Federal Reserve Bank of Minneapolis
Research Department Working Paper

Historical

APPROPRIATE POLICY DECISIONS WITH MULTIPLE VAGUE OBJECTIVES:

THE CASE OF PERSONNEL POLICIES

Michael J. Stutzer

Working Paper 311

June 1986

NOT FOR DISTRIBUTION
WITHOUT AUTHOR'S APPROVAL

*Federal Reserve Bank of Minneapolis

Normative studies are employed to determine appropriate conduct. They are used to guide what we have a right to do, what we should do, and/or what we must do. If decision makers have any free will to act at all, then nontrivial choices over alternative decisions must be made. It is thus legitimate to evaluate decisions within this normative context, for the alternatives will have different normative implications.

But do employers have sufficient free will to make normative studies worth worrying about? To help answer this, we must examine the possibility that the legal and competitive environments constrain firm decisionmaking severely. If so, there is little prospect that management could implement many recommendations from normative studies, other than those already dictated by those environments.

The legal environment in the U.S. doesn't severely constrain most important areas of personnel administration. In the area of employee/management relations, in the absence of explicit contracts, employers have traditionally been legally able to dismiss employees "at will," regardless of cause (Summers, 1980). Insubordination, regardless of its nature, could result in an employee being dismissed with no remedies against the employer or right to unemployment compensation. While the Restatement of Agency does not permit an employer to force an employee to perform acts which are "illegal or unethical," the employee's remedy to avoid having to do so is to quit, and even then is still obligated not to divulge this information (other than possibly criminal actions) in ways which may harm the former employer.

There are some legal constraints, though. Federal laws do prohibit dismissal for union membership activities and for reasons deemed discriminatory on the basis of race, creed, nationality, sex, or age (Summers, 1980). Further, the courts have granted additional rights to government employees (Blumberg, 1971). In Pickering vs. Board of Education (391 U.S. 563 (1968)), the Supreme Court protected a "whistle blowing" school teacher (Ewing 1978). And the Privacy Act of 1974 requires federal agencies to allow their employees to inspect their own personnel files.

Hiring and promotional practices of employers receiving federal funds are constrained to some extent by affirmative action requirements. But binding numerical quotas are usually not required to be in accord with them, and there is even a chance that the use of quotas may leave the employer open to a reverse discrimination suit. Further, the Reagan Administration, whose agencies are charged with monitoring compliance with affirmative action, is opposed to numerical quotas. As a result, good faith recruitment efforts seem to be sufficient in most cases for compliance.

More frequently, the competitive environment places constraints on personnel practices. Indeed, in extreme cases, competitive pressures may leave no room for principles other than profit maximization. To see this, consider the plight of an employer competing in perfectly competitive markets for its products and its inputs, including labor. In this environment, firms

which act to maximize profits wind up selling their products at
the lowest possible price. Firms which take actions conflicting
with profit maximization will not be able to match these low
prices, and will eventually be forced out of business due to loss
of demand. Accepting as axiomatic that the prime imperative of
employers is to ensure their own survival, the perfectly com-
petitive environment forces them to subordinate other principles
to the constraining force of profitability. While other prin-
ciples might not be violated, this would occur only when they
didn't conflict with profit maximization. In this sense, em-
ployers have no free will to do anything else but maximize prof-
its.

Because normative theories are attempts to provide
frameworks for making normative judgments about particular
actions, the argument in the preceding paragraph is really a
theory, which I'll call competitive profitism. Operationally, the
theory is expressed in the following way:

(1)   If an employer's profit maximizing decision
        takes place in a perfectly competitive en-
        vironment, then that decision is oblig-
        atory. This particular decision maximizes
        profits in a perfectly competitive environ-
        ment. Therefore, this decision is morally
        obligatory.

If this theory is both valid and sound, i.e., if the conclusion does indeed follow from the two premises, and the latter are true, then one can easily see how it might be put to use in determining many personnel policies. For example, in considering whether or not to implement proposed "pay equity" concepts in salary administration, the employer need only consider the effects on the present value of its own profits. This would be relatively easy to calculate in a perfectly competitive environment, because the employer would assume that its policies would not influence the policies of other competing firms. Concepts of equity and fairness would be relevant to discuss only insofar as they affected the "bottom line." Similarly, thorny questions such as whether or not to monitor employee use of business phones and other property, or to hire applicants convicted of crimes, would be relatively simple to answer. The latter, for instance, would be made simpler by needing to focus only on the relationship between former criminal behavior and employer profitability. In some cases, e.g., a security-conscious data processing firm interviewing a former hacker convicted of violating its computer system, it may even be obligatory to hire a former criminal.

But is the theory of competitive profitism valid and sound for the Federal Reserve Bank? Obviously, while the Fed does offer competitively priced services, it has other operations which don't sell services at all. Still, it must hire labor and other inputs in competitive markets. Perhaps competitive profitism can be modified to produce a sound theory applicable to the Federal

Reserve. However, even in idealized circumstances where the theory's soundness is not at issue, its validity may still be called into question. But before that deeper issue can be addressed, it will help to understand what modern normative studies can contribute to answering it.

Accordingly, we now turn to a brief survey of modern normative studies. Both its length and my lack of knowledge prevent this survey from being definitive in any dimension. Rather, it is intended to lead the reader through a thought process inherent in existing normative studies, and to illustrate the difficulties inherent in formulating a normative theory useful for personnel policy determination.

An Odyssey Through Normative Studies

Normative studies are extremely broad in scope--a trait not shared by most scientific investigations. So what can an economist contribute to this field of study? Should he/she even attempt to add to the mass of literature in this area, produced by such giants as Plato, Aristotle, John Stuart Mill, Kant, Nietzche, Rabbi Hillel, Woody Allen, St. Paul, and Norman Vincent Peale?[1] Most who cherish the notion of intellectual freedom believe that anyone has a right to express their views on any intellectual enquiry, but that does not imply that I should or must do so. But if one is capable of adding something new to such an important

_____

[1] Adlai Stevenson found St. Paul appealing and St. Peale appalling. You will be better equipped to judge them after finishing this section.

subject, isn't one obligated to do so, especially if little or no loss of personal welfare results from having done so?

Note that questions concerning the individual's right to do and obligation to do something arose in the last paragraph, even though it was concerned with the seemingly trivial issue of my writing that paragraph and the many to follow. Furthermore, in affirmatively answering the question about whether or not I had a right to write this paper, I employed a theory which can be summarized as follows:

> If an action results in an increase in knowledge, then I have an absolute right to do it; i.e., it is right. Writing this paper will increase our knowledge about normative studies, albeit infinitesimally. Therefore, it is right to write this paper.

The judgment that it is right to write this paper logically follows from the above two premises, by invocation of the familiar argument from classical logic known as modus ponens. The first premise is dubbed a principle, for it permitted us to derive a normative judgment concerning a particular action (writing this paper) from an if-clause stating a desideratum concerning a class of actions, i.e., the principle's objective or rule. This theory, comprised of a single principle, is called epistemism (Rosen, p. 33), and is one of many theories which have been proposed to deal with questions like those of the preceding paragraph.

While epistemism seems reasonable when applied to the question of my right to write this paper (although at this point some of you may still be in doubt of this), one doesn't have to think too hard to come up with other questions where its use seems immoral. For example, some academics have argued that the Federal Reserve experiment with novel means of monetary control, such as controlling the supply of total nonfinancial debt, arguing that even if such a policy had bad immediate effects, the knowledge gained by the experiment would aid in the construction of better policies in the future. Epistemism would justify this experiment even if the policy change caused massive economic disruption and drove despairing pensioners to suicide! Or, suppose that a personnel department implements an experimental piecework pay system for a select group of low productivity economists, in the hope that it might gain knowledge valuable in deciding whether to permanently extend the system to all employees. Furthermore, suppose the pilot employee group unanimously opposes being singled out for this treatment. Epistemism justifies the pilot program regardless of the feelings of the pilot employee group, or the effects it has on them--which could be quite traumatic.

At this point, you are probably itching to modify the theory of epistemism to get around these counterexamples. The problem seems to be that epistemism doesn't consider the consequences of the decisions it justifies, negative or positive. It states that knowledge production is justified for its own sake. Theories which don't require explicit mention of the consequences

flowing from decisions they justify, like epistemism, are called
<u>deontological theories</u>. To address the above counterexamples, one
might try replacing the theory with the principle that a decision
is right if those benefitting from it gain more than those who
lose from it lose, or by additionally insisting that the losers
actually be compensated by the gainers. Otherwise, the decision
is wrong. This would then be a compensation-based <u>utilitarian
theory</u>, for it depends on a quantification of the consequences
that decisions have on individuals' welfares. Theories which,
like utilitarianism, depend solely on the consequences that deci-
sions have, are called <u>teleological theories</u>.

But such teleological theories, like utilitarianism and
competitive profitism, are also plagued by counterexamples. With
regard to compensation-based utilitarianism, if compensation is
not paid to the losers, then they are prey to the tyranny of large
numbers of other people gaining by small amounts individually at
their large individual expense. There is thus no room for abso-
lute minority rights, like those guaranteed by the U.S. Constitu-
tion, which seems wrong. Even if compensation is to be paid, how
is it to be determined? Selfish, individuals have an incentive to
overstate the objectively unmeasurable harm done to them. Even if
these problems could be addressed, there are well-known paradoxes
from welfare economics that plague compensation-based principles
(see Quirk and Saposnik, pp. 120-123).

Counterexamples to Competitive Profitism

Suppose that slavery and/or indentured servitude were profitable in the posited perfectly competitive world of (1). Most citizens of modern private enterprise economies abhor these practices, yet competitive profitism would view them as obligatory. This counterexample is not merely hypothetical, because there is some evidence that slavery was quite profitable in the largely laissez-faire, pre-Civil War United States.

One could argue that the banning of slavery represents an efficient way for society to enforce a property right it has somehow agreed upon; i.e., every human's right to own their own labor endowment. As such, it would be analogous to the adoption of commercial codes and constitutional protections governing other property rights. These "rules of the game" are inevitably adopted in all private enterprise economies, yet it is doubtful that competitive profitism itself could explain their evolution, which has differed across countries.

The differences in explicit and implicit property rights across private enterprise economies are a fruitful source of other problems with competitive profitism. In the U.S., most employers refuse to issue guarantees of long-term employment to employees, implicitly invoking competitive profitism to argue that maximizing profits makes it obligatory to refuse to provide workers right to employment. Yet many large Japanese employers offer de facto guarantees of employment far in excess of the U.S. custom, believing it to be profit maximizing. It is possible that both American

and Japanese firms are right about their respective profitability
calculations, in which case advocates of competitive profitism
would have to admit that what is obligatory in Japan is wrong in
the U.S. But then competitive profitism is not of universal
validity, and must be augmented by qualifying language of some
kind to determine when particular decisions, like those granting
long-term employment contracts, are obligatory.

Those problems do not prove that there is no validity to
some theory incorporating the very real profitability constraint
many firms face in devising personnel policies in a competitive
environment. There may be a more sophisticated theory which
adequately addresses these problems, while still capturing the
crux of the argument where it seems reasonable.

The decision procedures developed later on this paper
should be of value in forming such a theory. The next section,
though, shows that merely adding additional principles to (1)
(e.g., if a decision forces someone to work against their will,
then that decision is wrong) is unlikely to produce a theory free
of problems.

Multiple Principles: The Problem of Logical Contradiction

One might hope that counterexamples could be eliminated
from single principle theories by augmentation with additional
principles. For instance, one of the problems with compensation-
based utilitarianism occurred because of the harmed individuals'
incentive to overstate the required compensatory payments. One
might argue that adding the principle, "If a person knows the

truth, then that person is obligated to tell the truth," to our theory would eliminate the problem of determining compensation, by forcing those harmed to state the correct level of compensation they need. But might not this rule itself also be plagued by counterexamples? Suppose your mother has a weak heart, and you know that she will die if you tell her that you support the experimental monetary policy. She then demands to know your opinion. Telling the truth in this instance seems wrong, even though it is mandated by that principle. Furthermore, even if one sees nothing wrong in killing one's mother to avoid violating a lofty principle, adopting that principle in addition to the compensation-based utilitarian rule leads to another problem. For had you lied so that your mother would live, she would no doubt be bettered by more than it harmed you to lie. (If you don't believe this, tell your mother you are opposed to Medicare. If she doesn't die, you'll wish you had.) In making the decision mandated by the rule of truth telling, you would have thus taken an action which violated the moral rule of compensation-based utilitarianism. In classical logic, a decision can't be both right and not right, so it dictates that there is a logical contradiction in this case, caused by insisting that more than one principle be adhered to simultaneously.

The Utterly Dismal Theorem

Unfortunately, common sense counterexamples seem to plague all known theories based on adherence to principles formulated in the following way:

(2)   For all x, if x is P, then x is M. This x is

           P. Therefore, this x is M.

where x is a particular decision, P is an objective that decisions
may or may not attain, (e.g., the decision produces knowledge, or
it benefits some people by more than it harms the rest), and M is
some normative criterion used to evaluate decisions (e.g., they
are right, or are obligatory). Further, all theories requiring
that multiple principles of the form (2) hold simultaneously seem
to be plagued by the possibility of common sense contradictions
(see Rosen (1978) for a nice presentation of all of this).

      As if this weren't bad enough, the vagueness inherent in
many principle-based theories' propositions makes it unnatural to
apply classical logic in deducing appropriate conduct. For exam-
ple, sometimes it is hard to say whether a decision is completely
definitely right or definitely wrong, yet we may feel confident
that some decisions seem more right (or wrong) than others. Some
other times it is hard to say whether we totally believe or dis-
believe that a particular decision has the property P. Still, one
might possess some intermediate degree of belief, rather than
total certainty, in the truth or falsity of a decision satisfying
the principles' objectives and normative criterion. But classi-
cal, two-valued logic is hard pressed to cope with this.

      If economics is the dismal empirical science,
principle-based ethics placed in the above form must then be the
utterly dismal normative science. An application of this bleak
finding follows.

Application of the Utterly Dismal Theorem: The Four-Way Test

For this paper, a particularly relevant application of this utterly dismal theorem is to the Rotary International Club's "Four-Way Test of the Things We Think, Say, or Do," which all Rotarians are pledged to uphold and, evidently, to hang prominently in sight of visitors to their offices. No doubt one or more of its tenets have made their way into many personnel policy determinations by non-Rotarians as well. Without further ado, The Test is:

1. Is it the truth?
2. Is it fair to all concerned?
3. Will it build good will and better friendships?
4. Will it be beneficial to all concerned?

All four of these objectives are easily translated into the form (2) above, so they can all be viewed as principles. Note the obvious vagueness in their interpretation.

The plight of your inquisitive mother showed that Principle 1 can lead to a common sense counterexample. And if we are supposed to interpret these rules categorically, i.e., as having to hold simultaneously, the dead mother left graphic proof that a logical contradiction occurs between Principle 1 and Principle 4, and thus did not die in vain--unless she died in a nursing home run by a Rotarian.

Of course, other less dramatic counterexamples and contradictions within the Four-Way Test could be constructed from

the daily decisions that personnel departments must make. For example, when a personnel department is contacted for information by a prospective employer of a former employee who was discharged (perhaps for smashing his supervisor's Four-Way Test plaque over the latter's head), should the personnel department truthfully divulge this information and/or answer all other relevant questions about the applicant? The personnel department could adopt a policy of not revealing any information about the nature of the employee's leaving, and claim not to be lying. But this might still be construed to violate Principle 1 if we read it to mean that we are obligated to tell the whole truth when asked. By hiding the whole truth under the cloak of a privacy protecting policy, the department may be technically not lying, but really isn't telling the whole truth. In any event, refusing to disclose that the employee was terminated doesn't seem very fair (Principle 2) nor beneficial (Principle 4) to the shareholders, management, and employees of the prospective employer, although it doesn't seem totally unfair, either. Telling the whole truth, to the contrary, seems to satisfy these three rules but violates Principle 3, for it is unlikely to build better friendships between management and discharged employees or current employees, who could fear that both their privacy and future income may be threatened by running afoul of some manager or management policy in the future. Thus, there is no decision which the department could take in response to this request which exactly satisfies all four of the Rotary's principles.

One attempt to escape from the problems of common sense counterexamples and logical contradictions inherent within theories comprised of several principles utilizes notions related to those of W. D. Ross (1930). Why not just decree that contradictions can be avoided by permitting one or more principles to override other conflicting rules when such conflicts arise? A morally troubled Rotarian personnel director might thus argue that all four principles of the Four-Way Test are _prima facie_, rather than categorical, principles. In the previous example, the director may decide that the duties to abide by Principles 1, 2, and 4 outweigh the duty to abide by Principle 3, thus justifying disclosure of employee terminations for this particular serious cause. Of course, had the personnel director placed a greater emphasis on maintaining employee good will, the decision may have been to not disclose, letting Principle 3 outweigh the other three rules.

Therein lies the rub with the concept of prima facie theories. In the absence of categorical principles assigning priority weights to the prima facie principles, the theory is incomplete. If the personnel director is free to assign weights to the prima facie principles on a case-by-case basis, then the normative judgments are really being made outside the theory. If one attempts to complete the theory by appending principles governing the weighting of the prima facie principles' priorities (e.g., building good will and better friendships always takes precedence over telling the truth), then those additional prin-

ciples must be categorical, for if they are prima facie, the theory would still be incomplete. But if the additional rules are categorical, then the theory is no longer a prima facie one, so it is once again going to become easy to find counterexamples and contradictions once again (see Rosen, pp. 126-131). For example, consider the personnel director who adopts the above priority ranking, and hence must not tell the whole truth about the discharge of employees, no matter how pertinent such information might be to their prospective employers.

Act-Based Theories

Another attempt to eliminate the problems of counterexamples, logical contradictions, and vagueness inherent in rulebased theories is to abandon the rules altogether. Why not examine each moral judgment on a case-by-case basis, formulating separate sets of _factual_ conditions necessary to justify each particular moral judgment? One would then have a _decision-based theory_ governing each moral judgment that needed to be made. For example, the personnel director might adopt a policy stating that:

1. If an employee is discharged for a cause other than committing a crime,

   and

2. that cause was something that did not expose the employer to the risk of a lawsuit,

3. then, we will not disclose that the employee was terminated,

4.   else, we will disclose that the employee was terminated

Note that no principles were explicitly stated in the above "theory," which seems better thought of as a policy statement on a particular issue. Of course, such principles may have been present in the mind of the person(s) who formulated it. The formulator(s) might actually have been applying the principle that if a person breaks the law on company time, then the company is obligated to inform all parties potentially affected by it (which includes prospective employers). If so, then counterexamples to other policies so formulated may occur due to that principle's unreasonableness in some circumstances.

Act theorists would then argue that a counterexample (i.e., a disagreement with the policy in some circumstance) can be resolved by identifying the key reasons for the disagreement, and incorporating them into a more elaborate policy statement. For example, suppose the employee was discharged for committing a crime which is no longer illegal; perhaps for driving a company vehicle out of a corporate parking lot into a lane solely dedicated for buses and taxis, which has since been opened to general use during rush hours (were it so!). All might argue that disclosure of an employee's termination is unwarranted in this circumstance, despite the fact that it did indeed involve a conscious criminal act at the time. To attain a consensus policy, the personnel director could just add the language ". . . which is still a crime" to the first condition in the policy. Hopefully, it would be possible to continue clarifying the issues and appro-

priately modifying the policy statement until it was unanimously
accepted.

A personnel department could painstakingly attempt in
this way to achieve a consensus for each decision that it must
make, although the resources devoted to the task could be enor-
mous. But even if this were a feasible framework for decision-
making, might not the whole look worse than the individual
parts? While each separate policy might seem reasonable when
viewed in isolation, it seems unlikely that coherency or consis-
tency across policies would result from this framework. Perhaps
due to bounded rationality, there is an irresistible urge to find
broad, overreaching objectives (e.g., the attainment of horizontal
equity in our treatment of employees) needed to make a set of
policies ethically coherent. The absence of such principles
leaves unfulfilled the very human need to categorize and system-
atize complexity.

Toward a Practical Framework for Coherent Personnel Policies

As we have seen, categorical principle-based theories
suffer from counterexamples, internal logical contradictions, and
vagueness. Prima facie principle-based theories help solve the
first two problems, but are incomplete, for they don't contain a
theory for determining the order in which the rules should be
applied. Act-based theories have the potential to avoid all three
problems, but the more they do so, the more they lose the sim-
plifying, organizing framework that inevitably drives people
toward principle-based theories. Is there any avenue left worth
exploring?

The rest of this paper details the development of a mechanism for constructing a theory which at least is consistent with a decision maker's own beliefs, and forces the decision maker to confront logical contradictions and vagueness inherent in those beliefs. It is based on the notion that decision makers can better construct coherent act-based theories for making a variety of decisions by forcing them to carry out a four step, computer-assisted procedure in constructing each act-based theory (i.e., particular personnel policy).

The first step of the procedure requires the decision maker to propose any number of plausible principles to use in evaluating the desirability of possible decisions. The decision maker must then classify the possible decisions into mutually exclusive categories, called rules or guidelines, having the property that all decisions falling within a guideline will have the same desirability.

The second step of the procedure elicits the decision maker's degrees of belief about "how much" decisions falling within a guideline satisfy the often vague objectives embodied in the principles that have been proposed.

The third step is to force the decision maker to give rough, quantitative meaning to their common sense, prior degrees of belief about the guidelines' satisfying the normative criterion.

The fourth step processes the information elicited during the above three steps through a microcomputer-based, inter-

active fuzzy logic analyzer, (see Whalen and Schott, 1983, for a survey of existing fuzzy expert systems), which helps the decision maker understand the logical implications of her/his beliefs. The proposed fuzzy logic analyzer uses the calculus of fuzzy restrictions (see Zadeh, 1973 and 1975) and the information elicited in the above three steps to determine logically consistent, implicit degrees of belief, i.e., prima facie weightings, attached to the various principles the decision maker purports to believe in. The system then uses a fuzzy inference procedure to calculate logical, revised evaluations of the degrees in which the guidelines satisfy the normative criterion. Using these revised evaluations, the system also updates the prima facie principle weights to be consistent with them.

The system thus presents the decision maker with the prima facie principle weights that must be adopted to logically support the decision maker's beliefs. Upon reviewing the weights, the user may choose to revise her/his prior beliefs, or propose new possible guidelines and/or principles for another pass through the system. The end result of these iterations between the decision maker and the system should be an act-based theory as consistent as possible with the decision maker's beliefs and desired weights.

By using this four step, iterative procedure in devising each important policy, the decision maker can determine whether or not the group of policies are reasonably coherent, in the sense of the relative invariance of prima facie weights across policies.

If not, further iterations can be conducted to reduce the weights' variability to acceptable tolerances.

The System: An Oversimplified Example

To facilitate the understanding of the proposed system, we will work through an oversimplified example of how it might work. To do so, let us return to the discussion of the disclosure of reasons for separation from employment. In response to a great number of types of requests for information, for purposes of simplicity let us suppose that in step one I thought of only three possible guidelines to use in classifying my possible decisions:

A1: disclosure when terminated for illegal behavior.

A2: disclosure when terminated for reasons other than illegal behavior (e.g. incompetence).

A3: disclosure when laid off in a budget cut.

Suppose, again for reasons of simplicity, that I only thought of two principles I believed in, or was otherwise interested in evaluating:

R1: If a decision is telling the whole truth, then that decision is right; else, that decision is not right.

R2: If a decision is fair to all concerned, then that decision is right; else, that decision is not right.

As you can see, I'm sort of a semi-Rotarian. Also, note that these principles are stated in a different form than (2), for they contain an "else" clause. It seems to me that this is the

normal way people think of normative principles—that if a deci-
sion doesn't have the stated property that makes it right, then it
does have some unstated property that makes it wrong, i.e., not
right. In fact, this formulation is still consistent with the
form (2) if we replace each if-then-else statement by two if-then
statements: if x is P then x is M and if x is not P then x is not
M. Symbolically, denote the two objectives by:

P1: a decision which is telling the truth

P2: a decision which is fair to all concerned

and the normative criterion to be:

M: a decision which is right

Then, the principles R1 and R2 take the normal form:

R1: If x is P1, then x is M; else x is not M.

R2: If x is P2, then x is M; else x is not M.

where x ranges over all decisions, which are classified into three
classes determined by the three guidelines A1, A2 and A3. It
should be noted, though, that else-clauses other than "x is not M"
could be built into the system.

Step two of the user input now commences. I, as user,
must now assign degrees of belief assessing the degree to which
decisions falling within each guideline satisfy each principle's
objective, i.e., the degrees to which each of A1, A2, and A3
satisfies P1 and P2. For now, let's assume that my beliefs are
not vague in this regard, so I must only determine if each of the
six possible beliefs are exactly true or exactly false. It will

be extremely useful in what is to follow if I use the Boolean
representation of logic, assigning the number "1" to an exactly
true statement and a number "0" to an exactly false statement.
The following table summarizes my beliefs:

	P1	P2
A1	1	1
A2	1	0
A3	1	1

(3)

Thus, the ones in the first row in the above table
indicate that I believe that disclosure of termination for illegal
behavior is both telling the truth and fair to all concerned
(perhaps the employee will clean up her act if she knows her
dismal record will follow her around, and will be better off in
the long run for having done so). The zero in the second column
of the second row means that I construe disclosure of termination
for other reasons as not fair. Others may disagree completely
with my belief, of course, believing this to be perfectly fair.

But there is a third possibility. Some may believe that
disclosure of termination for other than illegal behavior is
neither exactly fair nor unfair. Rather, they may believe that it
is "sort of" fair, being fair to the enquiring party but not too
fair to the employees. It is precisely this vagueness that we are
assuming away for the moment. To handle it, the system will
permit a user to assign an intermediate degree of belief to the
statement, i.e., a number somewhere between zero and one. Another
way to think about this is to visualize the set of decisions which

are fair as having an imprecise boundary. Some elements of the
set (i.e., decisions) are definitely in the set, i.e., definitely
fair, and hence have a grade of membership in the set equal to
one. Other decisions are definitely not in the set, and so have a
grade of membership equal to zero. Yet other elements aren't
exactly in the set or out of the set; rather, they have an inter-
mediate grade of membership strictly between zero and one. The
closer an element's grade of membership is to one, the closer that
element is to being definitely in the set, while the closer it is
to zero the closer the element is to being definitely not in the
set. A grade of membership equal to .5 then indicates total
vagueness about whether the element is in or out of the set.

A group of elements, coupled with a function assigning a
grade of membership between zero and one to each element is called
a fuzzy set. Thus, the system allows the user to represent her
beliefs about "x is P" as a fuzzy set, with decisions falling
within the guidelines as its elements, and her degrees of belief
in the various decisions satisfying "x is P", i.e., the column for
that P in table (3), as its grade of membership function. In our
oversimplified example, then, we have assumed that my degrees of
belief, which are listed in table (3) are all either zero or
one. In this special case of no vagueness, the associated fuzzy
sets (for "x is P1" and "x is P2") are said to be crisp.

The third step in the user input to the system is for
the user to assign prior degrees of belief to the principles'
normative criterion, which in this example is the rightness, of

the possible decisions. Once again, we'll assume for simplifica-
tion that I ,as user, have no prior vagueness about this, and will
hence assign a grade of membership of either a zero or one to each
element of the fuzzy set associated with "x is M". I have repre-
sented my beliefs below by augmenting table (3) with the addi-
tional input:

		P1	P2	M
	A1	1	1	1
(4)	A2	1	0	0
	A3	1	1	1

The third column of the table indicates that I defi-
nitely believe that disclosure of termination for illegal behavior
is right, as is disclosure when an employee is laid off due to a
budget cut. But I also definitely believe that disclosure when
termination is for reasons other than illegal behavior is not
right, i.e., wrong. These are my beliefs prior to the following
evaluation by the fuzzy logic analyzer.

The logic analyzer now takes over in step four of the
procedure. First, it determines the degree of belief I should
hold in each principle, if I want my prior beliefs to be logically
consistent with each principle. These are prima facie weightings
of the importance of these principles to me, which are implicit in
my prior beliefs represented in table (4). Because table (4) is
crisp, the logic analyzer applies classical logic to determine the
implied truth or falsity of the two if-then-else principles R1 and
R2. With respect to R2, the second and third columns of table (4)

indicate that whenever I thought that "x is P2" was true (i.e., when x lies within A1 or A3), I also thought that "x is M" was true. Further, whenever I thought that the former was false (i.e., when x lies within A2), I thought that the latter was false. There is no logical contradiction in any of this, for each of these three deductions about M follows logically by an application of modus ponens. The system figures this out and assigns the number one to all three possible uses (i.e., one use for each possible guideline) of R2, thus indicating that I implicitly completely believed in R2 when determining the desirability of each possible guideline. With the respect to R1, though, the first and third columns of table (4) indicate that I implicitly believed R1 in evaluating A1 and A3, but disbelieved it when evaluating A2. In the latter case, the zero in row 2 indicates that I believed that disclosure for reasons other than illegal behavior was telling the truth, yet still was not right, which contradicts R1. My implicit degree of belief in R1 when applied to A2 is zero, i.e., A2 is a counterexample to R1 implicit in my prior beliefs. Exhausting the six possible evaluations, the system produces the following prima facie weighting table:

		R1	R2
	A1	1	1
(5)	A2	0	1
	A3	1	1
	AVG	2/3	1

where the zero indicates the aforementioned implicit counter-example. The fact that the numbers in the second row are different indicates that a categorical interpretation of the rules leads to a logical contradiction implicit in applying them to my beliefs about A2 (the use of R1 implies that A2 is right, while the use of R2 implies that it is not right). To logically "rationalize" my belief about the desirability of A2, I must assign prima facie weights of one to R2 and zero to R1, i.e., I implicitly totally disbelieved the latter in coming to that particular judgment. On average, I acted as if I assigned a weights of 2/3 to R1 and 1 to R2. Because both numbers exceed 1/2, I implicitly accepted both principles on average. But, on average, I accepted R2 more than R1.

In real world applications with many more than three guidelines and two principles, the possibility of the system discovering counterexamples (i.e., zeroes in the weighting table) and contradictions (i.e., rows with different entries in the weighting table) implicit in the user's prior beliefs is much more likely than in this oversimplified example, even when vague beliefs are not present. When vague beliefs are present, the system will virtually always produce surprising information, as the calculation of the weights is slightly more complicated, and is thus relegated to the appendix to this paper.

Despite the possible presence of counterexamples and logical contradictions (under categorical interpretations of the principles), the system will also produce revised degrees of

belief in the desirability of the possible actions, which, in a certain sense, logically follow from the user's beliefs. The sense in which they are consistent is given by the calculus of fuzzy restrictions invented by Lofti Zadeh (1973 and 1975), and utilized by Ostergaard (1977) in constructing a fuzzy expert system to control a heat-exchanger process. To calculate the revised moral beliefs, the system represents the user's knowledge by the conjunction of the proposed principles and the conjunction of the principle's objectives. In this example, it forms the propositions:

R1 and R2

(6)   x is P1 and x is P2

The system applies the so-called compositional rule of inference (Zadeh, 1975) to (6) to infer revised degrees of belief about "x is M". In general, the calculation is a little complicated when both vagueness and large numbers of guidelines and/or principles are present, and is made feasible by computer calculation. But in this example, it is quite simple and can be quickly done by hand. Doing so produced the following belief table:

	Prior	Revision 1
A1	1	1
A2	0	0
A3	1	1

(7)

In this example, no revisions to the decision maker's prior beliefs about "x is M" were needed. However, the compositional rule of inference will generally yield revised beliefs

which differ from the prior beliefs. To illustrate this, suppose
that either now, or during step one of the procedure, I decided
that I really didn't believe that A2 (i.e., disclosure when ter-
minated for causes other than illegal behavior) was totally not
fair to all concerned. Instead, I believe that it is merely close
to being unfair--certainly closer to that than to being fair to
all concerned. Accordingly, I assign a degree of belief equal to
.2 to the statement "A2 is P2", rather than the value of zero in
table (3). Suppose I also decide that I really don't believe that
A2 is exactly telling the truth, but instead that it is mostly
telling the truth. I then assign a degree of belief to "A2 is P1"
equal to .8, rather than one in table (3). The system then cal-
culates the prima facie weighting table:

		R1	R2
	A1	1	1
(8)	A2.	.2	.8
	A3	1	1

which indicates that in forming your prior belief that A2 was
exactly wrong you implicitly attached the prima facie weights of
.2 and .8 to the principles R1 and R2, respectively. This makes
sense, for believing that A2 is wrong is much more consistent with
your belief that A2 is close to not being fair to all concerned,
and hence to R2, than it is to your belief that A2 is close to
telling the truth, and hence to R1. The system then applies the
compositional rule of inference to infer logically consistent,
revised degrees of belief about the desirability of the guide-
lines, contrasted below:

	Prior	M	Revision #1
	A1	1	1
(9)	A2	0	.2
	A3	1	1

The vagueness of my beliefs about the principles' objectives being satisfied by A2 prevents me from logically coming to an exact conclusion about its desirability. I should be somewhat unsure that A2 is wrong, and should accordingly assign a degree of belief of .2 to "A2 is right", rather than a degree of zero. In general, these revisions may also force revisions to the prima facie weighting matrix, although in this example, they do not.

In a real world application, of course, the weights and revised moral beliefs would be more difficult to understand. To aid user understanding of the properties of the system and the effects that different possible user beliefs have on the weights and the revised beliefs about "x is M", the analyzer will produce an analysis of their sensitivity to possible changes in user beliefs, should the user desire it. This sensitivity analysis helps determine which degree of belief assessments are crucial to the results, so that the user may focus her attention on more accurately assessing those degrees of belief.

At this point in the exercise, the user may decide one of two things. She may decide that her beliefs are adequately represented, that the prima facie weights reasonably reflect her views, and that the revised beliefs about "x is M" are reason-

able. Then, the user may recommend not to make any decision whose degree of belief in "x is M" is less than .5, for then it is closer to being undesirable that it is to being desirable. To help justify and clarify the recommendation to others, the user may also wish to present the system concepts, her belief assessments, and the revised weighting matrix she implicitly used. Alternatively, she may be surprised in some way by the weights and/or the revised beliefs. The nature of the surprise may suggest to her that her prior belief assessments were not reasonable, and she may decide to commence further rounds of the four steps, iterating until she is satisfied with the results. Hopefully, the process may lead the user to refine the guidelines into new, better guidelines to evaluate. It is in this way that the system facilitates the construction of an act-based theory for a particular policy.

In our example, after examining the analyzer's output, I felt uneasy that I was placing more implicit weight on fairness than I was on truth telling, both on average and when evaluating the morality of disclosure when terminated for reasons other than illegal acts. Perhaps another related guideline could be found that seems more right, while presenting a more even balance between truthfulness and fairness. An additional possible guideline occurred to me, namely:

A4:   disclosure when terminated for reasons other than illegal
       behavior <u>or gross incompetence.</u>

This seems a little less truthful than A2, for it per-
mits less relevant information to be conveyed to the inquiring,
prospective employer. But it also seems somewhat more fair than
A2, because it is more fair to employees unjustifiably accused of
gross incompetence. After dwelling on all this, I revised my
beliefs in the following way:

		P1	P2	M
	A1	1	1	1
(10)	A2	.8	.2	.2
	A3	1	1	1
	A4	.7	.6	.8

The logic analyzer would then calculate the following
implicit prima facie weights:

	R1	R2
A1	1	1
A2	.2	.8
A3	1	1
A4	.7	.6
AVG	.73	.85

and the revised beliefs:

		M	Revision 1	Revision 2
	A1	1	1	1
(11)	A2	0	.2	.3
	A3	1	1	1
	A4	–	–	.8

Note from (11) that the evaluation of the new guideline resulted in a slight increase in the logical degree of belief of A2's rightness. This is due to a so-called interference effect (Zadeh, 1975 p. 24) that the beliefs about A4 have on the rightness of A2. So, the elicitation of user beliefs over additional guidelines yields useful information applicable to other guidelines as well.

This second revision also results in revisions to the prima facie weighting matrix, yielding the logically consistent revised weights:

		R1	R2
	A1	1	1
	A2	.3	.7
(12)	A3	1	1
	A4	.7	.6
	AVG	.75	.83

Feeling comfortable with the revised, more nearly equal average weights, and getting tired of thinking and typing, I viewed the results of the exercise as successful, and ended the session. I did so fully knowing that should a particular case arise that doesn't fit A1, A2, A3, or A4 (e.g., disclosure about a person who quit), I can fire up the system again, evaluating additional possible guidelines that seem appropriate (e.g., disclosure when the person just quit without giving notice; disclosure when the quitter left without notice, but helped find a replacement; etc.). For now, though, the personnel guidelines should include A1, A3, and A4, but not A2.

The final prima facie weights can be stored and later compared with the weights used in adopting guidelines for policies other than disclosure. By doing so, one might be able to attain a rough invariance of weights used to justify most or all policy guidelines.

Summary

In summary, a system has been proposed to aid in the formulation of policies in the presence of vague and often conflicting multiple objectives. The system requires a user to tentatively specify principles the decision maker chooses to hold, including such traditional objectives as the obeyance of laws and the need to maximize profits. In constructing a particular policy, a user must first identify a preliminary set of mutually exclusive guidelines the decision maker could use to classify the possible decisions. The user then specifies degrees of belief about how much each guideline satisfies each principle's objective, and about how much each guideline satisfies the normative criterion. A fuzzy logic analyzer then calculates the prima facie principle weights logically implicit in the user's beliefs, and logically infers revised beliefs about each guideline's satisfying the normative criterion. After viewing the results, the user may wish to revise her earlier beliefs and/or evaluate other possible guidelines that may occur to her during the process. The result of this iterative process is a policy which is logically consistent with an explicit, weighted set of prima facie principles, which others may rationally understand and debate. After deter-

mining numerous policies in this way, the user can determine how consistent her implicit assignment of prima facie weights has been across policies. This global examination may suggest changes to existing policies which reduce the variability of the weights across policies.

Future Directions

The methods herein have potential applicability outside the area of personnel policy. Potential uses occur wherever inherently vague, conflicting multiple objectives are present, including other frequently discussed moral concerns of business. The experienced gained from real applications of these methods is essential in determining the value of this proposed system in the many contexts it could be used in. I recommend that the Federal Reserve initiate a pilot project to apply this methodology in formulating one complex policy.

Appendix

Zadeh's (1975) calculus of fuzzy restrictions is used to represent and draw inferences from the propositions:

R_1: If x is P_1, then x is M; else x is not M.

R_2: If x is P_2, then x is M; else x is not M.

(1)

R_n: If x is P_n, then x is M; else x is not M.

x is P_1,

x is P_2,

x is P_n,

In the current application, the if-then-else statements represent principles embodying objectives. The variable x is a name for a class of decisions partitioned into equivalence classes $X = \{A_j, j=1,...m\}$, called guidelines. Each P_k is a fuzzy set (see Zadeh (1965)) on the universe of discourse $\{A_j; j=1,...m\}$, and is defined by exemplifcation (Zadeh (1975), p.8), producing grade of membership functions $\mu_k : X \rightarrow [0,1]$, $k = 1, ..., n$. Each P_k describes an objective satisfied in varying degrees by the guidelines. M is also a fuzzy set defeated over the guidelines, representing the varying degrees in which the guidelines satisfy the normative criterion. It is also defined by exemplification, producing a membership function $\mu_M : X \rightarrow [0,1]$.

By Zadeh's maximin rule of conditional propositions, each R_k in (1) is a fuzzy binary relation on x, represented by the square, truth value matrix whose (i,j)th element:

(2) $R_k(A_i,A_j) = \max(\min(\mu_k(A_i),\mu_M(A_j)),\min(1-\mu_k(A_i),1-\mu_M(A_j));$

    $i,j=1\ldots m.$

is interpreted as the degree of belief in R_k implicit in $\mu_k(A_i)$ and $\mu_M(A_j)$. Its computation is diagrammatically represented in figure 1.

The <u>prima facie weighting matrix W</u> is defined to be the mxn matrix whose k^{th} column is the diagonal of (2), showing the degrees of belief in R_k when evaluated for each action:

(3) $W = \begin{pmatrix} R_1(A_1,A_1)\ldots R_n(A_1,A_1) \\ \vdots \quad \ddots \quad \vdots \\ R_1(A_m,A_m)\ldots R_n(A_m,A_m) \end{pmatrix}$

By Zadeh's rule of implied conjunction, the simultaneous belief in all the rules R_k, k = 1, ..., n is represented by the combined statement R_1 and R_2 and $\ldots R_n$, <u>denoted R</u>. Similarly, the simultaneous belief in the statements x is P_k, k = 1, ..., n is represented by the combined statement x is P_1 and x is P_2 and ...x is P_n, <u>denoted P.</u>

Zadeh's rule of maximal restriction and (2) imbues R with the representation:

(4) $R(A_i,A_j) = \min(R_1(A_i,A_j),R_2(A_i,A_j),\ldots,R_n(A_i,A_j));$

    $i, j = 1, \ldots, m.$

i.e., the componentwise minimum of the matrices R_k. Intuitively, the degree of belief in the conjunction is no stronger or weaker than its "weakest link". Similarly, the fuzzy set P has membership function:

(5) $\quad \mu_P(A_i) = \min(\mu_1(A_i), \mu_2(A_i), \ldots, \mu_n(A_i)); \; i \equiv 1, \ldots, m.$

The inference drawn from R and P is a fuzzy set determined by (4), (5), and Zadeh's compositional rule of inference. It is denoted PoR, and has the membership function.

(6) $\quad \mu_{PoR}(A_k) = \max[\min(\mu_P(A_1), R(A_1, A_k)), \min(\mu_P(A_2), R(A_2, A_k)), \ldots,$
$$\min(\mu_P(A_m), R(A_m, A_k))]$$

for $k = 1, \ldots, m.$

A compact representation of (6) computes μ_{PoR} as a row m-vector:

(7) $\quad \mu_{PoR} = [(\mu_P(A_1), \mu_P(A_2) \ldots \mu_P(A_m))] \begin{pmatrix} R(A_1, A_1) \ldots R(A_1, A_m) \\ \vdots \quad \ddots \quad \vdots \\ R(A_m, A_1) \ldots R(A_m, A_m) \end{pmatrix}$

calculated by an analog of vector-matrix multiplication, replacing each would-be multiplication of two numbers by their minimum, and replacing each would-be sum of m numbers by their maximum.

I interpret $\mu_{PoR}(A_k)$ as the logical, revised degree of belief in which A_k satisfies the normative criterion, to be contrasted with the prior degree of belief $\mu_M(A_k)$.

One may also compute a revised prima facie weighting matrix consistent with the revised beliefs (7), by substituting (7) for μ_M in computing (3).

Figure 1

Computation of $R_k(A_i, A_j)$
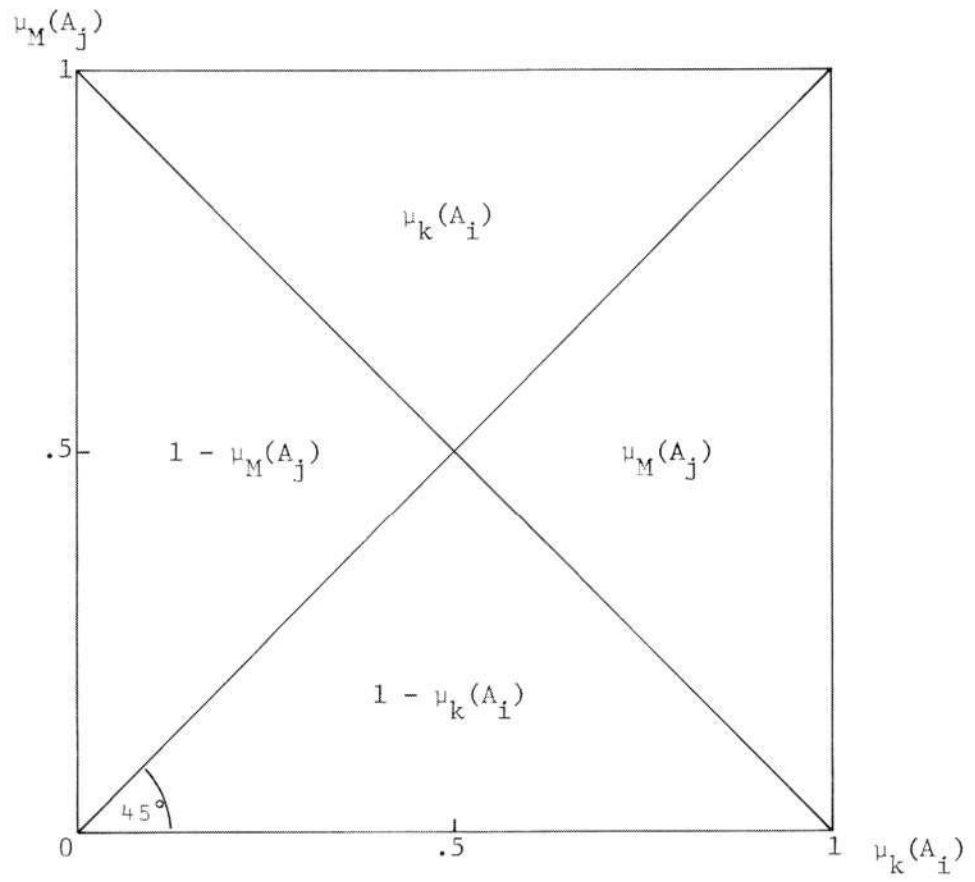
References

Beauchamp, T. L. and Bowie, N. E., "Ethical Theory and Business", Prentice-Hall, [1983].

Blumberg, P. I., "Corporate Responsibility and the Employee's Duty of Loyalty and Obedience," in Beauchamp, T. L. and Bowie, N. E., op. cit.

Ewing, D., "Civil Liberties In the Corporation," New York State Bar Journal, [1978]. Reprinted in Beauchamp, T. L. and Bowie, N. E., op. cit.

Ostergaard, J., "Fuzzy Logic Control of a Heat Exchanger Process," in Gupta, M., et. al., (eds), Fuzzy Automata and Decision Processes, North-Holland, [1977]

Quirk, J. and Saposnik, R., "Introduction to General Equilibrium Theory and Welfare Economics", McGraw-Hill, 1968.

Rosen, B., "Strategies of Ethics," Houghton Mifflin, [1978].

Ross, W. D., "The Right and the Good," Clarendon Press, Oxford, [1930].

Summers, C. W., "Protecting All Employees Against Unjust Dismissal," Harvard Business Review, [1980]. Reprinted in Beauchamp, T. L. and Bowie, N. E., op. cit.

Whalen, T. and Schott, B., "Issues in Fuzzy Production Systems," International Journal of Man-Machine Studies, 19 [1983], pp. 57-71.

Zadeh, L. A. "Fuzzy Sets," Information and Control, 8 [1965], pp. 338-353.

Zadeh, L. A. "Outline of a New Approach to the Analysis of Complex
    Systems and Decision Processes," IEEE Transactions On Systems,
    Man and Cybernetics, SMC-3, [Jan. 1973].

Zadeh, L. A., "Calculus of Fuzzy Restrictions," in Zadeh, L. A.,
    et. al, ed., Fuzzy Sets and Their Application to Cognitive and
    Decision Processes, Academic Press, [1975].