. Finally, the dissertation answered the call for the use of alternative (i.e., non-survey) research methods to investigate salesforce performance. The studies utilized two experiments conducted in a field setting with practicing sales managers as subjects.

The results revealed general support across the hypotheses. Table XXIII summarizes the results for each research hypothesis. The findings for Study One (Outcome Bias) and Study Two (Order Effects Bias) are discussed separately in the next two sections.

Analysis of Results: Outcome Bias

Strong evidence was generated overall for the robust nature of an outcome bias in which decision appropriateness and outcome interacted to impact performance ratings. The interaction took the form of a magnitude interaction in which ratings were relatively less impacted by outcome when the salesperson's decisions were appropriate versus when the decisions were inappropriate. Several studies across domains, utilizing both students and working professionals as subjects, have yielded similar results (c.f., Lipschitz 1989; Mowen and Stone 1982; and Mowen and Marshall 1983), although the present study did produce several results not seen before. These new findings will be highlighted next.

First, unlike in previous studies on the outcome

TABLE XXIII

SUMMARY OF RESEARCH RESULTS
BY HYPOTHESIS

H1, 1a, and 1b: Two-Way Interaction Between Decision Appropriateness and Outcome on DECQUAL

Result: Hypothesis supported. In the appropriate decision condition, however, ratings of DECQUAL were found to be significantly different between the bad and good outcome groups.

H2: Main Effect for Decision Appropriateness and Outcome on PEREVAL

Result: Main effect superseded by a two-way interaction between decision appropriateness and outcome.

H3: Three-Way Interaction Among Decision Appropriateness, Outcome, and Time on PEREVAL

Result: Three-way interaction found, but in the inappropriate decision/good outcome condition PEREVAL ratings over time did not increase as predicted. However, in the appropriate decision/bad outcome condition PEREVAL ratings did decline over time as predicted.

H4: Greater Level of Information Processing in the Inappropriate versus Appropriate Decision Conditions

Result: Hypothesis not supported. No difference found in the number of cognitive responses between the two decision conditions.

H5: Increased Internal Attributions in Outcome versus No Outcome Conditions

Result: Hypothesis not supported. Availability of outcome information did not increase internal attributions.

H6: A Recency Effect Overall, and

H7: An Interaction in which Recency is More Pronounced in the SbS Response Mode

Result: Hypotheses supported. A tendency toward recency was found overall, but in the EoS response mode this tendency was attenuated.

bias, the present results revealed a significant difference in ratings of DECQUAL in the appropriate decision condition when the outcome was bad versus good. The fact that DECQUAL ratings were significantly lower when the outcome was bad, even when salesperson Smith made the normatively appropriate decision, is disturbing. This result says that outcome is overwhelming the ratings of the correctness/competence of the employee's decisions. The finding represents a potentially dangerous threat to fair and equitable salesforce performance evaluations, and needs to be investigated further (it should be noted that a similar result in the appropriate decision condition was found for the attribution-based general performance evaluation measure, PEREVAL.)

In the present study, the interaction for PEREVAL was similar to that found for DECQUAL, except for the good outcome condition. Here, the pattern of the means for DECQUAL in the two-way interaction of decision appropriateness * outcome did not match the prediction. In the good outcome condition, ratings for DECQUAL were significantly lower when the decision was inappropriate than when the decision was appropriate. A similar finding by Marshall and Mowen (1993) prompted the authors to offer the explanation of a stronger manipulation of decision appropriateness versus prior studies. In Marshall and Mowen (1993), subjects were told specifically that the salesperson's choice was either consistent or inconsistent

with the recommendation of the sales planning department. In the earlier research, subjects had to infer the quality of the decision from the information given. Thus, the salience of the appropriateness information may have influenced the extent to which such information was used in evaluating the quality of the decision. That is, it is possible that as the appropriateness of the decision increased in clarity, the tendency toward an outcome bias decreased. An interesting and managerially relevant follow-up study would vary saliency of decision appropriateness information and test for an outcome bias. Evidence of an effect for salience of decision appropriateness information would signal to managers the need to clearly define the parameters of good and bad decisions across the gamut of job dimensions in personal selling. Examples of such decisions might be those related to prospecting, customer service strategy, selling approach, resource utilization, and a variety of ethical issues.

Another finding that differed from the prediction was a two-way interaction of decision appropriateness and outcome on PEREVAL. In the pilot study, only outcome impacted PEREVAL---decision appropriateness information was ignored. Based upon that finding, the prediction in the present study was a main effect for outcome only on PEREVAL. In the interaction, the mean for PEREVAL in the good outcome condition was the same regardless of the decision. Thus, when managers evaluated salesperson Smith on the

attribution-based general performance measure instead of on the decision quality measure, whether the decision was appropriate or not had no impact whatsoever on ratings. However, when the outcome of the decision was bad, PEREVAL ratings were significantly lower in the inappropriate versus appropriate decision condition. Future research needs to focus on this finding that inappropriate decisions can yield equal performance ratings to appropriate decisions so long as the results of those decisions are good. Such a process carries with it the very real potential of developing an "ends justify the means" mentality in the way salespeople approach the ethics of their decision making.

A key contribution of the present study was the demonstration of an outcome bias across multiple rating periods. A partial explanation for the pattern of the means across rating periods may be derived from Kelley's (1973) discounting/augmentation principles. In the inappropriate decision condition, raters may have viewed salesperson Smith as moving against the environment (i.e., making a decision contrary to the company line). Thus, when the outcome of that decision was bad, more internal attributions to Smith were made and ratings on the PEREVAL index suffered (recall that the PEREVAL scale measured primarily effort, ability, and performance). However, when the outcome of the inappropriate decision was good, subjects may have made more external attributions (i.e., territory factors or luck). As a result, ratings on the PEREVAL index did not change across

rating periods.

A similar explanation may be derived for the pattern of performance ratings across time in the appropriate decision condition. Here, when the outcome was good, raters likely discounted internal causes of performance because salesperson Smith was simply making the decision the company would advocate (an environmental cause for the action). However, when the decision was right but outcomes were bad across each of the rating periods, it is highly likely that raters were searching for internal causes for the poor results. Specifically, in the appropriate decision/bad outcome condition it would be difficult to justify an environmental attribution, since the environmentally favored decision was made each time but the results were continually poor. Thus, one can see how the rater might conclude that the problem must be the salesperson.

The fact that the pattern of PEREVAL ratings across rating periods by the no outcome/control group closely matched the pattern of ratings by the good outcome group in the appropriate decision condition and closely matched the pattern ratings by the bad outcome group in the inappropriate decision condition appears to have contributed to the three-way interaction. Because it is not an infrequent occurrence in professional selling that outcomes of decisions by salespeople may not be readily available for consideration by management, it is important that follow-up research be done to investigate the impact of lagging

outcome information on performance ratings.

The fact that the results did not support the predicted differences in level of information processing by subjects in the inappropriate versus appropriate decision conditions deserves mention. Both Mowen and Stone (1992) and Marshall and Mowen (1993) had sought to provide an information processing explanation for the magnification effect of available outcome information (for a discussion of the theoretical basis of this explanation, see Harkins and Petty 1981). The basic notion was that rater expectations of the ratee's behavior would be violated in the inappropriate decision condition, thus causing the rater to engage in more cognitive processing.

A possible explanation for the present finding of no difference in level of information processing was gleaned from a detailed review of the cognitive responses by decision condition. As was the case in the previous studies, subjects made numerous comments about possible reasons for salesperson Smith's targeting the inappropriate clients. However, in the present study, a number of respondents assigned to the appropriate decision condition queried why Smith was continually making the less risky choice across decision periods---an issue subjects in the prior studies could not raise since they were exposed to only one decision in the scenario. Common themes of the responses were as follows:

1) In business, one must take risks. Smith isn't taking

any risks.

2) Smith isn't seeking creative solutions. Smith is going for the obvious choices.

3) Smith needs to develop those customers who have been less likely to respond in the past.

Such comments are likely attributable in part to the relevance and salience of these kinds of decisions to the present subject pool versus the student sample used in the pilot study, and in part simply to the fact that (unlike prior studies on the outcome bias) subjects in the present study were able to witness Smith make multiple decisions, and thus were able to infer an overall pattern of decision making.

This finding merits additional research. The issue of just what is a normatively appropriate decision becomes fuzzy in personal selling when the sales organization's value system advocates risk-taking, innovativeness, and creativity but at the same time expects salespeople to follow prespecified rules for sales presentations, prospecting and new business development, call coverage, and other job dimensions.

Finally, why were no differences found in level of internal attributions based upon whether or not raters had outcome information? As discussed in Chapter III, this prediction was based on a finding by Mitchell and Kalb (1981) that evaluators who had information about both an individual's behavior and the outcome of the behavior were

more likely to make an internal attribution about the cause of the event than if no outcome information was available. If this were so, and a corresponding reduction in external attributions were demonstrated, evidence could be claimed for the "fundamental attribution error" (Ross 1977). When the fundamental attribution error occurs, contextual or background information is systematically ignored by the rater, and instead evaluations are based upon "person" factors such as perceived ability and effort (Marshall, Mowen, and Fabes 1992). However, in the present study no differential was found in internal versus external attributions for salesperson Smith's behavior based upon knowledge of outcome. As mentioned earlier, the general issue of availability/non-availability of outcome information on ratings of salespeople is quite relevant and deserves additional research efforts.

Analysis of Results: Order Effects Bias

The order effects study was more exploratory in nature than the outcome bias study. In fact, it is only the third known to exist utilizing predictions from Hogarth and Einhorn's (1992) model to test for primacy/recency effects. Here, the results supported the hypotheses.

As predicted, when subjects rated salesperson Smith once after the initial information (the stem) and once more after receiving two updates on Smith's performance (i.e., the EoS response mode), order of update information

(bad/good or good/bad) had no impact on ratings. However, when subjects performed an additional interim rating after receiving the middle informational update, order of information impacted ratings such that the most recent evidence presented significantly affected the evaluation. Thus, while the SbS mode yielded a recency effect, the EoS response mode tended to reduce the recency effect.

A key question is, do recency effects really represent a judgmental bias, or should managers weight evaluations more heavily based upon the latest level of performance of a salesperson? Hogarth and Einhorn (1992) describe recency as a bias in the context that, normatively, order in which information is received should not result in differential ratings. Most sales organizations operate with defined rating periods, typically sales quarters, half-years, or years. During these periods, evidence accumulates to indicate the level of performance of the sales personnel across a variety of performance dimensions. When the rating is made at the end of the period, the salesperson should expect a fair appraisal in which ratings are equal regardless of the order in which favorable or unfavorable performance evidence accrues within the rating period. One exception might be new salespeople. For them, overweighting recent information is probably appropriate.

Analysis of Covariate: LIKEME

While no hypothesis was developed for the variable

LIKEME, the construct bears mentioning again with regard to its performance as a covariate in several exploratory tests.

In analysis of covariance, variation in a response variable *y* that is associated with some covariate *x* is removed from the error variance, resulting in more precise estimates and more powerful tests. Group means of the *y* variable are adjusted to correspond to a common value of *x*, thereby producing an "equitable" comparison of the groups. By explicitly bringing into the model a concomitant variable---a covariate of the dependent variable of interest---error variance is reduced, thus increasing the precision of the model parameter estimates (Freund and Littell 1981).

The covariate LIKEME appears to be a textbook application of the above principles. In the present research, in applications across the numerous ratings, LIKEME accounted for a large amount of variance in the dependent measures. However, a series of applications of LIKEME as a blocking variable yielded no evidence of interactions with the independent variables in the studies. Because the impact of LIKEME was so pronounced on the models, further research needs to be undertaken specifically designed to extend the previous work on the effect, similar-to-me, which formed the conceptual basis for including this exploratory variable in the study.

In a number of studies, the effect, similar-to-me, has been shown to impact candidate ratings in selection

interviews (c.f., Anderson and Shackleton 1990; Baskett 1973; Dalessio and Imada 1984; Peters and Terborg 1975; Rand and Wexley 1975; Wexley and Nemeroff 1974). A consistent finding across the studies is that candidates with similar biographical backgrounds, attitudes, and perceived personalities to the interviewer are rated more favorably than those who differ in these respects. Thus, the tendency for interviewers to recruit in their own self-image may exert substantial influence upon the decision making process in selection. If, merely by demonstrating similarity to us in some way, persons with whom we interact in an employee/employer relationship can elicit positive feelings from the employer toward the employee strong enough to impact selection decisions, it is reasonable to assume that the potential exists for these positive feelings to bias decisions in performance evaluation as well. Additional research is needed to extend the work on the similar-to-me effect from the domain of employment selection to the domain of performance evaluation.

Implications

The results of the present studies suggest that an outcome bias and order effects bias may be pervasive in the evaluation of sales personnel. These biases, as well as a host of other judgmental heuristics and biases (e.g., representativeness, the availability heuristic, anchoring and adjustment, hindsight bias, framing error, ignoring

regression effects, underweighting base rate information, the fundamental attribution error, context effects, and others) have only recently begun to be addressed within the domain of marketing management decision making (c.f., Gentry, Mowen, and Tasaki 1991; Mowen and Gaeth 1992; Marshall, Mowen, and Fabes 1992; and Mowen and Marshall 1992). Nevertheless, much work from the field of behavioral decision theory suggests that many of the kinds of decisions made by marketing managers on a daily basis may be suboptimized by these judgmental biases (Hogarth 1987). Therefore, a key contribution of the present research is the new empirical evidence generated that, at least in one important domain of marketing decision making---salesforce performance evaluation---Hogarth and his colleagues in behavioral decision theory appear to be correct.

A key question becomes, what happens when a decision is suboptimized by interjection of one or more of the judgmental biases into the decision process? Within the context of salesforce performance evaluation, the result of such biases may be an ineffective performance appraisal system throughout the organization. For example, salespeople in the firm may find themselves "empowered" to utilize their own decision making skills and creativity at attaining sales during a rating period, only to be evaluated at the end of the period based overwhelmingly upon results (an outcome bias). Such an evaluation procedure would likely be viewed as duplicitous by a salesperson, who will

likely be left wondering why the rating system doesn't reflect the empowerment philosophy.

One final implication concerns the role of judgmental biases in marketing ethics. If salespeople in an organization know that their managers will systematically overweight outcomes and underweight the quality of the salespersons' decisions when evaluating their performance, managers should not be surprised later when salespeople turn to teleological (i.e., "the ends justify the means") approaches to business ethics. Account loading, payola, and selling without regard to customer needs are possible results. Then too, if salespeople are convinced that, during a given rating period, more recent performance information will be given substantially more weight by managers in ratings than earlier performance information, sales managers should not find it unusual that salespeople orchestrate their selling and customer contact activities accordingly. Unfortunately, such gamesmanship suboptimizes the use of company resources and compromises customer service and the building of long-term customer relationships. Organizations in which an outcome bias and an order effects bias dominate performance evaluations would appear likely to experience difficulty in implementing a relational approach to selling.

Clearly, the overall issue of judgmental biases in marketing management decision making deserves much more attention in the literature. From decisions about product

features and price to advertising campaigns to test marketing to strategic product planning, marketing decision making is fraught with the potential for decision biases to intrude. And in our profession, when decisions are suboptimal, the results of those decisions can be quite deleterious to the firm, running the gamut from sales organization turnover to new product introduction fiascos. One common thread in such results is that large sums of money are frequently lost by the organizations suffering from poor quality marketing management decision making.

It should be made clear that the author in no way suggests that sales organizations abandon outcome criteria for performance evaluation. Nor is the author suggesting that sometimes more recent performance information shouldn't be given relatively more weight than earlier performance information when rating salespeople. Rather, the issue is how to ensure that <u>both</u> decision appropriateness and outcome are taken into account when sales personnel are rated, and how to avoid succumbing to recency when order of receipt of performance information truly has no bearing on the evaluation.

Likewise, there is no implication from this research that sales organizations should switch from EoS response modes to SbS response modes in evaluating the performance of their salespeople. Rather, the point is to change the performance evaluation system where it can be changed in order to reduce the systematic introduction of judgmental

biases. But where the system can't be changed, the goal should be to train sales managers and salespeople to identify and avoid the biases.

Limitations

A key potential limitation of the research is that the subjects were drawn from a single company (albeit two separate sales divisions). This raises questions concerning generalizability of the results across other companies and industries. However, the basic results in Study One are highly similar to previous findings for an outcome bias across professional, student, and consumer samples. Likewise, the results in Study Two mirror those in the prior studies based upon the predictions of the Hogarth and Einhorn (1992) belief-adjustment model. Those studies also have utilized both student and business professionals as subjects.

The author is confident that the selection of a single company sample was appropriate and that the results can be generalized. Walker and Ruekert (1987) have suggested that studies employing single organizations are frequently worth some loss of external validity in order to increase internal validity. Also, Bonoma (1985) has argued that initial tests of a theory should be conducted using an in-depth case approach, which is analogous to a single-company study versus an interorganizational approach.

In fact, questions have been raised in the literature

as to the viability of prior studies within the domain of sales management that have incorporated broad-based samples across wide cross sections of industries. For example, in their meta-analysis on the determinants of salesperson performance, Churchill et al. (1985) found that the strength of the relationship between the major determinants and salespeople's performance was moderated by the type of products salespeople sell. Their analysis across 116 studies indicated that product type seemed to affect the relationship between five of the six main determinants of performance: personal factors, skill, role variables, motivation, and organizational/environmental factors. Only the predictor variable aptitude's correlation with performance was not moderated by type of product sold. Churchill et al. (1985) indicated that this finding provides powerful evidence for conducting salesforce performance research in job-specific settings if possible. Their argument is summarized as follows (Churchill, et al. 1985, p. 117):

> ...if there is a 'file drawer' problem inherent in academic researchers' inability to gain access to proprietary studies, it may be exactly the opposite of the problem usually encountered by social science researchers. Instead of the published studies having more positive results than those hidden away in people's files, in this case the hidden company studies may be more positive. The reason is that in the

published studies, the researchers often have tried to predict performance across salespeople in different kinds of sales jobs and in different firms and industries using the same set of performance criteria and the same set of predictor variables and measures.

Because the domain of the present study was the consumer packaged goods industry, and because many companies in that industry exhibit similar product lines, job requirements, training procedures, customers, and selling strategies, the results may be expected to be useful within a wide array of sales management applications.

Another potential limitation of the present study is based upon the research methodology and administration procedure used. Experiments conducted in a field setting as opposed to a lab sacrifice a modicum of experimental control for increased external validity. In the current experiment, a key threat to internal validity was potential hypothesis guessing by participants. Calder, Phillips, and Tybout (1981) have suggested carefully constructed cover stories and between-subjects designs as two ways to reduce the potential for hypothesis guessing. Both of these procedural suggestions were employed in the present study, and careful scrutiny of cognitive responses yielded no evidence of hypothesis guessing by subjects. Also, because for the most part these managers were geographically dispersed, little opportunity existed for them to discuss the project among themselves prior to mailing the packets back to the

researcher.

On a number of dimensions the studies revealed exceptionally good measurement properties. The scenarios were designed with the assistance of upper management of the participating company to be perceived as realistic to be as realistic as possible to the company's field sales managers. The overall response rate was quite high for a study of this type (78 percent), and thoroughness of completion of the questionnaires were exceptionally good---only two questionnaires that were returned were unusable. The scores on the ECOLVAL measure provided additional evidence that the managers believed the types of decisions and problems in the scenarios could actually happen to themselves or their salespeople. In all cases the manipulation checks gave evidence of strong manipulations, and all indications were that the subjects understood the task they were asked to perform, took the task seriously, and viewed it as realistic. Thus, as a whole the methodology selected appeared appropriate and the administration of the study successful.

One other limitation of the study bears mentioning---the strength of some of the ω^2's (omega squares, which represent the proportion of variance accounted for by each significant effect). In particular, the effect sizes for some of the interaction terms were relatively small (e.g., ω^2=0.9% for the interaction of decision appropriateness * outcome on DECQUAL). The variance explained by second order

and third order interaction terms will generally be low because of multicollinearity with main effect terms from which they are formed. In a discussion of variance accounted for across 33 regression analyses, Pedhazur (1982, p. 428) indicated that the variance explained by second-order terms ranged from 0.0 percent to 8.1 percent in the studies, with a median of 3.0 percent. Considering that the significant second-order interactions in the present study contributed from 0.9 to 3.6 percent of the total variance, these results compare favorably with effect sizes found in previous studies. Cautioning researchers against reading too much into the incremental variance findings, Pedhazur (1982) recommended that researchers not overemphasize incremental variance, but focus more instead on the substantive contribution made by theory that is being tested.

Managerial Action Steps

Based upon the results of the present studies, sales organizations are encouraged to take the following steps.

1) Invest in training of sales managers to debias performance evaluations. It is imperative that discussions of various rater errors be incorporated as part of all sales management training programs. The pervasive nature of biases such as an outcome bias and an order effects bias should be demonstrated to the managers. Role-play exercises and simulations should

address both good and bad examples of salesforce performance evaluation.

2) Ensure that sales managers work with their salespeople frequently and that they provide written performance feedback each time they work together. Sales organizations should insist that managers keep running files of performance evidence on each salesperson supervised in order to minimize the potential that evaluations will be overweighted toward the most recent evidence. It is unrealistic to suggest that sales organizations switch from an iterative approach to evaluation to an end of the period approach so as to reduce the potential for recency. It is, however, reasonable to suggest that sales organizations sensitize managers to the potential for recency and to the importance of taking into account both decision and outcome information in evaluations.

3) Provide salespeople with sufficient organizational support for effective empowerment. When a salesperson is told to make decisions in an entrepreneurial fashion (as is common today in many organizations practicing "Total Quality Management" and other similar business philosophies), but later is given differential performance feedback when an inappropriate decision was made depending upon whether the outcome was good or bad, the signal received by the salesperson is that the company is not serious about supporting the desired

entrepreneurial behavior. The result could be that salespeople reverting to a focus on seeking short-term, low risk customers, rather than targeting lower probability but higher potential payoff customers, in an effort to maximize the perception by the sales manager that successful outcomes are occurring. The importance of entrepreneurial behavior, creativity, and risk taking by salespeople was highlighted by the cognitive responses by subjects in the appropriate decision condition of the present study, as discussed previously.

4) Train salespeople about the potential problems of rater biases, and incorporate self-reporting procedures into the formal performance evaluation process. Such an approach provides another potential check against rater bias, since typically in such systems both the manager's and the salesperson's evaluations are synthesized into a consensus set of ratings. More and more organizations are utilizing self-report ratings, and in doing so will have to watch for yet another potential bias, self-serving attributional bias (Miller 1978). Marshall, Mowen, and Fabes (1992) found empirical evidence of such a bias in comparing self versus other salesforce performance ratings.

5) Design rating instruments that force sales managers to focus on the quality of decision making by sales personnel, as well as on quantitative results. Perhaps

a reconsideration of the benefits of behaviorally anchored rating scales (BARS) is in order. BARS are designed to increase the salience of performance information beyond mere outcomes. The unique aspect of BARS is that, in developing the scales, salespeople are actively involved in identifying important performance results and critical behaviors necessary to achieve those results. No other appraisal instrument incorporates this linkage between behaviors and results. (For more on the use of BARS in salesforce evaluation see Cocanougher and Ivancevich 1978; Ingram and LaForge 1992, pp. 555-557; Locander and Staples 1976; and Muczyk and Gable 1987.)

6) Design reward systems consistent with the expectations of entrepreneurial behavior by salespeople. As discussed in Chapter II, surprisingly little research has been conducted on salesforce reward structures. Rewards that are entirely appropriate in sales organizations where decision making follows a top-down approach and salespeople primarily <u>execute</u> decisions may be quite inappropriate in organizations where adaptive selling is practiced and much freedom in decision making by salespeople is encouraged (Weitz, Sujan, and Sujan 1986). At this point, the dearth of research in the area of reward structures means that even the range of possibilities for such alternative reward systems is quite speculative, although providing

a greater potential for attainment of intrinsic rewards seems a logical starting point. Approaches such as organizational behavioral modification (Scott et al. 1986) and behavioral self management (Sauers, Hunt, and Bass 1990) may prove useful in conceptualizing alternative reward systems in organizations practicing employee empowerment for decision making.

Actions such as those outlined above may be expected to contribute to a reduction in the amount of rater bias, and ultimately to more fair and accurate performance evaluation of sales personnel. When provided fair and accurate performance feedback and appropriate levels of organizational support, salespeople will be far more likely to increase efforts toward achieving the performance results that are desired by both the sales organization and the salespeople themselves.

REFERENCES

Abell, D.F. and J.S. Hammond (1979), Strategic Market Planning: Problems and Analytical Approaches, Englewood Cliffs, NJ: Prentice-Hall, Inc.

Adams, J. Stacy (1976), "The Structure and Dynamics of Behavior in Organizational Boundary Roles," in Handbook of Organizational Psychology, Marvin D. Dunnette, ed. Chicago: Rand Mc Nally, 1193-1199.

Adkins, Robert T. (1979), "Evaluating and Comparing Salesmen's Performance," Industrial Marketing Management, 5 (June), 207-212.

Anderson, Erin and Richard L. Oliver (1987), "Perspectives on Behavior-Based Versus Outcome-Based Salesforce Control Systems," Journal of Marketing, 51 (October), 76-88.

Anderson, Neil and Viv Shackleton (1990), "Decision Making in the Graduate Selection Interview: A Field Study," Journal of Occupational Psychology, 63 (March), 63-76.

Anderson, Norman H. (1971), "Integration Theory and Attitude Change," Psychological Review, 78, 171-206.

---------- (1981), Foundations of Information Integration Theory. New York: Academic Press.

Ashton, Alison Hubbard and Robert H. Ashton (1988), "Sequential Belief Revision in Auditing," The Accounting Review, 63 (October), 623-641.

Avila, Ramon A., Edward F. Fern, and O. Karl Mann (1988), "Unravelling Criteria for Assessing the Performance of Salespeople: A Causal Analysis," Journal of Personal Selling and Sales Management, 8 (May), 45-54.

Bagozzi, Richard P. (1978), "Salesforce Performance and Satisfaction as a Function of Individual Difference, Interpersonal, and Situational Factors," Journal of Marketing Research, 15 (November), 517-531.

---------- (1980a), "Salespeople and Their Managers: An Exploratory Study of Some Similarities and Differences," Sloan Management Review, 21 (Winter), 15-26.

---------- (1980b), "Performance and Satisfaction in an Industrial Sales Force: An Examination of Their Antecedents and Simultaneity," Journal of Marketing, 44 (April), 65-77.

Baron, Jonathan and John C. Hershey (1988), "Outcome Bias in Decision Making," Journal of Personality and Social Psychology, 54 (April), 569-579.

Baskett, G.D. (1973), "Interview Decisions as Determined by Competency and Attitude Similarity," Journal of Applied Psychology, 57, 343-345.

Bazerman, Max H. (1990), Judgment in Managerial Decision Making, New York: John Wiley and Sons.

Behrman, Douglas N. and William D. Perreault, Jr. (1982), "Measuring the Performance of Industrial Salespersons," Journal of Business Research, 10 (September), 355-370.

---------- and ---------- (1984), "A Role Stress Model of the Performance and Satisfaction of Industrial Salespersons," Journal of Marketing, 48 (Fall), 9-21.

Bettman, James C. (1979), An Information Processing Theory of Consumer Choice, Reading, MA: Addison-Wesley.

Bonoma, Thomas V. (1985), "Case Research in Marketing: Opportunities, Problems, and a Process," Journal of Marketing Research, 12 (May), 199-208.

Byrne, D. (1961), "Interpersonal Attraction and Attitude Similarity," Journal of Abnormal and Social Psychology, 62, 713-715.

Calder, Bobby J., Lynn W. Phillips, and Alice M. Tybout (1981), "Designing Research for Application," Journal of Consumer Research, 8 (September), 197-207.

Campbell, John P., Marvin D. Dunnette, Edward E. Lawler, III, and Karl E. Weick, Jr. (1970), Managerial Behavior: Performance and Effectiveness. New York: McGraw-Hill, Inc.

Chonko, Lawrence B. (1986), "Organizational Commitment in the Sales Force," Journal of Personal Selling and Sales Management, 6 (November), 19-27.

Churchill, Gilbert A., Jr., Neil M. Ford, and Orville C. Walker, Jr. (1976), "Organizational Climate and Job Satisfaction in the Salesforce," Journal of Marketing Research, 13 (November), 323-332.

----------, ----------, and ---------- (1979), "Personal Characteristics of Salespeople and the Attractiveness of Alternative Rewards," Journal of Business Research, 7, 25-50.

----------, ----------, Steven W. Hartley, and Orville C. Walker, Jr. (1985), "The Determinants of Salesperson Performance: A Meta-Analysis," Journal of Marketing Research, 22 (May), 103-118.

Cocanougher, A. Benton and John M. Ivancevich (1978), "'BARS' Performance Rating for Sales Force Personnel," Journal of Marketing, 42 (July), 87-95.

Cotham, James C., III and David W. Cravens (1969), "Improving Measurement of Salesman Performance: A Method for Evaluating Multiple Aspects," Business Horizons, 12 (June), 79-83.

Dalessio, Anthony and Andrew S. Imada (1984), "Relationships Between Interview Selection Decisions and Perceptions of Applicant Similarity to an Ideal Employee and Self: A Field Study," Human Relations, 37, 67-80.

DeNisi, Angelo S., Thomas P. Cafferty, and Bruce M. Meglino (1984), "A Cognitive View of the Performance Appraisal Process: A Model and Research Propositions," Organizational Behavior and Human Performance, 33 (June), 360-396.

Deshpande, Rohit and Fredrick E. Webster, Jr. (1989), "Organizational Culture and Marketing: Defining the Research Agenda," Journal of Marketing, 53 (January), 3-15.

Dubinsky, Alan J. and Thomas E. Barry (1982), "A Survey of Sales Management Practices," Industrial Marketing Management, 11 (April), 133-141.

----------, Roy D. Howell, Thomas N. Ingram, and Danny N. Bellenger (1986), "Salesforce Socialization," Journal of Marketing, 50 (October), 192-207.

----------, Steven J. Skinner, and Tommy E. Whittler (1989), "Evaluating Sales Personnel: An Attribution Theory Perspective," Journal of Personal Selling and Sales Management, 9 (Spring), 9-21.

Dunnette, Marvin D. and Wayne K. Kirchner (1960), "Psychological Test Differences Between Industrial Salesmen and Retail Salesmen," Journal of Applied Psychology, 44 (April), 121-125.

Early, P. Christopher, Gregory B. Northcraft, Cynthia Lee, and Terri R. Lituchy (1990), "Impact of Process and Outcome Feedback on the Relation of Goal Setting to Task Performance," Academy of Management Journal, 33 (March), 87-105.

Edwards, Ward, Istvan Kiss, Giandomenico Majone, and Masanao Toda (1984), "What Constitutes a 'Good Decision'?," panel discussion with introduction and postscript by Charles Vlek, Acta Psychologica, 56 (August), 5-27.

Evans, Kenneth R., Loren Margheim, and John L. Schlacter (1982), "A Review of Expectancy Theory Research in Selling," Journal of Personal Selling and Sales Management, 2, 33-40.

Feldman, D.C. (1976), "A Contingency Theory of Socialization," Administrative Science Quarterly, 21 (September), 433-450.

Feldman, Jack M. (1981), "Beyond Attribution Theory: Cognitive Processes in Performance Appraisal," Journal of Applied Psychology, 66 (April), 127-148.

Fern, Edward F., Ramon A. Avila, and Dhruv Grewal (1989), "Salesforce Turnover: Those Who Left and Those Who Stayed," Industrial Marketing Management, 18, 1-9.

Fischhoff, Baruch (1975), "Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty," Journal of Experimental Psychology: Human Perception and Performance, 1 (August), 288-299.

---------- and Ruth Beyth (1975), "'I Knew It Would Happen' ---Remembered Probabilities of Once-Future Things," Organizational Behavior and Human Performance, 13 (February), 1-16.

---------- (1982), "Debiasing," in Judgment Under Uncertainty: Heuristics and Biases, Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., New York: Cambridge University Press, 422-444.

Ford, Neil M., Orville C. Walker, Jr., and Gilbert A. Churchill, Jr. (1975), "Expectation-Specific Measures of the Inter -Sender Conflict and Role Anbiguity Experienced by Industrial Salesmen," Journal of Business Research, 3 (April) 95-112.

----------, Orville C. Walker, Jr., and Gilbert A. Churchill, Jr. (1976), "The Psychological Consequences of Role Conflict and Ambiguity in the Industrial Salesforce," in Marketing: 1776-1976 and Beyond, Kenneth L. Bernhardt, ed. Chicago: American Marketing Association, 403-408.

---------- Orville C. Walker, Jr., Gilbert A. Churchill, Jr., and Steven W. Hartley (1987), "Selecting Successful Salespeople: A Meta-Analysis of Biographical and Psychological Selection Criteria," in Review of Marketing, Michael J. Houston, ed. Chicago: American Marketing Association, 90-131.

Freund, Rudolf J. and Ramon C. Littell (1981), SAS for Linear Models: A Guide to the ANOVA and GLM Procedures, Cary, IN: SAS Institute, Inc.

Futrell, Charles M. and A. Parasuraman (1984), "The Relationship of Satisfaction and Performance to Salesforce Turnover," Journal of Marketing, 48 (Fall), 33-40.

Gentry, James W., John C. Mowen, and Lori Tasaki (1991), "Salesperson Evaluation: A Systematic Structure for Reducing Judgmental Biases," Journal of Personal Selling and Sales Management, 11 (Spring), 27-38.

Ghiselli, Edwin E. (1973), "The Validity of Aptitude Tests in Personnel Selection," Personnel Psychology, 26 (Winter), 461-477.

Goolsby, Jerry R. (1992), "A Theory of Role Stress in Boundary Spanning Positions of Marketing Organizations," Journal of the Academy of Marketing Science, 20 (Spring), 155-164.

Graen, G. and W. Schiemann (1978), "Leader-Member Agreement: A Vertical Dyad Linkage Approach," Journal of Applied Psychology, 63, 206-212.

Green, Stephen G. and Terence R. Mitchell (1979), "Attributional Processes of Leaders in Leader-Member Interactions," Organizational Behavior and Human Performance, 23 (June), 429-458.

Greenhouse, S.W. and S. Geisser (1959), "On Methods in the Analysis of Profile Data," Psychometrika, 24, 95-112.

Harkins, Stephen G. and Richard E. Petty (1981), "Effects of Source Magnification of Cognitive Effort on Attitudes: An Information-Processing View," Journal of Personality and Social Psychology, 40 (March), 401-413.

Hawkins, Scott A. and Reid Hastie (1990), "Hindsight Biased Judgments of Past Events After the Outcomes Are Known," Psychological Bulletin, 107 (May), 311-327.

Heider, Fritz (1958), The Psychology of Interpersonal Relations, New York: John Wiley and Sons.

Heneman, Herbert, Jr. (1975), "Research Roundup," The Personnel Administrator, as cited in Churchill, Gilbert A., Jr., Neil M. Ford, Steven W. Hartley, and Orville C. Walker, Jr. (1985), "The Determinants of Salesperson Performance: A Meta-Analysis," Journal of Marketing Research, 22 (May), 103-118.

Hogarth, Robin (1987), Judgment and Choice. New York: Wiley.

---------- and Hillel J. Einhorn (1992), "Order Effects in Belief Updating: The Belief Adjustment Model," Cognitive Psychology, 24, 1-55.

Hunt, Shelby D., Lawrence B. Chonko, and Van P. Wood (1985), "Organizational Commitment and Marketing," Journal of Marketing, 49 (Winter), 112-26.

Ilgen, Daniel R. and Jack M. Feldman (1983), "Performance Appraisal: A Process Focus," in Research in Organizational Behavior, Vol. 5, Barry M. Staw, ed., Greenwich, CT: JAI Press, Inc., 141-197.

Ingram, Thomas N., Keun S. Lee, and Steven J. Skinner (1989), "An Empirical Assessment of Salesperson Motivation, Commitment, and Job Outcomes," Journal of Personal Selling and Sales Management, 9 (Fall), 25-33.

---------- and Raymond W. LaForge (1992), Sales Management: Analysis and Decision Making, 2nd ed., Fort Worth: The Dryden Press.

Kelley, Harold E. (1967), "Attribution Theory in Social Psychology," in Nebraska Symposium on Motivation, Daniel Levine, ed. Lincoln, NE: University of Nebraska Press, 192-238.

---------- (1973), "The Process of Causal Attribution," American Psychologist, 28 (February), 107-128.

Kozlowski, S.W.J. and M.P. Kirsch (1987), "The Systematic Distortion Hypothesis, Halo, and Accuracy: An Individual-Level Analysis," Journal of Applied Psychology, 72, 252-261.

Jackson, Donald W., Janet E. Keith, and John L. Schlacter (1983), "Evaluation of Sales Performance: A Study of Current Practices," Journal of Personal Selling and Sales Management, 3 (November), 43-51.

Jaworski, Bernard J. and Ajay K. Kohli (1991), "Supervisory Feedback: Alternative Types and Their Impact on Salespeople's Performance and Satisfaction," Journal of Marketing Research, 28 (May), 190-201.

LaForge, Raymond W. and David W. Cravens (1981), "A Market Response Model for Sales Management Decision Making," Journal of Personal Selling and Sales Management, 2 (Fall/Winter), 10-16.

Lagace, Rosemary R. (1990), "Leader-Member Exchange: Antecedents and Consequences of the Cadre and Hired Hand," Journal of Personal Selling and Sales Management, 10 (February), 11-19.

Lamont, Lawrence M. and William J. Lundstrom (1977), "Identifying Successful Industrial Salesmen by Personality and Personal Characteristics," Journal of Marketing Research, 14 (November), 517-529.

Landy, Frank J. and James L. Farr (1983), The Measurement of Work Performance: Methods, Theory, and Applications, Orlando: Academic Press.

Latham, Gary P. (1986), "Job Performance and Appraisal," Review of Industrial and Organizational Psychology, Chichester, England: Wiley, 117-155.

----------, Diane Irvine, Daniel Skarlicki, and Jacob P. Siegel (1993), "The Increasing Importance of Performance Appraisals to Employee Effectiveness in Organizational Settings in North America," in International Review of Industrial and Organizational Psychology, C.L. Cooper and I. Robertson, eds.

LaTour, Stephen A. and Paul W. Miniard (1983), "The Misuse of Repeated Measures Analysis in Marketing Research," Journal of Marketing Research, 20 (February), 45-57.

Lawler, Edward E., III (1968), "A Correlational-Causal Analysis of the Relationship Between Expectancy Attitudes and Job Performance," Journal of Applied Psychology, 52 (December), 462-468.

Lipschitz, Raanan (1989), "'Either a Medal or a Corporal': The Effects of Success and Failure on the Evaluation of Decision Making and Decision Makers," Organizational Behavior and Human Decision Making, 44 (December), 380-395.

Locander, William B. and William A. Staples (1978), "Evaluating and Motivating Salesmen with the BARS Method," Industrial Marketing Management, 7 (February), 43-48.

Locke, Edwin A. (1970), "Job Satisfaction and Job Performance: A Theoretical Analysis," Organizational Behavior and Human Performance, 5, 484-500.

Lucas, George H., Jr., A. Parasuraman, Robert A. Davis, and Ben M. Enis (1987), "An Empirical Study of Salesforce Turnover," Journal of Marketing, 51 (July), 34-59.

Marshall, Greg W. and Stephen J. Miller (1991), "Total Quality Management and the Internal Customer: A Marketing Perspective on Employee Motivation," Southern Marketing Association Proceedings, 384-389.

----------, John C. Mowen, and Keith J. Fabes (1992), "The Impact of Territory Difficulty and Self Versus Other Ratings on Managerial Evaluations of Sales Personnel," Journal of Personal Selling and Sales Management, 12 (Fall), 35-47.

---------- and John C. Mowen (1993), "An Experimental Investigation of the Outcome Bias in Salesperson Performance Evaluations," Journal of Personal Selling and Sales Management, forthcoming.

Maurer, T.J. and R.A. Alexander (1991), "Contrast Effects in Behavioral Measurement: An Investigation of Alternative Process Explanations, Journal of Applied Psychology, 76, 3-10.

McKay, Sandra, Joseph F. Hair, Jr., Mark W. Johnston, and Daniel L. Sherrell (1991), "An Exploratory Investigation of Reward and Corrective Responses to Salesperson Performance: An Attributional Approach," Journal of Personal Selling and Sales Management, 11 (Spring), 39-48.

Miller, Dale T. (1978), "What Constitutes a Self-Serving Attributional Bias?," Journal of Personality and Social Psychology, 36 (November), 1221-1223.

Miner, John B. (1962), "Personality and Ability Factors in Sales Performance," Journal of Applied Psychology, 46 (February), 6-13.

Mitchell, Terence R., Stephen G. Green, and Robert E. Wood (1981), "An Attributional Model of Leadership and The Poor Performing Subordinate: Development and Validation," in Research in Organizational Behavior, L.L. Cummings and Barry M. Staw, eds., Greenwich, CT: JAI Press, Inc.

---------- and Laura S. Kalb (1981), "Effects of Outcome Knowledge and Outcome Valence on Supervisors' Evaluations," Journal of Applied Psychology, 66 (October), 604-612.

Mizerski, Richard W., Linda L. Golden, and Jerome B. Kernan (1979), "The Attribution Process in Consumer Decision Making," Journal of Consumer Research, 6 (September), 123-140.

Morris, Michael H., Duane L. Davis, Jeffrey W. Allen, Ramon A. Avila, and Joseph Chapman (1991), "Assessing the Relationships Among Performance Measures, Managerial Practices, and Satisfaction When Evaluating the Salesforce: A Replication and Extension," Journal of Personal Selling and Sales Management, 11 (Summer), 25-35.

Mosel, James N. (1952), "Prediction of Department Store Sales Performance from Personnel Data," Journal of Applied Psychology, 36, 8-10.

Mowen, John C. and James W. Gentry (1980), "Investigation of the Preference Reversal Phenomenon in a New Product Introduction Task," Journal of Applied Psychology, 65, 715-722.

----------, Stephen W. Brown, and Donald W. Jackson, Jr. (1981), "Cognitive Biases in Sales Management Evaluations," Journal of Personal Selling and Sales Management, 1 (Fall/Winter), 83-89.

----------, Janet E. Keith, Stephen W. Brown, and Donald W. Jackson, Jr. (1985), "Utilizing Effort and Task Difficulty Information in Evaluating Salespeople," Journal of Marketing Research, 22 (May), 185-191.

----------, Keith J. Fabes, and Raymond W. LaForge (1986), "Effects of Effort, Territory Situation, and Rater on Salesperson Evaluation," Journal of Personal Selling and Sales Management, 6 (May), 1-8.

---------- and Gary J. Gaeth (1992), "The Evaluation Stage in Marketing Decision Making," Journal of the Academy of Marketing Science, 20 (Spring), 177-187.

---------- and Greg W. Marshall (1992), "The Effects of Gain/Loss Frames on Product Management Risk-Taking: A Replication and Extension," Proceedings of the American Marketing Association Winter Educators' Conference, 189-195.

---------- and Thomas H. Stone (1992), "An Empirical Analysis of Outcome Biases in Constituent Evaluations of Public Policy Decision Makers," Journal of Public Policy and Marketing, 11 (Spring), 24-32.

Mount, M.K. and D.E. Thompson (1987), "Cognitive Categorization and Quality of Performance Ratings," Journal of Applied Psychology, 72, 240-246.

Mowday, Richard T., R.M. Steers, and L.W. Porter (1979), "The Measurement of Organizational Commitment," Journal of Organizational Behavior, 14, 224-247.

Muczyk, Jan P. and Myron Gable (1987), "Managing Sales Performance Through a Comprehensive Performance Appraisal System," Journal of Personal Selling and Sales Management, 7 (May), 41-52.

Murphy, K.R. and R. Jako (1989), "Under What Conditions are Observed Intercorrelations Greater or Smaller than True Intercorrelations?," Journal of Applied Psychology, 74, 827-830.

Nathan, B.R. and N. Tippins (1990), "The Consequences of Halo "Error" in Performance Ratings: A Field Study of the Moderating Effect of Halo on Test Validation Results,", Journal of Applied Psychology, 75, 290-296.

Nisbett, Richard and Lee Ross (1980), Human Inference: Strategies and Shortcomings of Social Judgment, Englewood Cliffs, NJ: Prentice-Hall, Inc.

O'Connell, William A. and William Keenan, Jr. (1990), "The Shape of Things to Come," Sales and Marketing Management, 142 (January), 36.

Oliver, Richard L. (1974), "Expectancy Theory Predictions of Salesmen's Performance," Journal of Marketing Research, 11 (August), 243-253.

Ouchi, William G. (1981), Theory Z, Reading, MA: Addison-Wesley.

Patton, W.E., III and Ronald H. King (1985), "The Use of Human Judgment Models in Evaluating Sales Force Performance," Journal of Personal Selling and Sales Management, 5 (May), 1-12.

Pedhazur, Elazar J. (1982), Multiple Regression in Behavioral Research: Explanation and Prediction, 2nd ed., New York: Holt, Rinehart, and Winston.

Peter, J. Paul, Gilbert A. Churchill, Jr., and Tom J. Brown (1993), "Caution in the Use of Difference Scores in Consumer Research," Journal of Consumer Research, 19 (March), 655-662.

Peters, L.H. and J.R. Terborg (1975), "The Effects of Temporal Placement of Unfavorable Information and of Attitude Similarity on Personnel Selection Decisions," Organizational Behavior and Human Performance, 13, 279-293.

Porter, Lyman W. and Edward E. Lawler, III, Managerial Attitudes and Performance, Homewood IL: Richard D. Irwin.

Posner, Barry Z., James M. Kouzes, and Warren H. Schmidt (1985), "Shared Values Makes a Difference: An Empirical Test of Corporate Culture," Human Resource Management, 24 (Fall), 293-309.

Pulakos, R.D., N. Schmitt, and C. Ostroff (1986), "A Warning About the Use of a Standard Deviation Across Dimensions Within Ratees to Measure Halo," Journal of Applied Psychology, 71, 29-32.

Rand, Thomas M. and Kenneth N. Wexley (1975), "Demonstration of the Effect, 'Similar to Me,' in Simulated Employment Interviews," Psychological Reports, 36 (April), 535-544.

Rhoads, Gary (1988), Selling Styles and Performance Effectiveness: The Relationship Between Interpersonal Selling Behaviors and Customer Satisfaction with the Exchange, unpublished Ph.D. dissertation, Lubbock, TX: Texas Tech University.

Rich, Leslie (1966), "Can Salesmen Be Tested?" Duns Review, 87 (March), 40-41.

Ross, Lee (1977), "The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process," in Advances in Experimental Social Psychology, Vol. 10, Leonard Berkowitz, ed., New York: Academic Press.

Ruekert, Robert W. and Orville C. Walker, Jr. (1987), "Marketing's Interaction with Other Functional Units: A Conceptual Framework and Empirical Evidence," Journal of Marketing, 51 (January), 1-19.

Ryans, Adrian B. and Charles B. Weinberg (1979), "Territory Sales Response," Journal of Marketing Research, 16 (November), 453-465.

Sager, Jeffrey K. and Mark W. Johnston (1989), "Antecedents and Outcomes of Organizational Commitment: A Study of Salespeople," Journal of Personal Selling and Sales Management, 9 (Spring), 30-41.

Sales and Marketing Management 1990 Survey of Selling Costs, February 26, 1990, p. 8.

Sauers, Daniel A, James B. Hunt, and Ken Bass (1990), "Behavioral Self-Management as a Supplement to External Sales Force Controls," Journal of Personal Selling and Sales Management, 10 (Summer), 17-28.

Saxe, Robert and Barton A. Weitz (1982), "The SOCO Scale: A Measure of the Customer Orientation of Salespeople," Journal of Marketing Research, 19 (August), 343-351.

Schein, Edgar H. (1984), Organizational Culture and Leadership, San Francisco: Jossey-Bass Publishers.

Scholl, Richard W. (1981), "Differentiating Organizational Commitment from Expectancy as a Motivating Force," Academy of Management Review, 6, 589-599.

Scott, Robert A., John E. Swan, M. Elizabeth Wilson, and Jenny J. Roberts (1986), Organizational Behavior Modification: A General Motivational Tool for Sales Management," Journal of Personal Selling and Sales Management, 6 (August), 61-70.

Shanteau, James C. (1970), "An Additive Model for Sequential Decision Making," Journal of Experimental Psychology, 85 (August), 181-191.

Smither, J.W., R.R. Reilly, and R. Buda (1988), "Effect of Prior Performance Information on Ratings of Present Performance: Contrast Versus Assimilation Revisited," Journal of Applied Psychology, 73, 487-496.

Spiro, Rosann L. and Barton A. Weitz (1990), "Adaptive Selling: Conceptualization, Measurement, and Nomological Validity," Journal of Marketing Research, 27 (February), 61-69.

Steiner, Dirk D. and Jeffrey S. Rain (1989), "Immediate and Delayed Primacy and Recency Effects in Performance Evaluation," Journal of Applied Psychology, 74, 136-142.

Teas, R. Kenneth (1981), "An Empirical Test of Models of Salesperson's Job Expectancy and Instrumentality Perceptions," Journal of Marketing Research, 18 (May), 209-26.

---------- and James C. McElroy (1986), "Causal Attributions and Expectancy Estimates: A Framework for Understanding the Dynamics of Salesforce Motivation," Journal of Marketing, 50 (January), 75-86.

Tubbs, Richard M., William F. Messier, Jr., and W. Robert Knechel (1990), "Recency Effects in the Auditor's Belief-Revision Process," The Accounting Review, 65 (April), 452-460.

Tversky, Amos and Daniel Kahneman (1974), "Judgment Under Uncertainty: Heuristics and Biases," Science, 185 (July-September), 1124-1131.

---------- and ---------- (1981), "The Framing of Decisions and the Psychology of Choice," Science, 211 (January-March), 453-458.

Tyagi, Pradeep K. (1982), "Perceived Organizational Climate and the Process of Salesperson Motivation," Journal of Marketing Research, 19 (May), 240-54.

---------- (1985a), "Organizational Climate, Inequities, and Attractiveness of Salesperson Rewards," Journal of Personal Selling and Sales Management, 5 (November), 31-37.

---------- (1985b), "Relative Importance of Key Job Dimensions and Leadership Behaviors in Motivating Salesperson Work Performance," Journal of Marketing, 49 (Summer), 76-86.

Umemura, George M. (1956), "Measuring Salesmen's Performance," in Conference Board Reports: Studies in Business Policy, No. 79, New York: National Industrial Conference Board, Inc.

VanMaanen, J. (1976), "Breaking In: Socialization to Work," in Handbook of Work, Organization, and Society, R. Dubin, ed. Chicago: Rand-McNally, Inc.

---------- and E.H. Schein (1979), "Toward a Theory of Organizational Socialization," in Research in Organizational Behavior, B.M. Staw, ed. Greenwich, CT: JAI Press, Inc.

Vroom, Victor (1964), Work and Motivation. New York: John Wiley and Sons.

---------- and E.L. Deci (1971), "The Stability of Post Decision Dissonance: A Follow-up Study of the Job Attitudes of Business School Graduates," Organizational Behavior and Human Performance, 6, 36-49.

Walker, Orville C., Jr., Gilbert A. Churchill, Jr., and Neil M. Ford (1972), "Reactions to Role Conflict: The Case of the Industrial Salesman," Journal of Business Administration, 3 (Spring), 25-36.

----------, ----------, and ---------- (1975), "Organizational Determinants of the Industrial Salesman's Role Conflict and Ambiguity," Journal of Marketing, 39, (January), 32-39.

----------, ----------, and ---------- (1977), "Motivation and Performance in Industrial Selling: Present Knowledge and Needed Research," Journal of Marketing Research, 14 (May), 156-168.

Weiner, Bernard (1974), Achievement Motivation and Attribution Theory. Morristown, N.J.: General Learning Corp.

Weitz, Barton A. (1979), "A Critical Review of Personal Selling Research: The Need for a Contingency Approach," in Critical Issues in Sales Management: State-of-the-Art and Future Research Needs, G. Albaum and G. Churchill, eds. Eugene, OR: University of Oregon, College of Business Administration.

---------- (1981), "Effectiveness in Sales Interactions: A Contingency Framework," Journal of Marketing, 45 (Winter), 85-103.

----------, Harish Sujan, and Mita Sujan (1986), "Knowledge, Motivation, and Adaptive Behavior: A Framework for Improving Selling Effectiveness," Journal of Marketing, 50 (October), 174-191.

Wexley, Kenneth N. and W.F. Nemeroff (1974), "Effects of Racial Prejudice, Race of Applicant, and Biographical Similarity on Interviewer Evaluations of Job Applicants," Journal of Social and Behavioral Sciences, 20, 66-78.

Williams, Michael R. (1992), "Organizational Culture as a Predictor of the Level of Salespersons' Customer Oriented Behavior," unpublished Ph.D. dissertation, Stillwater, OK: Oklahoma State University.

Yukl, Gary (1989), "Managerial Leadership: A Review of Theory and Research," Journal of Management, 15, 251-289.

APPENDIX A

COMPLETE SET OF MATERIALS MAILED IN STUDY ONE - OUTCOME BIAS

COVER LETTER FROM VICE PRESIDENT OF SALES
FOR EACH OF THE TWO SALES DIVISIONS
PARTICIPATING IN THE STUDY

(SENT ON COMPANY LETTERHEAD)

May 3, 1993

Enclosed please find a set of materials related to a national research project in which the *(insert division name)* is participating. This project is related to issues of evaluating field sales personnel. A research team from the College of Business Administration at Oklahoma State University is conducting the study. *(Insert division name)* has pledged full participation in the study by all managers.

May I ask that you follow the attached instructions exactly, fill out the forms completely, and return the materials to the researchers in the postage paid envelope provided. They need your responses back by **5/17/93**, therefore please complete all the materials and return the packet to them ASAP.

Let me assure you that no member of *(insert company name)* will ever see your individual responses. The research team will provide us with results only in aggregate (summarized) form---no names or territory numbers will be matched to your responses. You will receive a copy of the summarized results when they are available.

I have been asked to emphasize that there are no right or wrong answers to this exercise---only your views and opinions. Therefore, there is no reason to try to second guess the information or questions. The most important thing you can do is give the exercise your thoughtful consideration, follow the instructions exactly, fill out the questionnaires completely, and return the packet to the research team promptly.

Be sure to meet their due date of **5/17/93**. Thanks in advance for your efforts---our company will benefit from the research.

(Signature of appropriate vice president of sales)

TASK DESCRIPTION (INSTRUCTIONS)
(Please Review Carefully Before Proceeding)

On the following pages you will read some information in scenario form about a hypothetical salesperson engaged in several job-related activities over three different sales situations. These three sales situations span a time period of several weeks. After carefully reading each scenario, your task is to assume the role of the employee's sales manager (i.e., the salesperson's immediate supervisor) and respond to a few questions that follow each scenario.

For the results of this study to be meaningful, the instructions provided throughout these materials must be followed exactly. Note that there are no "right" or "wrong" answers to any of the questions—the questions ask only for your own views and opinions from the perspective of the role you are playing as the employee's sales manager.

Because your company will be provided results of this study only in aggregate form (overall summary numbers), your individual responses will never be seen by anyone in your organization. Thus, it is not necessary for you to identify yourself on these materials. The task you are about to perform has been designed to be interesting and fun, and the results have the potential to improve salesforce performance evaluation in all organizations.

A COUPLE OF NOTES: 1) It is very important that you completely answer the questions that follow each scenario before going on to the next scenario—please do not skip ahead. It is, however, permissible to refer back to each scenario individually as you answer the questions following it. 2) Please limit your evaluations to consideration only of the information provided in the scenarios. Please do not try to second-guess the information or find solutions not available via the facts provided.

INSTRUCTIONS FOR RETURNING COMPLETED MATERIALS: When you have completed the project please fold over this stapled set of materials once from top to bottom and return it immediately in the pre-addressed, postage-paid envelope provided. The research team needs your packet back no later than Monday, May 17 so we may begin tabulating the results and provide a report back to your organization in a timely manner.

Thanks for your participation!

Professor John C. Mowen
College of Business, Oklahoma State University
University Center at Tulsa
700 N. Greenwood Avenue
Tulsa, OK. 74106-0700
Phone: (405) 744-5112

PS: As a token of our appreciation for your participation in and prompt response to this research project, enclosed please find a small gift from Oklahoma State University.

GENERAL BACKGROUND INFORMATION ON SALESPERSON SMITH

Salesperson "Smith" is employed by your company as a territory representative. A large part of the job involves selling products to and servicing existing accounts, but effort is also put into developing new business when the opportunity arises. Your company trains and encourages its salespeople to make as many of their own decisions as possible regarding account strategies and territory management.

Smith was previously employed in a similar position with another company for 2 years, but has only worked for you and your company for 3 months. In terms of geographic size, number of customers, and number of competing salespeople from other companies, Smith's territory is comparable to most other territories within your company. Smith's time spent in call preparation is also about the same as that of most salespeople in your company and within the industry. Smith's previous employer provided a favorable recommendation during the employment interview process, and so far Smith seems to be generally well-liked by your company's management, other salespeople, and customers.

Over the last few years the overall economic climate in Smith's area has been slightly better than the economy of the U.S. as a whole, and that trend is forecasted to continue for the foreseeable future.

Please go on to the next page to read the first scenario about Salesperson Smith.

SCENARIO #1

Salesperson Smith recently faced a decision about which of two potentially large promotional orders for an allocated special pack item to pursue from two chain customers whose headquarters buying offices are in Smith's territory. The potential order would come from accounts referred to here as "Customer X" and "Customer Y." The headquarters buyer for each chain is a regular customer of Smith, and each has the authority to quickly make such buying decisions. The allocated pack is about to sell out in your company's distribution center, and there is only a sufficient quantity available for one of Smith's two potential customers to run a satisfactory promotion. Neither customer will buy the item for a promotion without the special pack.

Smith must direct selling efforts toward gaining only one of these orders because both customers would need to be pursued immediately and in person or the allocated pack will run out in your company's distribution center and Smith's prospective buyer will purchase an alternative item from a competitor. Neither of these customers can be dealt with effectively by phone, and they are too geographically separated to visit in person on the same day. There is absolutely no possibility of getting enough of the pack to satisfy both customers under any circumstances.

Historical account records indicate that both Customer X and Customer Y buy and promote allocated special packs. But, over the last two years Customer Y has shown a greater likelihood of responding quickly and favorably to such offers. Your company's salespeople are all trained such that when allocated packs are available in limited quantities, they should pursue customers who are most likely to buy and effectively promote the product. In the present case, based upon recent history it appears the odds are greater that Company Y will better fulfill this criteria.

Either of the potential orders would be for the same dollar value, approximately $10,000. Both potential customers are moderately large companies, both are equally financially sound, and both are a good credit risk. The preparation time required by Smith to pursue each order would be the same, and of course Smith can only pursue one order or the other (not both). Any service time required after the sale would be similar for each customer.

Smith's Actual Decision and the Result.

(At this point decision appropriateness and outcome were manipulated. Please refer to the last page in this Appendix for the manipulations.)

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low ability			Average			High ability
1	2	3	4	5	6	7

2. Smith made an excellent decision under the circumstances.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

3. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

4. I view Smith to be highly competent as a decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

5. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

6. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

7. I consider Smith to be a very <u>poor</u> decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

8. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

9. Smith made the <u>wrong</u> decision.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

11. Given the circumstances, the decision made by Smith was correct.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

13. I would describe the immediate <u>outcome</u> (result) of Smith's decision as:

Very unfavorable			Somewhat			Very favorable
1	2	3	4	5	6	7

14. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SCENARIO #2

Update on Salesperson Smith.

A couple of weeks have passed since the previous situation. Smith has just encountered an opportunity to sell a closeout item to one of two large customers in Smith's territory that actively seek closeouts: "Customer A" and "Customer B." They are the only two customers in Smith's territory that have a history of buying closeouts. This sale would represent some needed additional business for Smith's territory and district.

Because the closeout represents a new item and a new UPC number for both customers, when Smith contacted the buyers by phone both buyers said the closeout would be considered only if Smith would come by today, fill out some "new item forms," and give a brief presentation. If the buyer likes it, Smith will get an immediate order. Unfortunately, the two buyers are in cities far enough apart that they cannot both be called on by this afternoon. Your company's distribution center has indicated that they must have the order no later than this afternoon or they will offer the merchandise to another sales representative---they will hold it for Smith only until the end of the day.

Therefore, Smith must target only one of the two customers and visit that customer in person this afternoon to present the closeout. While both Customer A and Customer B have a history of active closeout purchasing, historical records and Smith's visits to the stores confirm that Customer A sells closeout product out of their stores much more rapidly than Customer B. Your company's training dictates that salespeople sell closeouts to customers who will more rapidly move it through the pipeline and into consumers' hands. In the current situation, it appears the chances are greater that Customer A will better fulfill this goal of quick movement.

If either customer buys the closeout they will buy all of it, which has a value of approximately $11,000. Both customers are equal in terms of financial soundness and creditworthiness. Call preparation time and any service time after the sale would be about the same for each customer. For the reasons described above, it will be impossible for Smith to target both customers for this order. Therefore, Smith must make a decision now to pursue one or the other customer only.

Smith's Actual Decision and the Result.

(At this point decision appropriateness and outcome were manipulated. Please refer to the last page in this Appendix for the manipulations.)

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low ability			Average			High ability
1	2	3	4	5	6	7

2. Smith made an excellent decision under the circumstances.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

3. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

4. I view Smith to be highly competent as a decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

5. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

6. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

7. I consider Smith to be a very poor decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

8. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

9. Smith made the wrong decision.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

11. Given the circumstances, the decision made by Smith was correct.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

13. I would describe the immediate outcome (result) of Smith's decision as:

Very unfavorable			Somewhat			Very favorable
1	2	3	4	5	6	7

14. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SCENARIO #3

Update on Salesperson Smith.

Another couple of weeks have passed. An important part of Smith's job is pulling merchandise from chain warehouses for building off-shelf promotional displays in retail stores. Often, the best time to do this is during store resets and new store set-ups when prime off-shelf space may be more readily available. Your company standard is to participate in such store set-ups only if merchandise can be pulled from the chain's warehouse for major off-shelf promotional displays.

Two supermarket chains in particular are important in Smith's area: "Chain M" and "Chain N." Each of these chains has a new store that is about to open in Smith's territory. According to Smith's route list, both markets are due for a visit, but they are in opposite directions and cannot both be visited in one day. Last Thursday evening, the local supervisors for each chain called Smith with a request to come to these two respective new stores the next day (Friday) to help with the new store set-up. Smith can rearrange schedules to go to one place or the other on Friday, but not both places. All promotional display space will be claimed by the various sales representatives who attend the sets.

At a recent meeting of Smith's district, a discussion of the district priorities on display efforts and a review of past account records resulted in a group consensus that there is a greater probability that a Chain N store will allow a salesperson to pull significant promotional display quantities of merchandise from the chain's warehouse during store set-ups. Chain M stores were viewed as having a lower probability of such promotional activity. Thus, it was decided by the district that Chain N should be given general priority over Chain M in new store set-ups and resets.

Smith knows that whichever of these two stores are pursued, the combination of travel time and time spent in the store will be about the same for each of the alternatives. As described above, due to the short notice, the scheduling conflict, and the distance factor only one or the other new store set-up (not both) scheduled for Friday can be attended by Smith. Unfortunately, no other sales representatives from your company are available to attend the store set-up that Smith does not attend, and the store supervisors will not conduct this type of business by phone. Therefore, it is up to Smith to decide which store set-up to attend.

Smith's Actual Decision and the Result.

(At this point decision appropriateness and outcome were manipulated. Please refer to the last page in this Appendix for the manipulations.)

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low ability			Average			High ability
1	2	3	4	5	6	7

2. Smith made an excellent decision under the circumstances.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

3. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

4. I view Smith to be highly competent as a decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

5. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

6. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

7. I consider Smith to be a very poor decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

8. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

9. Smith made the wrong decision.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

11. Given the circumstances, the decision made by Smith was correct.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

13. I would describe the immediate outcome (result) of Smith's decision as:

Very unfavorable			Somewhat			Very favorable
1	2	3	4	5	6	7

14. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SUMMARY - YOUR GENERAL VIEWS OF SALESPERSON SMITH

Please write down what factors you considered in making your previous ratings of Smith.

1. What is the likelihood you would consider Smith for promotion in the future?

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

2. Describe your expectation for Smith's performance in the future.

Low performance			Average			High performance
1	2	3	4	5	6	7

3. How much like yourself do you consider Smith to be?

Not like me			Somewhat			Like me
1	2	3	4	5	6	7

4. How important do you feel circumstances of the situation were in contributing to Smith's decisions?

Not important			Somewhat			Very important
1	2	3	4	5	6	7

5. How important do you feel Smith's personal characteristics were in contributing to Smith's decisions?

Not important			Somewhat			Very important
1	2	3	4	5	6	7

6. What is the likelihood Smith will leave your company in the near future?

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

7. Describe Smith.

Independent						Team player
1	2	3	4	5	6	7

Please indicate what you, as a manager, would do with Smith now in terms of training and development.

GENERAL INFORMATION ABOUT YOU

Sex: 1. Female ____ 2. Male ____ 3. Age: ____ years old.

4. How many years of sales management experience have you had?
____ years.

5. How many years of sales experience have you had?
____ years.

6. How many years of total work experience have you had?
____ years.

7. Give the best estimate of the total number of people for whom you have completed formal performance appraisals during your career.
____ people.

8. Give the best estimate of the number of times you personally have been evaluated by superiors during a formal performance appraisal process during your career.
____ times.

9. As a manager, I am willing to risk a small loss in order to achieve a large gain.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. I find that on the job I'm a highly conservative manager.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

11. As a manager, I am willing to take stands my boss may disapprove of.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. I find that most of my decisions on the job are made from my gut.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

13. I enjoy the thrill of taking chances in my decision making on the job.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

14. Rate the degree to which the types of decisions Smith faced could happen to you or your people on the job.

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

15. Rate the extent that the problems you've read in the previous scenarios could actually happen.

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

Thank you very much for your assistance in this project!

Please fold over this stapled set of materials once from top to bottom and return it immediately in the pre-addressed, postage-paid envelope provided. If your return envelope becomes separated, the address is on the front page of this packet. You will be provided a summary of the results when they are available.

Manipulations

All manipulations took place in the last paragraph of each scenario, under the heading "Smith's Actual Decision and the Result." The following decisions and outcomes were combined based upon the condition to which each subject was assigned.

Scenario #1
Inappropriate Decision:
Smith pursued the order from Customer X, which had the lower likelihood of buying the allocated special pack.
Appropriate Decision:
Smith pursued the order from Customer Y, which had the higher likelihood of buying the allocated special pack.

Bad Outcome:
Smith did not achieve the order, and the allocated special pack is no longer available in your company's distribution center. Smith has just given you (the manager) the news.
Good Outcome:
Smith achieved the desired order for all of the allocated special pack remaining in your company's distribution center, and has just given you (the manager) the news.
No Outcome (Control):
No information given.

Scenario #2
Inappropriate Decision:
Smith traveled to Customer B (which had the lesser chance of quickly selling the closeout merchandise through their stores), and gave a presentation that afternoon.
Appropriate Decision:
Smith traveled to Customer A (which had the greater chance of quickly selling the closeout merchandise sold through their stores), and gave a presentation that afternoon.

Bad Outcome:
Smith did not get the order. By the next morning the closeout was no longer available from your company's distribution center. Smith has just relayed this result to you (the manager).
Good Outcome:
Smith got the order for the entire quantity of the closeout. Your company's distribution center shipped the order to Smith's customer the next day. Smith's customer then sold the closeout merchandise through their stores very quickly. You (the manager) are reviewing these results.
No Outcome (Control):
No information given.

Scenario #3
Inappropriate Decision:
On Friday, Smith made the trip to Chain M's new store, which had the lower probability of allowing salespeople to order major off-shelf promotional displays during new store set-ups.
Appropriate Decision:
On Friday, Smith made the trip to Chain N's new store, which had the higher probability of allowing salespeople to order major off-shelf promotional displays during new store set-ups.

Bad Outcome:
Smith participated in the new store set-up, but did not achieve any off-shelf promotional displays in the store. You (the manager) have just received this news from Smith.
Good Outcome:
Smith participated in the new store set-up, and achieved several very impressive off-shelf promotional displays in the store. This resulted in some excellent movement of Smith's products out of the chain's warehouse. You (the manager) have just received this news from Smith.

No Outcome (Control):
No information given.

APPENDIX B

COMPLETE SET OF MATERIALS MAILED IN STUDY TWO - ORDER EFFECTS BIAS SbS RESPONSE MODE

COVER LETTER FROM VICE PRESIDENT OF SALES
FOR EACH OF THE TWO SALES DIVISIONS
PARTICIPATING IN THE STUDY

(SENT ON COMPANY LETTERHEAD)

May 3, 1993

Enclosed please find a set of materials related to a national research project in which the *(insert division name)* is participating. This project is related to issues of evaluating field sales personnel. A research team from the College of Business Administration at Oklahoma State University is conducting the study. *(Insert division name)* has pledged full participation in the study by all managers.

May I ask that you follow the attached instructions exactly, fill out the forms completely, and return the materials to the researchers in the postage paid envelope provided. They need your responses back by **5/17/93**, therefore please complete all the materials and return the packet to them ASAP.

Let me assure you that no member of *(insert company name)* will ever see your individual responses. The research team will provide us with results only in aggregate (summarized) form---no names or territory numbers will be matched to your responses. You will receive a copy of the summarized results when they are available.

I have been asked to emphasize that there are no right or wrong answers to this exercise---only your views and opinions. Therefore, there is no reason to try to second guess the information or questions. The most important thing you can do is give the exercise your thoughtful consideration, follow the instructions exactly, fill out the questionnaires completely, and return the packet to the research team promptly.

Be sure to meet their due date of **5/17/93**. Thanks in advance for your efforts---our company will benefit from the research.

(Signature of appropriate vice president of sales)

TASK DESCRIPTION (INSTRUCTIONS)
(Please Review Carefully Before Proceeding)

On the following pages you will first read some initial information in scenario form about a hypothetical salesperson. Then, two subsequent scenarios will describe the salesperson engaged in several job-related activities over two different sales situations. These two sales situations span a time period of several weeks. After carefully reading each scenario, your task is to assume the role of the employee's sales manager (i.e., the salesperson's immediate supervisor) and respond to a few questions that follow each scenario.

For the results of this study to be meaningful, the instructions provided throughout these materials must be followed exactly. Note that there are no "right" or "wrong" answers to any of the questions—the questions ask only for your own views and opinions from the perspective of the role you are playing as the employee's sales manager.

Because your company will be provided results of this study only in aggregate form (overall summary numbers), your individual responses will never be seen by anyone in your organization. Thus, it is not necessary for you to identify yourself on these materials. The task you are about to perform has been designed to be interesting and fun, and the results have the potential to improve salesforce performance evaluation in all organizations.

A COUPLE OF NOTES: 1) It is very important that you completely answer the questions that follow each scenario before going on to the next scenario---please do not skip ahead. It is, however, permissible to refer back to each scenario individually as you answer the questions following it. 2) Please limit your evaluations to consideration only of the information provided in the scenarios. Please do not try to second-guess the information or find solutions not available via the facts provided.

INSTRUCTIONS FOR RETURNING COMPLETED MATERIALS: When you have completed the project please fold over this stapled set of materials once from top to bottom and return it immediately in the pre-addressed, postage-paid envelope provided. The research team needs your packet back no later than Monday, May 17 so we may begin tabulating the results and provide a report back to your organization in a timely manner.

Thanks for your participation!

Professor John C. Mowen
College of Business, Oklahoma State University
University Center at Tulsa
700 N. Greenwood Avenue
Tulsa, OK. 74106-0700
Phone: (405) 744-5112

PS: As a token of our appreciation for your participation in and prompt response to this research project, enclosed please find a small gift from Oklahoma State University.

GENERAL BACKGROUND INFORMATION ON SALESPERSON SMITH

Salesperson "Smith" is employed by your company as a territory representative. A large part of the job involves selling products to and servicing existing accounts, but effort is also put into developing new business when the opportunity arises. Your company trains and encourages its salespeople to make as many of their own decisions as possible regarding account strategies and territory management.

Smith was previously employed in a similar position with another company for 2 years, but has only worked for you and your company for 3 months. In terms of geographic size, number of customers, and number of competing salespeople from other companies, Smith's territory is comparable to most other territories within your company. Smith's time spent in call preparation is also about the same as that of most salespeople in your company and within the industry. Smith's previous employer provided a favorable recommendation during the employment interview process, and so far Smith seems to be generally well-liked by your company's management, other salespeople, and customers.

So far, Smith has exhibited an acceptable level of skill and ability in fulfilling the requirements of the job. Smith seems to be putting forth a reasonable amount of effort toward selling and toward the various other company objectives. You have no particular complaints with Smith at this point, and the general consensus is that Smith's performance over the past 3 months could best be described as "Meets Expectations." You are, however, anxious for Smith to have the opportunity to face some challenging situations with customers so you will have additional evidence of just how effective Smith really is as a territory representative.

Over the last few years the overall economic climate in Smith's area has been slightly better than the economy of the U.S. as a whole, and that trend is forecasted to continue for the foreseeable future.

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low Ability			Average			High Ability
1	2	3	4	5	6	7

2. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

3. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

4. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

5. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

6. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

7. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

8. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

Now, please go on to the next page to read the first scenario about a sales situation faced by Salesperson Smith.

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SCENARIO #1

Update on Salesperson Smith.

Salesperson Smith has just encountered an opportunity to sell a closeout item to one of two large customers in Smith's territory that actively seek closeouts: "Customer A" and "Customer B." They are the only two customers in Smith's territory that have a history of buying closeouts. This sale would represent some needed additional business for Smith's territory and district.

Because the closeout represents a new item and a new UPC number for both customers, when Smith contacted the buyers by phone both buyers said the closeout would be considered only if Smith would come by today, fill out some "new item forms," and give a brief presentation. If the buyer likes it, Smith will get an immediate order. Unfortunately, the two buyers are in cities far enough apart that they cannot both be called on by this afternoon. Your company's distribution center has indicated that they must have the order no later than this afternoon or they will offer the merchandise to another sales representative—they will hold it for Smith only until the end of the day.

Therefore, Smith must target only one of the two customers and visit that customer in person this afternoon to present the closeout. While both Customer A and Customer B have a history of active closeout purchasing, historical records and Smith's visits to the stores confirm that Customer A sells closeout product out of their stores much more rapidly than Customer B. Your company's training dictates that salespeople sell closeouts to customers who will more rapidly move it through the pipeline and into consumers' hands. In the current situation, it appears the chances are greater that Customer A will better fulfill this goal of quick movement.

If either customer buys the closeout they will buy all of it, which has a value of approximately $11,000. Both customers are equal in terms of financial soundness and creditworthiness. Call preparation time and any service time after the sale would be about the same for each customer. For the reasons described above, it will be impossible for Smith to target both customers for this order. Therefore, Smith must make a decision now to pursue one or the other customer only.

Smith's Actual Decision and the Result.

(At this point the valence of the information was manipulated [either bad or good]. Please refer to the last page in this Appendix for the manipulations.)

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low ability			Average			High ability
1	2	3	4	5	6	7

2. Smith made an excellent decision under the circumstances.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

3. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

4. I view Smith to be highly competent as a decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

5. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

6. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

7. I consider Smith to be a very poor decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

8. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

9. Smith made the wrong decision.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

11. Given the circumstances, the decision made by Smith was correct.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

13. I would describe the immediate outcome (result) of Smith's decision as:

Very unfavorable			Somewhat			Very favorable
1	2	3	4	5	6	7

14. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SCENARIO #2

Update on Salesperson Smith.

A couple of weeks have passed since the previous situation. Smith recently faced a decision about which of two potentially large promotional orders for an allocated special pack item to pursue from two chain customers whose headquarters buying offices are in Smith's territory. The potential order would come from accounts referred to here as "Customer X" and "Customer Y." The headquarters buyer for each chain is a regular customer of Smith, and each has the authority to quickly make such buying decisions. The allocated pack is about to sell out in your company's distribution center, and there is only a sufficient quantity available for one of Smith's two potential customers to run a satisfactory promotion. Neither customer will buy the item for a promotion without the special pack.

Smith must direct selling efforts toward gaining only one of these orders because both customers would need to be pursued immediately and in person or the allocated pack will run out in your company's distribution center and Smith's prospective buyer will purchase an alternative item from a competitor. Neither of these customers can be dealt with effectively by phone, and they are too geographically separated to visit in person on the same day. There is absolutely no possibility of getting enough of the pack to satisfy both customers under any circumstances.

Historical account records indicate that both Customer X and Customer Y buy and promote allocated special packs. But, over the last two years Customer Y has shown a greater likelihood of responding quickly and favorably to such offers. Your company's salespeople are all trained such that when allocated packs are available in limited quantities, they should pursue customers who are most likely to buy and effectively promote the product. In the present case, based upon recent history it appears the odds are greater that Company Y will better fulfill this criteria.

Either of the potential orders would be for the same dollar value, approximately $10,000. Both potential customers are moderately large companies, both are equally financially sound, and both are a good credit risk. The preparation time required by Smith to pursue each order would be the same, and of course Smith can only pursue one order or the other (not both). Any service time required after the sale would be similar for each customer.

Smith's Actual Decision and the Result.

(At this point the valence of the information was manipulated [either bad or good]. Please refer to the last page in this Appendix for the manipulations.)

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low ability			Average			High ability
1	2	3	4	5	6	7

2. Smith made an excellent decision under the circumstances.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

3. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

4. I view Smith to be highly competent as a decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

5. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

6. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

7. I consider Smith to be a very poor decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

8. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

9. Smith made the wrong decision.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

11. Given the circumstances, the decision made by Smith was correct.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

13. I would describe the immediate outcome (result) of Smith's decision as:

Very unfavorable			Somewhat			Very favorable
1	2	3	4	5	6	7

14. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SUMMARY - YOUR GENERAL VIEWS OF SALESPERSON SMITH

Please write down what factors you considered in making your previous ratings of Smith.

1. What is the likelihood you would consider Smith for promotion in the future?

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

2. Describe your expectation for Smith's performance in the future.

Low performance			Average			High performance
1	2	3	4	5	6	7

3. How much like yourself do you consider Smith to be?

Not like me			Somewhat			Like me
1	2	3	4	5	6	7

4. How important do you feel circumstances of the situation were in contributing to Smith's decisions?

Not important			Somewhat			Very important
1	2	3	4	5	6	7

5. How important do you feel Smith's personal characteristics were in contributing to Smith's decisions?

Not important			Somewhat			Very important
1	2	3	4	5	6	7

6. What is the likelihood Smith will leave your company in the near future?

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

7. Describe Smith.

Independent						Team player
1	2	3	4	5	6	7

Please indicate what you, as a manager, would do with Smith now in terms of training and development.

GENERAL INFORMATION ABOUT YOU

Sex: 1. Female ____ 2. Male ____ 3. Age: ____ years old.

4. How many years of sales management experience have you had?
____ years.

5. How many years of sales experience have you had?
____ years.

6. How many years of total work experience have you had?
____ years.

7. Give the best estimate of the total number of people for whom you have completed formal performance appraisals during your career.
____ people.

8. Give the best estimate of the number of times you personally have been evaluated by superiors during a formal performance appraisal process during your career.
____ times.

9. As a manager, I am willing to risk a small loss in order to achieve a large gain.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. I find that on the job I'm a highly conservative manager.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

11. As a manager, I am willing to take stands my boss may disapprove of.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. I find that most of my decisions on the job are made from my gut.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

13. I enjoy the thrill of taking chances in my decision making on the job.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

14. Rate the degree to which the types of decisions Smith faced could happen to you or your people on the job.

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

15. Rate the extent that the problems you've read in the previous scenarios could actually happen.

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

Thank you very much for your assistance in this project!

Please fold over this stapled set of materials once from top to bottom and return it immediately in the pre-addressed, postage-paid envelope provided. If your return envelope becomes separated, the address is on the front page of this packet. You will be provided a summary of the results when they are available.

Manipulations

All manipulations took place in the last paragraph of each scenario, under the heading "Smith's Actual Decision and the Result." Decisions and outcomes were combined as shown below to create the negative evidence and positive evidence conditions.

Scenario #1

Negative Evidence:
Smith pursued the order from Customer X, which had the lower likelihood of buying the allocated special pack. Smith did not achieve the order, and the allocated special pack is no longer available in your company's distribution center. Smith has just given you (the manager) the news.

Positive Evidence:
Smith pursued the order from Customer Y, which had the higher likelihood of buying the allocated special pack. Smith achieved the desired order for all of the allocated special pack remaining in your company's distribution center, and has just given you (the manager) the news.

Scenario #2

Negative Evidence:
Smith traveled to Customer B (which had the lesser chance of quickly selling the closeout merchandise through their stores), and gave a presentation that afternoon. Smith did not get the order. By the next morning the closeout was no longer available from your company's distribution center. Smith has just relayed this result to you (the manager).

Positive Evidence:
Smith traveled to Customer A (which had the greater chance of quickly selling the closeout merchandise sold through their stores), and gave a presentation that afternoon. Smith got the order for the entire quantity of the closeout. Your company's distribution center shipped the order to Smith's customer the next day. Smith's customer then sold the closeout merchandise through their stores very quickly. You (the manager) are reviewing these results.

Scenario #3

Negative Evidence:
On Friday, Smith made the trip to Chain M's new store, which had the lower probability of allowing salespeople to order major off-shelf promotional displays during new store set-ups. Smith participated in the new store set-up, but did not achieve any off-shelf promotional displays in the store. You (the manager) have just received this news from Smith.

Positive Evidence:
On Friday, Smith made the trip to Chain N's new store, which had the higher probability of allowing salespeople to order major off-shelf promotional displays during new store set-ups. Smith participated in the new store set-up, and achieved several very impressive off-shelf promotional displays in the store. This resulted in some excellent movement of Smith's products out of the chain's warehouse. You (the manager) have just received this news from Smith.

APPENDIX C

COMPLETE SET OF MATERIALS MAILED IN STUDY TWO - ORDER EFFECTS BIAS EoS RESPONSE MODE

COVER LETTER FROM VICE PRESIDENT OF SALES
FOR EACH OF THE TWO SALES DIVISIONS
PARTICIPATING IN THE STUDY

(SENT ON COMPANY LETTERHEAD)

May 3, 1993

Enclosed please find a set of materials related to a national research project in which the *(insert division name)* is participating. This project is related to issues of evaluating field sales personnel. A research team from the College of Business Administration at Oklahoma State University is conducting the study. *(Insert division name)* has pledged full participation in the study by all managers.

May I ask that you follow the attached instructions exactly, fill out the forms completely, and return the materials to the researchers in the postage paid envelope provided. They need your responses back by **5/17/93**, therefore please complete all the materials and return the packet to them ASAP.

Let me assure you that no member of *(insert company name)* will ever see your individual responses. The research team will provide us with results only in aggregate (summarized) form---no names or territory numbers will be matched to your responses. You will receive a copy of the summarized results when they are available.

I have been asked to emphasize that there are no right or wrong answers to this exercise---only your views and opinions. Therefore, there is no reason to try to second guess the information or questions. The most important thing you can do is give the exercise your thoughtful consideration, follow the instructions exactly, fill out the questionnaires completely, and return the packet to the research team promptly.

Be sure to meet their due date of **5/17/93**. Thanks in advance for your efforts---our company will benefit from the research.

(Signature of appropriate vice president of sales)

TASK DESCRIPTION (INSTRUCTIONS)
(PLEASE REVIEW CAREFULLY BEFORE PROCEEDING)

On the following pages you will first read some initial information in scenario form about a hypothetical salesperson. Then, two subsequent scenarios will describe the salesperson engaged in several job-related activities over two different sales situations. These two sales situations span a time period of several weeks. After carefully reading each scenario, your task is to assume the role of the employee's sales manager (i.e., the salesperson's immediate supervisor) and respond to a few questions that follow the initial scenario and the two subsequent scenarios.

For the results of this study to be meaningful, the instructions provided throughout these materials must be followed exactly. Note that there are no "right" or "wrong" answers to any of the questions—the questions ask only for your own views and opinions from the perspective of the role you are playing as the employee's sales manager.

Because your company will be provided results of this study only in aggregate form (overall summary numbers), your individual responses will never be seen by anyone in your organization. Thus, it is not necessary for you to identify yourself on these materials. The task you are about to perform has been designed to be interesting and fun, and the results have the potential to improve salesforce performance evaluation in all organizations.

A COUPLE OF NOTES: 1) It is very important that you completely answer the questions that follow initial scenario before going on to the next scenarios—please do not skip ahead. It is, however, permissible to refer back to each scenario individually as you answer the questions following it. 2) Please limit your evaluations to consideration only of the information provided in the scenarios. Please do not try to second-guess the information or find solutions not available via the facts provided.

INSTRUCTIONS FOR RETURNING COMPLETED MATERIALS: When you have completed the project please fold over this stapled set of materials once from top to bottom and return it immediately in the pre-addressed, postage-paid envelope provided. The research team needs your packet back no later than **Monday, May 17** so we may begin tabulating the results and provide a report back to your organization in a timely manner.

Thanks for your participation!

Professor John C. Mowen
College of Business, Oklahoma State University
University Center at Tulsa
700 N. Greenwood Avenue
Tulsa, OK. 74106-0700
Phone: (405) 744-5112

PS: As a token of our appreciation for your participation in and prompt response to this research project, enclosed please find a small gift from Oklahoma State University.

GENERAL BACKGROUND INFORMATION ON SALESPERSON SMITH

Salesperson "Smith" is employed by your company as a territory representative. A large part of the job involves selling products to and servicing existing accounts, but effort is also put into developing new business when the opportunity arises. Your company trains and encourages its salespeople to make as many of their own decisions as possible regarding account strategies and territory management.

Smith was previously employed in a similar position with another company for 2 years, but has only worked for you and your company for 3 months. In terms of geographic size, number of customers, and number of competing salespeople from other companies, Smith's territory is comparable to most other territories within your company. Smith's time spent in call preparation is also about the same as that of most salespeople in your company and within the industry. Smith's previous employer provided a favorable recommendation during the employment interview process, and so far Smith seems to be generally well-liked by your company's management, other salespeople, and customers.

So far, Smith has exhibited an acceptable level of skill and ability in fulfilling the requirements of the job. Smith seems to be putting forth a reasonable amount of effort toward selling and toward the various other company objectives. You have no particular complaints with Smith at this point, and the general consensus is that Smith's performance over the past 3 months could best be described as "Meets Expectations." You are, however, anxious for Smith to have the opportunity to face some challenging situations with customers so you will have additional evidence of just how effective Smith really is as a territory representative.

Over the last few years the overall economic climate in Smith's area has been slightly better than the economy of the U.S. as a whole, and that trend is forecasted to continue for the foreseeable future.

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low Ability			Average			High Ability
1	2	3	4	5	6	7

2. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

3. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

4. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

5. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

6. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

7. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

8. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

Now, please go on to the next page to read two scenarios about sales situations faced by Salesperson Smith.

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SCENARIO #1

Update on Salesperson Smith.

Salesperson Smith has just encountered an opportunity to sell a closeout item to one of two large customers in Smith's territory that actively seek closeouts: "Customer A" and "Customer B." They are the only two customers in Smith's territory that have a history of buying closeouts. This sale would represent some needed additional business for Smith's territory and district.

Because the closeout represents a new item and a new UPC number for both customers, when Smith contacted the buyers by phone both buyers said the closeout would be considered only if Smith would come by today, fill out some "new item forms," and give a brief presentation. If the buyer likes it, Smith will get an immediate order. Unfortunately, the two buyers are in cities far enough apart that they cannot both be called on by this afternoon. Your company's distribution center has indicated that they must have the order no later than this afternoon or they will offer the merchandise to another sales representative—they will hold it for Smith only until the end of the day.

Therefore, Smith must target only one of the two customers and visit that customer in person this afternoon to present the closeout. While both Customer A and Customer B have a history of active closeout purchasing, historical records and Smith's visits to the stores confirm that Customer A sells closeout product out of their stores much more rapidly than Customer B. Your company's training dictates that salespeople sell closeouts to customers who will more rapidly move it through the pipeline and into consumers' hands. In the current situation, it appears the chances are greater that Customer A will better fulfill this goal of quick movement.

If either customer buys the closeout they will buy all of it, which has a value of approximately $11,000. Both customers are equal in terms of financial soundness and creditworthiness. Call preparation time and any service time after the sale would be about the same for each customer. For the reasons described above, it will be impossible for Smith to target both customers for this order. Therefore, Smith must make a decision now to pursue one or the other customer only.

Smith's Actual Decision and the Result.

(At this point the valence of the information was manipulated [either bad or good]. Please refer to the last page in this Appendix for the manipulations.)

SCENARIO #2

Update on Salesperson Smith.

A couple of weeks have passed since the previous situation. Smith recently faced a decision about which of two potentially large promotional orders for an allocated special pack item to pursue from two chain customers whose headquarters buying offices are in Smith's territory. The potential order would come from accounts referred to here as "Customer X" and "Customer Y." The headquarters buyer for each chain is a regular customer of Smith, and each has the authority to quickly make such buying decisions. The allocated pack is about to sell out in your company's distribution center, and there is only a sufficient quantity available for one of Smith's two potential customers to run a satisfactory promotion. Neither customer will buy the item for a promotion without the special pack.

Smith must direct selling efforts toward gaining only one of these orders because both customers would need to be pursued immediately and in person or the allocated pack will run out in your company's distribution center and Smith's prospective buyer will purchase an alternative item from a competitor. Neither of these customers can be dealt with effectively by phone, and they are too geographically separated to visit in person on the same day. There is absolutely no possibility of getting enough of the pack to satisfy both customers under any circumstances.

Historical account records indicate that both Customer X and Customer Y buy and promote allocated special packs. But, over the last two years Customer Y has shown a greater likelihood of responding quickly and favorably to such offers. Your company's salespeople are all trained such that when allocated packs are available in limited quantities, they should pursue customers who are most likely to buy and effectively promote the product. In the present case, based upon recent history it appears the odds are greater that Company Y will better fulfill this criteria.

Either of the potential orders would be for the same dollar value, approximately $10,000. Both potential customers are moderately large companies, both are equally financially sound, and both are a good credit risk. The preparation time required by Smith to pursue each order would be the same, and of course Smith can only pursue one order or the other (not both). Any service time required after the sale would be similar for each customer.

Smith's Actual Decision and the Result.

(At this point the valence of the information was manipulated [either bad or good]. Please refer to the last page in this Appendix for the manipulations.)

RESPONSE SHEET

Directions: Complete the following questions by circling the number that best represents your feelings and thoughts. *CAUTION: Do not go on to the next section until all these are complete.*

1. Please rate Smith's sales ability.

Low ability			Average			High ability
1	2	3	4	5	6	7

2. Smith made an excellent decision under the circumstances.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

3. How would you rate Smith's overall level of effort in obtaining sales?

Low effort			Average			High effort
1	2	3	4	5	6	7

4. I view Smith to be highly competent as a decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

5. Rate Smith's overall job performance.

Low performance			Average			High performance
1	2	3	4	5	6	7

6. How would you categorize Smith?

Not likeable			Somewhat likeable			Very likeable
1	2	3	4	5	6	7

7. I consider Smith to be a very poor decision maker.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

8. Do you think Smith deserves a promotion or bonus?

Not deserving			Somewhat			Very deserving
1	2	3	4	5	6	7

9. Smith made the wrong decision.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. How would you describe Smith as a person?

Bad			Average			Good
1	2	3	4	5	6	7

11. Given the circumstances, the decision made by Smith was correct.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. Rate Smith's skill level as a salesperson.

Very low			Average			Very high
1	2	3	4	5	6	7

13. I would describe the immediate outcome (result) of Smith's decision as:

Very unfavorable			Somewhat			Very favorable
1	2	3	4	5	6	7

14. On a scale of 0 - 100, give an overall rating of Smith, where 0 = Unacceptable and 100 = Far Exceeds Expectations. _____

⊗ STOP! VERY IMPORTANT! Please do not go on to this next section until you have completed all the questions on the previous page.

SUMMARY - YOUR GENERAL VIEWS OF SALESPERSON SMITH

Please write down what factors you considered in making your previous ratings of Smith.

1. What is the likelihood you would consider Smith for promotion in the future?

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

2. Describe your expectation for Smith's performance in the future.

Low performance			Average			High performance
1	2	3	4	5	6	7

3. How much like yourself do you consider Smith to be?

Not like me			Somewhat			Like me
1	2	3	4	5	6	7

4. How important do you feel circumstances of the situation were in contributing to Smith's decisions?

Not important			Somewhat			Very important
1	2	3	4	5	6	7

5. How important do you feel Smith's personal characteristics were in contributing to Smith's decisions?

Not important			Somewhat			Very important
1	2	3	4	5	6	7

6. What is the likelihood Smith will leave your company in the near future?

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

7. Describe Smith.

Independent						Team player
1	2	3	4	5	6	7

Please indicate what you, as a manager, would do with Smith now in terms of training and development.

GENERAL INFORMATION ABOUT YOU

Sex: 1. Female ____ 2. Male ____ 3. Age: ____ years old.

4. How many years of sales management experience have you had?
____ years.

5. How many years of sales experience have you had?
____ years.

6. How many years of total work experience have you had?
____ years.

7. Give the best estimate of the total number of people for whom you have completed formal performance appraisals during your career.
____ people.

8. Give the best estimate of the number of times you personally have been evaluated by superiors during a formal performance appraisal process during your career.
____ times.

9. As a manager, I am willing to risk a small loss in order to achieve a large gain.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

10. I find that on the job I'm a highly conservative manager.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

11. As a manager, I am willing to take stands my boss may disapprove of.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

12. I find that most of my decisions on the job are made from my gut.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

13. I enjoy the thrill of taking chances in my decision making on the job.

Strongly disagree			Somewhat			Strongly agree
1	2	3	4	5	6	7

14. Rate the degree to which the types of decisions Smith faced could happen to you or your people on the job.

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

15. Rate the extent that the problems you've read in the previous scenarios could actually happen.

Very unlikely			Somewhat			Very likely
1	2	3	4	5	6	7

Thank you very much for your assistance in this project!

Please fold over this stapled set of materials once from top to bottom and return it immediately in the pre-addressed, postage-paid envelope provided. If your return envelope becomes separated, the address is on the front page of this packet. You will be provided a summary of the results when they are available.

Manipulations

All manipulations took place in the last paragraph of each scenario, under the heading "Smith's Actual Decision and the Result." Decisions and outcomes were combined as shown below to create the negative evidence and positive evidence conditions.

Scenario #1

Negative Evidence:
Smith pursued the order from Customer X, which had the lower likelihood of buying the allocated special pack. Smith did not achieve the order, and the allocated special pack is no longer available in your company's distribution center. Smith has just given you (the manager) the news.

Positive Evidence:
Smith pursued the order from Customer Y, which had the higher likelihood of buying the allocated special pack. Smith achieved the desired order for all of the allocated special pack remaining in your company's distribution center, and has just given you (the manager) the news.

Scenario #2

Negative Evidence:
Smith traveled to Customer B (which had the lesser chance of quickly selling the closeout merchandise through their stores), and gave a presentation that afternoon. Smith did not get the order. By the next morning the closeout was no longer available from your company's distribution center. Smith has just relayed this result to you (the manager).

Positive Evidence:
Smith traveled to Customer A (which had the greater chance of quickly selling the closeout merchandise sold through their stores), and gave a presentation that afternoon. Smith got the order for the entire quantity of the closeout. Your company's distribution center shipped the order to Smith's customer the next day. Smith's customer then sold the closeout merchandise through their stores very quickly. You (the manager) are reviewing these results.

Scenario #3

Negative Evidence:
On Friday, Smith made the trip to Chain M's new store, which had the lower probability of allowing salespeople to order major off-shelf promotional displays during new store set-ups. Smith participated in the new store set-up, but did not achieve any off-shelf promotional displays in the store. You (the manager) have just received this news from Smith.

Positive Evidence:
On Friday, Smith made the trip to Chain N's new store, which had the higher probability of allowing salespeople to order major off-shelf promotional displays during new store set-ups. Smith participated in the new store set-up, and achieved several very impressive off-shelf promotional displays in the store. This resulted in some excellent movement of Smith's products out of the chain's warehouse. You (the manager) have just received this news from Smith.

APPENDIX D

DESCRIPTIVE STATISTICS FOR STUDY ONE AND STUDY TWO

DESCRIPTIVE STATISTICS FOR STUDY ONE - RATING ONE

	DECQUAL	PEREVAL	GLOBAL	OCMANIP	INFOPROC	EXATTRIB	INATTRIB	LIKEME	RISKTKNG	ECOLVAL
DECQUAL	4.25 1.71									
PEREVAL	.55 .0001 139	3.94 0.90								
GLOBAL	.59 .0001 142	.78 .0001 138	59.17 22.49							
OCMANIP	.31 .0019 95	.65 .0001 94	.62 .0001 95	3.61 1.97						
INFOPROC	-.02 .79 143	-.08 .37 139	-.01 .90 142	-.05 .62 95	2.08 1.41					
EXATTRIB	.41 .0001 142	.35 .0001 138	.26 .002 141	.09 .40 94	.05 .59 142	5.01 1.86				
INATTRIB	-.003 .97 142	.11 .20 138	.20 .02 141	.18 .07 94	.26 .002 142	-.13 .12 142	4.12 1.83			
LIKEME	.52 .0001 140	.57 .0001 137	.53 .0001 139	.44 .0001 92	.02 .77 140	.36 .0001 140	.04 .61 140	2.79 1.75		
RISKTKNG	.12 .14 141	-.16 .06 137	-.07 .38 140	.02 .87 93	-.15 .08 141	.03 .69 141	-.12 .17 141	-.03 .71 139	4.53 1.41	
ECOLVAL	.16 .06 141	.29 .0007 137	.17 .04 140	.20 .05 93	.03 .75 141	.27 .001 141	.02 .83 141	.31 .0002 139	.02 .82 141	4.82 1.80

Below the diagonal: Pearson correlations, p-values, and number of subjects in the cell.

On the diagonal: Means and standard deviations.

DESCRIPTIVE STATISTICS FOR STUDY ONE - RATING TWO

	DECQUAL	PEREVAL	GLOBAL	OCMANIP	INFOPROC	EXATTRIB	INATTRIB	LIKEME	RISKTKNG	ECOLVAL
DECQUAL	4.07									
	1.80									
PEREVAL	.64	3.79								
	.0001	1.13								
	139									
GLOBAL	.64	.83	55.14							
	.0001	.0001	24.69							
	138	134								
OCMANIP	.50	.80	.74	3.51						
	.0001	.0001	.0001	2.01						
	94	93	90							
INFOPROC	-.02	-.12	-.06	-.07	2.08					
	.78	.16	.50	.49	1.41					
	143	139	138	94						
EXATTRIB	.44	.33	.30	.16	.05	5.01				
	.0001	.0001	.0004	.13	.59	1.86				
	142	138	137	93	142					
INATTRIB	-.02	.09	.15	.14	.26	-.13	4.12			
	.84	.30	.09	.19	.002	.12	1.83			
	142	138	137	93	142	142				
LIKEME	.60	.65	.60	.59	.02	.36	.04	2.79		
	.0001	.0001	.0001	.0001	.77	.0001	.61	1.75		
	140	137	135	91	140	140	140			
RISKTKNG	.17	-.04	-.06	-.009	-.15	.03	-.12	-.03	4.53	
	.04	.63	.52	.93	.08	.69	.17	.71	1.41	
	141	137	136	92	141	141	141	139		
ECOLVAL	.15	.24	.17	.24	.03	.27	.02	.31	.02	4.82
	.08	.004	.04	.02	.75	.001	.83	.0002	.82	1.80
	141	137	136	92	141	141	141	139	141	

Below the diagonal: Pearson correlations, p-values, and number of subjects in the cell.

On the diagonal: Means and standard deviations.

DESCRIPTIVE STATISTICS FOR STUDY ONE - RATING THREE

	DECQUAL	PEREVAL	GLOBAL	OCMANIP	INFOPROC	EXATTRIB	INATTRIB	LIKEME	RISKTKNG	ECOLVAL
DECQUAL	3.91 1.88									
PEREVAL	.59 .0001 138	3.66 1.29								
GLOBAL	.63 .0001 137	.85 .0001 132	52.51 26.28							
OCMANIP	.43 .0001 95	.84 .0001 93	.77 .0001 90	3.37 2.07						
INFOPROC	-.11 .18 143	-.16 .06 138	-.07 .40 137	-.07 .48 95	2.08 1.41					
EXATTRIB	.47 .0001 142	.35 .0001 137	.28 .0008 136	.14 .19 94	.05 .59 142	5.01 1.86				
INATTRIB	-.08 .33 142	.11 .21 137	.12 .16 136	.13 .22 94	.26 .002 142	-.13 .12 142	4.12 1.83			
LIKEME	.57 .0001 140	.67 .0001 136	.62 .0001 134	.62 .0001 92	.02 .77 140	.36 .0001 140	.04 .61 140	2.79 1.75		
RISKTKNG	.18 .03 141	.02 .79 136	.01 .87 135	.01 .90 93	-.15 .08 141	.03 .69 141	-.12 .17 141	-.03 .71 139	4.53 1.41	
ECOLVAL	.14 .10 141	.23 .008 136	.14 .11 135	.18 .09 93	.03 .75 141	.27 .001 141	.02 .83 141	.31 .0002 139	.02 .82 141	4.82 1.80

Below the diagonal: Pearson correlations, p-values, and number of subjects in the cell.

On the diagonal: Means and standard deviations.

DESCRIPTIVE STATISTICS FOR STUDY TWO

	PEREVAL	DIFF	DECQUAL	GLOBAL	LIKEME	RISKTKNG	ECOLVAL
PEREVAL	3.75						
	0.81						
DIFF	.78	0.065					
	.0001	0.79					
	86						
DECQUAL	.57	.53	3.94				
	.0001	.0001	1.36				
	86	83					
GLOBAL	.71	.49	.39	55.71			
	.0001	.0001	.0003	14.83			
	85	82	81				
LIKEME	.52	.26	.20	.47	2.75		
	.0001	.02	.07	.0001	1.25		
	143	138	137	95			
RISKTKNG	.03	.00	-.14	.02	.03	4.29	
	.76	.99	.19	.87	.76	1.46	
	89	85	85	84	88		
ECOLVAL	-.06	-.11	-.03	.06	.06	.08	5.30
	.58	.31	.76	.57	.58	.47	1.57
	89	85	85	84	88	89	

Below the diagonal: Pearson correlations, p-values, and number of subjects in the cell.

On the diagonal: Means and standard deviations.

NOTE: The above statistics for PEREVAL, DECQUAL, and GLOBAL are for the last rating only, and include subjects in both the SbS and EoS response modes.

VITA

Gregory W. Marshall

Candidate for the Degree of

Doctor of Philosophy

Thesis: ASSESSING THE IMPACT OF AN OUTCOME BIAS AND AN ORDER EFFECTS BIAS ON PERFORMANCE EVALUATIONS OF FIELD SALES PERSONNEL BY SALES MANAGERS

Major Field: Business Administration

Biographical:

Personal Data: Born in Stillwater, Oklahoma, September 19, 1955, the son of Raymond B. and Maxine M. Marshall. Married August 5, 1989 to Patricia A. Marshall. Became the father of Justin S. Marshall on August 4, 1991.

Education: Graduated from Will Rogers High School, Tulsa, Oklahoma in May, 1973; received Bachelor of Science in Business Administration degree from the University of Tulsa, Tulsa, Oklahoma in May 1978; received Master of Business Administration degree from the University of Tulsa, Tulsa, Oklahoma in July 1983; completed the requirements for the Doctor of Philosophy degree at Oklahoma State University in December, 1993.

Professional Experience: Assistant Professor of Marketing, University of South Florida, Tampa, Florida, August 1993 to present; Assistant Professor of Marketing, Northeastern State University/University Center at Tulsa, Tulsa, Oklahoma, August, 1992 to May, 1993; Instructor of Marketing, Northeastern State University/ University Center at Tulsa, Tulsa, Oklahoma, August, 1989 to August, 1992; Instructor of Management, Tulsa Junior College, Tulsa, Oklahoma, January, 1986 to August, 1989; District Manager, Warner Lambert Company, Tulsa, Oklahoma, August, 1983 to December, 1985; Territory Representative, Warner Lambert Company, Tulsa, Oklahoma, August, 1981 to August, 1983; various management,

merchandising, and sales positions, Target Stores, Inc., Tulsa, Oklahoma, November, 1979 to August, 1981 and April, 1972 to July, 1976; Territory Representative, Senior Territory Representative, Assistant Product Manager, The Mennen Company, Tulsa, Oklahoma, Dallas, Texas, and Morristown, New Jersey, January, 1978 to June, 1979; Assistant Store Manager, May's Drug Stores, Inc./Oertle's Drug Department, July, 1976 to November, 1977.