MANAGERIAL MERIT PAY ALLOCATION: AN ANALYSIS OF ALTERNATIVE JUDGMENT MODELS USING REGRESSION AND FUZZY EXPERT SYSTEM TECHNIQUES

by

DAVID W. DORSEY

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychology
University of South Florida

April 1997

Major Professor: Michael D. Coovert, Ph.D.

UMI Number: 9724003

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Graduate School
University of South Florida
Tampa, Florida

CERTIFICATE OF APPROVAL

Ph.D. Dissertation

This is to certify that the Ph.D. Dissertation of

DAVID W. DORSEY

with a major in Psychology has been approved by
the Examining Committee on March 24, 1997
as satisfactory for the dissertation requirement
for the Doctor of Philosophy degree

Examining Committee:

Michael D. Coovert

Major Professor: Michael D. Coovert, Ph.D.

Walter C. Borman

Member: Walter C. Borman, Ph.D.

Lawrence O. Hall

Member: Lawrence O. Hall, Ph.D.

Edward L. Levine

Member: Edward L. Levine, Ph.D.

Sandra L. Schneider-Wright

Member: Sandra L. Schneider-Wright, Ph.D.

DEDICATION

This work is dedicated to my family and to the person who has been a constant source of support and friendship; very special thanks to my best friend Diana.

ACKNOWLEDGMENTS

I would like to acknowledge the great contributions that Dr. Michael Coovert has made in my professional growth. As my major professor, mentor, and friend, Mike has played a critical role in my intellectual and professional development. My research and work with Mike, along with our shared interests in nontraditional I/O research, has been a great and wonderful adventure – one that I hope we will continue to share.

Thanks also to my distinguished Ph.D. dissertation committee. Each of my committee members, including: Wally Borman, Ed Levine, Larry Hall, Sandy Schneider-Wright, and Bill Lee definitely enhanced the quality of this research project. Special thanks go to Wally and Ed for the distinct roles that they have played in my career, much to my benefit.

Thank you also to the many friends and colleagues that supported me during this research project – you know who you are!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

MANAGERIAL MERIT PAY ALLOCATION: AN ANALYSIS OF ALTERNATIVE JUDGMENT MODELS USING REGRESSION AND FUZZY EXPERT SYSTEM TECHNIQUES

by

DAVID W. DORSEY

An Abstract

Of a dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychology
University of South Florida

April 1997

Major Professor: Michael D. Coovert, Ph.D.

This research focused on the role of fuzzy expert systems, and the associated fuzzy systems theory, as a new methodology and framework for capturing and modeling important organizational judgments and decisions. It is proposed that fuzzy systems offer promise in terms of overcoming limitations in traditional policy capturing approaches. In this research effort, a judgment analysis study looking at a critical organizational decision task, managerial merit pay allocation, was conducted in which fuzzy system models were compared and contrasted with both linear and nonlinear regression models. Ten real-world managers served as participants. The results indicate that fuzzy systems are indeed a powerful tool for modeling judgments, generally outperforming both linear and nonlinear regression methods in terms of model fit. The results also yielded evidence that the managers, when completing a simulated merit pay allocation task, appeared to be using some types of nonlinear noncompensatory allocation strategies. These results imply that traditional policy capturing efforts based on linear regression may not be optimal for modeling these types of judgments. Although certain forms of nonlinear noncompensatory strategies were suggested, the exact judgment processes used by the managers could not be precisely specified due to deficiencies in both the fuzzy system and regression approaches. A discussion of the strengths and weaknesses of the various modeling approaches is offered, and areas for future research and development are noted.

Abstract Approved: Michael D. Coovert

Major Professor: Michael D. Coovert, Ph.D.
Professor, Department of Psychology

Date Approved: 3/24/97

4/97

INTRODUCTION

Compensation decisions have been shown to be consequential for organizational members in terms of motivation, job satisfaction, and performance (Lawler, 1981). Organizations have hoped to capitalize on the effects of various types of compensation by implementing a number of innovative pay schemes including skill-based pay, all-salaried work forces, lump-sum salary increases, performance-based pay systems, new performance appraisal systems (e.g., peer review), and gainsharing (Cummings & Huse, 1989). Performance-based pay systems or as they are commonly called "merit pay" systems, in particular, have been widely utilized as a form of incentive compensation. Evidence for the proliferation of this type of compensation plan can be found in two national surveys in which over 80% of the respondents indicated that their organization had a merit pay plan (Peck, 1984; Personnel Policies Forum, 1981). Moreover, a recent study suggests that despite a trend towards not relying solely on conventional individual incentive plans, as many as 90% of Fortune 1000 firms still use such incentives (Ledford, Lawler, & Mohrman, 1995). The implementation of merit pay plans is founded on the assumption that these plans both stimulate and support high levels of employee performance. The use of merit pay has generated a considerable amount of controversy and the need for further research on both the processes and outcomes associated with merit pay has been noted (Heneman, 1990).

Research on merit pay has traditionally used one of three methodologies (Sherer, Schwab, & Heneman, 1987). The first of these methodologies falls under the category of experimental studies, in which characteristics of the allocators, recipients, the organization, or the environment are manipulated in order to assess the impact of these variables on merit pay allocation. While generally demonstrating high levels of internal validity, these studies have questionable levels of external validity given the frequent use of student populations.

Also, individual differences on the part of the allocator are often analyzed as error because of sole reliance on between subject experimental designs.

A second popular methodology features actual salary raise decisions as captured in field research studies. While obviously increasing external validity, these studies are primarily inductive in nature, making statements concerning causality difficult if not impossible to assert (Sherer et al., 1987). In addition, these field studies typically feature between subject designs making it unfeasible to study the individual differences inherent in managerial merit pay allocation decisions.

The last type of traditional methodology involves the direct questioning of managers and supervisors. These studies, however, often focus on salary level decisions rather than salary raise decisions and the degree to which these studies capture meaningful individual differences is suspect given the extensive literature that documents the inability of decision makers to articulate their own decision making models (Sherer et al., 1987, Slovic & Lichtenstein, 1971).

Due to the limitations of these methods, a fourth type of methodology has been implemented. This methodology, known as "policy capturing", attempts to describe quantitatively the relation between judgments and the information or "cues" used in making those judgments (Stewart, 1988). Note that policy capturing generally deals with judgments rather than decisions (Stevenson, Busemeyer, & Naylor, 1990), although in the merit pay literature the term "decisions" is generally used. These terms are used interchangeably in this paper. The policy capturing approach attempts to maintain high levels of internal validity through the use of experimental manipulation, maintain high levels of external validity through the use of experienced judges, and increase the ability to describe meaningful individual differences through the application of quantitative models (Sherer et al., 1987).

Studies attempting to capture or model pay allocator strategies have the potential to address fundamental questions about merit pay allocation in organizations. Some issues to be resolved include: why subordinates who perform at similar levels receive different merit increases, and are managers who routinely make merit pay allocation decisions aware of the decision models that they use (Deshpande & Joseph, 1994)? Given the great importance attached to such issues, it is surprising that there have been few studies looking at merit pay allocation that have used a policy capturing approach. This may be due in part to the fact that pay allocation is often deeply embedded in and tied to specific organizational factors. This aspect of pay allocation remains a challenge to judgment researchers, and in response, many researchers have called for a more detailed analysis of non-performance factors (Deshpande & Schoderbek, 1993). Specifically, future policy capturing studies in the area of merit pay allocation need to build measurement models inclusive of non-performance factors that are representative of real-world organizational conditions.

One problematic aspect of policy capturing studies, including those done in the merit pay allocation domain, is the use of orthogonal cue sets due to the difficulty in analyzing intercorrelated cues with multiple regression. In fact, as cues become more intercorrelated, measures related to the relative importance of the cues in determining judgments go from being ambiguous to virtually useless (Stewart, 1988). The use of orthogonal cue sets may be especially questionable in the merit pay domain given the somewhat intuitive relationships that are likely to exist between variables affecting merit pay allocation. For example, one pair of variables that are likely correlated are current base salary and tenure. The principle of representative design dictates that cue intercorrelations should match those that exist in the environment (Stewart, 1988). Due to these issues, a challenge for judgment researchers is developing a methodology for studying judgment policies that yields information relative to cue importance and descriptive of a judge's policy that is not rendered uninterpretable under conditions of correlated cues.

It is also noteworthy that in the policy capturing studies conducted thus far looking at merit pay, researchers have relied solely on linear regression models without testing the extent to which nonlinear, noncompensatory, or configural strategies were used by the pay allocators. This fact is particularly troubling since current evidence does not favor the conclusion that the linear model is "the" model of the judgment process (Brehmer & Brehmer, 1988). In fact, the so called "pervasiveness of linearity" (Green 1968) attributed to human judgment may be more reflective of a lack of research on alternative models (Ganzach, 1995) and the use of weak nonlinear, noncompensatory models (Ganzach & Czackes, 1995).

Policy capturing studies are generally undertaken for three general purposes, which include: 1) examining the unique information processing behavior of raters; 2) comparing the stated rating policies with those actually used; and 3) training raters in the consistent use and application of a given policy (Hobson & Gibson, 1983). Within the merit pay allocation domain, the second of these purposes is especially interesting and potentially important. While employees may base their perceptions of equity and justice on explicitly stated policies and procedures, managers allocating pay may actually be using strategies that are contrary to stated organizational or managerial policies. In fact, general reviews of the human judgment literature, as well as studies focusing on managerial merit pay allocation, suggest that decision makers cannot accurately specify the importance that they attach to attributes when making judgments (Sherer et al., 1987). However, a great deal of concern has been expressed over the types of methods used to elicit subjective estimates relating to judgment policies (e.g., Deshpande & Shoderbek, 1993). Given this concern, it is surprising that little methodological research has been directed towards the elicitation and interpretation of participant's verbal descriptions of their policies (Brehmer & Brehmer, 1988).

Thus far, I have reviewed a number of problematic areas relating to the design and analysis of studies that focus on merit pay allocation decisions. These areas represent both conceptual issues specific to the merit pay domain as well as general methodological issues pertaining to policy capturing. These issues include the following:

1) a wider range of cues relating to merit pay allocation need to be studied, specifically, variables should be incorporated that represent non-performance factors;

2) the development and analysis of a methodology that can be used with correlated cue sets is necessary, especially for cases in which relatively high correlations between cues are reflective of a representative design;

3) further analyses of techniques that are sensitive to judgment strategies that have nonlinear and/or noncompensatory components are needed, in order to analyze the extent to which these strategies are used in allocating merit pay; and

4) the development and analyses of alternative methods for eliciting and interpreting judges verbal descriptions of their judgment policies are needed.

The issues noted above comprise the motivating factors for conducting the research discussed here. The strategy used to address these issues was to conduct a judgment analytic study that focused on a number of judgment cues hypothesized to be relevant to merit pay allocation. The design for this study followed the traditional design characteristics of a policy capturing study; however, a number of different analytic methods were used and contrasted. Included in this set of analytic techniques is a new methodology for policy capturing that has been derived from research in computer science, more specifically from research on "fuzzy systems". This methodology appears to hold promise for addressing some of the methodological concerns noted earlier.

The inclusion of a methodology that has emerged from computer science stands as one attempt to fulfill a trend towards multidisciplinary research methods that has been anticipated by decision making researchers. For example, the Nobel laureate Herbert A. Simon (1995) recently noted the following:

> Decision making is at the center of all of these human activities (and many more). You go to the philosophy of science to sharpen your methods and to computer science to find a formal language for expressing your theories. And in particular you go to psychology to study the underlying processes that enable people to make decisions, solve problems, and generally to think. (p.507).

Similarly, Stevenson, Busemeyer, and Naylor (1990) in their review chapter on judgment and decision making theory state: "The second development we are convinced is imminent is the increasing influence of an interface with cognitive science and, in particular, the application of models from the computer science field to issues of judgment and choice" (p.364).

In conducting this research, I hoped first to gain new insight into the effectiveness of various policy capturing techniques and secondly to contribute to knowledge about the merit pay allocation process through the analysis of individual pay allocator judgments. Before describing the specific design elements of this study, I will first attempt to establish a conceptual and methodological foundation by providing an overview of the literature associated with merit pay research, an overview of judgment analytic methodology, and an introduction to fuzzy expert systems and fuzzy systems theory.

An Overview of Merit Pay Research

In a review of empirical studies on merit pay, Heneman (1990) offers a useful conceptual framework for categorizing the variables that have been studied in merit pay research. In building this conceptual model Heneman defines merit pay as "individual pay increases based on rated performance of individual employees in a previous time period". While Heneman admits that this definition is debatable, it serves to contrast merit pay from other pay plans such as base-pay, profit sharing, and piece-rate plans (Heneman, 1990). Previous research has documented that within the concept of merit pay there is wide variation in the actual features of the pay plans. Heneman (1984) identified 12 different types of merit pay plans that were described in both research-oriented and practitioner-oriented journals. The variation within this broad range of merit pay plans can be characterized by the following five factors: 1) the degree to which these plans consider factors other than performance (e.g., tenure, cost of living, etc.), 2) the fact that pay is not always in the form of a salary increase but can include additional benefits, 3) differences in the way that merit increases are calculated, 4) the fact that merit pay may be distributed more than once a year, and 5) merit pay increases may or may not be permanent (Heneman, 1990). Additional variation in merit pay plans also occurs as a function of administrative differences that occur between public versus private sector organizations and variation in the actual sizes of merit increases (Heneman, 1990).

The conceptual framework offered by Heneman (1990) is shown graphically in Figure 1.

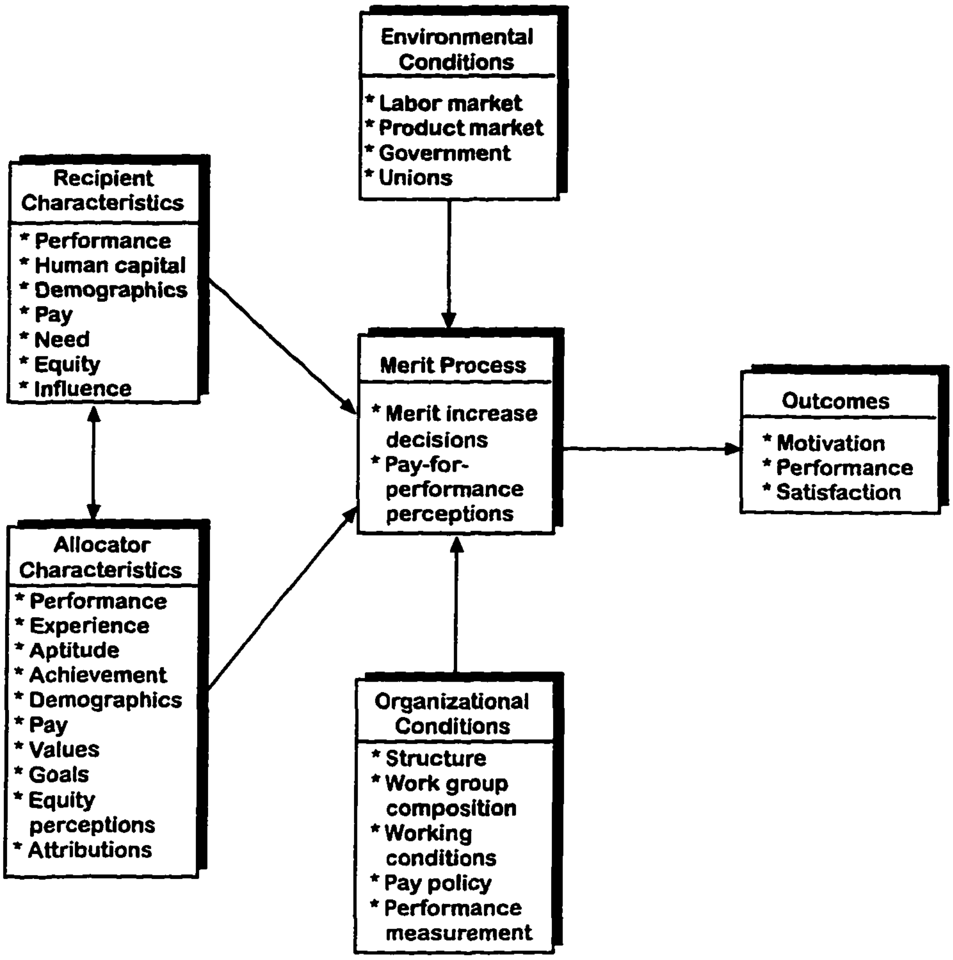


Figure 1. Merit pay framework.

This framework portrays merit pay as a process which can be viewed from both an individual and an organizational perspective. From an organizational perspective, decisions are made about the size of the pay increase for levels of rated performance. These merit increase decisions are perceived by the individual in the form of "pay-for-performance" perceptions, in which the individual forms an opinion regarding the extent to which the allocated pay raise is actually based on performance versus other extraneous factors. This framework also demonstrates that the merit pay process is assumed to be related to substantive outcomes such as motivation, performance, and satisfaction. Figure 1 portrays four individual factors that are presumed to be related to the merit pay process. The first factor contains characteristics of the merit pay recipient, including the recipient's values, attitudes, background characteristics, and behaviors. The second factor relates to the characteristics or attributes of the pay allocator, which are similar in nature to those of the recipient. The third factor contains aspects of the organization, including such elements of organizational functioning as the structure and policies of the organization. The fourth factor relates to environmental conditions and includes the role of governments, unions, labor and product markets. In the following sections, the individual relationships between these four factors and merit increase decisions will be reviewed. Note that the relevant dependent variable in this review is merit increase decisions or judgments concerning the size of the merit increase in pay.

Recipient Characteristics

Recipient characteristics represent the most frequently studied aspects of pay allocation decisions (Heneman, 1990). Figure 2 presents an overview of the types of recipient characteristics that have been studied and a general summary of the findings that relate these characteristics to merit increase decisions.

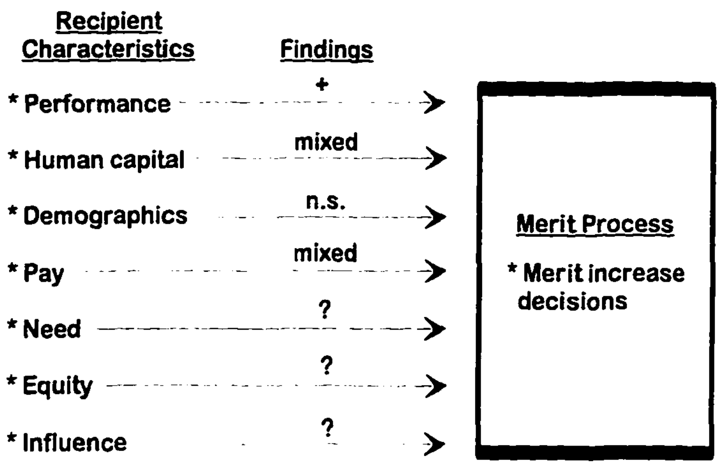


Figure 2. An overview of research findings relating recipient characteristics to merit increase decisions.

The first recipient characteristic shown in Figure 2, performance, represents one of the few variables for which there appears to be a well-documented relationship to merit increase decisions. The empirical literature on merit increase decisions documents a positive and significant relationship between these variables. Heneman (1990), in his review, cites 23 different samples in which this relationship has been supported. In terms of practical implications this finding would seem to be encouraging, however, at a conceptual level these results are more difficult to interpret given the lack of empirical findings concerning the causal relationship between performance and merit pay (Heneman, 1990). Specifically, it is currently equivocal whether performance causes pay, pay causes performance, or the relationship is reciprocal in nature. Also, the true magnitude of the relationship between pay and performance has not yet been established. It also appears based on the results of a policy capturing study by Beatty, McCune, and Beatty (1988) that there may be a boundary condition on the performance - pay relationship based on culture. Beatty and colleagues in comparing a sample of U.S. managers to a sample of Japanese managers, found that U.S. managers placed a heavier emphasis on performance than the Japanese managers who valued other aspects such as job worth as much or more than performance in making pay allocation decisions.

Another interesting aspect of the performance - merit pay relationship is the role of performance consistency. The role of performance consistency has been studied in policy capturing studies by Deshpande and Joseph (1994) and Sherer et al. (1987), resulting in somewhat mixed findings. However, for at least some allocators, performance consistency appeared to be an important determinant of pay allocation judgments.

A second recipient characteristic is noted by Heneman (1990) as "human capital". This label is reflective of variables such as tenure, education, and training time. The results relevant to these variables have been mixed. It appears that variables such as tenure and education may have differential effects, depending on variation in organizational pay policies (Heneman, 1990). Also, Bishop (1987) has suggested that human capital variables may have differential explanatory power depending on the time period of focus within one's tenure, with human capital variables having a greater influence earlier in the tenure period.

In terms of demographic characteristics, there has been very little support for the effects of sex, race, or age in merit pay decisions. One study by Heneman and Cohen (1988) did find a significant negative effect for age but these authors attributed this effect to the youthful status and closeness in age of the allocators and recipients.

A fourth recipient characteristic, noted as "pay" in Figure 2, corresponds to the recipient's position in the pay structure of the organization. This characteristic has been shown to have differential and complex relationships to merit pay increases depending on whether the focus is on job level versus comparative salary ratio. In terms of job level, there is evidence that individuals in higher level positions receive higher merit pay increases (Heneman, 1990). Possible reasons for this relationship include the belief that individuals in higher positions have higher levels of performance and the belief that individuals in higher positions are more critical to the organization (Heneman, 1990). It is also possible that the actual relationship between job level and merit increases may be more complex depending on the range of job levels studied. Work by Alexander and Barrett (1982) looking at low level jobs (characterized as "boring" and "dirty"), in conjunction with studies on job level for exempt jobs, suggests that the relationship between job level and merit increases may be U-shaped across the continuum of job levels, with larger merit increases being granted to the extreme ends of the job level spectrum (Heneman, 1990). The comparative salary ratio of merit pay recipients also may be significantly related to merit pay decisions, although at present this relationship has not been clearly established. Nonetheless, there are intuitive reasons to suspect a significant relationship between these variables. First, recipients below the midpoint of the salary range may receive larger increases due to a compensatory reaction to bring them higher up in the pay grade. Secondly, employees high in the salary range may receive lower increases in order to not exceed the maximum salary grade that is implicit in many organizational pay policies. A study by Heneman and Cohen (1988) lends some support to these hypotheses, in view of the fact that they found a significant negative correlation between comparative salary ratio and merit increases.

The status of need and equity as determinants of merit pay is as of yet unclear. However, some support for the effects of both need and equity has been indicated. In the case of perceived need, both Peters and Atkin (1980) and Fossum and Fitch (1985) found evidence of a relationship between need and pay, although only the Peters and Atkin study reported these results as significant. In both studies the direction of the relationship was positive, with larger perceived needs resulting in larger salary increases. Studies using actual salary decisions are needed to confirm these findings. In the case of equity, significant effects have been demonstrated in lab studies, with underpaid individuals receiving larger increases (Birnbaum, 1983; Freedman, 1978).

However, in a field study by Dreher (1981) neither internal nor external equity perceptions had a significant effect on salary increases.

The last recipient characteristic noted in Figure 2 is influence. Influence attempts by recipients to obtain larger salaries have only been investigated in a couple of studies, resulting in mixed findings (Freedman, 1978; Martin, 1987). As is the case with need and equity, more research is needed to solidify influence as a factor in merit pay decisions.

Some future research needs noted by Heneman (1990) within the domain of recipient characteristics include: 1) halting purely cross-sectional correlational studies, 2) gathering recipient perceptions of pay allocation factors, 3) more policy capturing research examining the weighting schemes used by allocators, 4) looking at possible in-group out-group distinctions, and 5) an increased focus on political behavior.

Allocator Characteristics

Figure 3 presents an overview of the findings relating characteristics of the merit pay allocator to merit pay decisions. As is evident from viewing this figure, very little is known about the role of allocator characteristics. Heneman (1990), in his review, noted only eight studies in which allocator characteristics had been considered. In reviewing these studies, Heneman noted the difficulty in summarizing the findings from these studies given the lack of statistically significant results, which may be due in part to the use of small sample sizes.

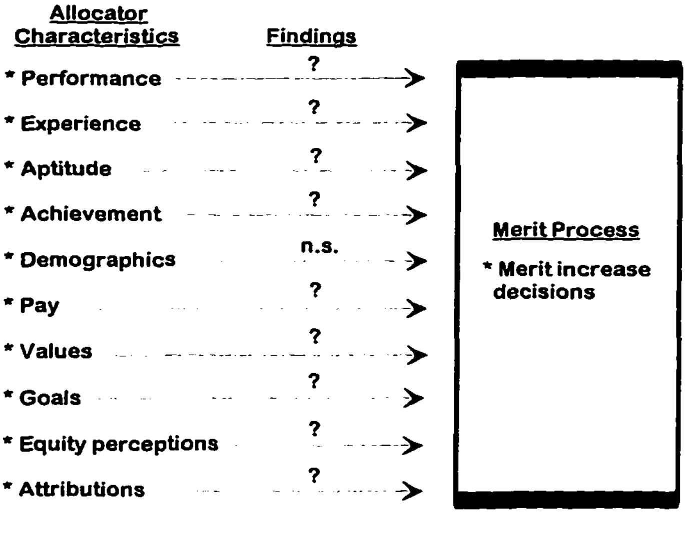


Figure 3. An overview of research findings relating allocator characteristics to merit increase decisions.

Despite the lack of significant findings, some trends noted by Heneman (1990) suggest that recipient increases are likely to be larger when the allocator has less experience, receives higher pay increases, is altruistic, and attributes recipient performance to internal rather than external causes. Heneman endorsed the continuation of research looking at allocator characteristics given the results of studies such as Heneman and Cohen (1988), in which allocator characteristics accounted for a respectable 11% of the variance in merit increases.

Another interesting line of research relating to allocator characteristics has been suggested by Bartol and Martin's (1988) proposition that merit pay increases are influenced by the allocator's dependence on the recipient. Some evidence in support of this proposition has been offered by Deshpande and Joseph (1994), who demonstrated in a policy capturing study that the importance of the subordinate's job in meeting the goals of the manager and the degree of disruption that would occur if the subordinate were to leave accounted for 12% of the variance in pay allocations. Additional evidence relating to the "dependence" hypothesis is found in studies by Bartol and Martin (1990) and Deshpande and Schoderbek (1993), who found that specialized skills/expertise on the part of the recipient as well as difficulty in replacing the recipient can affect managerial pay allocation.

Some important aspects of allocator characteristics that need to be addressed in future research include: 1) further exploration of the "dependence" proposition by Bartol and Martin (1988), 2) allocator interest in promoting group interdependence through pay allocation, and 3) the role of political behavior on the part of allocators (Heneman, 1990).

Organizational Conditions

Given the fact that merit increase decisions occur in a larger organizational context, organizational conditions should serve to establish norms and boundaries which impact merit pay allocation decisions (Freedman & Montanari, 1980). Heneman (1990) noted a number of possible organizational conditions relevant to merit pay which are noted in Figure 4. Organizational conditions or characteristics and their influence on merit pay decisions represents a relatively new area of research with results thus far suggesting that it is a promising line of inquiry (Heneman, 1990).

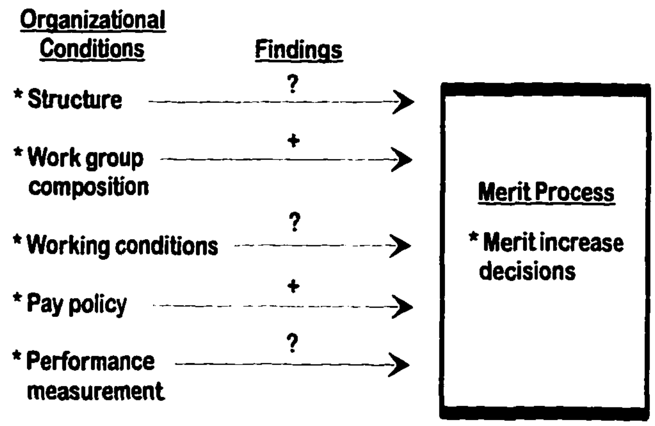


Figure 4. An overview of research findings relating organizational conditions to merit increase decisions.

The first characteristic noted in Figure 4 corresponds to organizational structure. While structure intuitively would impact merit pay allocation, very little is known at this point what the effects are.

The second characteristic, work group composition, represents the relationships between the allocator, recipient, and recipient coworkers (Heneman, 1990). Heneman notes three studies relevant to this characteristic which yield significant results. First, Ivancevich (1983) found evidence that a type of contrast effect impacted merit pay in a sample of engineers, with larger merit increases resulting from an increased proportion of unsatisfactorily performing coworkers. In a second study, Markham (1988) pointed out a possible level of analysis question, by demonstrating that for a sample of managers and professionals in a metals processing firm, the correlations between merit raises and individual performance was nonsignificant while a significant group effect was detected. In a final study, Turban and Jones (1988) looked at three types of allocator - recipient similarity and found that perceived similarity was related to recommended pay increases. It should be noted that in looking at variables such as similarity, the extent to which these factors represent job-relevant attributes can determine whether similarity should be looked at as true versus error variance in relation to performance (Heneman, 1990). Taken together, these three studies indicate that aspects of work group composition should be considered in future merit pay research.

Although it is too early to draw conclusions concerning the effect of working conditions on merit pay allocation, at least one study by Alexander and Barrett (1982) suggests that salary increase may be higher for jobs containing unpleasant working conditions. The effects of working conditions may factor into compensation systems differentially depending on organizational characteristics such as the presence of unions (e.g., hazard pay) or whether the conditions have been factored into base pay through the job evaluation process (Heneman, 1990). More research is needed to clarify these relationships.

Another organizational characteristic, the organization's pay policy, has been shown to affect merit pay. Examples of variables relating to pay policy include: pay secrecy, the treatment of nonproductive employees, the use of recommended salary guidelines, the permanency of pay raises, and work group harmony (Heneman, 1990). More research, especially in field settings, is needed to confirm which aspects of pay policies affect pay allocation the most. One recent study looking at three different organizational determinants of pay allocation decisions by Trahan, Lane, and Dobbins (1991) found that the size of the salary budget had a significant effect in two different studies, while organizational goals and openness of pay decisions affected decisions in one of the two studies.

The last organizational condition noted in Figure 4 is performance measurement. Two studies have looked at the effect of measurement properties of performance standards. Huber, Neale, and Northcraft (1987) found a positive but insignificant relationship between well-developed performance standards and the recommended size of the merit increase. Schwab and Olson (1988) found a significant decrease in the correlation between performance and pay when the reliability of the performance measure was .3 rather than .6. Given that only one significant finding has been reported thus far, more research is needed to assess the impact of performance measurement properties on the pay - performance relationship.

Organizational policies and procedures are potentially one of the strongest predictors of merit pay allocations (Heneman, 1990). Additional areas for research on organizational characteristics include: 1) the effect of merit guide charts, 2) the form of increases, 3) the role of communication of merit pay plans, 4) examining the choice of merit pay plans, 5) more research on budgeting concerns, and 6) further consideration of merit pay as contrasted with other compensation policies and practices.

Environmental Conditions

Just as pay allocation decisions are made in an organizational context, the environmental context also has the potential to affect pay decisions. Figure 5 highlights some of these environmental conditions.

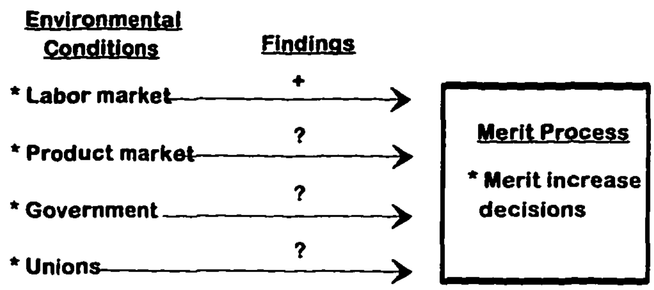


Figure 5. An overview of research findings relating environmental conditions to merit increase decisions.

Data associated with the labor market, the first condition shown in Figure 5, suggest that the labor market does play a role in merit pay allocation. Heneman (1990) cites support from three different samples to indicate that as economic theory would predict, allocators will grant larger increases to individuals in occupations that are difficult to replace. Similarly, there is support from an additional four studies that suggest that pay increases will be higher for those with another job offer (Alexander & Barrett, 1982; Deshpande & Schoderbek, 1993; Landau & Leventhal, 1976; Sherer, Schwab, & Heneman, 1987). However, it should be noted that in the studies using a within subject design (e.g., Sherer et al., 1987), allocators differentially reacted to a recipient with another job offer, suggesting the influence of individual differences on the part of the allocators.

Much less is known about the role of the other environmental conditions cited in Figure 5, although a study by Bishop (1987) points to the potential importance of unions, product markets, and government in the merit pay process. Certainly, more research is needed on all aspects of the environmental context.

Future Research

In the previous discussion, a number of factors were reviewed relating recipient, allocator, organizational, and environmental characteristics to merit pay increase decisions. It is obvious from this discussion that a significant amount of research is needed before a useful theoretical model can be developed that coherently addresses the interrelationships between the antecedents and consequences of merit pay. Heneman (1990) suggests three key developments that should be addressed in future merit pay research. These include: 1) to develop and draw upon a theory, bringing some order to the variables relating to merit pay, 2) to more carefully explicate the merit pay process, and 3) to more carefully explain the causal relationships between merit pay variables.

Continued work in the area of judgment analysis (i.e. policy capturing) holds promise for addressing the second concern noted above. Through the analysis of allocator judgments, a significant aspect of the overall merit pay process may be clarified. As Dyer and Schwab (1982) have noted, many researchers have only considered pay plans in the context of studying motivational theories, offering only ad hoc descriptions of the pay process. Dyer and Schwab challenged researchers to address the pay process itself. Using judgment analysis to study the pay allocation process stands as an attempt to meet this challenge.

An Overview of Judgment Analytic Methodology

Given the collection of judgments and decisions that affect organizations and their members on a daily basis, it is not surprising that there has been widespread application of judgment analysis for understanding organizational phenomena. Consider for example the growing literature associated with modeling the rating processes inherent in performance appraisal (Champagne & Stevenson, 1994; Hobson & Gibson, 1983). Based on the idea of mathematically or statistically modeling rater policies, researchers have provided a framework for assessing unique information processing behavior, assessing subjective versus objective policies, and expanding the knowledge base concerning the ability to train raters in the consistent use of a given policy (Hobson & Gibson, 1983).

Traditional Design Elements

As with all formal analytic methods, proper formulation and design are critical for conducting successful analyses (Stewart, 1988). In the case of judgment analysis, design elements of great importance include: defining the judgment of interest, identifying the cues, describing the context for judgment, defining the distributions of the cue values, and defining relations among the cues (Stewart, 1988).

In terms of defining the judgment, the judges must clearly understand the task and an appropriate response mode must be chosen (e.g., rating scales). The best method for gathering responses is generally the "one that is most acceptable to the judges" (Stewart, 1988 p. 43).

Another important element of a judgment analytic study is the identification of cues. All important cues should be included in the cue set but the number of cues should generally be kept relatively small. Specifically, in choosing the number of cues to manipulate, considerations must be taken into account which include sampling error, more specifically the number of cases to independent variables (Nunnally, 1978), and the fact that results from studies using different methods generally show that judges use a small subset of the cues available (Brehmer & Brehmer, 1988).

A context for judgment must also be established. The context provides information to the judges such as the purpose of the judgment, conditions leading up to the judgment, or other invariant characteristics of the judgment setting (Stewart, 1988).

Defining cue distributions can also be consequential for judgment analysis. Choices must be made in terms of measuring cue values on concrete versus abstract units and the distribution and range of cue values should be representative, that is, reflective of the distributions and ranges in the environment.

As previously discussed, cue intercorrelations are important in judgment analysis. This issue is often discussed in terms of orthogonal versus non-orthogonal cue sets. Once again, choosing correlated versus orthogonal cue sets involves choices on the part of the researcher that are based on emphasizing the merits of representative design principles versus facilitating statistical comparisons. Similarly, choosing the number of profiles or cases to present to judges involves balancing statistical concerns versus considering practical constraints (e.g., time, attention, etc.).

Policy Capturing

Two major types of judgment analysis are typically used. The first type follows the correlational paradigm (Slovic & Lichtenstein, 1971), which is highlighted by the use of multiple regression for what has been termed "policy capturing" or "judgment analysis".

The correlational paradigm is founded on establishing the correlations between information cues and psychological responses as a way of modeling information processing behavior. A main impetus for this approach has been Brunswik's "lens model" representation of the judgment process (Brunswik, 1952). In this framework, a regression model is interpreted as representative of the way cues are combined by the judge in making a judgment (Stevenson, Busemeyer, & Naylor, 1990). This approach has been popular and a host of researchers have chosen to represent judges' evaluation strategies as linear in nature (Hobson, Mendel, & Gibson, 1981; Naylor & Wherry, 1965; Taylor & Wilstead, 1974; Zedeck & Kafry, 1977). While being a useful and powerful tool, linear regression is potentially problematic for a number of reasons, which have been extensively reviewed in the judgment and decision making literature. These potential problems are tied to the some of the implicit assumptions of linear regression, and an understanding of these assumptions is critical. These assumptions are briefly reviewed below.

Assumption One - Linearity. The use of regression assumes that the relation between each cue and the judgment is linear, or stated another way, that the impact of one additional unit of cue value on the judgment does not change at different levels of the cue. For example, consider the relation between merit pay increase and recipient acceptability. A linear description would prescribe that a 3% change in merit pay would result in the same change in acceptability regardless if it was an increase from 0% to 3% or an increase from 7% to 10%. If however the change in acceptability due to a 3% increase depended on the level of increase, then the relation would not be linear.

Assumption Two - Additivity. The assumption of additivity states that the effects of one cue does not depend on the levels of others. Additive models are also called "compensatory", since increases in one variable can compensate for the lack of change in another variable due to the fact that the variables are added to produce an effect in a dependent variable. Some examples of nonadditive relations, as described by Stewart (1988), are shown below:

1) cue 1 is only considered when cue 2 is high;
2) if all cue values are high, then the judgment is high, otherwise the judgment is low (conjunctive model);
3) if any cue is high then the judgment is high, otherwise the judgment is low (disjunctive model);
4) the weight on cue 1 increases as the value of cue 2 increases (multiplicative model);
5) the judgment increases as the amount of discrepancy between cue 1 and cue 2 increases (absolute value of difference model);

It becomes obvious when reviewing the linearity and additivity assumptions that there are conceptually sound reasons to suspect that human judgments may often depart from linear additive strategies. For example, there is gathering evidence that performance evaluations tend to be conjunctive, or more specifically, raters attend more to negative attributes than to positive attributes (Ganzach, 1995). Support for this proposition was recently found in a variety of real world performance evaluation tasks (Ganzach, 1995). Note that a conjunctive strategy is non-additive because no amount of increase in other variables can compensate for a low value on one variable.

Linear additive models function well in terms of prediction (e.g., Dawes & Corrigan, 1974; Slovic & Lichtenstein, 1971). Specifically, a "pervasiveness of linearity" (Green 1968) has been cited when contrasting linear and nonlinear models in terms of predictability (i.e., differences in squared multiple correlation coefficients). However, this pervasiveness should be qualified by assessing first, the purpose of undertaking a judgment analytic study, and second, the choices involved in assessing nonlinear models. In terms of the purpose of judgment research, Brannick and Brannick (1989) noted that "when theoretical issues about judgment strategies are involved, a small difference in predictability can make a major difference in implication" (p.119).

Despite the importance of considering nonlinear and noncompensatory models, judgment researchers rarely consider departures from the linear additive model. Even when departures are considered, there are a number of difficulties associated with using current nonlinear noncompensatory modeling techniques. First, there are often a number of nonlinear models to consider and at times it can be difficult to determine which nonlinear model best represents the data at hand (Brannick & Brannick, 1989). One recent study compared four types of nonlinear models, including: the parabolic model and the hyperbolic model (Einhorn, 1970), the scatter model (Brannick & Brannick, 1989), and the true conjunctive-disjunctive model (Ganzach & Czaczkes, 1995). The results in this study suggested that the scatter model was the best model in terms of fit. There is also an added complexity in trying to interpret the importance of individual cues with nonlinear models and as the individual cues become intercorrelated it becomes increasingly difficult to establish meaningful interactions and evidence of noncompensatory processes (Brannick & Brannick, 1989; Goldberg, 1971; Stevenson, Busemeyer, & Naylor, 1990).

Policy Modeling

A second type of judgment analysis relies on experimental design and measurement theory for the purposes of "policy modeling" (Champagne & Stevenson, 1994; Stevenson, Busemeyer, & Naylor, 1990). Recently, Stevenson and colleagues (Champagne & Stevenson, 1994; Stuhlmacher & Stevenson, 1994) have extended earlier research (Birnbaum, 1976; Norman, 1976) that demonstrated the importance of manipulating the number of attributes presented in judgment tasks, by suggesting a policy modeling approach. Policy modeling has the same objective as policy capturing, however, policy modeling introduces subjective values for the attributes, assesses a wider variety of judgment strategies, and tests for the linearity of the judgment scale (Stuhlmacher & Stevenson, 1994). Policy modeling does not assume that raters use an additive model, but instead, a number of plausible strategies can be assessed including additive, averaging, and configural models. While policy modeling presents some distinct advantages over traditional policy capturing, it also appears to have some potentially serious constraints. First, the design of a policy modeling study involves varying systematically the type of information presented in terms of the presence or absence of certain cues as well varying the levels of each cue. For example, in the case of three cues, this approach involves presenting judges with profiles that include all three-way combinations of each cue level, all two-way combinations of each cue level, and profiles with each individual cue level present alone. In the case of judgment studies involving two or three cues the number of profiles may not be a critical issue but in terms of analyzing higher dimensional judgment situations (e.g., assessing performance appraisal judgments incorporating a number of dimensions greater than three) this aspect of policy modeling may be problematic given the practical constraints inherent in having judges allocate time, energy, and attention to the task. Another apparent problem is applying policy modeling techniques in more generalizable field settings where it may be exceedingly difficult to systematically manipulate cue presentation. Also difficult is the large number of possible linear and nonlinear terms that may be appropriate for a given model. For example, in the case of assessing a configural-averaging model or a multiplicative regression model, parameters are often added that are associated with the possible interactions of cues or attributes. As pointed out by Craiger and Coovert (1994), as the number of independent variables increases, the number of possible combinations of interactive terms rises exponentially.

Because of limitations in conventional methodologies, researchers such as Craiger and Coovert (1994) have suggested integrating techniques from computer science, more specifically the branch of computer science known as "artificial intelligence" (AI), into psychological research methods. In view of the previously described limitations of present judgment analytic methods, I propose to extend the use of techniques from AI to the domain of policy capturing. The foundations for the AI methodology discussed in this study lies in the pioneering work of Lotfi Zadeh (1965) on the concept of "fuzzy sets" and "fuzzy logic". In the following section, I provide an overview of fuzzy systems.

An Overview of Fuzzy Expert Systems and Fuzzy Systems Theory

Before describing the formal elements of fuzzy systems, the following introductory example will be used to illustrate the value of fuzzy systems. This example is an adaptation of an illustration used by Gulley and Jang (1995).

Introductory Example - Fuzzy vs. Non-fuzzy Modeling

The Non-fuzzy Approach. Consider the following problem: a manager is asked to allocate the "appropriate" amount of merit pay given an overall performance rating on a scale of one to seven (where seven indicates excellent performance); what is the appropriate merit increase? We might find that the allocation of pay by the manager can be modeled using a linear function, where the % merit increase increments linearly with performance. A model of this judgment strategy can be developed through a conventional technique such as regression. Such a linear model generated from the equation (% merit increase = .025 * performance - .025) is shown in Figure 6.

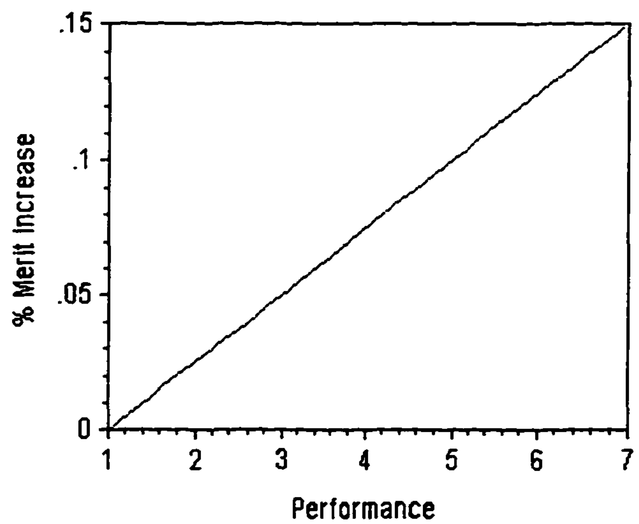


Figure 6. % merit increase as a linear function of performance.

Thus far the equation does what we want it to do, and it is fairly straightforward. But what if the manager is given another piece of information, for example - information concerning tenure, and the manager takes this variable into account along with performance in determining merit pay. Given a measure of tenure based on number of years, the manager's judgments may still be captured adequately using a linear additive model with an equation such as (% merit increase = (.0125 * performance) + (.0125 * tenure) - .025), which is shown graphically in Figure 7.

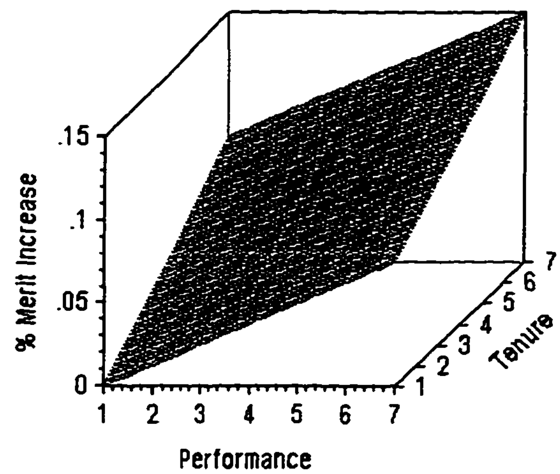


Figure 7. % merit increase as a linear function of performance and tenure.

So far we can fairly easily model this set of judgments, however, it may be the case that a manager would not weight both performance and tenure equally. For example, if the manager weights performance so that it accounts for 80% of the overall merit increase we must now change our model to indicate this emphasis using the following equation (% merit increase = .80 * (.0125 * performance) + .20 * (.0125 * tenure) - .0125) (shown graphically in Figure 8).

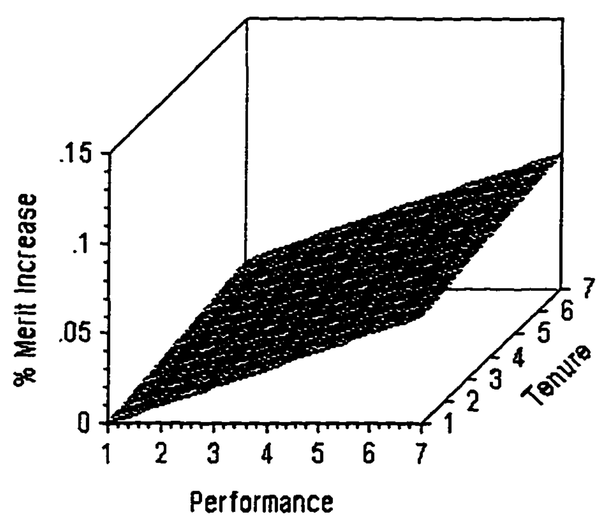
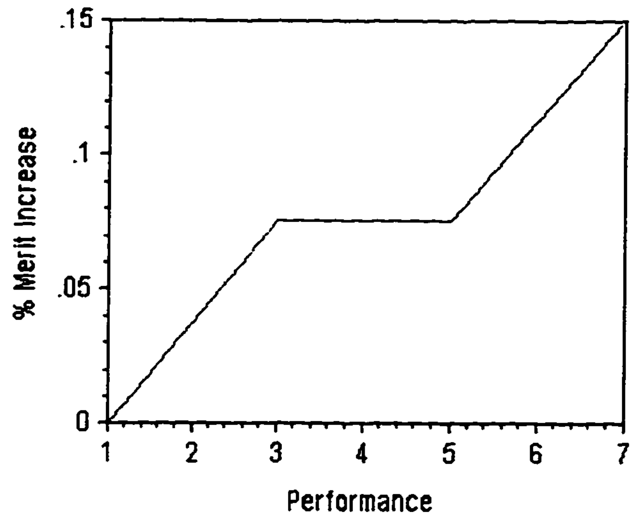


Figure 8. % merit increase as a linear function of performance and tenure with an increased weight on performance.

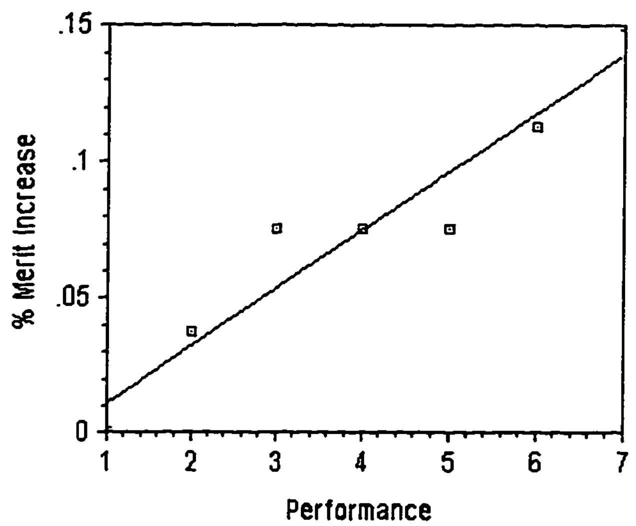
Up to this point we are able to model the manager's judgment strategy with little computational complexity by implementing a linear additive model. Now consider the case where further investigation leads to the hypothesis that a model of the manager's pay allocation judgments should have more of a flat response in the middle. That is, the manger will generally give a merit increase of around 8% and will only depart from this level when performance is exceptionally good or exceptionally bad. This hypothesis, which seems plausible and realistic, now requires that we abandon our use of simple linear relations. Due to the added complexity we must now attempt to model the pay allocation judgments using piecewise linear construction. First, returning to the one-dimensional problem of just considering performance, a model can be constructed using conditional statements as follows (shown graphically in Figure 9A):

```
if performance < 3 then
  % merit increase = (.0375 * performance) - .0375;

if performance > 2 and performance < 6 then
  % merit increase = .075;

if performance > 5 then
  % merit increase = (.0375 * (performance - 4)) + .0375;
```

One interesting facet of this model is that it is essentially nonlinear, however, a simple linear approximation of this function (shown in Figure 9B) yields a remarkable degree of fit (a coefficient of determination equal to .914). This serves as an illustration of the fact that nonlinear relations can often be approximated with linear models. However, despite the close fit of a linear approximation, the substantive inferences that we might make from this model concerning the manager's judgment policy would be very different if we looked only at the linear approximation versus the more accurate nonlinear representation.



9A.



9B.

<u>Figures 9A and 9B.</u> % merit increase as a nonlinear function of performance and a linear approximation of the nonlinear model.

If we now take the model represented in Figure 9A and extend this back out to two dimensions, where tenure is taken into account, a model such as the following might result (shown graphically in Figure 10):

if performance < 3 then
% merit increase = .80 * (.0375 * performance - .0375) + .20 * (.0125 * tenure - .0125);

if performance > 2 and performance < 6 then
% merit increase = (.8 * .075) + (.2 * (.0125 * tenure - .0125));

if performance > 5 then
% merit increase = .8 * (.0375 * (performance - 4) + .0375) + .2 * (.0125 * tenure - .0125);

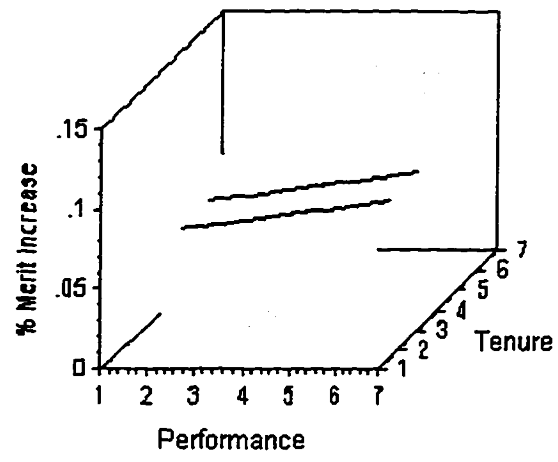


Figure 10. % merit increase as a nonlinear function of performance and tenure.

In looking at Figure 10 it is apparent how slight modifications to a model can result in a surprising degree of complexity. The complexity of this model made it difficult to capture using conventional linear techniques and further modifications to this model may be even more difficult. Moreover, it would be difficult to see how these equations work to someone who did not witness the model building process.

The Fuzzy Approach. A much less cumbersome approach to modeling or capturing the manager's pay allocation strategy in the preceding problem might involve capturing essential relationships while leaving aside somewhat arbitrary factors. For example, if we make a list of the commonsense elements of the manager's judgments we might include the following relations:

1) if performance is poor then the merit increase is low

2) if performance is moderate then the merit increase is moderate

3) if performance is excellent then the merit increase is high.

If we wanted to include the effect of tenure of the merit pay allocation then we might add the following relations:

4) if tenure is low then the merit increase is low

5) if tenure is high then merit increase is high.

If we combine these lists of "rules" together we might generate something like the following:

1) if performance is poor and tenure is low then the merit increase is low

2) if performance is moderate then the merit increase is moderate

3) if performance is excellent or tenure is high then merit increase is high.

The three rules above form the basis of our modeling solution. These three rules also happen to define the rules for a fuzzy logic system. If we can now give mathematical meaning to the linguistic variables (e.g., "high", "moderate", "low"), then we can generate a complete fuzzy system which will allow us to model and predict the manager's judgments. There are of course other elements to a fuzzy system and a more detailed description of this methodology is provided below, however, the essential mechanics of fuzzy logic are not exceedingly complex and the promise of this methodology lies in its adaptability, simplicity, and modeling power.

Fuzzy Systems Theory

Social scientists have adopted the language of mathematics for the following reasons: first, to handle logical complexity in areas of research, secondly, to make explicit and accurate relationships between phenomena of interest, and lastly, to employ mathematical models as decision aids (Smithson, 1985; Craiger & Coovert, 1994). An interesting problem that results from the use of mathematics in any area of scientific inquiry, is the need to move back and forth from qualitative language (i.e., natural language) to quantitative language (i.e., mathematics). Linguistic terms such as "excellent", "high", and "poor" reflect a degree of lexical imprecision that accompanies the type of language used in everyday conversation. This degree of imprecision results from the fact that almost all of everyday reasoning is approximate in nature (Zadeh, 1992). Because of this lexical imprecision, there is a basic mismatch between the levels of precision in quantitative analysis and the qualitative conclusions based on that analysis. For example, at what magnitude does a correlation change from indicating a bivariate relationship which can be characterized as "moderate" to a relationship which can be characterized as "strong"?

The problem of translation between qualitative language and quantitative language results from the fact that scientists often attempt to quantify, explain, and predict relationships or phenomena using precise, crisp categorizations. For example, as psychologists, we employ measurement systems that typically ask individuals to assign a single rating on a five- or seven-point Likert scale or we represent a person's attitude as a single summated score. In essence, these forms of measurement may be forcing a degree of precision that is not warranted given the vague and imprecise nature of many of the constructs we are intending to measure. In fact, methodological experts such as Cohen (1994) have decried the fact that statistical measures rich in information such as confidence intervals have not been employed, possibly, because they are often "embarrassingly large" (Cohen, 1994).

When attempts are made to deal with the vagueness or uncertainty inherent in modeling real-world relationships, psychologists, like many scientists, make use of probability. However, use of probabilistic models may necessitate making certain assumptions related to crisp logic (e.g., the law of excluded middle) that may not fit the real-world (Kosko, 1992). Also, it has been suggested that probability is not appropriate for representing all types of uncertainty. For example, regularly in human functioning uncertainty emerges due to the abstract or "subjective" nature of thoughts and concepts. This type of uncertainty may not be congruent with "randomness" as implied by probability theory (Jang & Sun, 1995; Smithson, 1988). Consider how different managers may implicitly define "good" job performance in different ways. There is a certain degree of subjectivity or uncertainty attached to these implicit definitions but the uncertainty may have little to do with randomness.

Fuzzy theory provides a set of tools for assessing real-world uncertainty. Lotfi Zadeh, the inventor of fuzzy set theory, suggested that as systems got more complex precise statements had less meaning. He later called this the <u>incompatibility principle</u>, which translates loosely into: "precision up, relevance down" (Kosko, 1993). The basic idea that conventional mathematics (e.g., precisely defined points, functions, sets, probability measures, etc.) are inadequate for describing complex systems (e.g., biological systems) prompted Zadeh to generate an alternative form of mathematics which began with his theory of fuzzy sets (Zadeh, 1965) and later generalized into fuzzy logic and fuzzy systems. Fuzzy set theory and fuzzy logic provide a system of mathematics that maps directly into natural language. This offers a method of capturing complex interactions between variables in qualitative descriptions that lend themselves to everyday reasoning.

It is worthwhile to note that fuzzy theory is not the only mathematical theory devised to deal with uncertainty. In fact, a number of theoretical frameworks have been proposed, including: classical probability, Bayesian probability, Hartley theory based on classical sets, Shannon theory based on probability, and Dempster-Shafer theory (Giarratano & Riley, 1994). A particularly zealous and long-standing debate has occurred between fuzzy theorists and statisticians (particularly Bayesian statisticians), concerning whether fuzzy theory captures aspects of uncertainty that cannot be captured with probability. While this debate currently rages on, the current stage of fuzzy systems development has not been driven by theoretical advances but instead has been driven by successful applications of fuzzy theory, particularly in Japan (McNeill & Thro, 1994). As noted by some authors (e.g., Giarratano & Riley, 1994), the major benefit being derived from such debates has been a re-examination of the foundations of probability theory and an increased interest in methods for dealing with uncertainty.

Model-free Estimation: A Case for Fuzzy Systems. Artificial intelligence (AI) can be looked at as both the study of intelligent computer behavior and as a name given to a number of advanced computing techniques (Munakata, 1994). Present AI developments that have been extensively applied or are promising areas of research include all of the following: knowledge engineering systems, machine perception, neural and fuzzy systems, and models of the brain and evolution. Techniques such as fuzzy systems, neural networks, and neuro-fuzzy systems (a combination of fuzzy system and neural network technologies) all can be characterized as "soft computing" techniques, which denotes their suitability to complex problems or applications that involve approximate reasoning (Zadeh, 1994). This connection to approximate reasoning is one of the main linkages between soft computing techniques and psychological and organizational research.

A key concept in soft computing and a concept potentially generalizable to psychological research is the idea of model-free estimation and prediction. Kosko (1992) in discussing soft computing techniques states the following: "Intelligent systems adaptively estimate continuous functions from data without specifying mathematically how outputs depend on input" (p. 19). Essentially this statement refers to the ability of fuzzy systems to map an input domain X (e.g., predictors) to an output range Y (e.g., criteria) without denoting the form of the function $f: X \rightarrow Y$. Moreover, it has been demonstrated that fuzzy systems are "universal approximators" of continuous functions of a rather general class (Klir & Bo Yuan, 1995). Because of this distinction as model free-estimators and universal approximators, modeling techniques such as fuzzy models have an innate freedom from a priori assumption of the type of relationships that may exist between variables (e.g., linear, nonlinear). Inherent in the claim that model-free estimation is an advantage is the belief that some of the relationships of interest to psychologists are more complex than the normally assumed linear form. Considering the wealth of information gathering in the physical sciences that many real life systems function through complex nonlinear interactions, the claim of ubiquitous linearity in psychology seems counter-intuitive (Abraham, Abraham, & Shaw, 1990; Barton, 1994). This same argument would seem to generalize to human judgment processes and as stated by Goldberg (1971), "some type of nonlinear and/or noncompensatory models should eventually prove superior to the linear model – at least for some judges with some sorts of inferential tasks" (p.459).

I noted earlier that mathematical models are often used as decision aids. Consider how decision aids are utilized in organizational contexts. Often, as in the case of personnel selection or assessment, an individual is given scores on a number of "dimensions" which reflect differentiable characteristics related to a criterion of interest (e.g., performance, potential, etc.). Within the context of decision making, it is necessary to combine several dimensions to compare applicants or candidates (Schmidt & Kaplan, 1971). Traditionally, linear combinations of these dimensions are employed, often in the form of weighted composites where the weights are derived from an empirical analysis using statistical models such as multiple linear regression. In applications such as assessment centers, statistical composites generally show higher validity than "clinical" judgments (Borman, 1982). Although linear regression is a useful tool for building statistical composites, linear regression does not always yield weights which optimize predictive power (i.e., maximize $\underline{R}^2$) when the functional form of the underlying relationships departs from the linear additive model. Mapping networks, such as neural networks and fuzzy systems, have more general functional forms than regression and particularly in high-dimensional spaces (input dimensions greater than 3) may outperform regression techniques which often fail to produce an appropriate approximation (Hecht-Nielson, 1990). In research looking at the performance of mapping networks in relation to the prediction of workplace behavior, mapping networks have shown mixed results (Dobbins & Coovert, 1992; Collins & Clark, 1993; Craiger & Coovert, 1994). It is likely that at least one reason for these mixed results is variation in the type of data used as input to the mapping networks. In response to this, we can ask a more fundamental question concerning the application of fuzzy systems to issues in psychology and organizational research; can fuzzy theory help us at a more foundational level of psychological measurement, for example, at the level of scale construction? This issue is addressed in following sections of this paper but first I review the fundamental concepts of "fuzzy sets".

Fuzzy Sets. A fuzzy set is basically a set whose members belong to it to some degree, for example, the set of highly effective managers at a plant. This set has fuzzy boundaries because individual plant managers have a graded membership in this set based on their level of effectiveness. Fuzzy sets, as opposed to crisp sets, have a gradual transition from membership to non-membership in the set. In mathematical terms, a fuzzy set is either a curve with a set of n fit values associated with it that indicate the degree of membership for each member, or a point in a hypercube where the vertices of the cube define nonfuzzy sets (Kosko, 1993). Formally, a fuzzy set can be represented as:

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

where X is a collection of objects denoted generically by x and where set A in X is defined by a set of ordered pairs. $\mu_A(X)$ is called the membership function of x in A. The membership function maps each element of X to a continuous membership value between 0 and 1.

Membership degree in a fuzzy set (i.e., fit value) is specified as a real-number on the interval [0, 1] with 0 indicating the element does not belong to the set and 1 indicating that the element belongs 100% to the set. In essence, the membership function defines the shape of the fuzzy set. Klir & Bo Yuan (1995) use as a simple illustration of fuzzy sets, a finite universal set that consists of seven levels of formal education, including:

0 - no education
1 - elementary school
2 - high school
3 - two-year college degree
4 - bachelor's degree
5 - master's degree
6 - doctoral degree

If we attempt to capture the concepts of little-educated, highly-educated, and very highly-educated, we can define three fuzzy sets that represent these linguistic concepts. Consequently, a person who has a bachelor's degree but no higher degree may be viewed as highly educated to the degree of 0.8 (having a membership degree of 0.8 in the highly educated fuzzy set) and very highly educated to the degree of 0.5 (having a membership degree of 0.5 in the very highly educated fuzzy set). Figure 11 shows how these three fuzzy sets are defined using triangular and trapezoidal membership functions, which are the most commonly employed types of functions used in current applications.

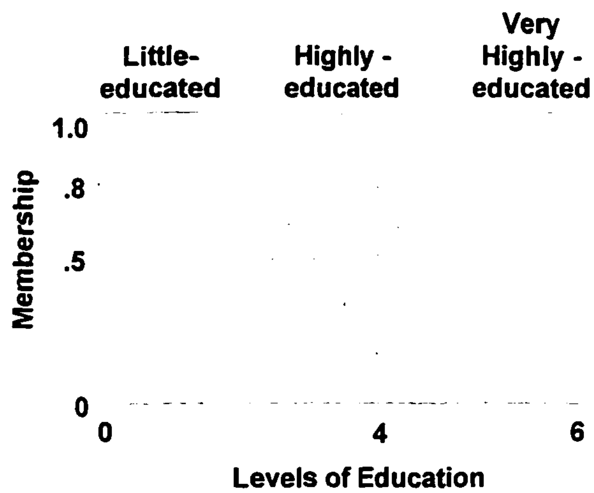


Figure 11. Levels of education as three fuzzy sets.

The Link Between Rating Scales and Fuzzy Sets. Early in the development of psychological and attitudinal rating scales, pioneers such as L. L. Thurstone put forth the idea that an individual's opinion could be characterized by more than a single point estimate response, for example, one could look at the range of opinions or options which a person was willing to endorse (Hesketh, 1988). This idea also appears in the discussion of "latitudes of acceptance and rejection" as discussed by Sherif and Sherif (1970). Historically, the measurement of ranges and asymmetries in dealing with rating scales has been somewhat neglected because of the lack of a parsimonious mathematical system for analyzing these types of measures. Lately, a number of authors have suggested that fuzzy set theory may be the key to addressing these earlier notions regarding the information content of ratings (Hesketh, Pryor, & Hesketh, 1988; Hesketh, Pryor, & Gleitzman, 1989; Pryor, Hesketh, & Gleitzman, 1989; Hesketh, Elmslie, & Kaldor, 1990; Hesketh, McLachlan, & Gardner, 1992; Hesketh & Gardner, 1993). Given the highly conceptual nature of many psychological constructs, the use of "fuzzy" rating scales may actually increase the congruence between the psychological representation of constructs (i.e., latent variables) and the associated mathematical mappings or measures.

Recently, researchers such as Hesketh, Pryor, Gleitzman, and Hesketh (1988) have applied fuzzy set theory to the traditional scale development methodologies of semantic differential and graphic rating scales. In an attempt to define measures related to Gottfredson's circumscription/compromise theory of career development, these researchers had subjects indicate a most preferred point on a graphic rating scale and then asked respondents to extend the rating to the left and right (if they wanted to). These extensions were interpreted as representing the imprecision or level of tolerance of an individual's attitude towards the item. An important point to note is that the researchers found a moderate to high degree of test-retest reliability (generally between .5 - .9) for subjects using these extensions. By making certain simplifying assumptions, not uncommon to fuzzy set theory, the researchers argued that the graphic rating could be viewed as a fuzzy variable (see Figure 12, adapted from Hesketh, Pryor, Gleitzman, & Hesketh, 1988).

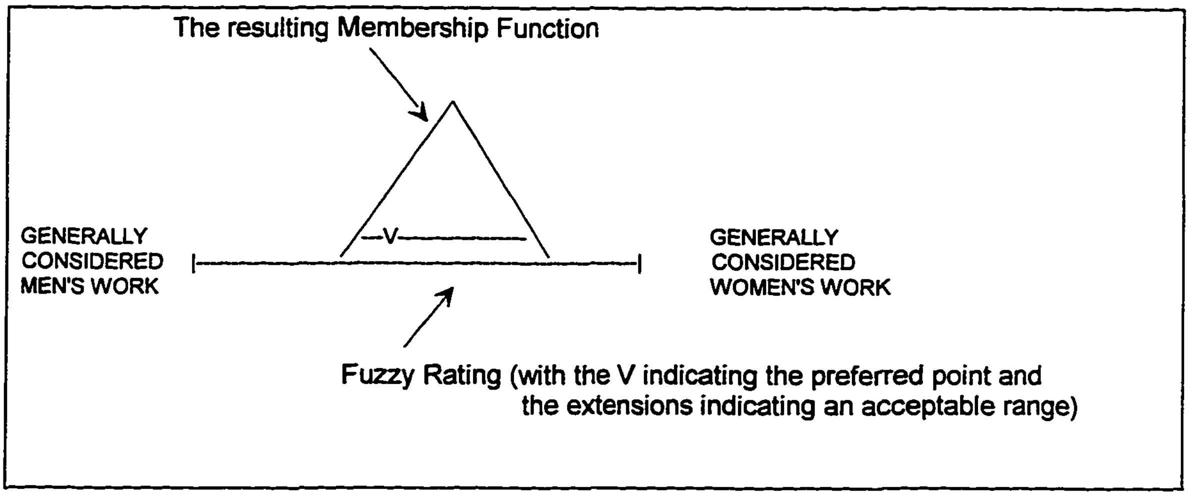


Figure 12. A fuzzy graphic rating scale.

Fuzzy set theory poses an interesting and important question to researchers interested in measuring real-world attributes. This question is: how do we appropriately quantize a variable, to reflect the appropriate level of coarseness in the real world attribute (Klir & Bo Yuan, 1995). For example, ever-present measurement error often makes finely tuned scales and measures unrealistic. Because of this, an important concept is "opted uncertainty" (Klir & Bo Yuan, 1995). Opted uncertainty does not come about due to information deficiency but comes about for pragmatic reasons. For example, consider the use of a seven point rating scale which specifies that seven "states" are available as markers for quantizing a judgment. If the actual human inference process utilizes a number of states and processes that are not congruent with the seven states we have made available, we have forced a certain amount of measurement uncertainty or "error" into our measurement. At times, it may be more pragmatic to reduce the information content of our measures to increase the approximation to the real-world, or stated another way we may "opt" for uncertainty.

Fuzzy Variables and Fuzzy Logic. A "fuzzy" or "linguistic" variable defines a "universe of discourse" containing a number of fuzzy sets (Kosko, 1992). For example, Craiger and Coovert (1994) described job experience as a fuzzy variable containing three fuzzy sets labeled low, moderate, and high (see Figure 13).

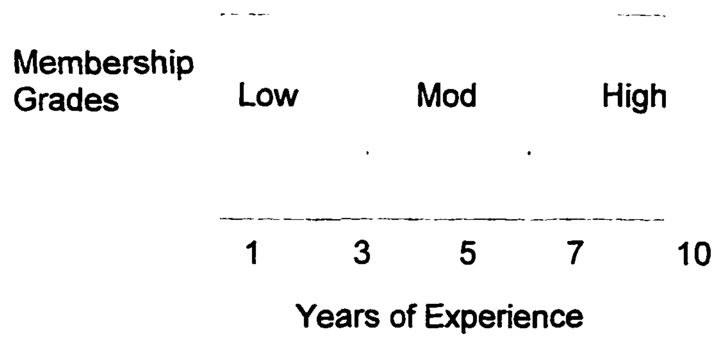


Figure 13. Job experience as a fuzzy variable.

Using this fuzzy variable, a person's job experience could be described in qualitative terms (i.e., low, moderate, and high) based on the fuzzy set that contains the highest membership function for the individual's scalar number of years. This mapping of experience into fuzzy sets allows us to talk about experience in natural language terms without specifying a point estimate at which low experience becomes moderate experience or a point at which moderate experience becomes high experience. This is because of the overlap of the fuzzy sets, which can be interpreted as the naturally occurring fuzzy boundaries inherent in a concept like job experience. A key concept that Figure 13 demonstrates is the idea of a linguistic variable. Zadeh (1973) proposed that key elements in human thinking are not numbers but fuzzy sets with verbal labels and that human reasoning is not based on traditional systems of logic but logic with fuzzy truths, fuzzy connectives (e.g., "and", "or", etc.), and fuzzy rules of inference. Zadeh argued that the mathematical definitions of fuzzy sets could be linked to natural language adjectives, which could be modified with verbal hedges (e.g., "very", "slightly", etc.), and that logical connectives such as "and" and "or" could be defined as mathematical operations on the fuzzy sets, hence the term "fuzzy logic".

Fuzzy If-Then Rules. Newell and Simon (1972), in their eminent work Human Problem Solving, demonstrated that much of human problem solving could be expressed as IF-THEN types of production rules. This finding not only had a significant impact on the field of psychology but also helped launch the field of intelligent expert systems. Expert systems have been implemented in a number of contexts and environments, including the increasing appearance of expert systems for many human resource functions ranging from recruitment to compensation (Greenlaw & Valonis, 1994).

Traditionally, expert systems have been rule-based. They are typically developed through the process of knowledge engineering whereby engineers acquire, store, and process rules as symbols through the use of a programming language such as LISP or PROLOG. These traditional expert systems suffer from several problems including the fact that human experts are often needed to articulate the propositional rules that approximate their expert behavior and the symbolic processing normally used in expert systems prevents direct application of numerical mathematics (Kosko, 1992).

Fuzzy systems also use rules but they incorporate the notions from fuzzy set theory that there is often uncertainty and imprecision in real-world reasoning. Fuzzy systems make use of fuzzy logic as a calculus for how fuzzy sets can be combined (Craiger & Coovert, 1994). An example of a fuzzy system is a fuzzy associative memory (FAM) (Kosko, 1992). In general, FAMs encode a bank of compound FAM rules that associate multiple output or consequent fuzzy sets with multiple input or antecedent fuzzy sets. These FAM rules can be treated as linguistic conditionals which allow us to interpret and obtain structural knowledge (Kosko, 1992). For example, suppose that we wish to develop a computerized system to help us make personnel selection decisions based on a structured interview. Assume that we have established three main dimensions on which the structural interview is scored, including: communication skills, personal impact, and persuasiveness. Then presuppose that these dimensions have been developed into rating scales which represent fuzzy variables, including fuzzy sets labeled as "very low", "low", "moderate", "high", and "very high". Also assume that a single output fuzzy variable has been defined for actual job performance, including fuzzy sets labeled as "high", "moderate", and "low". Following the definitions of the fuzzy variables, we employ the FAM methodology to create a mapping from the inputs (i.e., independent variables = communication skills, personal impact, and persuasiveness) to the outputs (i.e., dependent variable = job performance) (see Figure 14). This FAM model would require either sample data that includes values for both input and output data, for example a validation dataset, or verbal rules as articulated by experts. This system could be used to predict candidate future job performance, as well as allowing us to interpret the combination of characteristics or "rules" that result in highly effective job performers, for example: If Communication Skills are HIGH and Personal Impact is MODERATE and Persuasiveness is HIGH then Performance is HIGH.

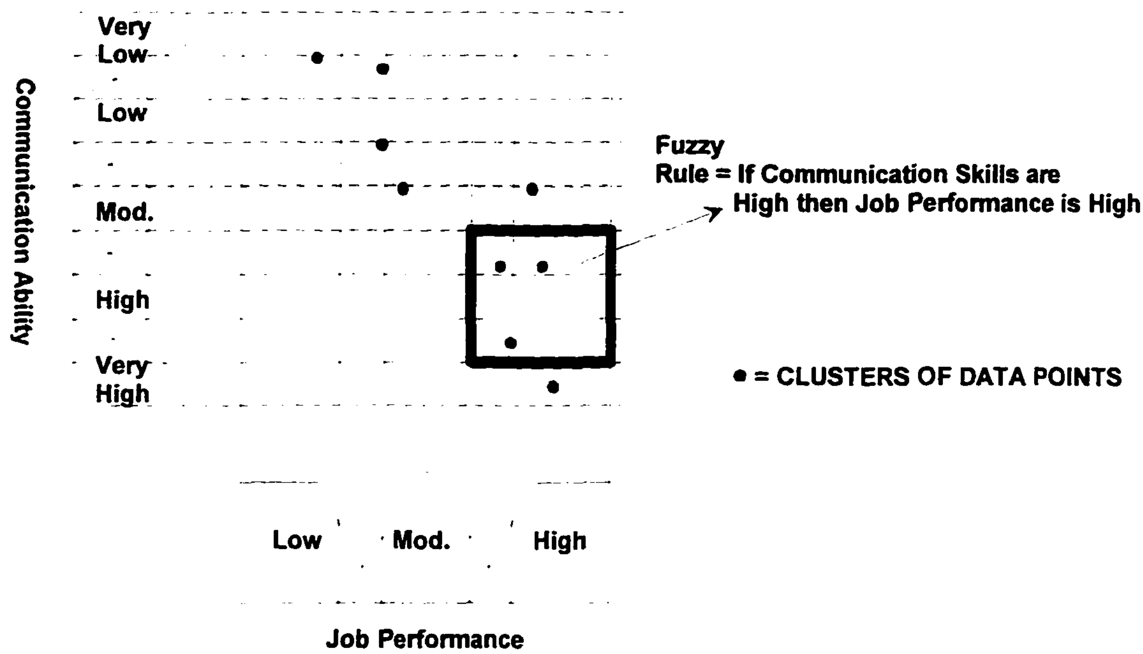


Figure 14. An example of a fuzzy if-then rule.

Another important feature of interpreting fuzzy if-then rules is the ability to look at one-to-many predictions versus one-to-one predictions. Consider the case where we have data on pairs of values for two variables X and Y. There are three possibilities for "if-then" and "if-and-only-if" propositions linking the phrases "X is high" and "Y is high". These propositions and corresponding diagrams of how X and Y might be related are shown in Figure 15 (adapted from Smithson, 1985). Inspection of this figure will reveal that popular measures of association such as correlation coefficients can only be used to test the "if-and-only-if" case while ignoring the other possible relations (Smithson, 1985). The rules generated from the calculus of fuzzy logic can be used to test one-to-many types of propositions.

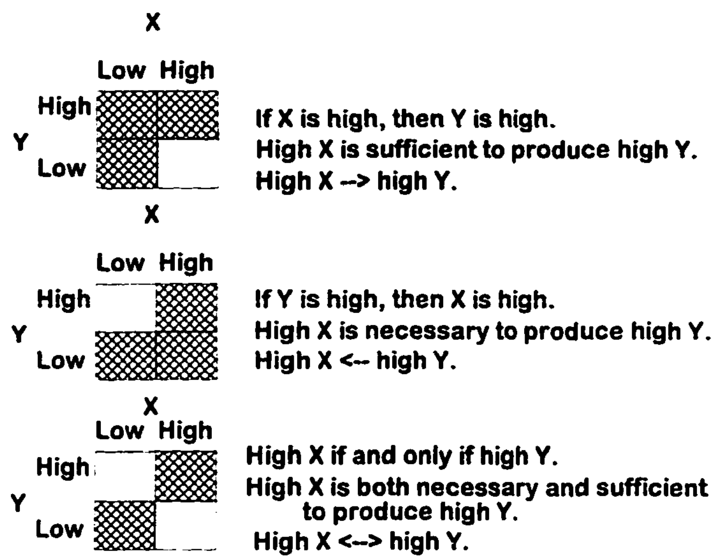


Figure 15. X-Y Relations (from Smithson, 1985).

Fuzzy Inferencing Using If-Then Rules. Although a complete treatment of the mathematics behind fuzzy sets and fuzzy logic is outside the scope of this paper, I will briefly review how fuzzy if-then rules can be combined in a process called "inferencing" to reach a conclusion. Inferencing is a fundamental process that underlies the functioning of a fuzzy expert system. Assuming that all variables have been converted to linguistic variable values, the fuzzy inference procedure is used to identify the fuzzy rules that apply to the current situation (i.e., current input values) and to compute the values of the output linguistic variable. Note that fuzzy sets are extensions of ordinary sets and fuzzy logic is an extension of ordinary logic and just as there are correspondences between ordinary sets and ordinary logic, similar relations exist within fuzzy logic (Munakata & Jani, 1994). In general, the membership degree of an element in a fuzzy set may be linked to a truth value of a proposition in fuzzy logic (Munakata & Jani, 1994).

The computation of a fuzzy inference consists of two parts, including the "aggregation" (computation of the IF part of the rules) and the "composition" (computation of the THEN part of the rules) (von Altrock, 1995). Each rule associated with a fuzzy system describes an action to be taken in the THEN part and the degree to which the action is taken depends on how valid the rule is for the current situation, defined in the IF part. The validity or how adequate the rule is for a situation is computed by the aggregation, or combining of antecedent conditions, in the IF part. In conventional logic, the combination of conditions can be computed by a Boolean AND. Similarly, operators such as AND and OR are defined in fuzzy logic. However, since fuzzy logic deals with degrees of truth, the connective AND and OR in fuzzy logic can correspond to the mathematical operations of MIN (minimum value) and MAX (maximum value) respectively. For example, consider the following if-then rules:

Rule 1: If Performance is HIGH and Tenure is HIGH then Merit Increase is HIGH

Rule 2: If Performance is LOW and Tenure is HIGH then Merit Increase is LOW.

Let the current input values for an employee be Performance = 6 and Tenure = 10 years, then membership values for the value 6 in fuzzy sets associated with performance might be .8 for the fuzzy set HIGH and .2 for the fuzzy set LOW, and membership values for the value 10 in fuzzy sets associated with Tenure might be .8 for the fuzzy set HIGH. To compute the degree to which each of these rules fit the current employee we combine the two antecedent fuzzy set membership values with the AND operator (correspond to the MIN operation), which results in the following:

Rule 1: $\min\{0.8, 0.8\} = 0.8$

Rule 2: $\min\{0.2, 0.8\} = 0.2$

These results can be described as the degrees of truth associated with each rule for the current situation, so Rule 1 is more valid for the current situation.

In the composition step (corresponding to the THEN consequent) we must calculate the action to be taken based on the previous aggregation results. Therefore, Rule 1 results in the consequence action Merit Increase is HIGH to the degree 0.8, and Rule 2 results in the action Merit Increase is LOW to the degree 0.2. Because multiple rules govern this system, these two results must be combined. To achieve this combination another operation is invoked, in this example utilizing the operator OR (corresponding to the MAX operation), to combine results as follows:

For the linguistic variable Merit Increase, the fuzzy inference result is:

HIGH	to the degree of 0.8	
MODERATE	to the degree of 0.0	(= MAX = 0.8)
LOW	to the degree of 0.2	

These results are used to define the height of the fuzzy consequent sets which are eventually combined and reduced to a single output value through a process called defuzzification.

Building Fuzzy Systems. Fuzzy systems have been developed and utilized in a wide variety of fields. In fact, because of success in industrial applications in countries such as Japan, fuzzy systems research and application has become big business, with research groups such as the Laboratory for International Fuzzy Engineering Research (LIFE) boasting boards of directors that contain the presidents of the Hitachi, Toshiba, Nissan, Minolta, Matsushita, and Fujitsu corporations (Kosko, 1993). Some rough generic categories of fuzzy system application would include: control (e.g., automobiles, consumer electronics, robotics), pattern recognition (e.g., OCR, audio, signal processing), quantitative analysis (e.g., operations research, statistics, management), inference (e.g., expert systems for diagnosis, planning, and prediction, intelligent interfaces, software engineering) and information retrieval (e.g., databases) (Munakata & Jani, 1994). Because of the multidisciplinary nature of fuzzy systems they have been labeled with many different names, including: fuzzy-rule-based system, fuzzy expert system, fuzzy model, fuzzy associative memory, fuzzy logic controller, and simply fuzzy system (Jang & Sun, 1995).

In general, a fuzzy expert system is composed of five functional blocks (see Figure 16), including: a rule base containing a number of fuzzy if-then rules, a database which defines the membership functions of the fuzzy sets used in the fuzzy rules, a decision making unit which performs the inference operations on the rules, a fuzzification interface which transforms the crisp inputs into degrees of match with linguistic values, and a defuzzification interface which transforms the fuzzy results of the inference into a crisp output (Jang, 1993).

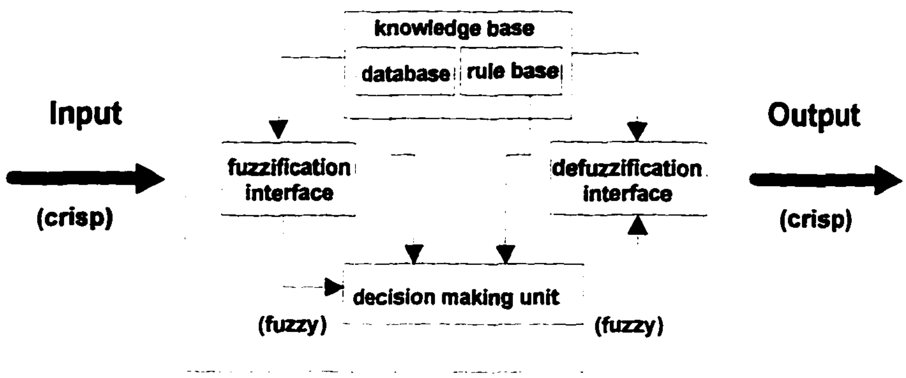


Figure 16. Fuzzy expert system (adapted from Jang, 1993).

A general overview of the steps needed to build a fuzzy system, as described by Klir and Bo Yuan (1995), are discussed below.

First, the relevant input and output variables are defined. In the case of policy capturing, these variables would correspond to the cues and subjective judgments, respectively. Next the ranges of both input and output variables are identified and each variable is expressed as a linguistic variable, where the continuum of each variable is expressed as an appropriate number of overlapping fuzzy sets. In most cases these fuzzy sets are fuzzy numbers which represent linguistic labels such as low, high, moderate, etc. Choosing the appropriate number of fuzzy sets for a linguistic variable is currently somewhat of a trial and error process. In many cases heuristics are used to establish a starting point for modeling efforts, such as starting with three terms for each linguistic input variable and five terms for each linguistic output variable (von Altrock, 1995). In fact, the "structure determination problem", which deals with the partitioning, the number of membership functions for each input, and the number of fuzzy if-then rules, is now an active area of research (Jang & Sun, 1995). Employing fuzzy sets allows a researcher to represent the implicit uncertainty or imprecision that accompanies human information processing. Often in practice, the membership functions associated with the fuzzy numbers for each linguistic variable are expressed as triangular membership functions, although the membership functions can take many different forms (trapezoidal, sigmoidal, bell-shaped, etc.). It is also possible to later modify the membership functions to maximize their representation of the data by using neural network learning methods (e.g., Jang, 1993).

In the second step, a fuzzification function is defined for each input variable to "express the associated measurement uncertainty" (Klir & Bo Yuan, 1995). The purpose of these functions is to interpret values for the input variables (i.e., expressed as real numbers) as realistic fuzzy approximations as defined by the linguistic variables.

Third, the knowledge relating inputs to outputs is formulated in terms of a set of fuzzy inference rules. There are two principal ways in which these rules can be determined; first, by eliciting them from a human expert, or second, by obtaining the rules from a set of statistically representative empirical data with the help of neural networks. In general, these inference rules take the form of: IF e = A and f = B then y = C, where A, B, and C are fuzzy numbers chosen from the set of fuzzy numbers that represent the linguistic states. The possible rule combinations that are derived from combining the various fuzzy numbers can be represented in matrix form. In estimating an output from input values, these rules fire to different degrees depending on the degree to which input values activate the rules in the rule bank. Because fuzzy systems allow knowledge to be expressed either from human experts or from empirical data, some interesting applications to judgment research can be inferred. For example, the specification of fuzzy if-then rules may be a new format in which to compare subjective and objective judgment policies.

The fourth step involves making inferences (i.e., predictions) regarding the output variables from the input variable values in combination with the relevant fuzzy information rules. A number of inferencing systems are possible, an example of which is Kosko's (1992) fuzzy associative memory (FAM).

In the last step, the researcher or system designer must select a suitable "defuzzification" method. The purpose of this method is to convert each conclusion obtained by the inference engine, expressed in terms of fuzzy sets, into a single real number, which acts as an estimate of the output value. Because a single real valued estimate is obtained, the results from fuzzy modeling can be compared with other modeling techniques such as multiple regression.

One area of research and development relating to fuzzy modeling methodologies is adaptive or automatic model generation. The ANFIS methodology (Adaptive-Network-based Fuzzy Inference System), developed by Jang (1993), is an example of such adaptive modeling tools. This methodology employs a hybrid learning procedure which can be used to refine fuzzy if-then rules obtained from human experts or, if experts are not available, clustering algorithms can generate a set of fuzzy if-then rules to approximate a desired data set and then ANFIS can be used for tuning. This methodology was demonstrated by Jang (1993), who successfully modeled a series of nonlinear functions and chaotic time series predictions with the ANFIS methodology. Adaptive development of fuzzy models holds much promise for the future of this technology.

Fuzzy Systems and Decision Making Research. Early in the history of fuzzy theory Bellman and Zadeh (1970) suggested a link between fuzzy set theory and human decision making. This link led to the translation of classical decision making paradigms involving actions, goals, and constraints into fuzzy models. This initial research has branched into several areas of decision making including individual decision making, multiperson decision making, multicriteria decision making, multistage decision making, and fuzzy ranking methods (Klir & Bo Yuan, 1995).

Despite extensive integration of fuzzy theory into decision making research, little has been done to establish a link between fuzzy theory and policy capturing. Because of this, I propose in this paper that a judgment policy can not only be represented in terms of numerical estimation but instead can be represented as structured knowledge, in the form of if-then types of rules. This proposal follows from research which demonstrates that through the use of linguistic labels and membership functions, fuzzy if-then rules can capture the "spirit of a rule of thumb used by humans" (Jang, 1993). This form of representation affords a tool for studying information processing behavior that combines the power of representation afforded by numerical estimation techniques while also facilitating interpretation in terms of natural language as typified by symbolic processing systems (Kosko, 1992). Or stated another way, this format may answer the call by researchers to combine statistical and rational approaches to defining judgment policies (Hobson & Gibson, 1983). Representing functional aspects of human information processing in a rule based format follows a line of reasoning originated by Newell and Simon (1972). Because of the power of fuzzy systems to model arbitrarily complex relations between inputs and outputs, fuzzy systems hold promise as a new methodology for judgment and decision making research.

Note that conventional statistical models also can be used to model complex judgment policies, for example, the scatter coefficient derived from the scatter model (Brannick & Brannick, 1989) can express evidence of a noncompensatory strategy in terms of a conjunctive-disjunctive continuum, where a zero coefficient indicates a linear strategy (Ganzach & Czaczkes, 1995). However, the complexity of using and interpreting conventional models is greater than the fuzzy system methodology proposed here and, conventional techniques are limited to detecting the types of functional relationships that are built into the modeling equation. For example, one term may be needed to indicate a multiplicative relation, another term may be needed for log-based relation, and still another term may be needed to indicate configural integration. Similarly, if policy capturing is used in conjunction with training efforts aimed at the consistent use of a given policy it may be necessary to give the judges feedback concerning aspects of their policies as captured from empirical data. There is an obvious difficulty in trying to explain to a judge, unfamiliar with statistics, aspects of a judgment policy which include statistical concepts and terminology. In fact, an entire area of research has emerged that addresses the effects of providing judges with "cognitive feedback", that is, information concerning relations in the environment, relations perceived by the person, and relations between the environment and the person's perceptions (Balzer, Doherty, & O'Conner, 1989). It may be the case that the type of information represented in fuzzy if-then linguistic rules can be used as a new form of cognitive feedback that can be easily understood by judges. Moreover, in a recent review of research looking at the effects of cognitive feedback on performance, Balzer and colleagues (1989) suggested that "another area in which research effort might make an impact is the development of methods for representing uncertainty" (p. 429). This stands as another impetus for research on the role of fuzzy theory in judgment research.

I have suggested that fuzzy if-then rules may serve as an alternative indication of a judge's policy, where certain combinations of antecedent and consequent fuzzy set labels can indicate either linear or nonlinear noncompensatory processes. For example, one plausible noncompensatory strategy that may be used in merit pay allocation is the conjunctive strategy. As described earlier, in a conjunctive process, a low value on any dimension leads to a low value on the criterion. Characterized in a different way, this is the "negativity bias" that has been documented in the performance evaluation literature (Ganzach, 1995). The conjunctive process is nonadditive, in that, a low value on any dimension cannot be offset by a high value on another dimension. For example, if we extract the following fuzzy if-then rule from a pay allocator's judgment set (assuming three fuzzy sets equivalent to the linguistic labels of HIGH, MODERATE, and LOW): if Tenure is HIGH and Recipient Importance is HIGH and Perceived Need is HIGH and Performance is LOW then Merit Increase is LOW, it is plausible that the allocator is using a conjunctive policy. This rule may denote a conjunctive policy because all of the factors are high or would favor a higher merit increase except for performance. Hence, the one low value on performance overrides the other cue values and the recipient is given a low merit pay increase. It is worthwhile to note that despite the form of a judgment model, either statistical weights or linguistic rules, these models represent "paramorphic" representations (Hoffman, 1960), which cannot tell us the exact process used by a judge but can only be used to predict responses thereby facilitating some interpretation on the part of the researcher. It is also worthwhile to note that inspection of a single empirically-derived linguistic rule may not yield sufficient insight into a judge's judgment strategy. Instead, inspection of a set of empirically-derived rules may be needed to draw meaningful inferences. It is critical to note at this point that any discussion of the types of inferences that can be made from fuzzy if-then rules concerning judgment policies is speculative. It may be that different types of fuzzy models (e.g., Mamdani, Sugeno, Tsukamoto) yield more appropriate forms of information in terms of interpretation. It is also equivocal at this point whether adaptive or data driven fuzzy models can yield sets of fuzzy if-then rules which are strictly interpretable. These factors all result in the need for exploratory analyses such as those discussed in this paper.

As noted earlier, another potential use of fuzzy systems is the exploration of linguistic if-then rules as a tool for eliciting subjective policies and comparing these rules to if-then rules derived from an empirical dataset. It is interesting to note that successful fuzzy models have been built from rules verbally obtained from human experts. For example, the Mamdani Fuzzy Model (Mamdani & Assilian, 1975) was first proposed as an attempt to control a steam engine and boiler combination by a set of linguistic control rules obtained from experienced operators. The ability to use linguistic information for modeling purposes is a specific feature of fuzzy systems. Note that use of linguistic information assumes that human operators can summarize their actions as a set of fuzzy if-then rules with approximately correct membership functions. Successful applications of human-determined models suggests that, at least for control actions, experienced operators can provide this roughly correct information (Jang & Sun, 1995). This stands in contrast to studies looking at human-determined policy capturing models which suggest that human judges cannot accurately articulate their decision making models. The disparity between these findings suggests that further research is needed on the nature of expert knowledge.

One potential advantage of fuzzy modeling approaches to eliciting and using subjective information is that the fuzzy approach does not confound cues with levels of cues. Specifically, traditional approaches to eliciting subjective impressions of judgment policies often ask judges to indicate the importance or influence of the specific cues or variables in making judgments. This is frequently accomplished by asking the judges to rate or rank the cues based on the perceived importance or influence of the individual cues in terms of their impact on judgment. In completing these tasks, judges focus on the cues as a whole rather than focusing on levels of the cues. Thus, for example, the impact of a cue like "individual performance" might be assessed in terms of the influence of performance information on judgments rather than asking the judge whether low performance or high performance had differential impact. Because a judge may potentially weight negative versus positive information differently, asking about the importance of a cue as a whole may not be as meaningful as asking about the influence of levels of a cue. In contrast, fuzzy if-then rules specify a grammar for mapping levels of cues to levels of judgments, in a manner which can incorporate variations in cue weightings schemes.

As previously discussed, another issue to be resolved in judgment research are the problems encountered in analyzing correlated cue sets. Because the fuzzy system methodology is not founded on the same assumptions as parametric statistics, the fuzzy methodology may not impose limitations to interpreting models with correlated predictors. However, much more research is needed to determine the effects of various input variable characteristics on fuzzy system development and performance.

The Current Study

Having provided a summary of previous merit pay research, a survey of conventional judgment analysis techniques, and an overview of fuzzy systems; the following section describes the key design elements of the current study.

The judgment of interest in this study is managerial merit pay allocation, that is, the amount of a pay increase that a judge feels a particular recipient profile warrants.

The intent in this study was to create fictional, yet representative and realistic, recipient profiles to be used as judgment stimuli. The use of these fictional profiles may incite criticism based on the "paper people problem" as it has been noted in the policy capturing literature (Gorman, Clover, & Doherty, 1978). The debate concerning the artificiality of using "paper people" has proponents on both sides of this argument. Justification for use of this approach in this study is founded on first, the almost insurmountable difficulties in designing a merit pay judgment study using actual recipients and actual allocators which has any degree of experimental control, and secondly, in a review of over thirty years of policy capturing studies, Brehmer and Brehmer (1988) conclude that "Thus far, then, it seems that the paper format as such does not lead to any important distortions in the policies obtaining in policy capturing studies..." (p. 89).

Choices concerning the inclusion of cues in a policy capturing study represent critical decision points (Stewart, 1988). Given the proliferation of variables believed to affect pay allocation decisions, the process of choosing variables is a difficult one. Typically, cues are included in a judgment study based on subjective criteria, often founded on judge verbal reports. While useful, this procedure does not ensure that variables of theoretical and conceptual interest are examined. Also, the use of verbal reports as criteria for variable inclusion may be suspect given the previously cited wealth of information concerning the inability of allocators, and judges in general, to accurately articulate key aspects of decision processes (Sherer et al., 1987, Slovic &

Lichtenstein, 1971). Due to these constraints, variables in this study were chosen on the basis of potential theoretical and conceptual importance as well as their importance in establishing representative and realistic profiles. Figure 17 shows the variables selected as cues for this study and a brief discussion of the cues follows.

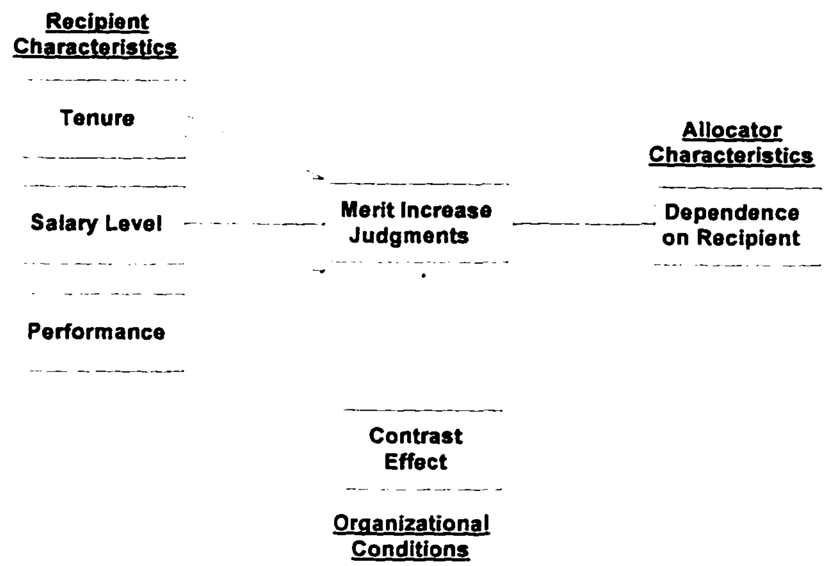


Figure 17. Judgment analysis model.

The set of variables shown in Figure 17 includes three of the four factors noted in the previous literature review and in the Heneman (1990) conceptual model of the merit pay process. The inclusion of variables representative of different factors increases the utility of this model in terms of contrasting the impact of different types of variables on pay judgments. The inclusion of variables from each of these factors also represents an advance in the design of policy capturing studies in the merit pay research area. The only factor included in the Heneman (1990) framework but not represented in this set of variables is environmental conditions. Environmental variables were not included given the difficulties in manipulating variables such as the presence or absence of unions, the effects of the labor market, and the effects of the product market in a single fictional organization, which was the context for this study. The choice of variables was also guided in part by methodological constraints associated with policy capturing. Specifically, in choosing the number of cues to manipulate, considerations must be taken into account which include sampling error, more specifically the number of cases to independent variables (Nunnally, 1978), and the fact that results from studies using different methods generally show that judges use a small subset of the cues available (Brehmer & Brehmer, 1988). Also, because this study involved exploratory analyses of fuzzy modeling techniques, high dimensionality would have been problematic because of the prohibitively large number of if-then rules needed for modeling efforts. This constraint is often characterized as the "curse of dimensionality" (Jang & Sun, 1995).

Representing recipient characteristics, Figure 17 shows that this study included: recipient tenure, the recipient's salary level, and the recipient's overall rated performance as cue manipulations. Recipient tenure was chosen as a manipulation based on its importance for establishing realism and representativeness in the profiles as well as the need to further explore the effects of tenure (Dreher, 1981; Kaun, 1984). Providing information relevant to tenure may increase profile realism due to the fact that tenure may be viewed in some organizations as a component of overall performance (Heneman, 1990). Also, tenure is often considered in compensation decisions due to the fact that tenure may serve as a cap on merit pay, due to the relationship between tenure and status in the pay grade (Heneman, 1990). Possible interactions between tenure and other relevant variables also need to be explored. Salary level was also included in this variable set due to its possible influence on merit pay decisions. The comparative salary ratio is an indication of the employee's status in their respective position's pay grade. As discussed earlier, previous research has established some relation between comparative salary ratios and merit pay, and there are intuitive and substantial reasons to expect these variables to be related. Also, information on comparative salary ratio seems highly relevant to establishing realism in the recipient profiles. Salary level in this study was indicated in terms of a comparative salary ratio (i.e., compa-ratio), which was chosen over just reporting the recipient's current salary level because of the increased information available to the judge in terms of a frame of reference as to how high or low the current salary is.

Overall recipient performance was also included in the cue set, due to its obvious relevance to making merit pay allocation decisions.

Given the lack of a knowledge base concerning the effects of allocator characteristics, a variable relating to the allocator was included in the variable set. The allocator characteristic noted in Figure 17 is allocator dependence on the recipient. Research on this aspect of merit pay allocation has emerged from Bartol and Martin's (1988) dependency perspective, which indicates that merit pay increases are influenced by the allocator's dependence on the recipient. This perspective has received some support (Deshpande & Joseph, 1994) but additional evidence is needed. The degree that allocator's depend on recipients has been manipulated in a previous policy capturing study by manipulating the importance of the recipient's job in meeting departmental goals, in terms of either "important" or not "important" (Deshpande & Joseph, 1994). Inclusion of this type of information in the recipient profiles may not be unrealistic, in that, this rating may be perceived as related to, but not the same as, overall performance.

The final variable noted in Figure 17, contrast effect, falls under the rubric of organizational conditions. The term "contrast effect" corresponds to characteristics of a work group's composition that may influence pay allocation. An example of this effect is given by Ivancevich (1983) who found that in a sample of engineers, merit increases were larger the greater the proportion of unsatisfactorily performing engineers. Because there has been very little research concerning this effect on pay allocators, this variable was included in the current study. Also, given the fact that context effects have been demonstrated in both personnel interview contexts (Wexley, Yukl, Kovacks, & Sanders, 1972) and in performance appraisal contexts (Grey & Kipnis, 1976), further demonstration of this effect in the merit pay domain may serve to document the ubiquitous nature of this effect.

Hypotheses

Having delineated a set of cues and discussed some of the design elements of the current study, I will now review research hypotheses that were directed specifically at the conceptual and methodological issues raised in the previous sections of this paper.

The first issue to be addressed is the comparative modeling power of the fuzzy system methodology versus models derived from conventional linear and nonlinear statistical approaches. To address this issue, the following hypotheses were offered:

H1 - In fitting various models to the judgment data generated from the merit pay allocation judgment task, empirical fuzzy system models will perform better, in terms of fit, than models generated from multiple linear regression.

H2 - Empirical fuzzy system models will also perform better than or equivalent to nonlinear regression models.

These two hypotheses reflect a research approach of comparative model estimation that is similar to Goldberg (1971), Brannick and Brannick (1989), and Ganzach and Czaczkes (1993). In hypothesizing that a fuzzy modeling approach will perform better in terms of model fit than a linear regression model, I am suggesting that the judges' merit pay allocation strategies may include nonlinear noncompensatory components that cannot be adequately captured in a linear additive model. A rationale for this hypothesis follows from the perspective that a number of particular types of nonlinearity and noncompensatory strategies appear to be theoretically and empirically viable as components in pay allocation strategies. For example, it is plausible that the weight attached to a performance measure will fluctuate depending on the level of performance in a referent work group (i.e., an individual by group performance interaction), or possibly, that a negativity/positivity bias will exist resulting in a conjunctive or disjunctive inference process. Note, that whenever the impact of one cue depends on the level of one or more other cues, the inference process is nonadditive or "configural" in nature. Also, it seems plausible that higher order function forms (e.g., quadratic, cubic terms) may be important components in modeling pay allocation judgment strategies. This follows from evidence emerging from the utility analysis literature that documents the existence of s-shaped curves which describe the relation between individual performance levels and the perceived value of that performance to the organization (Bobko, 1995). While plausible nonlinear noncompensatory processes in merit pay allocation can be specified, the exact nature of the actual strategies is hard to determine a priori. Therefore, this research is exploratory in nature.

A rationale for hypothesis two is evident when considering the nature of nonlinear regression approaches. Specifically, a common approach to capturing nonlinearities using regression requires the researcher to a priori specify the type of relation expected (e.g., curve components, interactions, scatter terms, etc.). Given the current lack of knowledge about nonlinear noncompensatory processes in areas such as merit pay, it is unlikely that an a priori identification of model terms will lead to a properly specified model. In contrast, fuzzy systems function as universal approximators, and as such, they should theoretically facilitate the modeling of both linear and nonlinear functions of arbitrary complexity. In terms of more direct evidence of nonlinear and noncompensatory judgment strategies, the following hypothesis was also specified:

> **H3** - Inspection of both fuzzy system rules (i.e., both empirical and subjective) and nonlinear regression models will reveal evidence of nonlinear noncompensatory allocation strategies for some of the participants (i.e., allocators).

Another issue addressed in this research is the role of fuzzy systems as tools for exploring subjective judgment policies. Specifically, the following hypothesis was specified:

> **H4** - Fuzzy system models constructed solely from subjective information directly elicited from the participants will perform more effectively than models based on a traditional subjective policy capturing approach.

A rationale for this hypothesis follows from the fact that the often cited inability of judges to identify aspects of their judgment policies (e.g., Brehmer & Brehmer, 1988; Zedeck & Kafry, 1977) is bound to the methods used to elicit subjective impressions of judgment policies. As noted earlier, traditional approaches to subjective policy capturing have inherent limitations which are not shared by fuzzy modeling methodologies. Specifically, it was expected that when subjective policy information was collected and implemented in a fuzzy system framework, models could be constructed that would outperform traditional approaches to estimating subjective policies.

METHOD

Participants

The participants for this study were 10 managers who had actual experience in making pay allocation decisions. This number of participants is similar to previous research efforts (e.g., Sherer et al., 1987) in which real-world managers have been used to study individual level judgments. Note that the number of participating individuals in these studies is constrained by the significant time commitment needed from the participants in order to make a large number of judgments across a series of profiles. Also, because the modeling efforts discussed here were targeted towards individual level analyses, using a large number of judges would have been computationally prohibitive.

All of the participating managers completed the study on a completely voluntary basis. Each of the participants were told that after completion of the study they would receive a customized executive summary of the research. Participants were chosen on the basis of convenience sampling, and as noted above, the only criterion for inclusion was having actual experience relevant to making pay allocation decisions.

In reviewing the demographic and background data on this set of participants, the following characteristics are revealed:

- 30% of the participating managers were female;
- the participating managers were highly experienced, with 40% of the managers having over 15 years of experience in allocating or making recommendations relevant to pay, an additional 40% having between 12 and 15 years of experience, 10% having between 9 and 12 years of experience, and 10% having between 6 and 9 years of experience;
- the participating managers held various positions and had varying organizational affiliations, including: a purchasing agent at a large manufacturing company located in the southeastern

United States, a vice-president of human resources at a large manufacturing company located in the southeastern United States, a vice-president of operations at a large manufacturing company located in the southeastern United States, a regional vice-president of a national property management company, an administrator in a public sector legal authority located in the southeastern United States, a vice-president of marketing at a large manufacturing company located in the southeastern United States, a vice-president of sales at a large manufacturing company located in the southeastern United States, the director of a public sector legal authority located in the southeastern United States, an executive vice-president of a printing company located in the southeastern United States, and a general manager for a company in the construction industry located in the southeastern United States.

Apparatus

Specialized software was used for the fuzzy expert system development. This software included the MATLAB computing environment and the associated MATLAB Fuzzy Logic Toolbox. The software was run on an IBM-compatible PC platform.

Procedures

All data were collected through research packages that included: a cover letter introducing the study, detailed instructions to the participants, the judgment task, and a post judgment questionnaire (see Appendix 1). All of the participants were given approximately one month to complete the research package. The research packages were personally delivered to the participants, and after completion they were returned through the mail. Participants were encouraged to contact the researcher if they had any questions. Several of the participants were interested in discussing the study after completing the research packages, and debriefing interviews were setup for those participants.

Judgment Task and the Judgment Context

Establishing a judgment scenario and context for the participants in this study involved instructing the participants to assume that they have just been hired into a new managerial position for a fictional company (i.e., "the Personnel Solutions Corporation", a management consulting firm, see Appendix 1). They were then instructed to look at a set of personnel profiles (i.e., the judgment or recipient profiles) and make merit pay increase recommendations for each employee to their direct superior. Justification for making these judgments was provided to the participants in the scenario by suggesting that their superior was interested in seeing "how they handled compensation issues", and that their superior was interested in establishing a rough estimate of the cost of the merit pay program in their department. Instructions and information were provided to the participants that included an overview of the company, a discussion of the type of data included in the personnel profiles, a brief statement about the company's pay policy, and the necessary instructions for making the merit pay allocations (see Appendix 1). Included in this information was an overview of the company's merit pay program. Information about the program included: the form of the increases (i.e., increases were based on a percentage of base salary), the frequency of the increases (i.e., annual merit reviews), a typical average merit increase for the company (i.e., an average of 4% was cited, which was based on realistic merit pay increase values (Flannery, Hofrichter, & Platten, 1996)), the role of cost of living factors, and the role of managerial discretion in merit pay allocation. A set budgetary constraint, in terms of the total amount of money the participants could allocate, was not imposed due to the fact that the participant's allocation judgments were framed in terms of "recommendations" to their direct superior and as such, the total amount that they allocated was to be viewed as a "bottom-up" estimate of the total cost for the merit pay plan.

The specific judgment task the participants were asked to complete involved allocating to each of the fictional employees a specific amount of a merit pay increase. Each of the participant's merit pay allocation decisions were indicated on a scale of 0 - 15% increase, which is typical of the ranges used in both research and in organizational settings (Deshpande & Schoderbek, 1994).

Cue Variables

As discussed earlier, five variables were used as cues in this study. These variables included: performance, group performance (i.e., contrast), importance (i.e., dependency), tenure, and salary level (i.e., compa-ratio). Because some of these cues had naturally occurring concrete units of measurements, such as years of tenure and salary level, establishing quantitative scale values for these cues was not problematic. In the case of the two variables requiring abstract scale units, performance (of both the individual and the referent group) and importance, the advice of Stewart (1988) was followed in terms of establishing meaningful anchors for the judges. Past policy capturing studies looking at pay allocation have used somewhat restricted cue values, generally using only "presence/absence" or "high/average" types of labels (Deshpande & Joseph, 1994; Sherer et al., 1987). Use of these labels is often done to facilitate orthogonal coding of the cues; however, use of these labels would appear to limit the information available to the judges, potentially impacting how judges weight and integrate the cues in forming an overall judgment. Cues in this study were continuously valued.

Following the suggestions of Stewart (1988), graphic bar chart scales were used to display cue values (see Appendix 1). Graphic cue presentation formats hold advantages such as the fact that they are clearly and easily readable, and they show cue values relative to the cue's total range and relative to the values of the other cues (Stewart, 1988).

The cues in this study were operationalized as follows:

1) <u>Performance</u> - Performance was manipulated as ratings on a seven point graphic/behavioral expectation type of rating scale. The use of this scale was intended to anchor performance information, in terms of providing a common frame of reference, while maintaining a high degree of realism due to the fact that rating scales are frequently used in organizational performance measurement systems (Murphy & Cleveland, 1991). Because many companies that implement merit pay plans attempt to tie pay increments to a goal setting or management by objectives program, anchors for the performance scale reflected relative degrees of goal and objective accomplishment (see Appendix 1).

2) <u>Group Performance</u> - A contrast effect was manipulated within the scope of this study by including in the recipient profiles an indication of the overall performance of the employee's referent work group (i.e., the mean individual overall performance rating for the recipient's work group or department). The scaling of this variable followed that of the individual performance measure (see Appendix 1).

3) Importance - Similar to Deshpande and Joseph (1994), dependency was manipulated by assigning a rating on a graphic/behavioral expectation type of rating scale that was associated with the recipient's importance in accomplishing managerial/departmental goals. Anchors for this scale were generated in congruence with the sources of dependence identified by Bartol and Martin (1988) (e.g., specialized skills) (see Appendix 1).

4) Tenure - Tenure was operationalized as the number of years the employee had worked for the company. The tenure variable was scaled to have a general range between 1 and 10 years. Tenure information was also presented using a graphic scale (see Appendix 1).

5) Salary Level - Salary levels are often defined in terms of an index such as a comparative salary ratio, which is the current salary level of the employee divided by the recommended midpoint for the employee's salary grade or position. In a similar approach, salary levels for the mock employees in this study were indicated on a graphic scale, which showed the employee's current salary level graphically contrasted against the recommended midpoint for that employee's position (see Appendix 1). For use in numerical analyses, the salary level was defined as the difference between the current salary and midpoint salary values, which is functionally equivalent to a compa-ratio. Current salaries in the study were scaled to have a general range between $27,000 and $39,000, which is similar to previous research studies (Deshpande & Joseph, 1994). In terms of recommended midpoints, five salary midpoints were chosen ($29,000, $31,000, $33,000, $35,000, and $37,000). These five midpoints essentially defined five different pay grades. The fictional employees were evenly divided among these five pay grades. The use of different salary midpoints was done to increase profile realism.

Profile Construction

In determining the number of profiles generally needed for a policy capturing study, Stewart (1988) recommended the following design: judging 50 profiles followed by a break (five minutes to one week), followed by judging 25 cross validation cases, followed by judging 25 repeated cases from the first 50 (for the purpose of estimating reliability). Thus, based on this design, 100 profiles would be presented to the judges. Roughly following this design, and in accordance with the need to have sufficient modeling power, the need for over-determined models in terms of the number of fitting parameters to data points, and considering the attention and time demands potentially placed on the participants, a target number of 110 profiles was established for the current study. This number of profiles included the repetition of 20 profiles, scattered throughout the research package, for the purposes of estimating the consistency or reliability of the participant judgments. Because of the relatively large number of profiles to judge, participants were encouraged in the judgment task instructions to take breaks as needed.

Given that there were five cue variables to present in each profile, a key design question in this study was how to combine cue values in order to generate cue profiles. As noted earlier, policy capturing studies often feature profiles with orthogonal cues due to the difficulty in analyzing intercorrelated cue sets with multiple regression. The use of orthogonal cue sets may be problematic in the merit pay domain, as it is in other domains, given the somewhat intuitive relationships between pay allocation cues. For example, one probable pair of correlated cues is current base salary and tenure. The principle of representative design dictates that cue intercorrelations should match those that exist in the environment (Stewart, 1988). This is often a problematic design aspect, in that, estimates of the environmental correlations are not often available. To address this issue in the current study, a profile generation approach was used that follows from work done by Naylor and Wherry (1965) and Wherry, Naylor, Wherry, and Fallis (1965). This approach entails defining a referent structure, which in the case of the current study was a correlation matrix, and generating any number of scores following the population parameters specified in the referent structure. Below I outline the steps involved in generating profiles for the current study.

Step One. As an initial step in this approach, three highly experienced Industrial/Organizational psychologists were used as subject matter experts to estimate the correlation between each of the variables to be used as cues in the current study. For example, they were asked to generate a general estimate of the correlation between individual performance and group performance. They were told to make these judgments in terms of non-specific or meta-analytic estimates, thinking of their estimates as values generalized across organizational contexts. These individual correlational judgments were then averaged to yield the following cue variable correlation matrix.

Table 1

Referent Cue Correlation Matrix

	Performance	Group Performance	Importance	Tenure	Salary Level
Performance	--				
Group Performance	.433	--			
Importance	.600	.433	--		
Tenure	.353	.250	.517	--	
Salary Level	.567	.283	.683	.517	--

The average inter-judge correlation among the three subject matter experts in making these correlational judgments was .58 (alpha = .80).

Step Two. The above referent correlation matrix was then used as input to a program written by Aguinis (1994), which was used to generate multivariate random normal scores with the given intercorrelations noted in the referent matrix. One hundred and ten of the random normal score vectors served as the basis for the judgment profiles used in the current study.

Step Three. After generating the profiles, each of the cue values that were initially in the form of normal scores (i.e., $\underline{z}$ - scores) were transformed into the appropriate scales by employing linear transformations using the following means and standard deviations: performance mean = 4.0, performance standard deviation = 1.0; group performance mean = 4.0, group performance standard deviation = 1.0; importance mean = 4.0, importance standard deviation = 1.0; tenure mean = 5.5, tenure standard deviation = 1.5; salary level mean = 33,000, salary level standard deviation = 2,000. Essentially, the means used in these transformations corresponded to median values for the expected cue value ranges, and the standard deviation values were chosen so the that the majority of scores (roughly out to three standard deviations) would fall into the expected cue value ranges. While the profile generation procedures used in this study in no way guarantee accurate representation, they are superior to the a priori assumption of cue orthogonality.

Model Development

The key analyses in the current study focused on comparing and contrasting various models of the merit pay allocation judgments. Below, an overview of the various modeling approaches is provided. Note that each of these modeling approaches involved building a model for each individual participant, which attempted to capture the idiosyncratic features of that manager's merit pay allocation judgment strategy.

Modeling Approach 1 - Linear Regression. Following the traditional approach to policy capturing, each manager's set of merit pay allocation judgments was regressed on to the set of cues using linear regression models. For each of the regression models used in this study, information concerning the fit of the model was determined through inspection of $\underline{R}^2$ and $\underline{R}^2$ change values, which are generally interpreted in terms of variance accounted for in a set of judgments. Also, as is typical in traditional policy capturing studies, indices relating to relative cue importance were assessed through the analysis of standardized regression coefficients.

Modeling Approach 2 - Nonlinear Regression. In the current study, four different regression models employing terms used for capturing nonlinear or noncompensatory judgment strategies were included. Note that in this section, I use the term "nonlinear" in a general sense indicating the presence of additional terms used in regression equations. While not critical in this study, in other settings in may be consequential to draw distinctions between nonlinear models that are "intrinsically linear" versus those that are "intrinsically nonlinear" (Jang, Sun, & Mizutani, 1997).

The first of these nonlinear regression models involved the use of polynomial terms for the performance cue variable. Specifically, higher order performance terms including performance^2 and performance^3 were included in regression models. As discussed earlier, the possible importance of these factors was based on previous research, which suggests a role for second and third order polynomial terms in modeling systems similar to merit pay allocation such as utility analysis, where performance is assessed as a covariate to perceived value. Also, focus was placed on higher order terms for the performance cue rather than focusing on other cue variables, due to the likelihood that performance information would have the greatest impact on the managers' merit pay allocation judgments, as has been demonstrated in previous studies (e.g., Deshpande & Schoderbek, 1993). To test the impact of the higher order terms, the significance of the polynomial terms was evaluated hierarchically; that is, first Y was regressed on X, then on X and X^2, then on X, X^2, and X^3 (Cohen, 1978).

A second nonlinear noncompensatory regression model used in the current research involved the use of the scatter model (Brannick & Brannick, 1989). The basic version of this model is defined as:

$$\mathrm{Y}s = b_0 + \sum_{i=1}^{k} b_i X_i + b_{k+1}[\sum_{i=1}^{k} (Z_i - \overline{Z})^2]^{\frac{1}{2}}.$$

This model emerged from the idea that more information is contained in patterns of cue values than is contained in weighted sums of the values (Brannick & Brannick, 1989). The scatter term, which is added to a regression model, captures the deviation of profile scores about the profile mean. Because the scatter term assesses the impact of information concerning particularly low or high cue values (low or high in contrast to the mean of the other cue values within a given profile), when significant, this term is interpreted as indicating reliance on disjunctive or conjunctive judgment rules (Ganzach, 1995). Specifically, when a scatter term is positively related to judgments then a disjunctive rule is thought to be operating; whereas, with a negative scatter term a conjunctive rule is indicated (Ganzach, 1995). Because the scatter model takes into account deviations for each cue value, all of the cue dimensions are influential in assessing the impact of the scatter term. In evaluating the fit of a scatter model, the significance of the scatter term coefficient is assessed. The test for the scatter coefficient informs the researcher as to whether the scatter term has any unique contribution to the prediction of a set of judgments, over and above all other predictors. Note that the scatter model is a viable alternative model for merit pay allocation judgments, given evidence from closely aligned areas of research such as performance evaluation, which suggest reliance upon disjunctive/conjunctive judgment strategies.

A third nonlinear regression modeling approach employed the use of interaction terms. This approach represents one of the most common uses of nonlinear terms in the social sciences (Bobko, 1995). Interaction terms correspond to the product of two variables, which suggests that the linear function between a given X_1 and Y depends on levels of a second variable, X_2. Use of such interactive terms is fundamental to moderator analyses (Bobko, 1995). Within a judgment analysis context, the inclusion of interaction terms speaks to the issue of configural judgment. Interaction terms seem viable as components in merit pay allocation judgment models, especially given the likely dependencies that exist between the cues of interest. For example, as previously cited, individual performance levels may have a differential impact on pay allocation strategies depending on factors such as the performance of a referent work group. In the current study, interaction terms were included representing interactions between individual performance and each of the other cues. The focus on interactions with the performance cue follows from the previously discussed expectation that performance information would play a dominant role in influencing merit pay allocation judgments. In this study, as suggested by Lubinski and Humphreys (1990), interaction terms were assessed concurrently with higher order polynomial terms, in an incremental stepwise fashion, to avoid interpreting spurious moderator effects.

In addition to the nonlinear noncompensatory regression models noted above, a last modeling approach was employed involving saturated regression models. The term "saturated" denotes that, in these models, all of the linear and nonlinear cue terms (noted above) were used in combination as predictors. These models assessed the possible increase in predictive power that might accompany using combinations of terms within a single model.

Modeling Approach 3 - Subjective Regression. A great deal of research has been directed towards attempts to assess the congruence between subjective estimates of the relative influence of different cues and statistical weights derived from procedures such as multiple regression (Cook & Stewart, 1975). The most typical procedure used for assessing subjective weights is to have each judge in a policy capturing study distribute 100 points across a set of cues in a manner to indicate the influence that each of the cues or sources of information had on their judgments (Cook & Stewart, 1975). These weights can then be used as if they were regression weights to yield predicted judgments. Summers, Taliaferro, and Fletcher (1970) used such a procedure to show that with a sample of graduate students engaging in a four-cue task, subjective weights accounted for 20% less of the variance in subjects' judgments than use of optimal linear regression weights.

In the current study, subjective estimates for cue weights were obtained using a 100 point distribution task, which was included in the post judgment questionnaire that was completed by each of the participating managers (see Appendix 1). Models were developed by essentially using these subjective weights as beta weights and obtaining predicted judgments by multiplying subjective weights by the standardized values of the cues (Cook & Stewart, 1970).

Modeling Approach 4 - Subjective Fuzzy Systems. In order to compare the subjective regression models noted above to a fuzzy system approach to building subjective models, fuzzy systems were built using data obtained directly from the participating managers. These models were of the Mamdani type, which uses rules of the following format: if x is A then y is B, where A and B are linguistic values defined by fuzzy sets (Jang & Sun, 1995). Participants in the current study were asked to complete a knowledge elicitation questionnaire (i.e., the post judgment questionnaire shown in Appendix 1), which was used in the construction of these fuzzy models. This questionnaire yielded information relevant to both membership function definitions and construction of a fuzzy logic rule base. Below, specific aspects of these models are reviewed.

First, in order to fuzzify each of the cue variables, a decision was made regarding how many fuzzy sets were needed for each variable. This decision rule specified three terms (i.e., fuzzy sets) for each input cue variable and five terms for the judgment output variable (von Altrock, 1995). Note that the decision to use three fuzzy sets for each cue variable was also driven by the number of combinations of membership functions for each input that might result in fuzzy if-then rule combinations. For example, with a large number of inputs, or with a large number of membership functions per input variable, the number of possible if-then rules grows prohibitively large. Consider a fuzzy model with 10 input variables and two membership functions (i.e., fuzzy sets) on each input. This model would result in $2^{10} = 1024$ fuzzy if-then rules (Jang & Sun, 1995).

Since human-determined membership functions may contain important knowledge obtained through experience and may be reflective of important subjective information, the participants were prompted to define membership functions, following a four step process described by von Altrock (1995). This process asked the participants to define for each linguistic term (e.g., "high", "moderate", "low") the value along a scale that best fits the linguistic meaning of the term (or the most "typical value") (see Appendix 1). From this information, standard membership functions were defined for each of the variables. The specific steps involved in defining the membership functions were as follows (von Altrock, 1995):

> 1) for each term, define the value that best fits the linguistic meaning of the term. This most typical value for each term gets the membership degree $\mu = 1$;
>
> 2) for each term set the members
>
> hip degree $\mu = 0$ where the terms next to it have their most typical value;
>
> 3) connect the point $\mu = 1$ with the points $\mu = 0$ by straight lines. This results in membership functions of Λ - type for the inner terms;
>
> 4) no terms lie beyond the rightmost term nor below the leftmost term; thus, points in these regions belong to the respective membership functions with $\mu = 1$ (p. 229-230).

In terms of using the knowledge elicitation questionnaire for developing a fuzzy logic rule base, the participants were asked to respond to a number of questions regarding the pay allocation actions that they took for a particular series of antecedent conditions. This follows an elicitation technique described by von Altrock (1995). This technique required the participants to choose the most plausible term for the "then" parts of rules (see Appendix 1). One question of this type was defined for each plausible combination of terms from the input variables (i.e., cues). Plausible combinations of terms were defined by examining the cue profiles. Specifically, the range of each cue variable was divided evenly into thirds, with each third being mapped to either the term "high", "moderate", or "low". For example, on the individual performance scale, the scale point "1" would be characterized as "low". Using these decision rules, 30 plausible rule antecedents were defined, and the participants were prompted to choose the appropriate consequent that would best represent their response to the given antecedent cue profile (see Appendix 1).

Note that because fuzzy systems can yield predicted values through the operation of an inference engine, indices such as $\underline{R}^2$ (i.e., the correlation between actual and predicted values) can be and were evaluated for these models. Some exploratory analyses were also conducted with the subjective fuzzy models involving the use of different types of membership functions (e.g., Gaussian shapes), different fuzzy operators or logical connectives (e.g., product, probor), different implication methods (e.g., product) by which the fuzzy set in the consequent is shaped, and different aggregation methods (e.g., sum) by which the consequents of each fuzzy rule are combined. Figure 18 shows a schematic overview of the structure of the fuzzy system models built using subjective information.

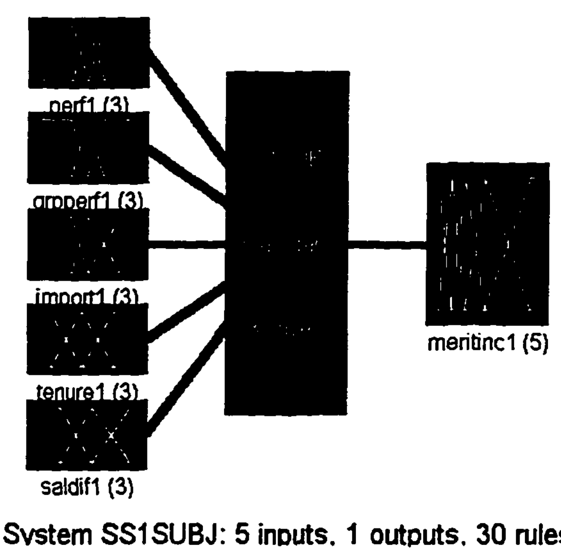


Figure 18. An overview of a subjective fuzzy system model.

Modeling Approach 5 - Empirical Fuzzy Systems. In addition to the fuzzy system models built solely from subjective data, a set of fuzzy system models were generated for the participants using the actual cue values as input variables and the actual managers' judgments as the output variable. Consequently, these fuzzy models were similar to the linear and nonlinear regression approaches in that they were data-driven or based on empirical data.

The empirical fuzzy system models developed in this study used a different type of fuzzy inferencing system from the subjective fuzzy models that were based on the Mamdani style of inference. The empirical models were instead based on the Sugeno style of inference, due to the particular methods used for extracting and tuning the fuzzy rules. The Sugeno fuzzy model (also known as the TSK fuzzy model) was proposed as a way to systematically generate fuzzy rules from a given input-output data set (Jang & Sun, 1995). A rule in a Sugeno style system has the form: if x is A and y is B then $z = f(x,y)$, where A and B are fuzzy sets in the antecedent, while $z(x,y)$ is a crisp function in the consequent. Note that it is in the consequent that the Sugeno systems differ from the Mamdani system discussed previously. In the current study, the output function $f(x,y)$ took the form of a first order polynomial, thus yielding first-order Sugeno fuzzy models, with rules of the form: if x is A and y is B then $z = p*x + q*y + r$. It is noteworthy that these rules look remarkably like linear regression equations, and in fact, the Sugeno method is highly effective for smoothly interpolating multiple linear models (i.e., multiple Sugeno rules) in order to model nonlinear systems (Gulley & Jang, 1995).

In developing the empirical fuzzy system models, two major steps were involved that are highlighted below.

In the first step, a set of clustering algorithms that are part of the MATLAB Fuzzy Logic Toolbox were used to extract an initial set of fuzzy rules directly from the input (cues) - output (judgments) data for each manager. The clustering algorithms used include "subtractive clustering", which is a fast, one-pass algorithm used to estimate the number of clusters and the cluster centers in a data set, and "fuzzy c-means clustering", which defines data points as belonging to clusters to a degree specified by membership grades (Gulley & Jang, 1995). The idea of using clustering algorithms to define an initial set of fuzzy if-then rules is based on the assumptions that a) similar inputs to a system should produce similar outputs, and b) that these similar input-output pairs are grouped in terms of clusters in the data (Jang, Sun, & Mizutani, 1997).

After defining initial sets of rules for each participant, the ANFIS (Adaptive-Network-based Fuzzy Inference System) routine was used as a hybrid learning algorithm to tune the initial

Sugeno fuzzy if-then rules to more closely approximate a given data set. ANFIS uses the least squares method and the backpropagation gradient descent for linear and nonlinear parameters, respectively (Gulley & Jang, 1995). Because the ANFIS architecture is a powerful tool for modeling, a test or cross-validation data set was used as part of the ANFIS training. To create the test data set, the available data was divided using a 2/3 (training) to 1/3 (test) division, which is typical for this type of research (Flexer, 1996). Generally, what is desired is not only a fuzzy system that closely approximates a set of training data (used to develop the model) but a model with great generalization capability for data that the system has not been exposed to.

It should be noted that in developing the empirical fuzzy system models, a number of parameters associated with both the clustering and ANFIS routines had to be adjusted using a trial and error process. Even between managers, certain modeling parameters had to be changed to increase model fit. Specifically, for the clustering programs, radii values between .6 and .8 were used. The radii values correspond to a vector that specifies cluster center ranges of influence. For the ANFIS training, the number of training epochs was generally between 40 and 150 and step-size increase and decrease rates fell between 1.5 and 1.8. Also, the empirical Sugeno based fuzzy models used Gaussian membership functions. Figure 19 shows a schematic overview of the structure of the fuzzy system models built using the empirical data.

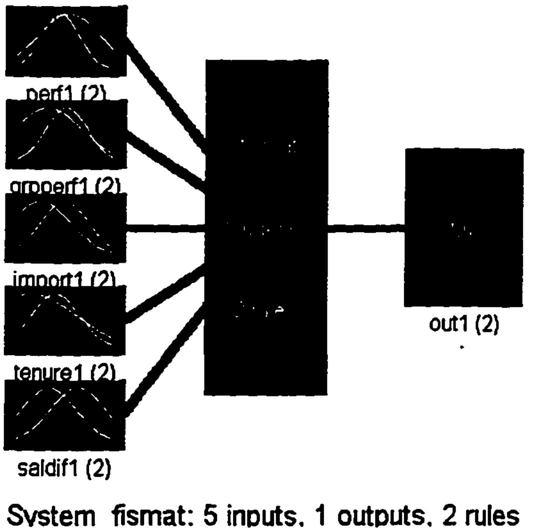


Figure 19. An overview of an empirical fuzzy system model.

Analyses

A number of descriptive analyses were conducted, describing features of the participants' judgments and characteristics of the models developed from the different approaches.

In order to test hypotheses one and two, model fit comparisons were conducted between the linear regression, nonlinear regression, and empirical fuzzy models using bootstrapped estimates of cross-validated multiple correlations. Typically, procedures for comparing model fit involve calculating the correlation between predicted and actual values for a model in a "holdout" or "test" data set. This correlation stands as an estimate of the cross-validated multiple correlation. As noted above, the key fit measure in the current study was also a cross-validated correlation between actual and predicted values. However, instead of relying on a single holdout sample, an approach known as "statistical bootstrapping" (Cooksey, 1996) was used in order to more accurately assess model fit. Statistical bootstrapping is a nonparametric technique which allow inferences to be made about parameters without the need to satisfy traditional statistical assumptions such as normality and homogeneity of variance (Cooksey, 1996). Thus, bootstrapping is appropriate for situations where formulas for standard errors do not exist (Cooksey, 1996). In the case of judgment analysis, bootstrapping offers a way to empirically estimate the sampling distribution of judgment policy characteristics through a process called "resampling". Resampling refers to repeatedly drawing random samples (with replacement) of profiles from a given data set and calculating the associated model characteristics (e.g., regression weights, fuzzy rules, $\underline{R}^2$ values). Because of the large number of samples drawn (e.g., usually around 1000), an empirical sampling distribution is formed that allows estimates of centrality parameters, the calculation of confidence intervals, and making statistical comparisons between statistics of interest. Moreover, when assessing generalization or cross-validation through bootstrapping, it is possible to avoid random influences on cross-validated statistics that may result from use of a single arbitrary division of a data set into test versus training data. In the current study, bootstrapping was used for estimating the cross-validated multiple correlations for both regression and fuzzy system models. A bootstrapping program written in MATLAB facilitated 1000 resamples using 2/3 - 1/3 data splits (for training versus test data) for the regression models, and 100 resamples also using 2/3 - 1/3 splits for the fuzzy system models. The reason that only 100 resamples could be achieved for the fuzzy system models has to do with the computational complexity of the clustering algorithms and ANFIS routine used to derive the empirical fuzzy models. However, even with only 100 resamples, a more stable estimate of cross-validated parameters can be computed than with traditional single sample cross-validation. Because the bootstrapping program yielded empirical sampling distributions, direct comparisons of the mean cross-validated multiple correlations (e.g., the centrality parameters of the empirical sampling distributions) were facilitated. Mean differences in cross-validated multiple correlations between models were also summarized for the entire group of managers using an inferential test (i.e., a t-test) after transforming the correlations using Fisher's $\underline{r}$ to $\underline{z}$ transformation (Ganzach, 1995).

It is important to note that when using linear regression, cross-validation does not have to be accomplished empirically. Instead, formula estimates of the validity "shrinkage" can be computed, and in fact, these estimates are generally accurate (Murphy, 1984). However, these formulas currently exist only for linear regression type models. For more complex nonlinear models or models based on methodologies outside the realm of parametric statistics, empirical cross-validation using techniques such as bootstrapping is the only alternative.

The evaluation of hypothesis three involved both qualitative analyses of the fuzzy system rules and quantitative analyses (i.e., significance tests and $\underline{R}^2$ increments) of the nonlinear regression models.

Testing hypothesis four also involved the comparison of estimated multiple correlations (i.e., correlating actual with predicted values), in this case comparing the fit or predictive power of the subjective fuzzy models to the fit of the subjective regression models. For the subjective models, multiple $\underline{R}$ and $\underline{R}^2$ values could be computed using the entire original data set, without having to cross-validate the results, since the subjective models did not use optimization methods (e.g., least squares optimization) that are sample dependent.

RESULTS

Descriptive Results - Merit Pay Allocation Judgments

Table 2 shows descriptive statistics for the managers' merit pay allocation judgments. Evident in this table is the fact that the mean amounts of merit increases hovered around the 4% value that was cited in the judgment scenario as representing the typical mean increase for the fictional company. Even though the mean increases across managers were in fairly close proximity, the differences in mean allocations were significantly different, with F (9,1088) = 13.478, p < .001, possibly suggesting differences in leniency/severity among the managers. One manager skipped two profiles, so these profiles were coded as missing data for that participant. Reliability estimates for the managers suggested a high degree of consistency in judgment, with an average reliability across the managers (computed on the 20 repeated profiles) of r = .95.

Table 2

Descriptive Statistics for Merit Pay Allocation Judgments (Across 110 Profiles)

Judge	Mean	SD	Min.	Max.
Manager 1	4.10	2.17	.00	14.00
Manager 2	4.97	3.46	.00	12.00
Manager 3	4.04	1.45	.00	10.00
Manager 4	3.85	1.48	.00	10.00
Manager 5	4.22	1.12	.00	8.00
Manager 6	5.76	3.21	.00	13.50
Manager 7	3.81	2.14	.00	12.00
Manager 8	4.07	1.62	.00	12.00
Manager 9	2.60	3.58	.00	15.00
Manager 10	3.73	1.61	.00	10.00

* these values represent percentages (of recipient salaries)

Table 3 reports correlations between cue values and the resulting merit pay allocation judgments. These correlations generally indicated positive relationships between the cues and resultant judgments. What these correlations or "cue dependencies" (Stewart, 1988) do not indicate is the importance of cues in determining judgments, due to the existence of cue intercorrelations.

Table 3

Correlations of Cues with Merit Pay Allocation Judgments Among Managers (Across 110 Profiles)

	Performance	Grp. Performance	Importance	Tenure	Salary Level
Manager 1	.950**	.383**	.586**	.437**	.461**
Manager 2	.738**	.581**	.673**	.510**	.593**
Manager 3	.902**	.570**	.647**	.511**	.544**
Manager 4	.810**	.221*	.646**	.507**	.513**
Manager 5	.906**	.588**	.761**	.628**	.615**
Manager 6	.873**	.542**	.578**	.450**	.489**
Manager 7	.851**	.296**	.692**	.506**	.513**
Manager 8	.901**	.248**	.489**	.325**	.327**
Manager 9	.639**	-.025	.306**	.262**	.117
Manager 10	.957**	.368**	.456**	.335**	.342**

** $p < .01$, * $p < .05$

Results - Hypothesis One

Hypothesis one dealt with the comparative modeling power of empirical fuzzy system and linear regression models. Specifically, it was hypothesized that the fuzzy systems would outperform linear regression models. This hypothesis was supported. As shown in Table 4, for every participant, fuzzy system modeling resulting in a superior mean cross-validated (bootstrapped) R value, with an average increase in R^2 of .04 across the 10 participants.

Table 4

A Comparison of Mean Cross-validated R Values for Linear Regression and Fuzzy Systems

Manager	1	2	3	4	5	6	7	8	9	10
Linear Regression										
Mean Bootstrap Cross-validated R	.965	.829	.934	.885	.985	.892	.920	.928	.740	.962
95% Conf. Intervals	.946 .980	.744 .896	.887 .966	.778 .949	.971 .996	.847 .930	.870 .953	.928 .957	.624 .846	.935 .976
Fuzzy System										
Mean Bootstrap Cross-validated R	.978	.863	.944	.940	.990	.946	.930	.935	.756	.969
95% Conf. Intervals	.953 .994	.726 .920	.884 .977	.862 .976	.975 .997	.891 .976	.863 .963	.870 .969	.556 .872	.952 .983

Table 4 also reports confidence intervals constructed using the "percentile method" (Cooksey, 1996), which consists of finding the upper and lower percentiles in the empirical bootstrap distributions that correspond to 5% of the scores (2.5% for each tail of the distribution). Note that these confidence intervals may differ from traditionally calculated confidence intervals to the degree that there are violations of parametric assumptions (e.g., normality requirements). Also, the confidence intervals for the fuzzy system models should be interpreted cautiously, given that only 100 resamples were used. While there appears to be some overlap between the confidence intervals for the linear regression and fuzzy system models, the centrality parameters (i.e., means) of the distributions are different for every manager, and the differences between the means from the two modeling approaches are large enough to meet a traditional criterion for significance, $t(9) = 4.30$, $p < .01$ (based on transforming the correlations to Fisher z scores). Another feature of these models is the extremely high R^2 values achieved by the modeling efforts. The magnitude of these values is similar to other efforts attempting to capture merit pay allocation decisions (e.g., Sherer et al., 1987). Note that the sample size for the correlations reported in Table 4 is 37, the sample size of the test data sets.

In order to further explore the fuzzy system and linear regression models, an analysis of the regression coefficients and Sugeno fuzzy rule parameters was undertaken. Tables 5 and 6 report model characteristics for one random training/test data split and resulting models, which are representative of the models used in the bootstrapping analyses. Table 5 presents the coefficients for the regression equations and fuzzy rules in their standardized form (i.e., based on z scores). Table 6 presents the coefficients in a standardized form in which the weight vectors for both the linear regression models and the fuzzy system models have been normalized. Normalizing a vector is achieved by dividing each component of a vector by that vector's length, which is equal to the square root of the sum of the squares of all the vector's components (Wasserman, 1989). Normalization was done to facilitate comparing the regression weights to the fuzzy rule weights by compressing the range of the coefficients to fall within the interval [0,1], while still indicating the general influence of the different cue variables.

Table 5

Fuzzy Rule and Linear Regression Standardized Weights (Computed on N = 73)

Manager		Performance	Group Performance	Importance	Tenure	Salary Level	R^2 For Model
1	Fuzzy Rule 1	1.883	-.436	.035	.074	-.373	.994
	Fuzzy Rule 2	-.294	.402	.346	-.217	.429	
Regression Beta Weights		.912 **	-.090 *	.211 **	.006	-.066	.932
2	Fuzzy Rule 1	17.701	-4.231	-42.392	-39.837	44.710	.909
	Fuzzy Rule 2	-2.471	-.727	.140	-1.932	.709	
	Fuzzy Rule 3	.167	-.766	-.307	.244	-1.288	
Regression Beta Weights		.468 **	.223 **	.315 **	-.007	.076	.725
3	Fuzzy Rule 1	.855	-.058	.934	-.529	.274	.976
	Fuzzy Rule 2	.794	.145	.014	.192	-.033	
Regression Beta Weights		.793 **	.104 *	.177 **	.062	-.042	.904
4	Fuzzy Rule 1	2.580	18.870	-188.091	-68.842	-11.110	.954
	Fuzzy Rule 2	1.0e-017* .302	-.0009	.696	.820	.991	
	Fuzzy Rule 3	.761	-.157	.458	.045	.010	
Regression Beta Weights		.681 **	-.225 **	.441 **	.094	-.035	.870
5	Fuzzy Rule 1	2.667	-.331	.841	.111	-.466	.997
	Fuzzy Rule 2	.835	.085	.277	.175	.077	
Regression Beta Weights		.689 **	.077 **	.300 **	.154 **	-.030	.979
6	Fuzzy Rule 1	1.870	.200	1.150	-1.005	.408	.866
	Fuzzy Rule 2	.256	-.063	.098	.001	-.087	
	Fuzzy Rule 3	.162	-.015	.153	-.105	.001	
Regression Beta Weights		.784 **	.127 *	.125	.008	-.020	.818
7	Fuzzy Rule 1	7.753	-.451	.272	.518	-1.768	.888
	Fuzzy Rule 2	5.302	-.741	-.030	-.577	.227	
Regression Beta Weights		.707 **	-.218 **	.462 **	.104	-.117	.870
8	Fuzzy Rule 1	.432	.051	-.168	.237	-.177	.957
	Fuzzy Rule 2	1.092	-.363	.427	-.292	-.048	
Regression Beta Weights		.955 **	-.204 **	.226 **	-.026	-.138 *	.881
9	Fuzzy Rule 1	.011	-.001	.007	.001	.005	.821
	Fuzzy Rule 2	.929	-.500	.217	.201	-.438	
	Fuzzy Rule 3	.112	.027	.192	-.215	-.001	
Regression Beta Weights		.792 **	-.394 **	.282 *	.115	-.327 **	.632
10	Fuzzy Rule 1	176.665	175.199	40.872	49.864	-264.798	.973
	Fuzzy Rule 2	.967	-.031	.071	-.004	-.151	
Regression Beta Weights		1.02 **	-.049	.078	-.033	-.150 **	.929

** regression coefficient significant $p < .01$, * regression coefficient significant $p < .05$

Table 6

Fuzzy Rule and Linear Regression Normalized Weight Vectors (Computed on N = 73)

		Performance	Group Performance	Importance	Tenure	Salary Level
Manager 1	Regression	.967	-.096	.224	.006	-.070
	Fuzzy Rule 1	.956	-.221	.018	.038	-.189
	Fuzzy Rule 2	-.380	.520	.447	-.280	.554
Manager 2	Regression	.766	.365	.515	-.011	.124
	Fuzzy Rule 1	.234	-.056	-.561	-.527	.592
	Fuzzy Rule 2	-.749	-.220	.042	-.586	.215
	Fuzzy Rule 3	.107	-.492	-.197	.157	-.827
Manager 3	Regression	.964	.127	.215	.075	-.051
	Fuzzy Rule 1	.610	-.041	.667	-.378	.195
	Fuzzy Rule 2	.956	.175	.017	.231	-.040
Manager 4	Regression	.803	-.265	.520	.111	-.041
	Fuzzy Rule 1	.013	.094	-.933	-.342	-.055
	Fuzzy Rule 2	.202	-.0006	.466	.549	.664
	Fuzzy Rule 3	.843	-.174	.507	.050	.011
Manager 5	Regression	.893	.100	.389	.199	-.039
	Fuzzy Rule 1	.934	-.116	.294	.039	-.163
	Fuzzy Rule 2	.923	.094	.306	.194	.085
Manager 6	Regression	.975	.158	.155	.010	-.025
	Fuzzy Rule 1	.761	.082	.468	-.409	.166
	Fuzzy Rule 2	.870	-.213	.334	.004	-.295
	Fuzzy Rule 3	.657	-.059	.620	-.426	.003
Manager 7	Regression	.798	-.246	.521	.117	-.132
	Fuzzy Rule 1	.971	-.056	.034	.065	-.221
	Fuzzy Rule 2	.984	-.138	-.006	-.107	.042
Manager 8	Regression	.944	-.201	.223	-.026	-.137
	Fuzzy Rule 1	.782	.092	-.304	.429	-.321
	Fuzzy Rule 2	.865	-.288	.338	-.231	-.038
Manager 9	Regression	.799	-.397	.285	.116	-.330
	Fuzzy Rule 1	.776	-.081	.476	.088	.396
	Fuzzy Rule 2	.787	-.424	.184	.170	-.371
	Fuzzy Rule 3	.361	.088	.620	-.691	-.002
Manager 10	Regression	.985	-.048	.076	-.032	-.145
	Fuzzy Rule 1	.479	.475	.111	.135	-.718
	Fuzzy Rule 2	.985	-.031	.073	-.004	-.154

A review of the standardized regression coefficients (i.e., beta weights) in Table 5 suggests that first, performance has a dominant influence on merit pay allocation decisions across all of the managers. However, besides this trend, there is a great deal of discrepancy among the managers in terms of the suggested relative importance of the other cues. For example, in looking at the group performance variable, the linear regression results suggest a positive influence of group performance on merit pay allocation for some managers, while with other participants, a negative weight is indicated. Similar variations in coefficients are suggested for the tenure and salary level cues also. The importance cue generally yielded a positive regression coefficient. Note that any interpretation of the regression weights should be viewed cautiously due to the impact of multicollinearity. As mentioned earlier, high multicollinearity leads to a high degree of imprecision and slight fluctuations in correlations due to sources of error may lead to large changes in the estimation of regression coefficients (Pedhazur, 1982). Also, the superiority of the fuzzy system models demonstrated in the bootstrapping results and in the sample $\underline{R}^2$ values in Table 5 suggests that there are systematic characteristics of the cue to judgment mappings that are not adequately captured with linear regression. In terms of an overview of the empirical fuzzy system models, an initial result that is striking is the finding that only a small number of rules are extracted from the data when using the adaptive model generation methods (i.e., clustering and ANFIS). In general, only two or three Sugeno style rules tuned with the ANFIS algorithm were sufficient to create a highly accurate mapping of the merit pay allocation judgments. This supports the idea that Sugeno rules are generally more compact and computationally efficient than other types of fuzzy and non-fuzzy expert system rules (Cox, 1995; Gulley & Jang, 1995). Inspection of the rule parameters in Tables 5 and 6 show that different rules within the same model have very different parameter values, and hence emphasize different cues. This is in fact an advantageous property of Sugeno style inferencing. Different rules may serve to model different planes or linear functions within a complex response surface. As an example of the influence of the different rules, Figure 20 shows the differential outputs of two fuzzy if-then rules for manager number five along different levels of merit increase judgments. The standardized version of the parameter weight values for these two rules are shown in Table 5 (under Manager 5), and Figure 21 shows the membership functions derived for a fuzzy system model of this manager's judgments.

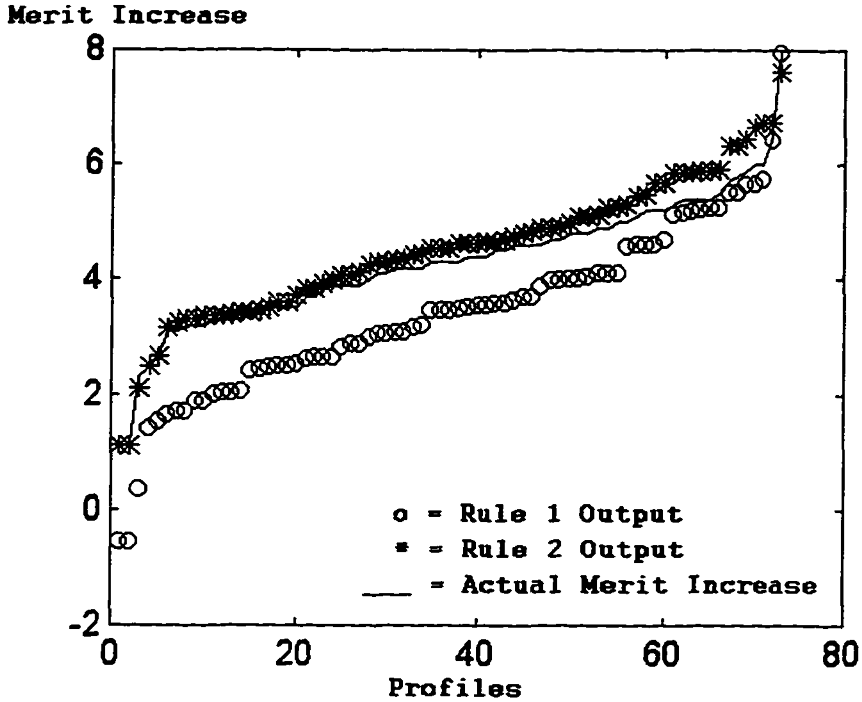


Figure 20. Fuzzy rule outputs and actual merit increase judgments for manager five.

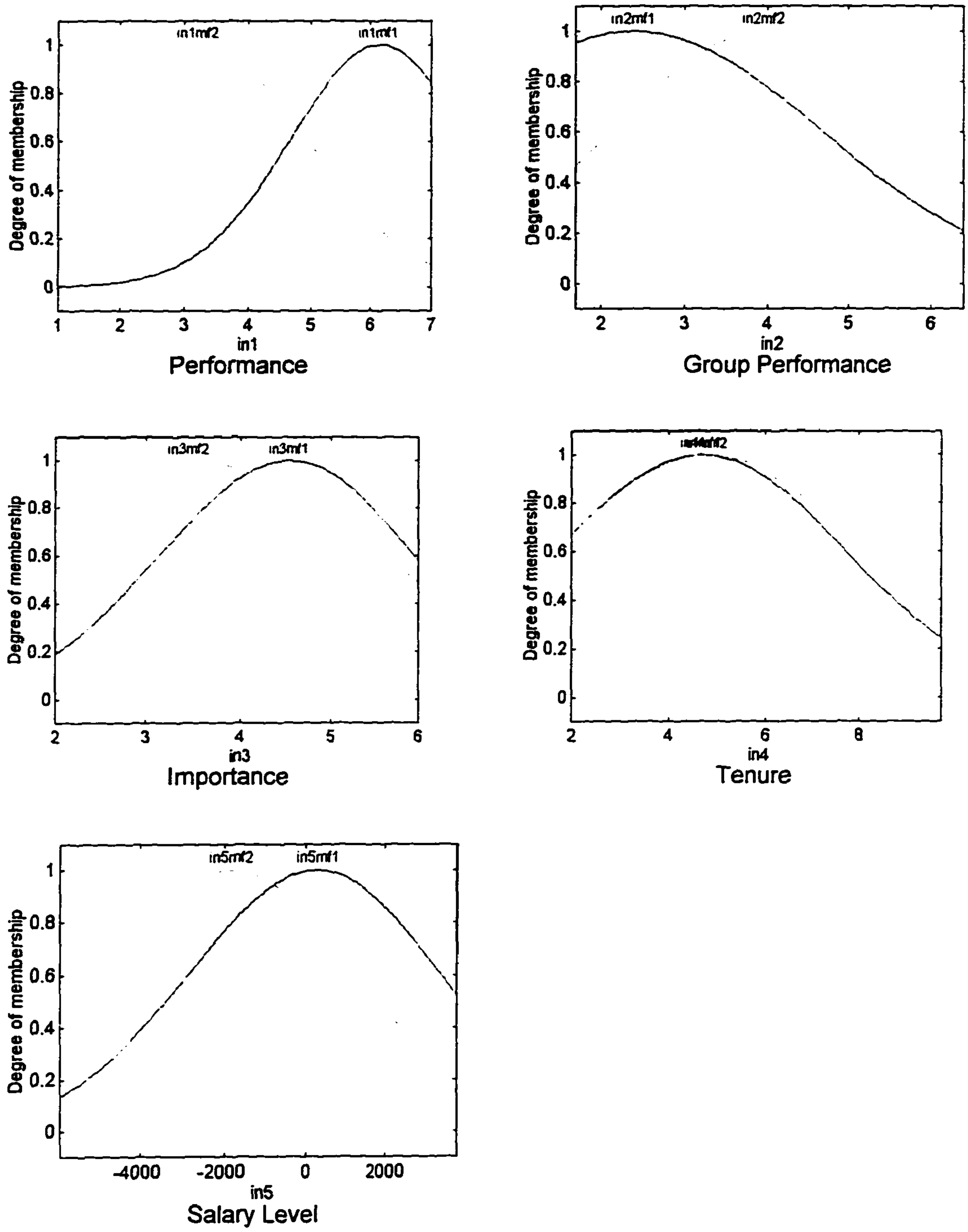


Figure 21. Fuzzy membership functions for cue variables in the manager five fuzzy model.

Figure 20 reveals that for the majority of the merit pay allocation judgments, fuzzy rule number two yields outputs that closely approximate the actual judgments. However, this is only true up to a certain level of merit increases because at the highest levels of merit pay allocation, the first fuzzy rule's outputs yield a closer approximation. This figure show the interplay between fuzzy system rules that facilitate modeling complex systems. In viewing the rule parameters for the fuzzy system model for manager number five in Tables 5 and 6, it becomes evident that the two different rules actually involve different weightings of the cues. For example, rule one weights both group performance and salary level negatively, while rule two does not. This suggests a possible discounting effect in fuzzy rule one where group performance and current salary levels serve to constrain merit increases. An important factor in analyzing the Sugeno fuzzy if-then rules is understanding when each of the rules fires to the greatest degree. Returning to the fuzzy system model for manager number five, the fuzzy rules for this system can be verbally written as follows:

1. If (performance is in1mf1) and (group performance is in2mf1) and (importance is in3mf1) and (tenure is in4mf1) and (salary level is in5mf1) then use rule 1,
2. If (performance is in1mf2) and (group performance is in2mf2) and (importance is in3mf2) and (tenure is in4mf2) and (salary level is in5mf2) then use rule 2.

Although both rules will always fire in parallel for each set of inputs, the rules will fire to different degrees depending on how the input cue values match the antecedent fuzzy sets listed in the rules. For example, the membership function labeled "in1mf1" generally covers the upper part of the performance measurement scale. Since this membership function is listed in rule number one, higher levels of performance will instantiate rule one to a higher degree. The problem with the membership functions shown in Figure 21 is the high degree of overlap between the membership functions, which is an artifact of the data-driven methods used to extract and tune the rules. This leads to limitations in interpreting the membership functions and fuzzy rules since the membership functions do not easily map onto meaningful linguistic terms such as "high", "moderate", or "low". As suggested by Jang, Sun, and Mizutani (1997), these issues involve a dilemma between "precision" and "interpretability", which is elaborated on in a later section.

Results - Hypothesis Two

Hypothesis two proposed that empirical fuzzy system models would also perform better than or equivalent to nonlinear regression models. This hypothesis was supported. Specifically, the empirical fuzzy system models outperformed the entire set of nonlinear regression models across all of the participants. Table 7 shows results which suggest that for every participant, fuzzy system modeling resulting in a superior mean cross-validated (bootstrapped) R value to any of the nonlinear regression models. Note that each of the nonlinear regression elements listed in this table were computed using 1000 resamples, and a 2/3 - 1/3 data split for training versus test data.

In summarizing these results across managers, the findings reveal that the fuzzy system models have an average model superiority over the polynomial models, in terms of an increment in R^2, of .05. Also, the mean cross-validated R's, when tested using Fisher z scores, meet the criterion for a significance mean difference between the two modeling strategies, with $t(9) = 3.03$, $p < .05$. When looking at the interaction regression models, the fuzzy system models are also superior with an average increment in R^2 across managers of .03, and a significant difference between the mean cross-validated R's, with $t(9) = 3.99$, $p < .01$. Similar results are indicated for the scatter and saturated regression models, with average fuzzy model R^2 increments of .04 and .05, respectively, and significant mean R differences, with $t(9) = 4.90$, $p < .01$ and $t(9) = 3.03$, $p < .05$, respectively.

Table 7

A Comparison of Mean Cross-validated R Values for Nonlinear Regression and Fuzzy System Models

Manager	1	2	3	4	5	6	7	8	9	10
Fuzzy System										
Mean Bootstrap Cross-validated R	.978	.863	.944	.940	.990	.946	.930	.935	.756	.969
95% Conf.	.953	.726	.884	.862	.975	.891	.863	.870	.556	.952
Intervals	.994	.920	.977	.976	.997	.976	.963	.969	.872	.983
Polynomial Regression (higher order performance terms with linear terms)										
Mean Bootstrap Cross-validated R	.970	.839	.922	.879	.988	.819	.928	.928	.732	.956
95% Conf.	.917	.737	.861	.756	.971	.479	.877	.868	.569	.923
Intervals	.986	.907	.963	.941	.995	.935	.966	.960	.846	.977
Scatter Regression (scatter term with linear terms)										
Mean Bootstrap Cross-validated R	.963	.826	.931	.887	.986	.891	.918	.923	.736	.960
95% Conf.	.942	.741	.880	.781	.970	.848	.863	.877	.591	.937
Intervals	.981	.894	.965	.951	.995	.927	.954	.955	.847	.976
Interactive Regression (performance interaction terms with linear terms)										
Mean Bootstrap Cross-validated R	.967	.855	.939	.934	.984	.889	.925	.911	.716	.951
95% Conf.	.946	.778	.891	.887	.965	.829	.875	.846	.567	.914
Intervals	.984	.915	.969	.967	.996	.926	.961	.951	.838	.971
Saturated Regression (all nonlinear and linear terms)										
Mean Bootstrap Cross-validated R	.972	.844	.937	.923	.985	.827	.921	.914	.727	.945
95% Conf.	.917	.748	.878	.833	.960	.442	.867	.844	.497	.872
Intervals	.988	.914	.969	.960	.995	.947	.962	.956	.862	.973

Results - Hypothesis Three

Hypothesis three proposed that inspection of both fuzzy system rules (i.e., both empirical and subjective) and nonlinear regression models would reveal evidence of nonlinear noncompensatory allocation strategies for some of the participants (i.e., allocators). This hypothesis was supported. Below, I review supporting results from the nonlinear regression models, the empirical fuzzy system models, and the subjective fuzzy system models.

First, Table 8 provides an overview of the nonlinear noncompensatory regression terms that were significant across the set of participating managers.

Table 8

The Significance of Nonlinear Noncompensatory Terms in Regression Models Across Managers

	Polynomial Model Terms [a]		Scatter Model Term	Interactive Model Terms [b]				Saturated Model Terms						
Term No. >	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Manager 1	**	**	*	*		*	**	**	**		*		*	**
2	**	**												
3	*	**			**	*		**	**			**	*	
4	**	**			**					*		**		
5		**	*	**			**	*	**		**			**
6		**			*	**		**	**			*	**	
7	**													
8		**						**	**					
9	**			**				**	*		**			
10	*	*						*	*	*				
mean 1	.010		-.004	.004				.014						
change 2	.017		-.005	.044				.025						
in R^2 3	-.022		-.006	.009				.006						
over 4	-.011		.004	.089				.069						
linear 5	.006		.002	-.002				.000						
model 6	-.125		-.002	-.005				-.112						
(cross- 7	.015		-.004	.009				.002						
valid- 8	.000		-.009	-.031				-.026						
ated 9	-.012		-.006	-.035				-.019						
results) 10	-.012		-.004	-.021				-.032						

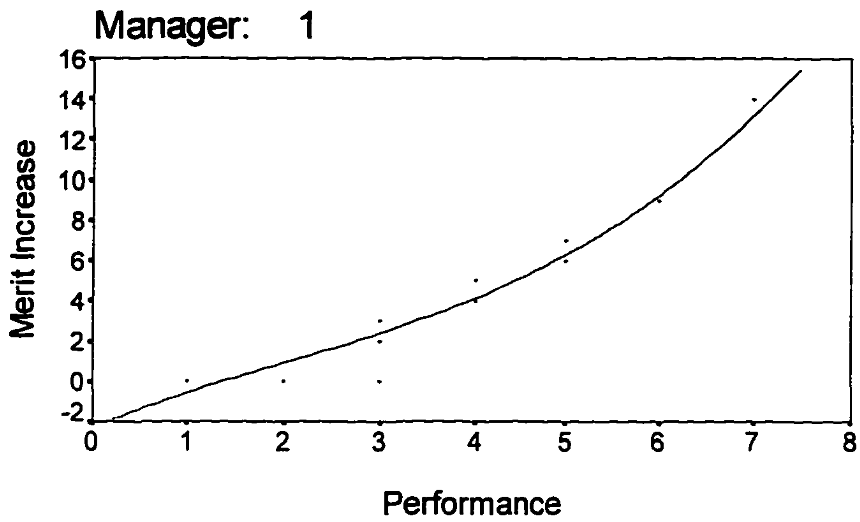
* coefficient significant, $p < .05$ ** coefficient significant, $p < .01$ N = 110 (total number of profiles)

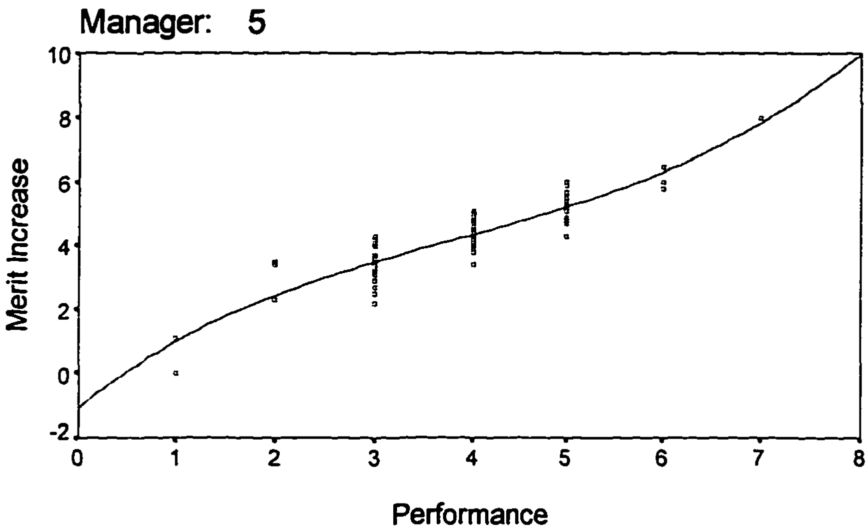
Term No.'s
1 - PERFORMANCE ^{2} TERM
2 - PERFORMANCE ^{3} TERM
3 - SCATTER TERM
4 - PERFORMANCE X GROUP PERFORMANCE TERM
5 - PERFORMANCE X IMPORTANCE TERM
6 - PERFORMANCE X TENURE TERM
7 - PERFORMANCE X SALARY LEVEL TERM

[a] = terms entered hierarchically, [b] = terms entered after controlling for polynomial effects

An overview of Table 8 suggest that 100% of the managers had significant polynomial terms when fitting these factors to their data, 60% of the managers had significant interaction terms, and 20% of the managers had significant scatter coefficients (which were both positive indicating disjunctive type processing). Although the cross-validated $\underline{R}^2$ values indicate that generally only a small increment (and in some cases a decrement) in variance accounted for is associated with the nonlinear noncompensatory terms, several of the increments represent substantial increases and suggest that inclusion of the nonlinear noncompensatory terms does improve model fit. Moreover, small increments in variance accounted for may have large implications for understanding the psychological fidelity of mathematical models in the area of decision making (Brannick & Brannick, 1989). A particularly interesting finding in terms of the nonlinear regression analyses is the somewhat ubiquitous nature of the higher order performance terms. Previously, I suggested that higher order performance factors have theoretical significance given their role in areas such as utility analysis. As is evident in Figure 22, which shows cubic regression models for three of the managers in the current study, there does appear to be an interesting parallel between these function forms and utility functions, since both have similar curve orientations.

Another important point is that although a number of the managers appear to be using some types of nonlinear judgment strategies, it is difficult to exactly specify the processes used due to the variety of possible nonlinear terms and the presence of multicollinearity.





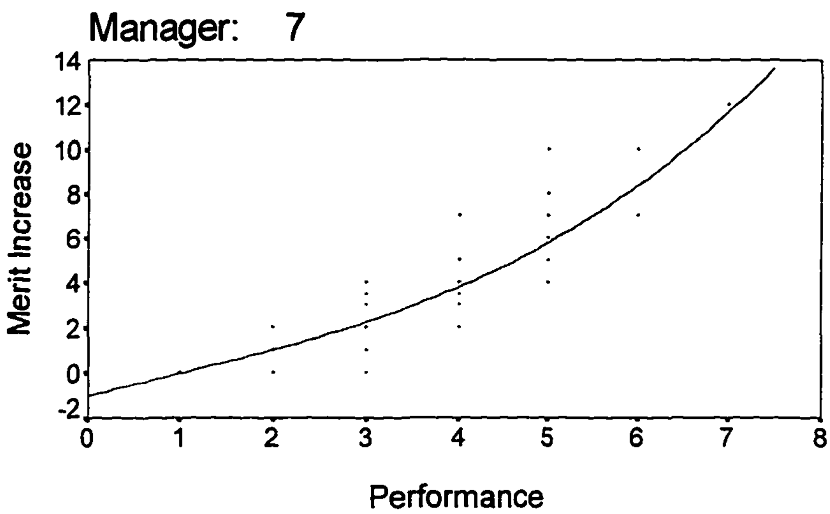


Figure 22. Cubic regression models of performance and merit increases.

Evidence from the empirical fuzzy system models also supports the presence of nonlinearities in the managers' allocation strategies, thereby lending support to hypothesis three. First, inspection of the rule parameters in Tables 5 and 6 suggests that within the same fuzzy system, cue variables are weighted differentially (e.g., in one rule, group performance receives a positive weight, while in another rule, group performance receives a negative weight), which is consistent with the interpretation that nonlinearities exist in the judgment response surfaces. Moreover, visual inspection of the empirical fuzzy system response surfaces yields evidence of complex nonlinear components (e.g., see Figure 23).

There is also evidence directly from the managers' questionnaire responses, which suggests nonlinear noncompensatory judgment strategies. Specifically, in viewing the responses to the post judgment questionnaire used for building the subjective fuzzy systems (where managers were asked to choose a consequent action based on a set of antecedent cue levels), certain responses suggested nonlinear noncompensatory responses. As an example, Figure 24 shows one of the actual responses to a questionnaire. Here the manager appears to react to a low level of individual performance by assigning a very low merit increase. This low level of an increase is allocated even though the recipient has a moderate level of tenure and a moderate level of importance. Consequently, a noncompensatory reaction is indicated, in that, moderate levels of tenure and importance do not compensate for the low performance. A number of such nonlinear noncompensatory responses were noted in the questionnaire data, and these responses essentially became fuzzy rules in the subjective models. It should be noted that these responses represent subjective evaluations of judgment policies and may not be indicative of actual policies.

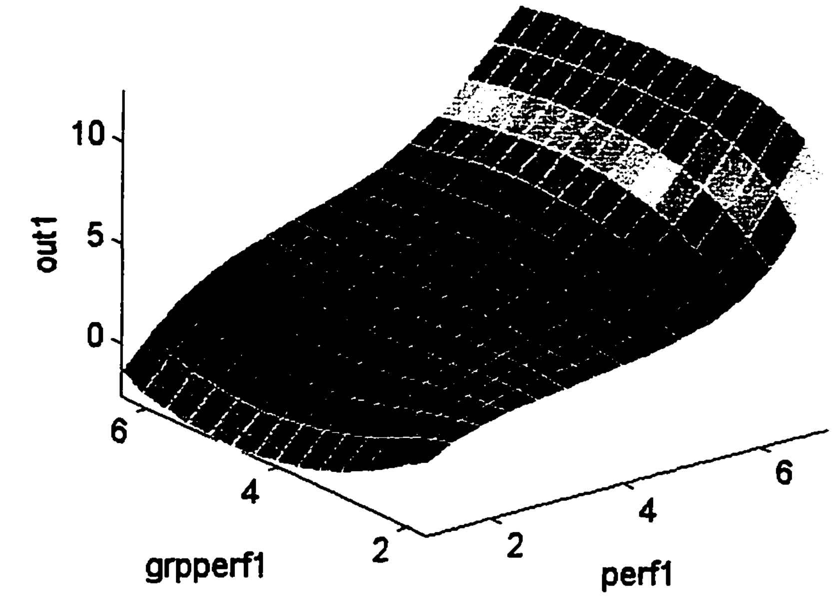


Figure 23. Fuzzy system output surface for manager number one (performance x group performance x predicted merit increase judgment).

28) **If the employee profile values were as follows:**

Employee's Performance Rating - LOW
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level - MODERATE

Your merit pay allocation would be (circle one) -

VERY LOW LOW MODERATE HIGH VERY HIGH

Figure 24. An example noncompensatory response in a post judgment questionnaire.

Results - Hypothesis Four

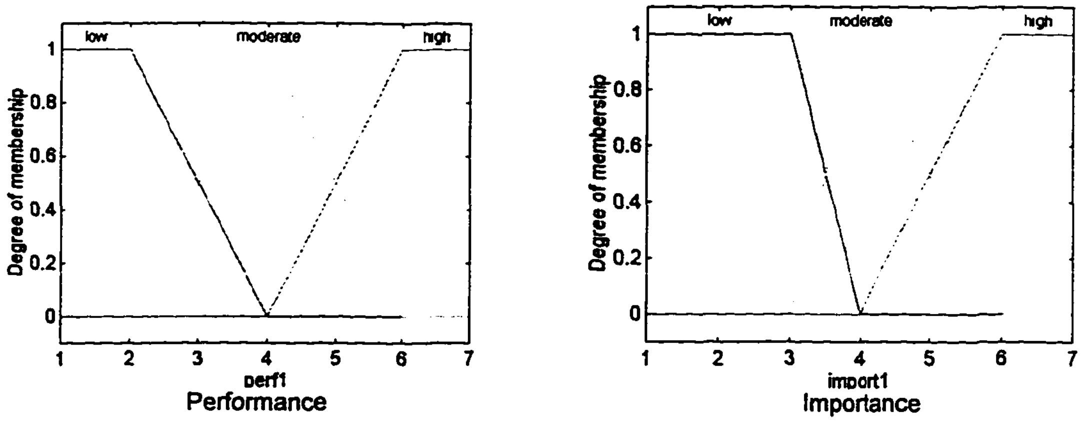
Hypothesis four suggested that fuzzy system models constructed solely from subjective information directly elicited from the participants would perform more effectively than models based on a traditional subjective policy capturing approach. This hypothesis was not supported. As shown in Table 9, the subjective regression approach (based on subjectively estimated weights) outperformed the subjective fuzzy models for 80% of the managers.

Table 9

Comparisons Between Subjective Regression and Subjective Fuzzy System Models (Based on Predictions Across Entire 110 Profiles)

Manager	1	2	3	4	5	6	7	8	9	10
Subjective Regression										
Multiple R	.944	.837	.885	.744	.969	.853	.873	.669	.404	.906
Subjective Fuzzy Systems										
Multiple R	.880	.439	.840	.700	.925	.754	.815	.704	.464	.859

Figure 25 presents example subjectively elicited membership functions for the cue variables and some of the Mamdani style rules for one of the managers. The subjective fuzzy models, based on Mamdani style fuzzy inference systems, seemed to perform poorly even when based on what appeared to be the best combination of membership function types and implication and aggregation methods. Note that while the subjective regression models performed well, they did not match the statistical regression results. On average, the subjective regression models accounted for approximately 16% less variance in judgments than when using the optimal least squares regression weights.



1. If (performance is moderate) and (group performance is moderate) and (importance is moderate) and (tenure is moderate) and (salary level is moderate) then (merit increase is moderate).

2. If (performance is moderate) and (group performance is moderate) and (importance is moderate) and (tenure is high) and (salary level is moderate) then (merit increase is moderate).

3. If (performance is moderate) and (group performance is low) and (importance is moderate) and (tenure is moderate) and (salary level is low) then (merit increase is high).

Figure 25. Example Mamdani rules and membership functions from a subjective fuzzy system model for manager number eight.

DISCUSSION

Motivating factors for conducting this research revolved around the search for a new policy capturing/modeling approach for studying merit pay allocation judgments under conditions of possible multicollinearity, where nonlinear and/or noncompensatory strategies might play a role (including configural processes involving non-performance factors), and where subjective impressions of judgment policies could be elicited and interpreted in a framework that extends beyond obtaining subjective estimates of linear weights. To address these issues, a proposal for using fuzzy systems for judgment research was made, thus attempting to bridge the disciplinary gap between expert system technologies and research on judgment processes.

Given the previously reviewed results, several general conclusions can be offered in summarizing this research effort. First, fuzzy system methodologies offer a powerful alternative to traditional statistical methods for conducting judgment research. Support for this statement is found in the evidence affirming the first two research hypotheses. Specifically, fuzzy system models outperformed both linear and nonlinear regression models in terms of mapping precision and model fit. However, before the full utility of fuzzy systems for judgment and decision research can be realized, the search must continue for fuzzy modeling techniques that concurrently maximize and balance mapping precision with interpretability (Jang, Sun, & Mizutani, 1997). In the current study, the interpretability of the empirically derived fuzzy system rules was restricted due to the inability to map the Sugeno style rules to meaningful linguistic values, and due to constraints on identifying the role that individual rules played in defining the overall judgment policies. Thus, in this study, the fuzzy systems did not fulfill the potential role of being a methodology capable of yielding highly interpretable policies under conditions of highly correlated judgment cues (i.e., multicollinearity). One important area for research on fuzzy systems, especially for psychological applications, is in the area of developing adaptive tuning methods such as the ANFIS methodology that maintain linguistic integrity. Jang, Sun, and Mizutani (1997) suggest several potential approaches for addressing the interpretability - precision dilemma.

A second general conclusion emerging from this research is that organizational decision tasks such as merit pay allocation may involve nonlinear and noncompensatory judgment strategies. This statement is based on evidence that the participating managers in the current research appeared to be using judgment strategies which incorporated nonlinear and noncompensatory components. This would seem to challenge continued reliance on linear regression for policy capturing studies. Implicit in the use of linear regression is the a priori assumption that the linear additive model is sufficient for accurately modeling the judgment of interest. The evidence presented here suggests that this assumption should receive increased attention. In fact, one of the most probable reasons for the superiority of the fuzzy systems over the regression approach is the proven ability of fuzzy systems to function as universal approximators, which facilitates handling complex, nonlinear, and noisy systems. Potential nonlinear noncompensatory judgment components identified in this research effort include: higher order function forms (e.g., quadratic and cubic terms), configural strategies involving cue interactions, and noncompensatory strategies. Nonlinear components potentially have theoretical significance when similar function forms are identified across research areas. For example, in the current study there was evidence relating higher order performance terms to merit increase judgments, in a manner similar to s-shaped utility functions. In the utility analysis literature, utility functions suggest that for very low performance levels the value of that performance drops abruptly, and similarly, for very high performance levels, the value is perceived in an extremely positive way. (Bobko, 1995). It appears that managers allocating merit pay may perceive and reward performance in an analogous manner.

A third general conclusion that can be offered is that, in the current study, the fuzzy system framework did not function well in terms of building models based solely on subjective information generated from a minimum of knowledge engineering. This statement relates to research hypothesis four and the superiority of the subjective regression modeling approach. One possible reason for this finding has to do with the "structure determination" problem mentioned earlier (Jang & Sun, 1995). A useful framework for thinking about the structure determination issue emerges when borrowing the concept of "model specification" from the literature on structural equation modeling (e.g., Long & Trivedi, 1993). Subjectively defined membership functions and fuzzy rules based on an assumed number of antecedent and consequent fuzzy sets will probably lead to a fuzzy system model that is incorrectly specified in terms of optimizing the mapping precision of the system. Note that a fuzzy model can be misspecified in a number of ways, including: improper decomposition of the variable into meaningful fuzzy sets, incorrect overlap in membership functions, a lack of rules covering one or more fuzzy regions, or incorrect weight parameters on the rules (Cox, 1995). Similarly, the subjective regression models were not perfectly specified either. However, when both fuzzy models and regression models are misspecified based on subjective parameters, it appears that the subjective regression models are more robust, which fits with research by Dawes (1972) who showed that even with a policy of equal weights, linear models are robust predictive models. Despite this, as evidenced by other results in the current research, linear models may have robustness at the price of modeling precision. Also, it is unknown whether more elaborate knowledge engineering efforts would have resulted in better subjective fuzzy models.

A last general conclusion is that there appears to be substantial individual differences in managerial merit pay allocation strategies. In both the regression and fuzzy system models, the patterns of relative cue importance (as specified by variable weight parameters) and indices relating to variance accounted for were different among the managers. In the literature, variation between managers in merit pay allocation policies has been characterized as both possible systematic individual differences (Sherer et al., 1987) and as variation accounted for by statistical artifacts such as sampling error and unreliability in the dependent variable (Deshpande & Joseph, 1994). To definitively decide which perspective is more accurate, further research is needed. Researchers attempting to account for variance in policy characteristics across managers should consider the possible existence of complex judgment policy components (e.g., nonlinear components) and reflect on the implications of alternative definitions of inconsistency and error (e.g., inconsistency as fuzziness).

An important point to note in summarizing this research effort is that the major limitations of this study involve the artificiality of the judgment task and the use of the "paper people" approach. Research efforts are needed to replicate the results of this research in field settings, with actual merit pay judgments. Issues that are currently unresolved include the role of nonlinear noncompensatory components in actual pay decision environments and the degree to which there are organizational contextual factors also existent that help explain the presence of nonlinear strategies. When viewed in specific organizational environments, judgment policies that appear complex may be viewed in a new light as adaptive, functional strategies. Consider an example from the current study where a specific fuzzy rule appeared to approximate more precisely the judgment output surface at higher levels of merit increases. This rule was associated with a negative weight on salary level. This weighting is consistent with a perceived need to constrain costs, so that recipients likely to receive large merit increases, due to high performance levels, might receive a lower increase than warranted to the degree that they were already extremely high in a given pay grade or band.

Emerging from this research project are several ideas with broad ranging implications for decision making research. One of these ideas is that the traditional definition of a "judgment policy" may not be the most appropriate definition. Traditionally, judgment policies have been viewed as linear equations thought to be descriptive of the relative importance of judgment cues. These policies have been of the "one size fits all" variety, in that a single policy equation is assumed to be characteristic of the judgment process. Whether this view is appropriate has recently come into question. For example, McIntyre and James (1995) in looking at how individual raters combine performance information suggested that raters may use different rating policies for different targets. These authors suggest that this situation poses a problem for traditional policy capturing studies due to the fact that ratings are pooled across targets to derive policy equations and that in combining information across targets a loss in accuracy and a potential distortion of the judgment process occurs. Proposed in this research effort is the idea that a judgment policy can be thought of as a set of fuzzy rules, rather than as a single equation. This idea carries with it the implication that certain rules may fit better with certain combinations of cue values. Consequently, the idea forwarded by McIntyre and James of target by rater effects is handled within a fuzzy system framework simply by considering that the different interactions correspond to different rules. If targets influence the way that raters combine information then experience with a certain type of target may cause an individual judge or rater to update their set of policy rules or to add a new rule. Future research efforts are needed to clarify such issues.

The idea of a judgment policy as a set of fuzzy rules can actually be viewed from a number of perspectives. From one perspective, this concept may suggest that judges are not consistent in judgment because they are not implementing a single set judgment policy. However, from a different perspective, a set of fuzzy rules may be viewed as a highly adaptive and functional cognitive mechanism that facilitates handling different types of cases or situations.

Fuzzy systems theory has much to offer decision making researchers. As well as facilitating the development of powerful modeling tools, fuzzy theory may offer insight and different perspectives on what is meant by "measurement error" in judgment analysis and rating research. From the fuzzy point of view, a potential source of error is the mismatch between initial cognitive impressions formed by judges in considering a target individual, which may exist in a qualitative or linguistic form, and the translation of this impression into precise quantitative scales. An initial impression of a pay recipient as a "great performer" may, due to slight disturbances in cue configurations, ultimately be translated into a 6% merit increase rather than 7%. However, the pay recipient may be a "great performer" to essentially the same degree as a recipient who received a 7% increase. By building this type of uncertainty into modeling efforts, the result may be an increase in variance accounted for and a new framework for thinking about error. Note that the idea of directly representing error or uncertainty in modeling efforts is consistent with current trends in social science methodological development (e.g. structural equation modeling, meta-analysis).

Expert system technologies and decision support systems are likely to play an increasingly salient role in organizational functioning and in industrial and organizational psychological research, especially as new frameworks for representing expert knowledge develop. Ultimately, these technologies may be evaluated on pragmatic grounds. The idea of providing employees throughout an organization access to the same expertise is a powerful one. In decision environments such as pay allocation, expert systems may be used as managerial tools for evaluating business scenarios and providing expert quality solutions to end users.

As recently noted by several researchers (Lawler, 1992; Sturman, Hannon, & Milkovich, 1996), there needs to be an increasing research focus not only on how these systems are built but on attitudinal and behavioral reactions to such systems. Also, there is a definite need for studies in line with the research presented here. Specifically, studies are needed that go beyond describing system features to evaluate whether new technologies and methodologies have strengths that are orthogonal to existing methods. Fuzzy system technology is one such methodology that should receive additional attention. In addition, research in these areas should be multidisciplinary in nature. For psychologists and organizational researchers to gain maximum benefit from advanced technologies, they will have to play an active role in research and development efforts. In this way, the vision of luminaries such as Simon (1995) may be realized, and researchers in areas such as decision making and judgment will have a rich multidisciplinary framework to draw from and build upon.

REFERENCES

Abraham, F.D., Abraham, R.H., & Shaw, C.D. (1990). A visual introduction to dynamical systems theory for psychology. Santa Cruz, CA: Aerial Press.

Aguinis, H. (1994). A Quickbasic program for generating correlated multivariate random normal scores. Educational and Psychological Measurement, 54(3), 687-689.

Alexander, R.A., & Barrett, G.V. (1982). Equitable salary increase judgments based upon merit and non-merit considerations: A cross-national comparison. International Review of Applied Psychology, 31, 443-454.

Balzer, W.K., Doherty, M.E., & O'Conner, R. (1989). Effects of cognitive feedback on performance. Psychological Bulletin, 106(3), 410-433.

Bartol, K.M., & Martin, D.C. (1988). Influences on managerial pay allocations: A dependency perspective. Personnel Psychology, 41, 361-378.

Bartol, K.M., & Martin, D.C. (1990). When politics pay: Factors influencing managerial compensation decisions. Personnel Psychology, 43, 599-614.

Barton, S. (1994). Chaos, self organization, and psychology. American Psychologist, 49, 5-14.

Beatty, J.R., McCune, J.T., & Beatty, R.W. (1988). A policy-capturing approach to the study of United States and Japanese managers' compensation decisions. Journal of Management, 14(3), 465-474.

Bellman, R.E., & Zadeh, L.A. (1970). Decision making in a fuzzy environment. Management Science, 17, 141-164.

Birnbaum, M.H. (1983). Perceived equity of salary policies. Journal of Applied Psychology, 68, 49-59.

Bishop, J. (1987). The recognition and reward of employee performance. Journal of Labor Economics, 5, 36-55.

Bobko, P. (1995). Correlation and regression, Principles and applications for industrial/organizational psychology and management. New York: McGraw-Hill.

Borman, W.C. (1982). Validity of behavioral assessment for predicting military recruiter performance, Journal of Applied Psychology, 67, 3-9.

Brannick, M.T., & Brannick, J.P. (1989). Nonlinear and noncompensatory processes in performance evaluation. Organizational Behavior and Human Decision Processes, 44, 97-122.

Brehmer, A., & Brehmer, B. (1988). What have we learned about human judgment from thirty years of policy capturing? In B. Brehmer & C.R.B. Joyce (Eds.), Human Judgment: The SJT View. North Holland: Elsevier Science Publishers.

Brunswik, E. (1952). The conceptual framework of psychology. Chicago: University of Chicago Press.

Champagne, M., & Stevenson, M.K. (1994). Contrasting models of appraisal judgments for positive and negative purposes using policy modeling. Organizational Behavior and Human Decision Processes, 59, 93-123.

Cohen, J. (1978). Partialed products are interations; Partialed powers are curve components. Psychological Bulletin, 85(4), 858-866.

Cohen, J. (1994). The earth is round (p < .05). American Psychologist, 49, 997-1003.

Collins, J.M., & Clark, M.R. (1993). An application of the theory of neural computation to the prediction of workplace behavior: An illustration and assessment of network analysis. Personnel Psychology, 46, 503-524.

Cook, R.L., & Stewart, T.R. (1975). A comparison of seven methods for obtaining subjective descriptions of judgmental policy. Organizational Behavior and Human Performance, 13, 31-45.

Cooksey, R.W. (1996). Judgment analysis: Theory, methods, and applications. New York: Academic Press.

Cox, E. (1995). Fuzzy logic for business and industry. Rockland, Massachusetts: Charles River Media.

Cox, E. (1994). The fuzzy systems handbook. San Diego, CA: AP Professional.

Craiger, J.P., & Coovert, M.D. (1994). A model of the relationship between job experience and job sample test performance: Application of a neuro-fuzzy system. Technical Report. Armstrong Laboratory, Brooks AFB, Texas.

Cummings, T.G., & Huse, E.F. (1989). Organization Development and Change. (4th ed.). New York: West Publishing Company.

Darlington, R.B. (1968). Multiple regression in psychological research and practice. Psychological Bulletin, 69, 161-182.

Dawes, R.M. (1972). Slitting the decision maker's throat with Occam's Razor: The superiority of random linear models to real judges. Oregon Research Institute Research Bulletin, 12, 13.

Dawes, R.M., & Corrigan, B. (1974). Linear models in decision making. Psychological Bulletin, 81, 95-106.

Deshpande, S.P., & Joseph, J. (1994). Variation in compensation decisions by managers: An empirical investigation. Journal of Psychology, 128(1), 41-50.

Deshpande, S.P., & Schoderbek, P.P. (1993). Pay-allocations by managers: A policy-capturing approach. Human Relations, 46(4), 465-479.

Dobbins, A. & Coovert, M.D. (1992). In search of the holy grail for job classification. Paper presented at the Eighth Annual Meeting of the Society for Industrial and Organizational Psychology, Montreal, Canada.

Dreher, G.F. (1981). Predicting the salary satisfaction of exempt employees. Personnel Psychology, 34, 579-589.

Dyer, L., & Schwab, D.P. (1982). Personnel/human resource management research. In T.A. Kochan, D.B. Mitchell, & L. Dyer (Eds.), Industrial relations research in the 1970s: Review and appraisal. Madison, WI: Industrial Relations Research Association.

Flannery, T.P., Hofrichter, D.A., & Platten, P.E. (1996). People, performance, and pay: Dynamic compensation for changing organizations. New York: The Free Press.

Flexer, A. (1996). Statistical evaluation of neural network experiments: Minimum requirements and current practice (Technical Report oefai-tr-95-16). Vienna, Austria: The Austrian Research Institute for Artificial Intelligence.

Fossum, J.A., & Fitch, M.K. (1985). The effects of individual and contextual attributes on the sizes of recommended salary increases. Personnel Psychology, 36, 587-602.

Freedman, S.M. (1978). Some determinants of compensation decisions. Academy of Management Journal, 21, 379-409.

Freedman, S.M., & Montanari, J.R. (1980). An integrative model of managerial reward allocation. Academy of Management Review, 5, 381-390.

Ganzach, Y. (1995). Negativity (and positivity) in performance evaluation: Three field studies. Journal of Applied Psychology, 80(4), 491-499.

Ganzach, Y., & Czaczkes, B. (1995). On detecting nonlinear noncompensatory judgment strategies: Comparison of alternative regression models. Organizational Behavior and Human Decision Processes, 61(2), 168-176.

Giarratano, J., & Riley, G. (1994). Expert systems: Principles and programming (2nd ed.). Boston: PWS Publishing Company.

Goldberg, L.R. (1971). Five models of clinical judgment: An empirical comparison between linear and nonlinear representations of the human inference process. Organizational Behavior and Human Decision Processes, 6, 458-479.

Gorman, C.D., Clover, W.H., & Doherty, M.E. (1978). Can we learn anything about interviewing real people from "interviews" of paper people? Two studies of the external validity of a paradigm. Organizational Behavior and Human Performance, 22, 165-192.

Green, B.F. (1968). Descriptions and explanations: A comment on papers by Hoffman and Edwards. In B. Kleinmuntz (Ed.) Formal Representation of Human Judgment. New York: Wiley.

Greenlaw, P.S., & Valonis, W.R. (1994). Applications of expert systems in human resource management. Human Resource Planning, 17, 27-42.

Grey, R.J., & Kipnis, D. (1976). Untangling the performance appraisal dilemma: The influence of perceived organizational context on evaluative processes. Journal of Applied Psychology, 61, 329-335.

Gulley, N., & Jang, J. -S. R. (1995). Fuzzy logic toolbox user's guide. Natick, Mass.: The Mathworks, Inc.

Hecht-Nielsen, R. (1990). Neurocomputing. New York: Addison-Wesley.

Heneman, R.L. (1990). Merit pay research. In G. R. Ferris & K. M. Rowland (Eds.), Research in personnel and human resources management: Vol 8. Greenwich, CN: JAI Press.

Heneman, R.L., & Cohen, D.A. (1988). Supervisory and employee characteristics as correlates of employee salary increases. Personnel Psychology, 41, 345-360.

Hesketh, B., Pryor, R., Gleitzman, M., & Hesketh, T. (1988). Practical applications and psychometric evaluation of a computerized fuzzy graphic rating scale. In T. Zetenyi (Ed.), Fuzzy sets in psychology. North-Holland: Elsevier Science Publishers.

Hesketh, B., Pryor, R., Gleitzman, M. (1989). Fuzzy logic: Toward measuring Gottfredson's concept of occupational social space, Journal of Counseling Psychology, 36, 103-109.

Hesketh, B., Elmslie, S., & Kaldor, W. (1990). Career compromise: An alternative account to Gottfredson's theory. Journal of Counseling Psychology, 37, 49-56.

Hesketh, B., McLachlan, K., & Gardner, D. (1992). Work adjustment theory: An empirical test using a fuzzy rating scale. Journal of Vocational Behavior, 40, 318-337.

Hesketh, B., & Gardner, D. (1993). Person-environment fit models: A reconceptualization and empirical test. Journal of Vocational Behavior, 42, 315-332.

Hobson, C.J., & Gibson, F.W. (1983). Policy capturing as an approach to understanding and improving performance appraisal: A review of the literature. Academy of Management Review, 8, 640-649.

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

Hoffman, P.J. (1960). The paramorphic representation of clinical judgment. Psychological Bulletin, 57, 116-131.

Huber, V.L., Neale, M.A., & Northcraft, G.B. (1987). Judgment by heuristics: Effects of rater and ratee characteristics and performance standards on performance-related judgments. Organizational Behavior and Human Decision Processes, 40, 149-160.

Ivancevich, J.M. (1983). Contrast effects in performance evaluation and reward practices. Academy of Management Journal, 26, 465-476.

Jang, J. -S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665-684.

Jang, J. -S. R., & Sun, C. -T. (1995). Neuro-fuzzy modeling and control. Proceedings of the IEEE, 83(3), 378-406.

Jang, J. -S. R., Sun, C. -T., & Mizutani, E. (1997). Neuro-fuzzy and soft computing. Upper Saddle River, NJ: Prentice Hall.

Kaun, D.E. (1984). Faculty advancement in a nontraditional university environment. Industrial and Labor Relations Review, 36, 592-606.

Klir, G.J., & Yuan, B. (1995). Fuzzy set and fuzzy logic. New Jersey: Prentice-Hall.

Kosko, B. (1992). Neural networks and fuzzy systems: A dynamical systems approach to machine intelligence. Englewood Cliffs, NJ: Prentice-Hall.

Kosko, B. (1993). Fuzzy thinking. New York: Hyperion.

Landau, S.B., & Leventhal, G.S. (1976). A simulation study of administrators' behavior toward employees who receive job offers. Journal of Applied Social Psychology, 6, 291-306.

Lawler, E.E. III (1981). Pay and development. Reading, MA: Addison-Wesley.

Ledford, G.E., Lawler, E.E., & Mohrman, S.A. (July-August, 1995). Reward innovations in Fortune 1000 Companies. Compensation & Benefits Review.

Long, J.S., & Trivedi, P.K. (1993). Some specification tests for the linear regression model. In K.A. Bollen, & J.S. Long (Eds.), Testing Structural Equation Models (pp. 66-110). Newbury Park, CA: Sage Publications.

Lubinski, D., & Humphreys, L.G. (1990). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial and mathematical ability. Psychological Bulletin, 107(3), 385-393.

Mamdani, E.H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies, 7(1), 1-13.

Markham, S.E. (1988). Pay-for-performance dilemma revisited: Empirical example of the importance of group effects. Journal of Applied Psychology, 73, 172-180.

Martin, D.C. (1987). Factors influencing pay decisions: Balancing managerial vulnerabilities. Human Relations, 7, 417-430.

McNeill, F.M., & Thro, E. (1994). Fuzzy logic: A practical approach. New York: AP Professional.

Munakata, T. (1994). Commercial and industrial AI. Communications of the ACM, 37, 23-25.

Munakata, T., & Jani, Y. (1994). Fuzzy systems: An overview. Communications of the ACM, 37, 23-25.

Murphy, K.R. (1984). Cost-benefit considerations in choosing among cross-validation methods. Personnel Psychology, 37, 15-22.

Murphy, K.R., & Cleveland, J. (1991). Performance appraisal: An organizational perspective. Boston: Allyn & Bacon.

Naylor, J.C., & Wherry, R.J., Sr. (1965). The use of simulated stimuli and the "JAN" technique to capture and cluster the policies of raters. Educational and Psychological Measurement, 25, 969-986.

Newell, A., & Simon, H.A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Norman, K. (1976). Theoretical note: A solution for weights and scale values in functional measurement. Psychological Review, 83, 80-84.

Nunnally, J.C. (1978). Psychometric theory. New York: McGraw-Hill.

Peck, C. (1984). Pay and performance: The interaction of compensation and performance appraisal. New York: The Conference Board.

Pedhazur, E.J. (1982). Multiple regression in behavioral research. New York: Holt, Rinehart, and Winston.

Personnel Policies Forum. (1981). Wage and salary administration (PPF Survey 131). Washington, DC: Bureau of National Affairs.

Peters, R., & Atkin, R. (1980). The effect of open pay systems on allocation of salary increases. Academy of Management Proceedings, 40, 293-297.

Pryor, R.G.L., Hesketh, B., Gleitzman, M. (1989). Making things clearer by making them fuzzy: Counseling illustrations of a fuzzy graphic rating scale. The Career Development Quarterly, 38, 136-147.

Schmidt, F.L., & Kaplan, L.B. (1971). Composite vs. multiple criteria: A review and resolution of the controversy. Personnel Psychology, 24, 419-434.

Schwab, D.P., & Olson, C.A. (1988). Pay-performance relationship as a function of pay for performance policies and practices. Academy of Management Best Papers Proceedings, 48, 287-291.

Sherer, P.D., Schwab, D.P., & Heneman, H.G. III. (1987). Policy capturing managerial raise decisions. Personnel Psychology, 40, 27-38.

Simon, H.A. (1995). The information-processing theory of mind. American Psychologist, 50(7), 507-508.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Decision Processes, 22, 251-263.

Smithson, M. (1985). Translatable statistics and verbal hypotheses. Quality and Quantity, 19, 183-209.

Smithson, M. (1988). Possibility theory, fuzzy logic, and psychological explanation. In T. Zetenyi Ed.), Fuzzy sets in psychology. North-Holland: Elsevier Science Publishers.

Stewart, T.R. (1988). Judgment analysis: Procedures. In B. Brehmer & C.R.B. Joyce (Eds.), Human judgment: The SJT view. North Holland: Elsevier Science Publishers.

Stevenson, M.K., Busemeyer, J.R., & Naylor, J.C. (1990). Judgment and Decision-making Theory. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology (pp. 283 - 374). Skokie, IL: Rand-McNally.

Stuhlmacher, A.F., & Stevenson, M.K. (1994). Predicting the strength of preference for labor contracts using policy modeling. Organizational Behavior and Human Decision Processes, 57, 253-289.

Summers, D.A., Taliaferro, J.D., & Fletcher, D.J. (1970). Subjective versus objective description of judgment policy. Psychonomic Science, 18, 249-250.

Taylor, R.L., & Wilstead, W.D. (1974). Capturing judgment policies: A field study of performance appraisal. Academy of Management Journal, 17, 440-449.

Trahan, W.A., Lane, I.M., & Dobbins, G.H. (1991). Organizational determinants of managerial pay allocation decisions. Journal of Social Psychology, 131(1), 93-106.

Turban, D.B., & Jones, A.P. (1988). Supervisor-subordinate similarity: Types, effects, and mechanisms. Journal of Applied Psychology, 73, 228-234.

von Altrock, C. (1995). Fuzzy logic and neurofuzzy applications explained. New Jersey: Prentice Hall.

Wasserman, P.D. (1989). Neural computing theory and practice. New York: Van Nostrand Reinhold.

Wexley, K.N., Yukl, G.A., Kovacks, S.Z., & Sanders, R.E. (1972). Importance of contrast effects in employment interviews. Journal of Applied Psychology, 56, 45-48.

Wherry, R.J., Naylor, J.C., Wherry, R.J., & Fallis, R.F. (1965). Generating multiple samples of multivariate data with arbitrary population parameters. Psychometrika, 30(3), 303-313.

Zadeh, L.A. (1965). Fuzzy sets. Information Control, 8, 338-353.

Zadeh, L.A. (1973). The concept of a linguistic variable and its application to approximate reasoning, Memorandum ERL-M 411, University of California at Berkeley.

Zadeh, L.A. (1992). Fuzzy logic: Advanced concepts and structures. IEEE Educational Activities. IEEE Inc., Piscataway, NJ.

Zadeh, L.A. (1994). Fuzzy logic, neural networks, and soft computing, Communications of the ACM, 37, 77-84.

Zedeck, S., & Kafry, D. (1977). Capturing rater policies for processing evaluation data. Organizational Behavior and Human Performance, 18, 269-294.

APPENDICES

APPENDIX 1. RESEARCH QUESTIONNAIRE

Greetings,

Thank you for your participation in this research. The focus of this study will be the policies and decision making strategies that managers use when allocating merit pay to their subordinates. This research effort is part of my doctoral work in the Industrial/Organizational Psychology program at the University of South Florida. The results from this study will contribute to current knowledge about how individual managers make critical organizational decisions such as those regarding pay. Moreover, the results of this study will be relevant to developing new technologies for studying and training individuals in effective organizational decision making.

You will be asked to consider a hypothetical situation and to provide judgments on how you would allocate merit pay to a set of hypothetical employees. You will then respond to a series of questions related to how you made your pay allocation judgments. You have been chosen for inclusion in this study based on your real-world experience in making pay allocation decisions. It should be emphasized that there are no right or wrong judgments or answers in this study. I am only interested in how you approach the hypothetical pay allocation situation.

Your participation in this study is completely voluntary and your answers to the research questionnaires will be kept entirely confidential. If you choose to participate, you will be provided a tailored executive summary of the research project upon completion of the study. This summary will include an overview of the methodology used in the study as well as descriptions of the policies that you and other managers used in making the pay allocation judgments.

If you have any questions about this study or about your participation in it, please contact me at (813) 664-8758 or by e-mail at *dorsey@luna.cas.usf.edu*. Also, please feel free to contact my faculty advisor, Dr. Michael Coovert at 974-0482. Without the support of individuals such as yourself who are generous with their time and expertise, research efforts such as this one could not be conducted. Your help is vastly appreciated!

Thank you for your time and interest,

David W. Dorsey

GENERAL INSTRUCTIONS:

The following research questionnaire contains two parts. The first part contains the Merit Pay Allocation Judgment Task. After completing the judgment task you are then asked to complete a second part which asks you questions relevant to the strategies you used in allocating merit pay to the hypothetical employees. When completing the questionnaire, take breaks as needed, and please make your responses to the questionnaire as thoughtful as possible. Because of time constraints on collecting this data, I would ask that you complete and return the research packet by **September 2**. You should have also received a self-addressed, stamped envelope in which you can return the research packet. Thank you again for your efforts.

Before proceeding to part one, please answer the general demographic questions listed below, which will be used to describe the characteristics of the participants used in this study.

NAME: ____________________________

MAILING ADDRESS:
(for mailing summary reports) ____________________________

____

____

CURRENT JOB TITLE: ____________________________

CURRENT EMPLOYER: ____________________________

Using the following scale, please estimate your level of experience in allocating pay or making recommendations relevant to pay (Please Circle One):

1 - Little Experience (less than 1 year)

2 - Some Experience (more than 1 year up to and including 3 years)

3 - Moderate Experience (more than 3 years up to and including 6 years)

4 - Considerable Experience (more than 6 years up to and including 9 years)

5 - Quite A Lot of Experience (more than 9 years up to and including 12 years)

6 - A Great Deal of Experience (more than 12 years up to and including 15 years)

7 - A Vast Amount of Experience (more than 15 years)

– NOW PROCEED TO PART 1 –

PART 1 - Merit Pay Allocation Judgment Task

INSTRUCTIONS FOR JUDGMENT TASK:

For this judgment task assume the following:

You have just been hired as a manager at the *Personnel Solutions Corporation (PSC)*, one of the largest management consulting companies in the southeastern United States.

Founded in 1977, PSC has diversified consulting services, offering consultation in many areas of management and administration, including such areas as employee development, personnel selection, and legal consultation.

Having established yourself in your new management position, your boss approaches you and asks you to review some of the personnel records from your department. Specifically, you are asked to offer recommendations on how much merit pay, if any, certain employees should receive. Your boss explains that you have been given this assignment in order to assess "how you handle compensation issues", and your boss is also hoping to use your recommendations as a rough estimate of the cost of the merit pay program for your department.

In order to help you make your recommendations, an administrative assistant has put together a personnel profile for each of the employees that you have been asked to recommend merit pay for. A description of the information provided in these profiles and a brief description of the Personnel Solutions Corporate merit pay program are provided below.

The Personnel Solutions Corporate Merit Pay Program

Merit pay at PSC has been defined as "individual pay increases based on performance related factors for individual employees in a previous time period." The merit pay program at PSC allocates to employees an annual pay increase, which is a percentage of the individual's current base salary. For example, if an employee is currently making $30,000 base pay and the employee is given a 10% merit increase, the actual amount of their merit pay raise would be $3,000. At PSC the merit pay increase is built into the base salary for subsequent years. The average merit increase at PSC has typically been around 4% and the range of the increase is specified to be between 0 and 15%.

Note that at PSC, merit pay increases are considered separate from aspects of compensation such as cost of living increases and other aspects of base pay. For this judgment task, your allocations should be related only to annual merit increases, excluding aspects such as cost of living increases.

Managers are in charge of allocating merit pay at PSC and they are given full discretion, hence, they may take various factors into account when allocating the merit increases.

Information in the Employee Profiles

The employee profiles, which you will review in order to make your allocation decisions, contain bar charts that specify five important types of information relevant to the each employee. This information includes the following:

1) **Employee's Performance Rating** - This is the employee's overall performance rating as assigned in the employees latest annual performance review (see the performance scale used for this rating on the top of the following page). Each employee is given a rating between 1 and 7 based on the progress that the employee has made in achieving the goals and objectives that were jointly set by the employee and their direct supervisor as part of a goal setting/management by objectives type of program. Definitions for the 1, 4, and 7 scale points are provided to assist you in understanding the performance scale.

EMPLOYEE'S PERFORMANCE RATING

1	2	3	4	5	6	7

UNACCEPTABLE

- Could **not** be expected to achieve **any** of specified goals or objectives

ACCEPTABLE

- Could be expected to achieve an **acceptable** number of specified goals and objectives

EXCEEDS PERFORMANCE EXPECTATIONS

- Could be expected to achieve **all** specified goals and objectives and accomplish significant achievements **beyond** those expected

2) **Average Performance Rating in Employee's Work Group** - Many employee's actually work in different work groups or sub-departments, and the various groups may have different current levels of individual performance (for example, some work groups may have a large number of poor performers). Information is provided in the profiles that specifies the average individual performance rating in an employee's specific work group. This information is provided on the same scale as the individual performance ratings described above.

3) **Employee's Importance Rating** - Because employees may differentially contribute to the department, this rating (taken from a recent managerial review) indicates how important the employee is to their direct supervisor and department, in terms of the employee having specialized knowledge, skills, expertise or organizational connections that contribute to accomplishing departmental/managerial goals and the employee being more or less difficult to replace. Similar to the performance rating, this information is provided on a 7-point scale with definitions provided for the 1, 4, and 7 scale points (scale is shown below).

EMPLOYEE'S IMPORTANCE RATING

1 2 3 4 5 6 7

NON-ESSENTIAL	CONTRIBUTES	ESSENTIAL
- Could **not** be expected to be **essential** in achieving departmental / managerial goals; Employee would **not** be difficult to replace	- Could be expected to **contribute** in achieving departmental / managerial goals due to specialized knowledge, skills, expertise, or organizational connections; Employee **might** be difficult to replace	- Could be expected to be **essential** in achieving departmental / managerial goals due to specialized knowledge, skills, expertise, or organizational connections; Employee would be **extremely** difficult to replace

4) **Employee's Tenure** -Information regarding tenure is provided in terms of the number of years the employee has been with the company (see the tenure scale below).

EMPLOYEE'S TENURE

1 2 3 4 5 6 7 8 9 10

Number of years with the company

5) **Employee's Current Salary Level** - Information about the employee's current salary level is indicated on a scale (shown below). This scale shows the employee's current base salary as compared to the recommended midpoint salary for employees in that position. The midpoint generally represents the desired average that the organization wishes to pay for a particular job. This information is similar to the comparative salary ratio often used by companies to show how high a given employee's salary is in the pay grade for a specific position.

Midpoint for Position **$32,000**

Employee's Current Salary **$28,000**

$27,000 $29,000 $31,000 $33,000 $35,000 $37,000 $39,000

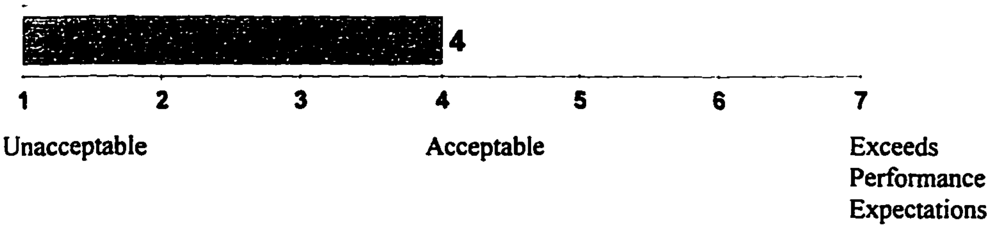
EMPLOYEE'S SALARY LEVEL

Having read and understood the previous instructions, please proceed to the following page and begin the merit pay allocation judgment task. Using the information provided in the instructions, the employee profiles, and your best judgment, allocate a specific amount of merit pay to each of the following employees. Please assign an amount of merit pay to each employee, making sure that you don't skip any. Each employee has been assigned an employee number between 001 and 110. Your merit pay allocation should be between 0 and 15% (corresponding to a percentage of the employee's current base salary). As I suggested earlier, there are no right or wrong answers. Please take breaks and review the scales when needed to ensure that your responses are as thoughtful and realistic as possible. Now please begin!

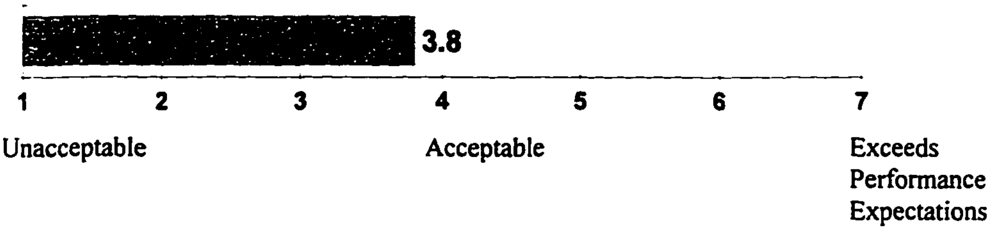
APPENDIX 1. (Continued)

Employee Number: 001

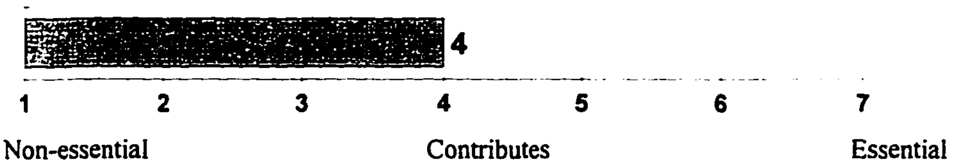
EMPLOYEE'S PERFORMANCE RATING



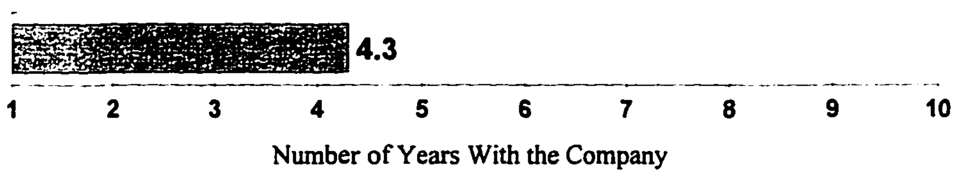
AVERAGE PERFORMANCE RATING IN EMPLOYEE'S WORK GROUP



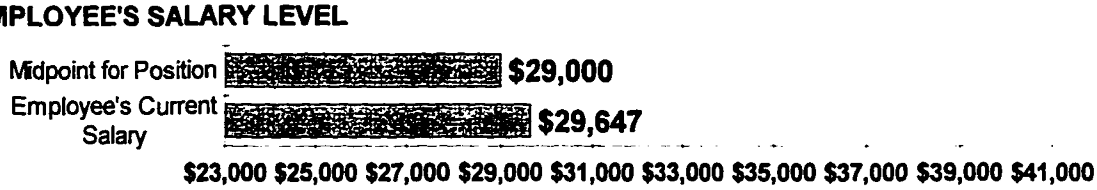
EMPLOYEE'S IMPORTANCE RATING



EMPLOYEE'S TENURE



EMPLOYEE'S SALARY LEVEL



YOUR MERIT PAY ALLOCATION TO THIS EMPLOYEE = _____________ (between 0-15%)

END PART 1.

PART 2 - Post Judgment Questionnaire

INSTRUCTIONS FOR QUESTIONNAIRE:
Following the instructions listed for each question, please provide answers for each of the following items.

1) In the previous merit pay allocation judgment task you looked at five pieces of information or variables for each hypothetical employee. Please rate the relative importance that you placed on each of the variables (listed below) by dividing **100 points** among the five variables, with higher points assigned to the variables that influenced you most in allocating merit pay.

EXAMPLE:

Employee's Performance Rating -	35
Average Performance Rating in Employee's Work Group -	10
Employee's Importance Rating -	30
Employee's Tenure -	10
Employee's Current Salary Level -	15

Your Ratings:

Employee's Performance Rating	________
Average Performance Rating in Employee's Work Group	________
Employee's Importance Rating	________
Employee's Tenure	________
Employee's Current Salary Level	________
TOTAL =	100 Points

2) Consider the variable **Employee's Performance Rating** which was used in the employee profiles:

- What whole number rating between 1 and 7 would you consider **most typical** for someone with (fill in an answer for each of the following):

a "low" performance level ? ________
a "moderate" performance level ? ________
a "high" performance level ? ________

3) Consider the variable **Average Performance Rating in Employee's Work Group** which was used in the employee profiles:

- What value between 1.0 and 7.0 would you consider **most typical** for a group with (fill in an answer for each of the following):

a "low" average performance level ? ________
a "moderate" average performance level ? ________
a "high" average performance level ? ________

4) Consider the variable **Employee's Importance Rating** which was used in the employee profiles:

- What whole number rating between 1 and 7 would you consider **most typical** for someone with (fill in an answer for each of the following):

a "low" importance level ? ________
a "moderate" importance level ? ________
a "high" importance level ? ________

5) Consider the variable **Employee's Tenure** which was used in the employee profiles:

- What value between 1.0 and 10.0 years would you consider **most typical** for an individual with (fill in an answer for each of the following):

a "low" level of tenure ? ________
a "moderate" level of tenure? ________
a "high" level of tenure ? ________

6) Consider the variable **Employee's Current Salary Level** which was used in the employee profiles:

- Assuming a recommended salary midpoint of $33,000 for an individual's position, what salary value between $23,000 and $41,000 would you consider **most typical** of (fill in an answer for each of the following):

a "low" salary level ? ________
a "moderate" salary level ? ________
a "high" salary level ? ________

7) Consider the **merit increases** which you allocated in the previous judgment task. What amount of an increase between 0 and 15% would you consider **most typical** of (fill in an answer for each of the following):

a "very low" merit increase ? ________
a "low" merit increase ? ________
a "moderate" merit increase ? ________
a "high" merit increase ? ________
a "very high" merit increase ? ________

For items 8 - 37, consider how you allocated merit pay to the hypothetical employees in the previous judgment task, and describe your strategy for allocating merit pay by responding to the following items. Each item lists a general pattern of profile values for an employee (in terms of each variable being either generally HIGH, MODERATE, or LOW). **Respond to each item by circling the general amount of merit pay that you would allocate for an employee with that particular pattern of general profile values.**

For example, consider the case below. Here the "VERY HIGH" label has been chosen to indicate that the allocator would give a relatively VERY HIGH merit pay allocation for an employee having a relatively HIGH performance rating, a relatively MODERATE group performance rating, a relatively MODERATE importance rating, a relatively HIGH tenure level, and a relatively MODERATE current salary level.

EXAMPLE: **If the employee profile values were as follows:**

Employee's Performance Rating - HIGH
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - HIGH
Employee's Current Salary Level - MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

Following the previous example, please provide answers for each of the items below:

If the employee profile values were as follows:

8)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

9)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - HIGH
Employee's Current Salary Level- MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

10)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - LOW
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- LOW

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

11)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- LOW

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

12)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - LOW
Employee's Current Salary Level- MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

13)
Employee's Performance Rating - LOW
Average Performance Rating in Employee's Work Group - LOW
Employee's Importance Rating - LOW
Employee's Tenure - LOW
Employee's Current Salary Level- LOW

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

If the employee profile values were as follows:

14)
Employee's Performance Rating - LOW
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - LOW
Employee's Tenure - LOW
Employee's Current Salary Level- MODERATE
Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

15)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - LOW
Employee's Current Salary Level- LOW
Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

16)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - HIGH
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- MODERATE
Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

17)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - LOW
Employee's Importance Rating - MODERATE
Employee's Tenure - LOW
Employee's Current Salary Level- LOW
Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

18)
Employee's Performance Rating - HIGH
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - HIGH
Employee's Current Salary Level- MODERATE
Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

19)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - HIGH
Employee's Importance Rating - MODERATE
Employee's Tenure - HIGH
Employee's Current Salary Level- MODERATE
Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

If the employee profile values were as follows:

20)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - LOW
Employee's Tenure - LOW
Employee's Current Salary Level- LOW

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

21)
Employee's Performance Rating - HIGH
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - HIGH
Employee's Tenure - MODERATE
Employee's Current Salary Level- HIGH

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

22)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - HIGH
Employee's Current Salary Level- LOW

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

23)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - HIGH
Employee's Importance Rating - HIGH
Employee's Tenure - HIGH
Employee's Current Salary Level- HIGH

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

24)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - HIGH
Employee's Current Salary Level- HIGH

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

25)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - LOW
Employee's Tenure - MODERATE
Employee's Current Salary Level- LOW

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

If the employee profile values were as follows:

26)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - HIGH
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- LOW

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

27)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - LOW
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

28)
Employee's Performance Rating - LOW
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

29)
Employee's Performance Rating - HIGH
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

30)
Employee's Performance Rating - LOW
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- HIGH

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

31)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - HIGH
Employee's Importance Rating - MODERATE
Employee's Tenure - LOW
Employee's Current Salary Level- MODERATE

Your merit pay allocation
would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

If the employee profile values were as follows:

32)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - HIGH
Employee's Importance Rating - LOW
Employee's Tenure - MODERATE
Employee's Current Salary Level- LOW

Your merit pay allocation would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

33)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - LOW
Employee's Importance Rating - LOW
Employee's Tenure - LOW
Employee's Current Salary Level- LOW

Your merit pay allocation would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

34)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - HIGH
Employee's Tenure - HIGH
Employee's Current Salary Level- MODERATE

Your merit pay allocation would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

35)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - LOW
Employee's Tenure - MODERATE
Employee's Current Salary Level- MODERATE

Your merit pay allocation would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

36)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - MODERATE
Employee's Importance Rating - MODERATE
Employee's Tenure - MODERATE
Employee's Current Salary Level- HIGH

Your merit pay allocation would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

---

37)
Employee's Performance Rating - MODERATE
Average Performance Rating in Employee's Work Group - LOW
Employee's Importance Rating - LOW
Employee's Tenure - MODERATE
Employee's Current Salary Level- LOW

Your merit pay allocation would be (circle one) - VERY LOW LOW MODERATE HIGH VERY HIGH

YOU HAVE NOW FINISHED! THANK YOU AGAIN FOR YOUR EFFORTS!

VITA

David Dorsey received a B.A. in Psychology from the University of South Florida and an M.A. and Ph.D. in Industrial/Organizational Psychology, also from the University of South Florida. David completed a graduate minor in Computer Science as part of his graduate studies.

During his graduate tenure, David was active as a researcher, publishing in journals such as Personnel Psychology and the Journal of Applied Psychology and presenting papers at both domestic and international conferences. In conducting research, David has specialized in integrating research and methodologies from a multidisciplinary perspective to address substantive issues ranging from the use of expert systems and computer simulations in Industrial/Organizational Psychology to investigations focusing on cognitive measures of knowledge acquisition relevant to training design and evaluation. While attending graduate school, David was also active as a teacher, consultant, and practitioner.