ARI Research Note 88-46

Decomposition Strategies for Eliciting Expert Knowledge: Judgements of Dangerousness

Sarah Lichtenstein and Paul Slovic
Perceptronics, Inc.

for

Contracting Officer's Representative
Michael Drillings

ARI Scientific Coordination Office, London
Milton S. Katz, Chief

Basic Research Laboratory
Michael Kaplan, Director

U. S. Army
Research Institute for the Behavioral and Social Sciences

June 1988

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Research accomplished under contract
for the Department of the Army

Perceptronics, Inc.

Technical review by

Dan Ragland

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ADA197913

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARI Research Note 88-46	2. GOVT ACCESSION NO. - -	3. RECIPIENT'S CATALOG NUMBER - -
4. TITLE (and Subtitle) Decomposition Strategies for Eliciting Expert Knowledge: Judgements of Dangerousness		5. TYPE OF REPORT & PERIOD COVERED Final Report April 83 - April 87
		6. PERFORMING ORG. REPORT NUMBER - -
7. AUTHOR(s) Sarah Lichtenstein and Paul Slovic		8. CONTRACT OR GRANT NUMBER(s) MDA903-83-C-0198
9. PERFORMING ORGANIZATION NAME AND ADDRESS Perceptronics, Inc. 6271 Variel Avenue Woodland Hills, CA 91367		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q161102B74F
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600		12. REPORT DATE June 1988
		13. NUMBER OF PAGES 29
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) - -		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE - -

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

- -

18. SUPPLEMENTARY NOTES

Michael Drillings, contracting officer's representative

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decomposition Principles, Expert Judgements,
Knowledge Elicitation, Problem Solving. (SDC)
Decision Analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This research note discusses one important finding from decision-aiding research, that people often have relevant knowledge that they do not use effectively when making a judgement or decision.

Research has shown, however, that simple "wholistic" judgements can be improved upon through an approach that breaks up or decomposes the problem into a series of sub-problems, or components, each of which can be understood more easily and judged separately. The components are then assembled according to

(OVER)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ARI RESEARCH NOTE 88-46

20. Abstract (continued)

a logically prescribed set of combination rules to yield a solution, estimate, or prediction. In the present paper, we outline how a decomposition approach may help a large consortium of expert judges to utilize their own knowledge base more effectively, in an extremely difficult and important judgement task -- (the task examined is assessment of dangerousness among people who have threatened to assassinate the President of the United States).

Psychologists and other scientists have been studying judgment and decision processes extensively during the past three decades. One stream of this research has attempted to model the ways in which people process information in judgment and decision-making tasks. A second stream has examined the deficiencies of judgments and decisions and explored ways to overcome them through "decision-aiding" procedures.

One important finding from decision-aiding research is that people often have relevant knowledge that they do not use effectively when making a judgment or decision. For example, if one's task is to provide an estimate of some uncertain quantity (e.g., "How much money will our family spend on food next year?") the simplest and most natural approach is to think hard about the problem and intuitively produce an estimate that seems reasonable in light of whatever knowledge comes to mind. Research has shown, however, that simple "wholistic" judgments can be improved upon through an approach that breaks up or decomposes the problem into a series of sub-problems or components, each of which can be understood more easily and judged separately. The components are then assembled according to a logically prescribed set of combination rules to yield a solution, estimate, or prediction. Decomposition is a divide-and-conquer approach that assumes the components of a problem to be more tractable than the undecomposed problem.

Decomposition techniques have been employed in a wide variety of problem areas. For example, decision analysis, a methodology for use in situations involving uncertainty (Raiffa, 1968), partitions a decision problem into actions and outcomes. Each outcome has an

associated payoff amount and probability which are analyzed to determine the optimal course of action.

The decomposition principle has also been applied to human judgment, through the use of algorithms. An algorithm is a series of steps or operations that, when sequentially applied, produce a solution to a problem. Algorithms work by providing an unambiguous procedure for solving problems. They help structure what is known about a problem, point out what is not known, and specify the rules by which information should be combined. Since the combination of information is mechanical, algorithms have the potential for high reliability; different individuals using the same algorithm should arrive at very nearly the same solution.

Singer (1971) illustrated the use of algorithmic decomposition to estimate the amount of money per year taken in muggings, robberies, and burglaries by heroin addicts in New York City. Using an algorithm with components he could estimate more accurately, such as the population of the city, the number of reported burglaries, and the number of addicts in prison, he arrived at an estimate of $250 million stolen per year, far smaller than estimates previously suggested. Subsequent experimental studies by Armstrong, Denniston and Gordon (1975), MacGregor, Lichtenstein, and Slovic (1984) and Lichtenstein and MacGregor (1984) have demonstrated the superiority of algorithmically aided judgments for a wide variety of problems in which the objective was to estimate an uncertain quantity of some sort.

Judging Dangerousness

In the present paper, we shall propose a somewhat different

application of the principle that a structured, logical, rule-based decomposition can elicit and combine knowledge more effectively than unaided, wholistic judgment. Specifically, we propose to outline how a decomposition approach may help a large consortium of expert judges utilize their own knowledge base more effectively in an extremely difficult and important judgment task. The task we shall examine is assessment of dangerousness among persons who have threatened to assassinate the President of the United States. The judges are several hundred agents of the intelligence division of the U. S. Secret Service, who must respond to and evaluate 5,000 to 10,000 threats per year (IOM, 1984).

Dangerousness is extraordinarily hard to predict, even for common street violence (Monahan, 1981; Stone, 1985). The difficulties of prediction seem to derive from the fact that violent behavior is determined by a complex interaction between many personality, situational, and contextual factors (Megargee, 1976). The problem of predicting dangerousness to the President is even more difficult because it is such a rare form of violence that we cannot acquire sufficient data to build an empirical assessment system. Of course, the main reason assassination is so rare is because society views it as intolerable and directs the Secret Service to prevent it at all costs. The intolerability of assassination also means that we cannot perform controlled experiments on the population of potential assassins, letting some roam uncontrolled in order that we might test theories about their dangerousness; hence, an important avenue of learning is precluded (Moses, 1981). Some have proposed studies of "proxy events"

such as violence toward famous people not protected by the Secret Service (IOM, 1984), but even these events are too infrequent for reliable statistical analyses to be conducted. Also, it is by no means evident that violence toward movie or rock stars or governors is determined by the same factors as is violence toward the President.

In sum, the rarity of assassination attempts precludes the acquisition of sufficient data on which to build an empirical assessment procedure. Controlled experimentation is likewise ruled out. Psychometric theory further demonstrates the folly of attempting to predict phenomena so rare as an assassination attempt on the basis of less than perfect assessment measures (Meehl & Rosen, 1955).

The inadequacy of traditional empirical methods for assessing dangerousness leads us to propose a purely judgmental approach in which decompositional strategies are used to build a rule-based system for "defining dangerousness." This approach uses the experience and wisdom of experts to develop a prescriptive algorithm to guide the assessment process. It assumes that experts (including the Secret Service's own agents) have developed implicit theories of dangerousness. The proposed approach would make these theories explicit in the form of algorithms which can be used to guide information gathering, analysis, and decision making.

The approach proposed here has much in common with the concept of "expert systems" that has developed within the field of artificial intelligence. Expert systems are designed to represent and apply factual knowledge from specific areas of expertise in order to solve a wide range of problems involving interpretation, prediction, diagnosis,

design, monitoring, and planning. One of the first of these models to be applied was DENDRAL (Buchanan, Sutherland, & Feigenbaum, 1969), which was designed to determine chemical structure on the basis of mass-spectograph data. DENDRAL was able to determine a narrow range of possible structures out of millions of possibilities. For some families of compounds it could outperform the best human experts. Models for medical diagnosis came along in the mid-1970's. An early system named MYCIN (Shortliffe, 1976) was able to diagnose bacterial infections about as well as physicians. INTERNIST was developed to make diagnoses in the domain of internal medicine and PUFF interprets the results of pulmonary tests (Waldrop, 1984). PROSPECTOR guides exploration for minerals based on geologic data (Duda, Geschnig, & Hart, 1979). It discovered molybdenum deposits in Canada worth many millions of dollars. The electric power industry is currently developing rule-based models to help diagnose malfunctions in nuclear reactors (Moore, 1985). The enthusiasm of the designers of these models knows no bounds. It is likely that expert systems will affect all facets of our society in years to come (Hayes-Roth, Waterman, & Lenat, 1983).

Although the focus of our discussion is on dangerousness of assassination threats, the approach we describe should be equally applicable to other situations in which the stakes are high and expert judgment is the only recourse. For example, our proposed approach could be applied to the task of assessing the credibility and seriousness of a terrorist threat involving nuclear weapons.

Defining Dangerousness

In this section we shall describe how one might build a decision aid based on an algorithmic definition of dangerousness in the context of assassination threats. A complete decision aid for this task might involve four separate components:

1) An early closure model indicating which cases could be closed with minimal investigation.

2) A first-contact model indicating how one should decide whether to classify a subject as dangerous.

3) A repeated-contacts model indicating how one should decide whether to continue or discontinue a subject's classification as dangerous during the course of investigations.

4) A Dangerousness Index: This Index would yield, for each investigated subject, a numerical score indicating that subject's level or degree of dangerousness.

For simplicity, we shall focus here upon the first-contact model and the Dangerousness Index.

The development of these decision aids would require extensive efforts by one or more skilled analysts. Indeed, for these analysts prior knowledge of judgment and decision-making research and skill in building expert models would be more important than detailed prior knowledge about the Secret Service or about assassins. Secret Service agents are the experts on dangerousness; the analysts must be expert in extracting and codifying the agents' knowledge.

Steps in the Process

We envision that the analyst(s) would proceed through the following stages:

Stage one: individual models. A first-contact model would be developed for each of several Secret Service agents. These agents should be those recognized by their peers and by the Service as experienced and skilled in making dangerousness decisions. In selecting these agents, some attention should also be given to the inclusion of a diversity of beliefs about dangerousness. As few as five or as many as twelve agents might be selected for modeling.

Each model would be built through a series of interviews and discussions between the agent and the analyst. Fryback and Keeney (1983) reported that eliciting one expert model (a model of trauma severity) from one physician required three interviews totalling 12 hours. The resulting model, although complex in form, included only 13 items of information, far fewer than agents will likely wish to include in their models.

Stage two: the Consensus Model. From these individual models, the analyst, in consultation with agents, USSS staff, and possibly with outside experts, would then build a Consensus Model. The Consensus Moel would be an amalgam of the individual models; it would represent the collective wisdom of the Secret Service and would constitute the Service's definition of dangerousness.

Stage three: the Dangerousness Index. Only after the Consensus Model is completed can the Dangerousness Index be built. Whereas the Consensus Model primarily expresses an underlying theory of

dangerousness, the Index should have a directly practical orientation. The Index would be a scoring system whereby any set of facts about a subject can be translated into a single numerical score, such that more dangerous subjects receive a higher score than less dangerous subjects. The Consensus Model shows which facts are to be included and how much weight each fact should receive in the Index.

The First-Contact Model

The first-contact model is the central model of dangerousness. It represents answers to these questions: How should an agent decide on the dangerousness of a subject? What facts are important? How should these facts be combined in reaching a decision?

The primary use of the consensus first-contact model is in the development of the Dangerousness Index. The Dangerousness Index is like a checklist designed for ease of use. The first-contact model, in contrast, is the expression of a "theory" of dangerousness; it lays bare the rationale underlying the Index. Such a theory might be quite complex, with many causal links and sub-components.

This separation between an expanded theory and a practical tool is, we believe, one of the key strengths of the approach here recommended, because it may facilitate a more sensible approach to the validation problem. For the reasons described earlier, it is not possible to establish the predictive validity of the dangerousness index. But it may be possible to test the validity of some of the causal links in the first-contact model. The validity of specific components might be examined in a broader context than of assassination of protectees, thus circumventing the low-base-rate problem. For

example, one component of the model may deal with the effect of habitual use of a certain drug on the ability of a subject to carry out a plan. This relationship can be explored by research or by consultation with drug experts. If studies have shown the drug strongly impairs many kinds of intellectual functioning, this finding would lend consensual validation to at least that component of the model.

What If No Consensus Emerges?

We suspect that a consensus dangerousness model will emerge through the research here recommended, at least for the more important elements. If one thinks of the model as a tree, with important branches and minor twigs, we expect agreement on the important branches, which might be:

- interest in or motivation toward harming the protectee,
- possession of weapons and knowledge of how to use them,
- ability to plan an attack,
- ability to carry out an attack, and
- inhibitory factors (e.g., a good job, strong family ties)

We expect there would be more disagreement about the "twigs." For example, the following quotes, taken from interviews with agents, indicate some genuine disagreements about how homosexuality relates to dangerousness:

"Homosexuality is not important."

"There is some evidence that homosexuals have a lot of pent-up hostility."

"Homosexual males are to a great extent hysterical people and hysterical people are dangerous because they are out of control."

"Homosexuals are usually not that aggressive, but they can be."

"If a person admits to being a homosexual, at least he has set up some standard for himself."

"If he is a homosexual, you know he already has one strike against him."

"In general, I do not make any distinction between homosexual and heterosexual. A lot of heterosexuals are just as wacky as anybody else."

"I have often heard or read that homosexuals would tend to be assassins. To me homosexuals are quiet types who want to be left alone. If they use whips, chairs, leathers, then I take another look."

Note that the disagreements are not about dangerousness per se, but about the typical characteristics of homosexuals. As previously discussed, such topics could be the focus for research and consultation with outside experts (Do homosexuals have a lot of pent-up hostility? Are they often hysterical people, and are hysterical people out of control? Are they usually passive, quiet types who want to be left alone?)

Despite our optimism, the possibility remains that no consensus model could be found. Such a result would be discouraging, for if experienced agents cannot agree on who is dangerous, we see little hope

for improvement in the Service's investigative mission. But if widespread disagreement in fact exists, it is better for the Service to know that than not to know it.

Model Forms

The analysts who will elicit the agent's models shoud not hold any preconceptions about what forms the models will take. To do so would be to deny some aspects of the agents' expertise. In this section we will describe several different model forms. In addition, we will sometimes speculate on agent's possible models, not to prejudge such issues, but to provide examples that may clarify the differences among the model forms.

Algebraic models. Algebraic models are models that can be expressed in an algebraic equation. The elements of the equation are of two sorts. The first is the scale value for some fact, attribute, or dimension considered relvant to dangerousness. For example, weapon possession might be scaled from 0 to 5, with 0 meaning no weapons and 5 meaning an extensive arsenal of a wide variety of potent weapons. The other sort of element in the equation is a constant that rescales the scale values so that there is consistency across attributes. These constants are often interpreted as measures of the relative importance of the different attributes in determining dangerousness. For example, if an agent believes that both weapons knowledge and weapons possession are aspects of dangerousness, but that weapons knowledge is three times more important than weapons possession (because one can easily obtain a weapon), then the constant applied to the scale value for weapons knowledge would be three times as large as the constant applied to the

scale value for weapons possession. However, these constants can not always be directly interpreted as importance weights. They are sometimes used for technical rescaling, for example, to correct for the fact that one attribute is scaled either 0 or 1, while the other is scaled from 1 to 100.

The simplest form of algebraic model is the additive model. Each attribute value is weighted by its constant and then added to the rest:

$$\text{Dangerousness} = c_1X_1 + c_2X_2 + \ldots + c_kX_k,$$

where the c's are the constants and the X's are the scale values of the attributes. This model has the same equation as the linear model often used as the basis of a technique for studying judgment (Dawes & Corrigan, 1974). However, in that application the scale values were chosen by the researcher and the constants were arrived at by statistical maximization techniques based on correlational analysis. Here, the agent chooses the X's and c's directly. Moreover, here the X's need not be linear, in the sense that there need not be a straight-line relationship between the scale values and some underlying measure. For example, supposing X_1 was a scale of the number of violent acts the subject is known to have performed. The scale might be nonlinearly related to the number of acts, as shown in Figure 1. The curve in Figure 1 expresses the belief that dangerousness increases faster for each additional act when there are many acts than when there are few.

Insert Figure 1 about here

The other most common form of an algebraic model is a multiplicative one. It is used when two or more attributes interact in their effect upon dangerousness. Thus, if a case in which two attributes took intermediate values was deemed far more dangerous than a case in which one of those values was very high and the other was very low, the multiplicative term, $c_{ij}X_iX_j$, might appear in the equation. With such a term, at the extreme, when either X_i or X_j is equal to zero (totally non-dangerous), the term is equal to zero.

An example of a possible multiplicative relationship in dangerousness is the interaction between intent and ability. Dangerousness requires both. No matter how able the subject is to devise and carry out a plan, if the subject totally lacks the intent to harm a protectee, the subject is not dangerous. Likewise, if the subject is totally incapable of carrying out a plan, the subject is not dangerous regardless of intent.

Additive and multiplicative terms can both appear in the equation, even for the same attributes:

$$c_iX_i + c_jX_j + c_{ij}X_iX_j.$$

This form expresses the belief that two attributes make both independent and interactive contributions to dangerousness.

Hierarchical models. Hierarchical models are used when the decision situation can best be described with two or more levels of attributes, the levels differing in degree of abstractness or generality. The "top" of the hierarchy contains the most abstract attributes; the "bottom" of the hierarchy contains the most specific attributes.

A hierarchical model is a special case of an algebraic model, in the sense that it can also be expressed in an equation rather than in tree form. At each level there is a cluster of one or more attributes which is directly linked to one attribute at the next higher level. Each such cluster is a mini-model; the weights for the attributes and the combination rule are specified. For the lowest level, the scale values for each model must also be specified. Thus, when the model is complete it can be "folded up" algebraically to yield an equation.

The advantages of the hierarchical model are: (a) For many problems, people's knowledge and beliefs are structured hierarchically; for such situations, the model is easier to elicit in this form. (b) The model, when elicited, is easier for other people to understand when displayed in this form.

Figure 2 shows an exemplar portion of a possible hierarchical dangerousness model. In this model, dangerousness is determined by three general factors (Level 1): (a) intent to harm, attack, or embarrass a protectee, (b) the ability to do so, and (c) a factor, here called "Inhibitors," which contains all the factors (e.g., warm, strong family ties) that might deter an otherwise intent and able subject. The combination rule for Level 1 is Intent times Ability minus Inhibitors, indicating a belief, as previously discussed, that <u>both</u> intent <u>and</u> ability are necessary for dangerousness. The relative weights of the three factors are not shown. One could speculate that the Inhibitors factor is less important than the other two.

Insert Figure 2 about here

Level 2 is explicated in Figure 2 only for the Ability factor of Level 1. It is here supposed that Ability is composed of four parts: (a) the ability to plan an attack, (b) the ability to carry out such a plan, (c) the tendency towards violence, and (d) weapons. Again the combination rule is partly multiplicative (one needs both planning ability and execution ability to receive a high score on the Level 1 ability factor) and additive. Again relative weights would be needed to complete the model.

Level 3 shows the breakdown of Violence into two components, the potential for violence exhibited via the subject's behavior during a personal interview and the subject's history of past violence. There may be several boxes all labeled "Interview Behavior" in the complete hierarchy, because agents learn many different things from subjects' behavior. For example, they also may learn quite a bit about a subject's intent to harm a protectee. The present box refers only to violence, as is made clear by the boxes above and below it.

Level 4 is the lowest level for the branch involving interview behavior. It might be composed of scales or a simple checklist of agent observations. In the latter case differential weighting of the items on the checklist would be accomplished by assigning a different number of points to each item (e.g., 10 points for threatening the agent but only 3 points for yelling at the agent).

In the full model, each Level 1 factor would be further broken down into more specific components. However, there is no need for each branch to have the same number of levels. For example, Intent may end after branching into three additive parts in Level 2: intent to harm, intent to attack, and intent to embarrass, each to be scaled judgmentally by the agent.

Rule-based models. It must not be assumed that agents' dangerousness models will be algebraic, with scale values that add or multiply to produce a single final number the size of which reflects the degree of dangerousness. It may be that an agent's "theory" of dangerousness can best be expressed by a series of if-then rules that result in a categorical decision (e.g., close the case now vs. continue the investigation). Two special cases of rule-based models are the conjunctive model and the disjunctive model.

The conjunctive model says: Make decision A only if all of the answers to a series of questions are "Yes." Otherwise, make decision not-A. It involves the listing of all necessary conditions and denies both the possibility that there is more than one avenue to a decision and the possibility that some facts can trade off against others in their effect upon the decision. We have not been able to think of a convincing example of a conjunctive model of dangerousness, perhaps because our implicit bias is to believe that dangerousness is graduated, not categorical.

The disjunctive model says: Make decision A if any one or more answers to a series of questions is "Yes." Make decision not-A only if all answers are "No." It involves listing the sufficient conditions

for a decision. Like the conjunctive model, it denies the possibility that some facts can trade off against other facts in their effect upon the decision.

In general, rule-based models can be lengthy, complex combinations of rules. A rule-based model can look quite a bit like a hierarchical model. But when using a rule-based model to evaluate a subject's dangerousness, one starts at the top and follows only one path in the model. That path leads to a decision. In contrast, when evaluating a subject using a hierarchical model, one starts at the bottom and follows all paths. The result is not a decision but a dangerousness score.

A rule-based model is particularly appropriate when some of the questions asked in one case are irrelevant to another case. For example, an agent might have quite different models for sane subjects and insane subjects. The first node in the path model then might be "Sane vs. Insane." The "Insane" paths might include rules like: "If the subject hears voices, then find out if the subject will obey the voices' commands," or "If the subject is on medication, then find out if the subject is always medicated. If not, interview again when the subject is not medicated." The possibilities explored by these rules are not relevant for sane subjects; using a rule-based model one would enter these paths only when appropriate.

Model combinations. It is not necessary that a model be purely one type. Rule-based models could have additive components (e.g., "If X is true, score the subjects on scales A, B, and C, and add the scores"). When designing the partial hiearchical example in Figure 3,

it occurred to us that a subject who has the intent to embarrass a protectee (e.g., the subject plans to publicly spit in the face of a protectee) poses a quite different threat from one who intends to harm the protectee. In such a case, information about weapons may become irrelevant; that is, a "spitter" who owns many weapons may not be more dangerous than a "spitter" who has neither knowledge nor possession of weapons. Mobility and creative intelligence may take on more importance. If so, a single hiearchical model won't be adequate, because it does not have different weights for an attribute as a function of the score on another attribute. This, then, is a situation in which two or more algebraic models would be linked by rules specifying which model is to be used. The models would differ in which attributes are included and in the relative weights of the attributes.

Forms for the Consensus Models. If different agents have models with different forms, devising a satisfactory Consensus Model may be difficult. Two considerations suggest that, nonetheless, the task can be accomplished. First, sometimes models that look quite different are, at heart, quite similar. For example, a multiplicative model is a close approximation of a disjunctive model (Einhorn, 1970). However, if most agents' models are primarily algebraic or hierarchical, it would be difficult to accommodate into the Consensus Model an entirely rule-based model, because such a model has no scale values or relative weights. If a few such models emerge, the analyst may have to return to those agents to see if an algebraic model was acceptable as a second choice.

Second, variations in models are often unimportant in the determination of the total dangerousness score. Moreover, there is a technique (called "sensitivity analysis") for determining which aspects of a model can be changed (including scale values, relative weights, combination rules, and addition or elimination of attributes without seriously altering the total score (Behn & Vaupel, 1982; Howard, 1968).

What is most important is that, to the fullest extent possible, the Consensus Models express, in richness and detail, the agents' shared underlying "theory" of dangerousness. Comprehensiveness and theoretical consistency, not simplicity, should guide the formulation of the Consensus Models.

Forms for the Dangerousness Index. The Dangerousness Index should be designed primarily for ease of use rather than for elucidating an underlying theory of dangerousness. Our preference would be an additive model, perhaps with some multiplicative terms.

As the preceding discussion of model forms has suggested, there is no guarantee that the Consensus Models will be in forms suitable for direct translation into a Dangerousness Index. Here, the tools of sensitivity analysis can be used to see how much simplification is possible without a serious loss of fidelity. It is likely, we believe, that the Consensus Model will be largely algebraic or hierarchical. Such models can be translated quite directly into a scorable Index.

Using the Models

The decision-aiding model resulting from the approach advocated above should be reviewed and revised as experience dictates. The model would have many uses. First, it would guide the development of a

system for collecting, coding, and storing data. It could be used to automate agents' dangerousness decisions, or at least to provide one form of input to those decisions. The model and Dangerousness Index could also be used to increase communication and facilitate case discussions among agents, supervisors, and headquarters staff. In a more general way, the models would provide a shared set of vocabulary and concepts with which to answer such questions as "Why did you make that decision?" or "What do you think is going on with that subject?"

Finally, the model and Index would be excellent training tools for new agents. The model expresses the Secret Service's theory of dangerousness and could thus serve as a focus for training discussions. In addition, the elements of the model and of the Index show clearly and in detail what information should be gathered during the course of an investigation. Agents who are relatively inexperienced in protective investigations may also profit from this form of sharing their more experienced colleagues' knowledge.

Research, too, could employ the model and Index. For example, the Dangerousness Index could be validated to some extent by examining subjects' subsequent behavior. A subject with a high Index score who does not now possess a weapon should be more likely to obtain a weapon in the future than would a subject with a low Index score and no weapon. In addition, the relationship between changes in Index scores and changes in subjects' dangerousness classification could be explored. On what kinds of occasions are large changes in Index scores not accompanied by changes in dangerousness status? How often and for

what reason does status change without a significant change in the subject's Index score?

The model and Index would also provide a means for studying several different kinds of consistency in agent's decisions. First, studies could be done on the consistency in agent's decsions, both for one agent over time and for comparisons across agents. To the extent that the Index reflects a shared theory of dangerousness, one would expect a fairly systematic relationship between Index scores and decisions. However, inconsistencies should not necessarily be viewed as "agent errors." Any of the following might account for inconsistencies:

- Omissions in the Index.
- Changes in the world environment. For example, an extremist group may suddenly adopt terrorist tactics, leading an agent to reclassify a subject who is a member of that group as dangerous although the subject has a low Index score.
- Genuine disagreements about the definition of dangerousness.

The existence of the detailed and explicit models and Index should greatly facilitate the untangling of these different sources of inconsistency. Such studies would also provide the evidence needed to improve the decision aids and might suggest ways to improve agent training for pinpointing the areas of agent disagreement.

Conclusion

Structural models, which explicitly codify expertise and elicit experts' knowledge in decomposed form, appear to provide a promising way to improve upon unaided, wholistic judgments--particularly when

there is no real alternative to the use of expert judgment. In this paper, we have sketched the development of such models in a specific applied domain involving great uncertainty, high stakes, and little opportunity to learn from experience.

References

Armstrong, J. S., Denniston, W. B., Jr., & Gordon, M. M. (1975). The use of the decision position principle in making judgments. Organizational Behavior and Human Performance, 14, 257-263.

Behn, R. D., & Vaupel, J. W. (1982). Quick analysis for busy decision makers. New York: Basic Books.

Buchanen, B. G., Sutherland, G. L., & Feigenbaum, E. A. (1969). Heuristic DENDRAL: A program for generating exploratory hypotheses in organic chemistry. In B. Meltzer and D. Michie (Eds.), Machine Intelligence (Vol. 4). Edinburgh: Edinburgh University Press.

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

Duda, R. O., Gasching, J. G., & Hart, P. E. (1979). Model design in the PROSPECTOR consultant system for mineral exploration. In D. Michie (Ed.), Expert systems in the micro-electronic age. Ediburgh: Edinburgh University Press.

Einhorn, H. (1970). The use of nonlinear, noncompensatory models in decision making. Psychological Bulletin, 73, 221-230.

Fryback, D. G., & Keeney, R. L. (1983). Constructing a complex judgmental model: An index of trauma severity. Management Science, 29, 869-883.

Hayes-Roth, R., Waterman, D. A., & Lenat, D. B. (1983). Building expert systems. Reading, MA: Addison-Wesley.

Howard, R. A. (1968). The foundations of decision analysis. IEEE Transactions on Systems Science and Cybernetics (Vol. SSC-4), 393-401.

IOM. (1984). Research and training for the Secret Service: Behavioral science and mental health perspectives. Institute of Medicine, Washington, DC: National Academy Press.

Lichtenstein, S., & MacGregor, D. (1984). Structuring as an aid to performance in base-rate problems (Report No. 85-1). Eugene, OR: Decision Research.

MacGregor, D., Lichtenstein, S., & Slovic, P. (1984). Structuring knowledge retrieval (Report No. 84-14). Eugene, OR: Decision Research.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficacy of psychometric signs, patterns, or cutting scores. Psychological Bulletin, 52, 194-216.

Megargee, E. (1976). The prediction of dangerous behavior. Criminal Justice and Behavior, 3, 3-21.

Monahan, J. (1981). Predicting violent behavior. Beverly Hills: Sage.

Moore, T. (1985, June). Artificial intelligence: Human expertise from machines. EPRI Journal, 6-15.

Moses, L. E. (1981). Learning from experience, the Secret Service, and controlled experimentation. In J. Takeuchi et al. (Eds.), Behavioral science and the Secret Service. Washington, DC: National Academy Press.

Raiffa, H. (1968). Decision analysis. Reading, MA: Addison, Wesley.

Shortliffe, E. H. (1976). Computer-based medical consultation: MYCIN. New York: American Elsevier.

Singer, M. (1971). The vitality of mythical numbers. The Public Interest, 23, 3-9.

Stone, A. A. (1985). The new legal standard of dangerousness: Fair in theory, unfair in practice. In C. D. Webster, et al. (Eds.), Dangerousness. New York: Cambridge University Press.

Waldrop, M. M. (1984). Artificial experts. Science, 223, 1281.

# of acts	X_i value
0	0
1	.25
2	.8
3	1.5
4	3
5 or more	5

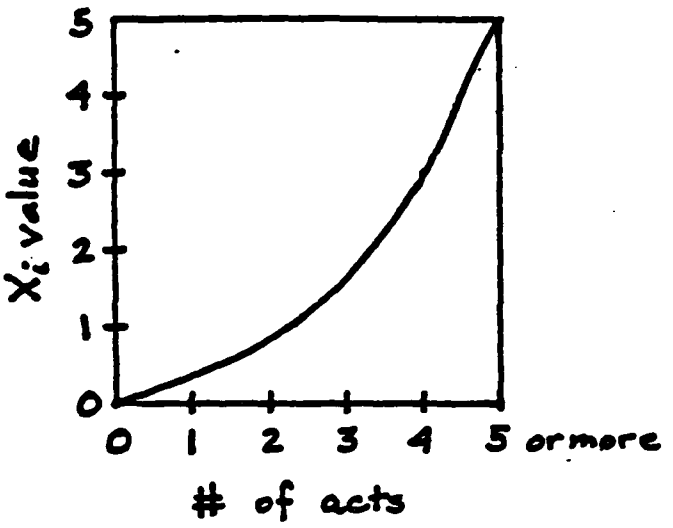

Figure 1. Proposed non-linear relationship between number of violent acts and a component of the dangerousness index.

Figure 2

Possible (Incomplete) Hierarchical Dangerousness Model