Appendix D. A Socio-technical Approach to Assessing Human Reliability (STAHR)

of

A PRESSURIZED THERMAL SHOCK EVALUATION OF THE CALVERT CLIFFS UNIT 1 NUCLEAR POWER PLANT

written by

Lawrence D. Phillips and Patrick Humphreys
Decision Analysis Unit
London School of Economics and Political Science

and

David Embrey
Human Reliability Associates
Lancashire, England

for

The PTS Study Group
of
Oak Ridge National Laboratory

Date of Draft: August 22, 1984

NOTICE: This document contains information of a preliminary nature. It is subject to revision or correction and therefore does not represent a final report.

---

*Research sponsored by U.S. Nuclear Regulatory Commission under Contract No. DE-AC05-840R21400 with the Martin Marietta Energy Systems, Inc. with U.S. Department of Energy.

A PRESSURIZED THERMAL SHOCK EVALUATION OF THE CALVERT CLIFFS UNIT 1 NUCLEAR POWER PLANT

List of Chapters

D.0. APPENDIX D. A SOCIO-TECHNICAL APPROACH TO ASSESSING HUMAN RELIABILITY (STAHR)

LIST OF FIGURES

APPENDIX D. A SOCIO-TECHNICAL APPROACH TO ASSESSING HUMAN RELIABILITY (STAHR)

D.1. Introduction

This appendix describes the status, as of June 1983, of a new approach for assessing human reliability in complex technical systems such as nuclear power plants. This approach was utilized in the present PTS study for Calvert Cliffs Unit 1, the results of which are described in Appendix E.

The new approach includes both a social component and a technical component. To help keep this in mind and also to provide an easily recognized acronym, we are calling our methodology a "socio-technical assessment of human reliability" - or the STAHR approach.

It is important to emphasize that the approach described here does not provide the definitive technical fix to a problem on which a great deal of effort has already been spent. It does, however, provide regulators and risk assessors with another methodology that has certain advantages and disadvantages compared to existing approaches. How useful it proves to be in practice is yet to be determined, but work to date indicates that additional research on this approach is warranted.

A key feature of the approach is that it draws on two fields of study: decision theory and group processes. Decision theory provides the form of the model that allows the desired error rates to be determined, while group processes provide the input data through the group interaction of experts who are knowledgeable about the factors influencing the event whose error rate is being assessed. The different perspectives of these experts, if managed effectively by the group, can lead to informed, useful inputs to the model. Thus, the validity of any error rates that are produced by the model depends not only on the technical model itself, but also on the social processes that help to generate the model inputs.

The impetus for the socio-technical approach began in 1982 at an Oak Ridge, Tennessee, meeting addressing methods for assessing human reliability in the PTS studies. One of us (Phillips) introduced influence diagram technology[1] as a potentially easier modeling tool then event trees or fault trees. The main advantage of an influence diagram from a technical perspective is that it capitalizes on the independence between events and models only dependencies; that is, the influence diagram organizes the dependencies as a system of conditional probabilities, as explained in Section D.2. By the early spring of 1983, the Decision Analysis Unit at the London School of Economics and Human Reliability Associates, Lancashire, England, together had developed a human reliability assessment technology utilizing influence diagrams to the point that it could be tested in the field. In late May a field test was carried out at Hartford, Connecticut, to address operator actions associated with potential pressurized thermal shock events that could occur at the Calvert Cliffs Unit 1 nuclear power station, the results of which are described in Appendix E.

In the paragraphs that follow, a general discussion of influence diagrams is first presented, followed by a description of how group processes work to provide specific diagrams and the input data. Finally, the particularized STAHR approach is described.

D.2. General Description of the STAHR Approach

D.2.1. The Technical Component: The Influence Diagram

As stated above, STAHR consists of both a social component and a technical component. The technical component is the influence diagram. Influence diagrams were developed in the mid-70's by Miller et al.[2] at the Stanford Research Institute and then were applied and further developed at Decisions and Designs, Inc.[3] for intelligence analysis, all without a single paper being published in a professional journal. In 1980, Howard and Matheson[1] extended the theory and showed that any event tree can be represented as an influence diagram, but not all influence diagrams can be turned into event trees unless certain allowable logical transformations are performed on the linkages between the influencing events.

F. D.1

The key principles of influence diagram technology are illustrated by the simple diagrams shown in Figure D.1. Diagram (a) shows the simplest kind of influence. Here Event A is influenced by Event B; that is, the probabilities that one would assign to the occurrence or non-occurrence of Event A are conditional on whether or not Event B has occurred. Shown with the influence diagram is an equivalent event tree representation, where Events A and B are assumed to have only two outcomes, A and $\overline{A}$, B and $\overline{B}$. In the event tree the probability of B occurring is given by p_1. The probability of A occurring, given that B has occurred, is shown by p_2, and the probability of A occurring, given that B has not occurred, is given by p_3. The

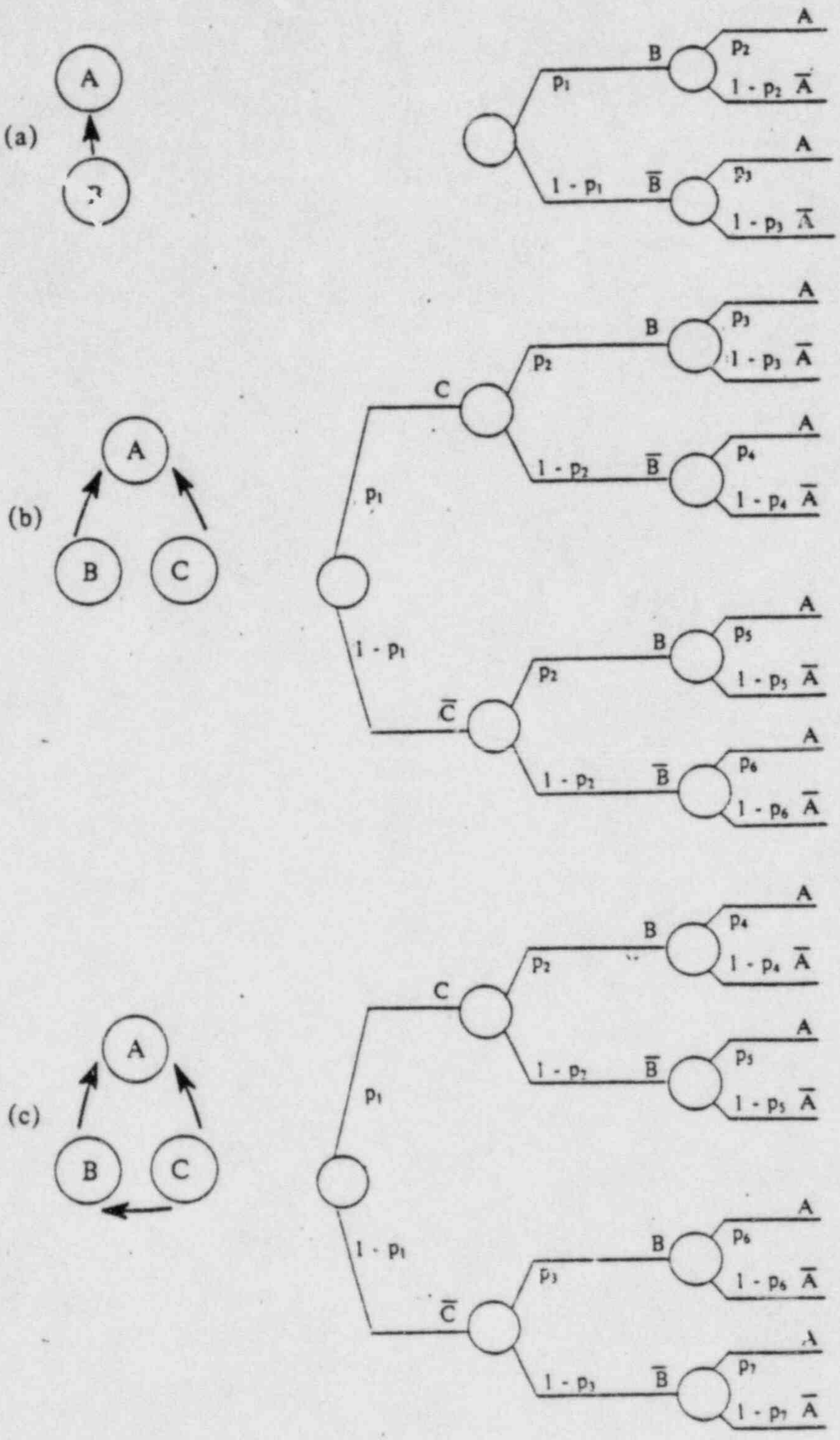

Figure D.1. Influence diagrams and their corresponding event trees. (a) Event A is influenced by event B; (b) event A is influenced by events B and C; and (c) event A is influenced by events B and C, and B is influenced by C.

point here is that p_2 is not equal to p_3. If p_2 and p_3 were equal, then the influence diagram would show two circles unconnected by any influencing link.

Diagram (b) shows a slightly more complex influence. Here, Event A is influenced by both Event B and Event C. The comparable event tree consists of three tiers because the probability assigned to A at the extreme right depends upon the previous occurrence or non-occurrence of both B and C. These probabilities for A, conditional on previous events, are shown by p_3 through p_6. Note that p_2 appears in two places, indicating that the probability assigned to B is the same whether or not C occurs.

Finally, diagram (c) shows the same influences on A as diagram (b), but now Event C influences not only Event A but also Event B. Note that the event trees for diagrams (b) and (c) have the same structure, but for diagram (c) the probability assigned to B conditional on C is no longer the same as the probability of B conditional on C. Thus, while there are six different probabilities in the event tree for diagram (b), there are seven different probabilities in the event tree for diagram (c). It is easy to see that the influence diagram representation not only is compact, but also contains more information than the structure of the event trees without any probability assignments.

In practical situations for which an influence diagram has many nodes, it is typical for the actual number of influencing paths to be far fewer than the maximum that could occur if every node were linked to every other node. Any assessment procedure based on the influence diagram will require only the minimum number of probability assignments. For example, an influence diagram procedure for (b) in Figure D.1 would require only six probabilities and would recognize that the same probability is assigned to Event B whether or not Event C occurs. In an event tree representation of the same problem, dependencies between events are not obvious until probabilities have been associated with each branch, and keeping track of independent events within a large tree can be a tedious housekeeping chore.

In applying influence diagram technology, Event A is taken as the target event and assessments are made of only the necessary and sufficient conditional probabilities that enable the unconditional probability of the target event outcomes to be calculated. For example, in diagram (a) of Figure D.1, the probability of A is given by calculating the joint probabilities of all paths on which an A occurs and then summing the joint probabilities, i.e., $p_1p_2 + (1 - p_1)p_3$. For more complex influence diagrams, successive application of the addition and multiplication laws of probability are sufficient to enable the unconditional probability of the targeted event to be calculated.

It is, of course, important to recognize that no probability is ever unconditional. All events shown on an influence diagram occur within some context, and it is this context that establishes conditioning events that are not usually shown in the notation on the influence diagram. Thus, in applying this technology, it will be important to establish at the start of every assessment procedure what these common conditioning events are.

D.2.2. The Social Component: Human Judgments

The preceding discussion has illustrated how the influence diagram provides the technical means for organizing the conditional probability assessments that are required for calculating the unconditional probability of the target event. But where does the specific influence diagram needed come from, and how are the conditional probability assessments obtained? The answer is that they are developed mainly through human judgments obtained from experts working in groups, and it is these judgments that comprise the "socio" component of the STAHR approach.

The theory behind the socio component was developed and illustrated with a case study by Phillips.[4,5] The key idea is that groups of experts are brought together to work in an iterative and consultative fashion to create a requisite model of the problem at hand. A judgmental model is considered requisite if it is sufficient in form and content to solve the problem. A requisite model is developed by consulting "problem owners," people who have the information, judgment and experience relevant to the problem.

The process of creating a model is iterative, with current model results being shown to the problem owners who can then compare the current results with their own holistic judgments. Any sense of discrepancy is explored, with two possible results: intuition and judgment may be found lacking or wrong, or the model itself may be inadequate or incorrect. Thus, the process of creating a requisite decision model uses the sense of unease felt by the problem owners about current model results, and this sense of unease is used to develop the model further and to generate new intuitions about the problem. When the sense of unease has gone and no new intuitions emerge, then the model is considered requisite. The aim of requisite modeling is to help problem owners toward a shared understanding of the problem, thus enabling decision makers to act, to create a new reality.

A requisite model usually is neither optimal or normative, is rarely descriptive, and is at best conditionally prescriptive. A requisite model is about a shared social reality, the current understanding by the problem owners. Requisite models are appropriate when there is a substantial judgmental element that must be made explicit in order to solve a problem.

Because judgment, intuition and expertise are important ingredients of requisite models, there can be no external reality that can serve as a criterion against which optimality would be judged. Thus, requisite models are not optimal models. Nor are requisite models normative models in the sense that they describe the behavior of idealized, consistent decision makers; that claim would be too strong. Neither can they be considered as descriptive models in the sense that they describe the behavior of actual people. Requisite models are stronger than that; they serve as guides to action, though they may not themselves model alternative courses of action. A requisite model attempts to overcome limitations on human processing of information due to bounded rationality.

Requisite modeling seems ideally suited for the determination of human error rates in complex technical systems. The human operator in a complex system cannot, for the purpose of determining error rates, be treated as an unreliable machine component. In determining error rates for machines, two fundamental assumptions are made. First, that all machines of a particular type are identical as far as error rates are concerned, and second, that all machines of a particular type will be operating within environmental bounds over which the error rate remains unchanged. Neither of these assumptions is true for the human operator. Each person is different from the next, and not even requiring certain standards of training and competence can ensure that other factors, such as those affecting morale and motivation, will not have over-riding effects on the error rates. Moreover, environmental factors can have a substantial impact on human error rates. The same operator may perform differently at a new plant of the same design, if, for example, teams function differently in the two plants. In short, people are different, and the environments they operate in are different, not only from plant to plant but also, from time to time, within a plant. Human error rates are not, then, unconditional figures that can be assigned to particular events. Rather, they are numbers that are conditional on the individual, and on the social and physical environment in which he is operating.

The effective assessment of error rates should take these conditioning influences into account. Technically, the STAHR approach does this by using the influence diagram to display the conditioning influences, and by using the educated assessments of experts to provide judgments that can take account of the uniqueness of the influences for a particular plant.

As yet, it is not known when the STAHR approach should be used in preference to other approaches. It is not even clear whether the STAHR approach should be considered as a competitor to other methods, for it may well turn out that different methods are called for in different circumstances.

Clearly, the STAHR approach focuses on the process of obtaining assessments and in this respect it differs considerably from the handbook approach (the THERP approach) of Swain and Guttman.[6] At this stage of research, it can only be said that the STAHR approach is different from THERP. Our guess is that both STAHR and THERP, and possibly other approaches as well, will each find their own uses, depending on the circumstances. Research is needed to identify those circumstances.

Finally, can experts provide assessments that are valid? Our view is that given the right circumstances people can provide precise, reliable and accurate assessments of probability. This viewpoint is elaborated in Phillips,[7] but some authorities believe that bias is a pervading element in probability assessment.[8] Unfortunately, virtually none of the research that leads to the observation of bias in probability assessments has been conducted under circumstances that would facilitate good assessments. Many of these circumstances are explained in Stael von Holstein and Matheson.[9]

Recent research by the Decision Analysis Unit with insurance underwriters suggests that two additional factors contribute to obtaining good probability assessments. One is the structure of the relationships of events whose probabilities are being assessed, and the other is the use of groups in generating good assessments. In the STAHR approach, the influence diagram presents a well-understood structure within which groups of experts generate assessments.

The success of the STAHR approach depends, in part, on the presence of a group facilitator who is acquainted with the literature on probability assessment and who is experienced in using techniques that facilitate good assessments. How crucial this role is we do not yet know, but we are sure that the necessary expertise and skills can be acquired with reasonable effort by potential group facilitators. In any event, there is nothing in the research literature to suggest that people are incapable of making good assessments. In the United States, weathermen do it now. For example, a review of weather predictions showed that when weathermen predicted a 60% chance of rain within 24 hours, 60% of the time it rained within 24 hours. Thus, weather forecasts are said to be "well-calibrated"; the STAHR approach tries to arrange for circumstances that will promote "well-calibrated" probability assessments. However, calibrating the very low probabilities that emerge from the STAHR approach, or indeed any other approach, is technically difficult because of the low error rates implied. There are simply too few opportunities to determine whether the weathermen's low probability of rain in the desert is realistic.

D.3. Design of the STAHR Influence Diagram

After several revisions, the influence diagram as of June 1983 for events that are influenced by operator actions in nuclear power stations is shown in Figure D.2. We do not yet know whether this influence diagram is generic in the sense that it can handle all events in which operators are expected to take actions. Possibly parts of the diagram are generic and others need to be developed to fit the specific situation. The STAHR approach is sufficiently flexible that modifications to the influence diagram can be made to suit the circumstances, or entirely different influence diagrams could be drawn.

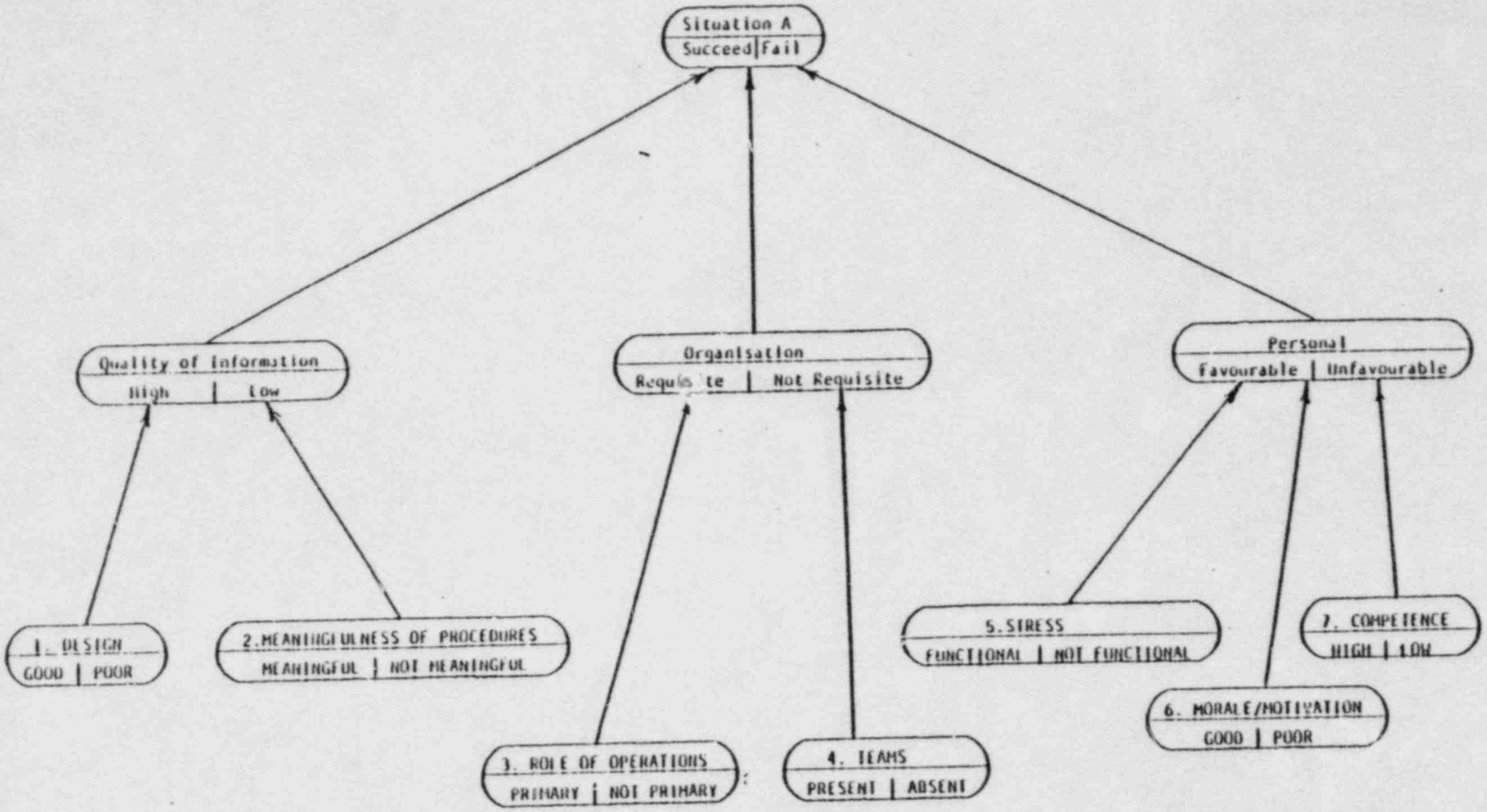

Figure D.2. The STAHR influence diagram (as of June 1983).

The top node in Figure D.2 indicates the target event. For example, if an alarm in the control room signals that some malfunction has occurred and the operator attempts to correct the malfunction by following established procedures, one target event might be that the operator correctly performs a specified step in the procedures. The influence diagram shows three major influences on the target event. One is the quality of information available to the operator, the second is the extent to which the organization of the nuclear power station contributes to getting the work done effectively, and the third is the impact of personal and psychological factors pertaining to the operators themselves. Another way of saying this is that the effective performance of the target event depends on (A) the physical environment, (B) the social environment, and (C) personal factors.

Each of these three major factors is itself influenced by other factors. The quality of information available is largely a matter of good design of the control room and of the presence of meaningful procedures. The organization is requisite; i.e., it facilitates getting the required work done effectively if the operations department has a primary role at the power station and if the organization at the power station allows the effective formation of teams. Personal factors will contribute to effective performance of the target event if the level of stress experienced by operators is helpful, if morale and motivation of the operators are good, and if the operators are highly competent. In other words, the following seven "bottom-level" influences actually describe the power station, its organization and its operators:

(A) Physical Environment

(1) Design of control room (good vs. poor).

(2) Meaningfulness of procedures (meaningful vs. not meaningful).

(B) Social Environment

(3) Role of Operations Department (primary vs. not primary).

(4) Effectiveness of teams (team work present vs. absent).

(C) Personal Factors

(5) Level of stress (helpful vs. not helpful).

(6) Level of morale/motivation (good vs. bad).

(7) Competence of operators (high vs. low).

These seven influences are discussed in more detail in Appendix E with respect to their application during the field testing of the STAHR methodology at Calvert Cliffs. Suffice it to say here that in considering the impact of these seven influences, most nuclear power stations will be found to have mixtures of "good" vs. "poor," "high vs. low," etc.

D.4. Application of the STAHR Influence Diagram

Using the STAHR influence diagram is a matter of applying the following ten steps:

(1) Describe all relevant conditioning events.

(2) Define the target event.

(3) Choose a middle-level event and assess the weight of evidence for each of the bottom-level influences leading into this middle-level event.

(4) Assess the weight of evidence for this middle-level influence conditional on the bottom-level influences.

(5) Repeat steps 3 and 4 for the remaining middle- and bottom-level influences.

(6) Assess probabilities of the target event conditional on the middle-level influences.

(7) Calculate the unconditional probability of the target event and the unconditional weight of evidence of the middle-level influences.

(8) Compare these results to the holistic judgments of the assessors; revise the assessments as necessary to reduce discrepancies between holistic judgments and model results.

(9) Iterate through the above steps as necessary until the assessors have finished refining their judgments.

(10) Do sensitivity analyses on any remaining group disagreements; report either point estimates if disagreements are of no consequence, or ranges if disagreements are substantial.

In step 1, participants would describe the general setting in which the target event might occur, as well as all conditions leading up to the target event. Assessors are reminded that this description and statement of initial conditions form a context for their subsequent assessments and that these assessments are conditional on this context.

In the second stage, the target event is defined in such a way that its occurrence or non-occurrence is capable, at least theoretically, of confirmation without additional information. Thus, "rain tomorrow" is a poorly defined event, whereas "less than 0.1 mm of precipitation falls in a range gauge located at weather station x" is a well-defined event.

In carrying out step 3, the assessors might begin by focusing attention on the left-most middle node, quality of information, and assess weights of evidence for the two bottom influences, design and procedures. This is done with reference to the specific definitions of these bottom influences.* For example, with respect to the design influence, the group of assessors must decide whether, on balance, the design of the particular power station is more similar to the good definitions or to the poor definitions (see items A1 and A2 in list of bottom-level influences in Section D.3). The assessors may find it helpful to imagine a continuous dimension between good and poor and then try to determine where on this dimension this particular power station lies with respect to the event in question. In short, the assessors are judging numbers that reflect the relative weight of evidence as between the poles of the design influence. The weight of evidence would also be judged for the next bottom node, meaningfulness of procedures, but here six different factors, from realism to format, must be taken into account in making the judgment.

The weights of evidence placed on the poles of each dimension are assigned as numbers that sum to 1. Thus, by letting w_1 represent the weight of evidence on the design being good and w_2 represent the weight of evidence on the procedures being meaningful, the assessments for these two bottom nodes can be represented as follows:

	Good	Poor
Design	w_1	$1 - w_1$

	Meaningful	Not Meaningful
Procedures	w_2	$1 - w_2$

---

*For specific definitions, see Table E.2 in Appendix E.

Step 4 requires the assessment of probabilities for the quality of information, a middle-level influence, conditional on the lower-level influences. The poles of the two bottom-level influences combine to make four different combinations: good design and meaningful procedures; good design and not-meaningful procedures; poor design and meaningful procedures; and poor design and not-meaningful procedures. Each of these four combinations describes a hypothetical power station of the sort under consideration, and these hypothetical stations are kept in mind by the assessors when they determine the weight of evidence for the quality of information. This can be set out as follows:

If DESIGN & PROCEDURES		then QUALITY OF INFORMATION is HIGH	 LOW	 JOINT WEIGHTS
Good	Meaningful	w_3	$1 - w_3$	$w_1 w_2$
Good	Not meaningful	w_4	$1 - w_4$	$w_1(1 - w_2)$
Poor	Meaningful	w_5	$1 - w_5$	$(1 - w_1)w_2$
Poor	Not meaningful	w_6	$1 - w_6$	$(1 - w_1)(1 - w_2)$

For example, w_3 is the weight of evidence that the quality of information is high, given that design is good and the procedures are meaningful. Here high quality of information does not mean an ideally perfect power station; instead it means a power station in which both the design and the procedures are of a high, yet practically realizable standard. Neither does low quality of information mean some abysmally bad standard, but rather a standard that is minimally licensable. The assessments w_3 through w_6

capture possible interactions between design and procedures. This is a key feature of the influence diagram technology and experience to date suggests that it is an important feature for human reliability assessment. For example, in some power stations good design may compensate to some extent for procedures that are not very meaningful, whereas if the design were poor the additional burden of procedures that were not meaningful could be very serious indeed.

At this point, a brief technical diversion from describing the ten-step procedure is warranted because it is now possible to illustrate the calculations that are involved in using influence diagrams. The weights are assessed in such a way that they are assumed to follow the probability calculus. Thus, the overall weights of evidence that would be assigned to those four hypothetical stations described at step 4 can be obtained by multiplying the two relevant weights of evidence. For example, the weight of evidence assigned to the actual power station under consideration being both good in design and meaningful in procedures is given by the product of w_1 and w_2. These are shown above as joint weights. Note that the product rule for probabilities is applied. The next stage in the calculation is to multiply these four joint weights by the weights w_3 through w_6 and then to add these four products to obtain the overall weight of evidence that quality of information is high for the power station under consideration. That is,

$$w(\text{HIGH}) = w_3 w_1 w_2 + w_4 w_1 (1 - w_2) + w_5 (1 - w_1) w_2 + w_6 (1 - w_1)(1 - w_2) \ .$$

Note that this calculation makes use of both the product and the addition laws of probability. It is the repeated application of these two laws that allows unconditional weights at higher nodes to be determined. The unconditional weights now determined for the quality of information will serve as weights on the rows of the matrix for the next higher level event, and the types of calculations just illustrated are repeated to obtain the unconditional probabilities for the target event.

Returning now to the ten-step procedure, step 5 requires that steps 3 and 4 be repeated for the rest of the middle- and bottom-level influences. Thus, weights of evidence would be assessed for the role of operations and for teams; then a matrix of conditional probabilities would be assessed for the organizational influence conditional on the lower-level influences. The same procedures would then be followed in making the necessary assessments for the personal factors.

Step 6 requires, for the first time, assessments of probabilities. However, these probabilities are for the target event conditional on the middle-level influences. In a sense, what is being assessed is conditional error rates; that is, assessors are giving their judgments about what the error rates would be under the assumption of particular patterns of influences. Since the quality of information can be either high or low, the organization can either be requisite or not, and personal factors can be favorable or unfavorable. There are eight possible combinations of these influences. A separate error rate associated with the target event is assessed for each of those eight combinations. This is not a particularly easy job for assessors because they must keep in mind three different influences as well as their possible interaction. Favorable personal factors, for example, may well save the day even if the organization is not requisite, and may even compensate to some extent for low quality of information. Insofar as the middle-level influences interact, this stage in the assessment process is important, for it allows assessors to express the effect on error rates of these interactions.

Step 7 is best carried out by a computer which can apply the multiplication and addition laws of probability to determine the unconditional probability of the target event as well as the next-lower influences.

In step 8, the unconditional probabilities and weights of evidence for the middle-level influences are given to the group of assessors who then compare these results to their own holistic judgments. Discrepancies are usually discussed in the group and revisions made as necessary to any assessment.

Step 9 indicates that iteration through the first 8 steps may occur as individual assessors share their perceptions of the problem with each other, develop new intuitions about the problem, and revise their assessment. Eventually, when the sense of unease created by discrepancies between current model results and holistic judgments disappear, and when no new intuitions arise about the problem, model development is at an end, and the model can be considered requisite.

Since individual experts may still disagree about certain assessments, it is worthwhile in step 10 to do sensitivity analyses to determine the extent target event. An easy, but not entirely satisfactory, way to this is first to put in all those assessments that would lead to the lowest probability for the target event and see what its unconditional value is and then to put in all assessments that would lead to the largest probability, thus determining a range of possible results. The difficulty with this is that no individual in the group is likely to believe all of the most pessimistic or all of the most optimistic assessments, so the range established by this approach to sensitivity analysis is unduly large. It should not be too difficult, however, to develop easy and effective procedures for establishing realistic ranges for the probability of the target event, ranges that accommodate the actual variation of opinion in the group.

This has been only a very brief description of the stages that appear to be necessary for applying the influence diagram technology. As experience is gained in the STAHR approach, these steps no doubt will be modified and elaborated. The steps are certainly not intended as a rigid procedure to be followed without deviation. Instead, they should be thought of as an agenda that will guide the work of the group.

D.5. Group Processes

So far, little has been said about the group processes that form the "socio" component of the STAHR approach. A key assumption here is that many heads are better than one for probability assessments. Particularly for human reliability assessment in complex systems, there is unlikely to be any single individual with an unbiased perspective on the problem. Although each individual may be biased in his view, the other side of the coin is that each person has something worthwhile to contribute to the overall assessment. It is within the context of the group that different perspectives of the problem can most effectively be revealed and shared with others, so that the group's main function is the generation of assessments that take into account these different perspectives.

To ensure that all perspectives on the problem are fairly represented, it is important that a group climate be established within which information is seen as a neutral commodity to be shared by all regardless of an individual's status or investment in the problem. The role of group consultant can be established to help create this climate. This individual needs to be conversant with the technical aspect of influence diagrams and with probability assessment and to have a working knowledge of group processes. The group consultant should be seen by the group as an impartial facilitator of the work of the group, as someone who is providing structure to help the group think about the problem but is not providing any specific content. Although the group consultant needs some minimal acquaintance with the principles of nuclear power generation and with the key components in the plant itself, it is probably desirable that he not be a specialist in nuclear power; otherwise he might find it more difficult to maintain a neutral, task-oriented climate in the group. Thus, a major role for the group consultant is not to tell people what to think about the problem but how to think about it.

The other major role for the group consultant is to attend to the group processes and intervene to help the group maintain its task orientation. The group can easily become distracted from its main task because viewpoints in the group will often be divergent. The cognitive maps that a design engineer and a reactor operator have of the same system may be quite different, yet each will at times insist on the validity of his particular viewpoint. The group consultant must help the group to legitimize each of these viewpoints and to explore them in generating useful assessments.

To a certain extent, adversarial processes may even operate in these groups. Operators will openly criticize certain aspects of design, and design engineers may well be contemptuous of procedures that they deem to be unnecessary if only people would operate the system properly. Trainers may be somewhat sceptical of the optimistic "can-do" attitude of the operators, while operators may feel that anyone who has not had "hands-on" experience in the real control room rather than just simulator experience is out-of-date at best and simply out of touch at worst. Unless the group consultant manages the group processes effectively, minor squabbles can easily turn into major confrontations that seriously divert the group from its effective work.

This discussion is not meant to imply that the group should be composed so as to reduce adversarial processes. On the contrary, an underlying assumption of the STAHR approach is that diversity of viewpoint is needed if good assessments are to be generated. Differences are to be confronted openly in the group and to be taken seriously regardless of the status of the holder of the viewpoint. Thus, diversity of viewpoint is a key criterion in composing the groups. As yet, we are not certain about the roles that should be represented in the groups but it would appear that at least the following are necessary: group consultant, technical moderator to help direct the discussion on technical issues, trainer of nuclear power station operators, eligibility and systems analyst, thermohydraulics engineer, possibly one or two other engineers with specialized knowledge of the power station, and, of course, reactor operators. Further work is needed to establish exactly who the problem owners are for these human reliability assessments.

D.6. Summary Statement

This appendix has described the STAHR approach as it was originally conceived for application to the assessment of the reliability of operator actions at a nuclear power station during potential PTS events. As will be apparent from Appendix E, the first field test of the methodology resulted in some modifications of the detailed definitions of the bottom-level influences, and further revisions are anticipated as the approach is more generally applied.

REFERENCES

1. Howard and Matheson, 1980.
2. Miller et al., 1976.
3. Selvidge, 1976.
4. Phillips, 1983.
5. Phillips, 1982.
6. NUREG/CR-1278.
7. Phillips, 1981.
8. Fischhoff & MacGregor, 1982.