

The role of error in organizing behaviour*

J Rasmussen

Qual Saf Health Care 2003;12:377–385

During recent years the significance of the concept of human error has changed considerably. The reason for this has partly been an increasing interest of psychological research in the analysis of complex real life phenomena, and partly the changes of modern work conditions caused by advanced information technology. Consequently, the topic of the present contribution is not a definition of the concept or a proper taxonomy. Instead, a review is given of two professional contexts for which the concept of error is important. Three cases of analysis of human–system interaction are reviewed: (1) traditional task analysis and human reliability estimation; (2) causal analysis of accidents after the fact, and (3) design of reliable work conditions in modern sociotechnical systems. It is concluded that “errors” cannot be studied as a separate category of behaviour fragments; the object of study should be cognitive control of behaviour in complex environments.

reliability analysis is focused on the less well structured and stable tasks of diagnosis and contingency planning. Focus of error analysis is moved back from overt acts to decision functions and further on to psychological mechanisms (fig 1).

It is a remarkable fact that, given a particular sequence of human acts, taxonomies of error analysis resulting from detailed analysis of actual cases of incidents and accidents and from psychological laboratory research show definitive convergent properties.² When a particular task sequence can be taken as reference (i.e. a sequence which is functionally constrained by the equipment to be operated or firmly established by training), a failure-mode-and-effect analysis is a very feasible approach to identify the hazards presented by human error. It will be effective during design to ensure error tolerance, even if quantitative reliability prediction may not be realistic.³

A necessary precondition is, however, that the sequence in which the “error” is analysed can be taken for granted. This is the case only when we are involved in a local analysis focused on the immediate human–machine interface: we then try to predict the risk involved in the operation of some particular technological system of a known design. The acceptable work procedure is identified from the functional requirements of equipment, given a definite goal. This is, as noted, a reasonable assumption if the task is repetitive, which was the normal case in established technology. In addition, we are dealing with a human link in an extended chain of events; the “error” is a link in the chain, in most cases not the origin of the course of events. This kind of analysis and, consequently, definition of error is completely inadequate when we are dealing with design or improvement of large scale sociotechnical systems. In general, we do not have a simple causal trace deflected from its intended course toward one goal. Actually, such a separate trace is the manifestation of the dynamic flow of events in a complex network involving several goals and side effects and many side branches. Previous flows of events along these branches serve to precondition the “riverbed” in which the dynamic flow is found.

CAUSAL ANALYSIS OF ACCIDENTS AFTER THE FACT

In this case, we are analysing an accidental chain of events upstream from an accident in order to understand *why* it happened; to find somebody to blame, *who* did it; or to find out how to improve the system. We are trying to describe a particular course of events and to identify the particular causal trace in which human error is embedded.

Accidents are normally analysed in terms of accidental chains of events, i.e. causal representations. Since no two accidents will be identical,

TRADITIONAL TASK ANALYSIS AND HUMAN RELIABILITY ESTIMATION

Human activities in traditional work environments can be described in terms of repetitive tasks, i.e. sequences of acts in control of some equipment or tool. Manufacturing systems were normally planned for effective and economic operation over long periods of time. Planned or normal work sequences had time to settle into stable patterns which could be identified during design by analysis of the task, to control tools and equipment, or afterwards by field studies. Since successful operation during production or a mission was of fundamental interest, technical and human reliability analysis became important design tools both for military and high hazard industrial operations.

In this situation, human errors can easily be defined; normative sequences of proper acts are available for reference and errors can be identified and recorded. As long as actors enter the proper sequence at all, errors caused by lack of resources or proper intention are of minor importance and errors can be studied in terms of their overt effects (Swain’s Therp method).¹ In modern workplaces people are frequently moved to supervisory tasks and decision making. In that case,

Correspondence to:
J Rasmussen, Risø National
Laboratory, P O Box 49,
DK-4000 Roskilde,
Denmark

*This is a reprint of a paper that appeared in *Ergonomics* 1990, Volume 33, pages 1185–99. Copyright © 1990 Taylor & Francis Ltd.

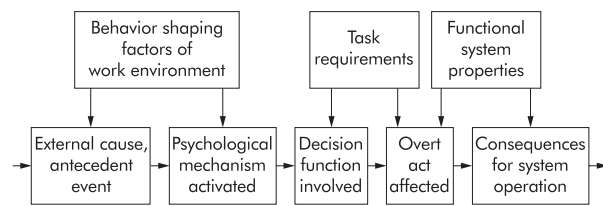


Figure 1 An illustration of the human involvement in a causal sequence of events. The event of human error is decomposed to identify the cognitive task element and the psychological mechanism involved in the error. At this level of detail, an event in the work context activates a particular psychological mechanism which influences the immediate decision task required by the work. A decision error in turn introduces an error in the overt action sequence, with unacceptable consequences for the work goal. Two aspects are essential in the present context. One is that human "error" very frequently will be a link in a sequence, not the origin. Secondly, "errors" can be categorized at different states in the flow. This representation is well suited for failure-mode-and-effect analysis in the interface of technical systems: the various psychological "error mechanisms" are folded on to the cognitive task from which the effects, in turn, are folded on to functional system properties for evaluation of the acceptability of consequences.

accident analysis will depend on prototypical categories of causes, events, and consequences.⁴ An explicit representation of elements in the physical world makes causal analysis a very effective technique for identifying and representing accidental conditions. It is, however, important to consider the implicit frame of reference of a causal analysis.

The behaviour of the complex, real world is a continuous, dynamic flow which can only be explained in causal terms after decomposition into discrete events. The concept of a causal interaction of events and objects depends on a categorization of human observations and experiences. Perception of occurrences as events in causal connection does not depend on categories which are defined by lists of objective attributes, but on categories which are identified by typical examples, prototypes.⁵ This is the case for objects as well as for events. Everybody knows perfectly well what "a cup" is. To define it objectively by a list of attributes that separates cups from jars, vases, and bowls is no trivial problem. It has, for instance, been faced in many attempts to design computer programs for picture analysis. The problem is that the property of being "a cup" is not a feature of an isolated object, but depends on the context of human needs and experience. The identification of events in the same way depends on the relationship in which they appear in a causal statement. An objective definition, therefore, will be circular.

In the analysis of accidents, decomposition of the dynamic flow of changes will normally terminate when a sequence is found including events which match the prototypes familiar to the analysis. The resulting explanation will take for granted his or her frame of reference and, in general, only what he or she finds to be unusual will be included: the less familiar the context, the more detailed the decomposition. By means of the analysis, a causal path is found upstream from the accidental effect. This path will be prepared by resident conditions which are latent effects of earlier events or acts. Also, the resident conditions can be explained by causal back tracking, in this case branches in the path are found. To explain the accident, these branches are also traced backward until all conditions are explained by abnormal, but familiar, events or acts. The point is: How does the degree of decomposition of the causal explanation and the selection of the side branches depend on the circumstances of the analysis? Another question is: What is the stop rule applied for termination of the search for causes? Ambiguous and implicit stop rules will make the results of analysis very sensitive to the topics discussed in the professional community at any given time (fig 2). There is a

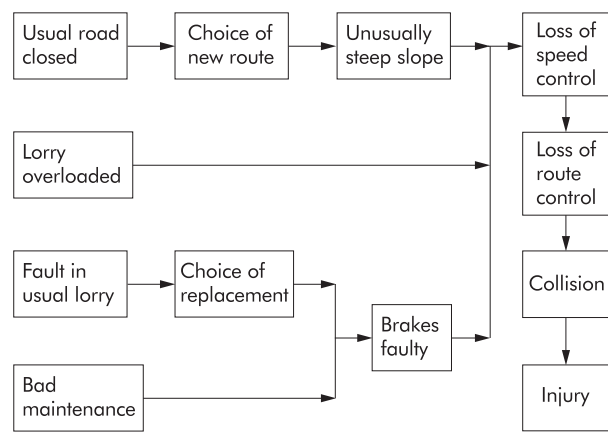


Figure 2 An illustration of a causal explanation of a driving accident. The flow of behaviour is decomposed into chains of events. Note that only abnormal or unusual events together with violations of rules are included. The normal activities conditioning the path are not included. Furthermore, decomposition and causal backtracking stop at events which are taken to be "reasonable explanations". Adapted from Leplat and Rasmussen.⁶

tendency to see what is expected; during one period technical faults were in focus as causes of accidents, then human errors predominated, while in the future focus will probably move upstream to designers and managers.

The perception of stop rules is very important in the control of causal explanations. Every student knows the relief felt when finding a list of solutions to mathematical problems. Not that it gave the path to solution to any great extent, but it gave a clear stop rule for the search for possible mistakes, overseen preconditions, and calculation errors. The result: hours saved and peace of mind. A more professional example to the same point is given by Kuhn.⁷ He mentions the fact that chemical research was able to come up with whole number relations between elements of chemical substances only after the acceptance of John Dalton's chemical atom theory. There had been no stop rule for efforts to refine the experimental technique until the acceptance of this theory.

Stop rules are not usually formulated explicitly. The search will typically be terminated pragmatically in one of the following ways: (a) an event will be accepted as a cause and the search terminated if the causal path can no longer be followed because information is missing; (b) when a familiar, abnormal event is found to be a reasonable explanation; or © if a cure is available. The dependence of the stop rule upon familiarity and the availability of a cure makes the judgement very dependent upon the role in which a judge finds himself. An operator, a supervisor, a designer, and a legal judge may very likely reach different conclusions.

To summarize: identification of accident causes is controlled by pragmatic, subjective stop rules. These rules depend on the aim of the analysis, i.e. whether the aim is to explain the course of events, to allocate responsibility and blame, or to identify possible system improvements in order to avoid future accidents.

Analysis of explanation

In an analysis to explain an accident, the backtracking will be continued until a cause is found which is familiar to the analysts. If a technical component fails, a component fault will only be accepted as the prime cause if the failure of the particular type of component appears to be "as usual". Further search will probably be made if the consequences of the fault make the designer's choice of component quality unreasonable, or if a reasonable operator could have terminated the effect had he been more alert or been trained better. In such a

case, a design or manufacturing error, respectively, or an operator error can be found.

In most recent reviews of larger industrial accidents it has been found that human errors played an important role in the course of events. Frequently, errors are attributed to operators involved in the dynamic flow of events. This can be an effect of the very nature of the causal explanation. Human error is, particularly at present, familiar to analysts: to err is human, and highly skilled people will frequently depart from normative procedures, as we will see subsequently.

Analysis for allocation of responsibility

In order to allocate responsibility, the stop rule of the backward tracing of events will be to identify a person who has made an error and at the same time was "in control" of his or her acts. The very nature of the causal explanation will focus attention on people directly and dynamically involved in the flow of abnormal events. This is unfortunate because they may very well be in a situation where they do not have "control". Traditionally, a person is not considered responsible if physically forced to act by another person or when subject to disorders such as epileptic attacks. In such cases, acts are involuntary^{8,9} from a judgement based on physical or physiological factors. It is, however, a question as to whether cognitive psychological factors also should be taken more into account when judging responsibility. Inadequate response of operators to unfamiliar events depends very much on the conditioning taking place during normal work. This problem also raises the question of the nature of human error. The behaviour of operators is conditioned by the conscious decisions made by work planners or managers. They will be more "responsible" than an operator in the dynamic flow of events. However, their decisions may not be considered during a causal analysis after an accident because they are "normal events" which are not usually represented in an accident analysis. Furthermore, they can be missed in analysis because they are to be found in a conditioning side branch of the causal tree, not in the path involved in the dynamic flow.

Present technological development toward high hazard systems requires a very careful consideration by designers of the effects of "human errors" which are commonplace in normal daily activities, but unacceptable in large scale systems. There is considerable danger that systematic traps can be arranged for people in the dynamic course of events. The present concept of "responsibility" should be reconsidered from a cognitive point of view, as should the ambiguity of stop rules in causal analysis.

Analysis for system improvements

Analysis for therapeutic purpose, i.e. for system improvement, will require a different focus with respect to selection of the causal network and of the stop rule. The stop rule will now be related to the question of whether an effective cure is known. Frequently, cure will be associated with events perceived to be root causes. In general, however, the effects of accidental courses of events can be avoided by breaking or blocking any link in the causal tree or its conditioning branches. Explanatory descriptions of accidents are, as mentioned, focused on the unusual events. However, the path can also be broken by changing normal events and functions involved. The decomposition of the flow of events, therefore, should not focus solely on unusual events, but should also include normal activities.

The aim is to find conditions sensitive to improvements. Improvements imply that some person in the system makes decisions differently in the future. How do we systematically identify persons and decisions in a (normal) situation where it would be psychologically feasible to ask for a change in behaviour, when reports from accidents focus only on the flow of unusual events? An approach to such an analysis for improving safety has been discussed elsewhere.⁶

In conclusion, the choice of stop rules for the analysis of accidents is normally left to the subjective judgement of the analyst, depending heavily on the aim of his analysis. Analyses made for one purpose may therefore be misleading for other purposes.

DESIGN OF RELIABLE WORK CONDITIONS AND SOCIOTECHNICAL SYSTEMS

Modern work conditions

A number of problems are met when attempts are made to improve the safety of sociotechnical systems from analyses tied to particular paths of accidental events. This is due to the fact that each path is a particular token shaped by higher order relational structures. If changes are introduced to remove the conditions for the existence of a particular link in the chain, odds are that this particular situation will never occur again. We should be fighting types of accident causation, not these individual tokens. Only in the immediate interface with technical systems is human behaviour constrained in a way that makes the chain of events reasonably predictable. The farther away from the technical core we are, the more degrees of freedom agents will have in their mode of behaviour. Consequently, the less certain is the reference in terms of normal or proper behaviour for judging "errors". This problem is becoming increasingly important as modern manufacturing systems and organizations are forced to respond to increasingly dynamic market requirements, technological innovations, and legal constraints.

Given this situation, improvements of safety features of a sociotechnical system depend on a global analysis: no longer can we assume the time course of human behaviour to be predictable. Tasks will be formed for the occasion, and design for improvements must be based on attempts to find means of control at higher levels than that of particular task procedures. If, for instance, sociotechnical systems have features of adaptation and self-organization, changes which aim to improve safety at the individual task level might well be compared with attempts to control the temperature in a room with a thermostat controlled heater by opening the window. In other words, it is not sensible to try to change the performance of a feedback system by alterations inside the loop; you have to identify mechanisms that are sensitive, i.e. related to the control reference itself.

Some basic high level features of "human error" in a flexible sociotechnical system are related to the dependence of human performance on features such as: (1) learning and adaptation; (2) conflicts among cognitive control structures; (3) resource limitations; and finally, (4) stochastic variability. An attempt to develop guidelines for the design of human-work interfaces has been presented elsewhere.¹⁰

Human adaptation

In all work situations constraints are found which must be respected in order to perform satisfactorily. There are, however, also many degrees of freedom which have to be resolved at the worker's discretion. In stable work conditions, know-how will develop which represents prior decisions and choice and the perceived degrees of behavioural freedom will ultimately be very limited, i.e. "normal ways" of doing this will emerge, and the process of adaptation will no longer confuse the concept of error. By contrast, in modern, flexible, and dynamic work conditions, the immediate degrees of freedom will have to be continuously resolved. This implies that effective work performance includes continuous awareness of the available degrees of freedom together with effective strategies for making choices, ahead of the task of controlling the chosen path to a goal. This changes the concept of error in a very fundamental way.

The behaviour in work of individuals (and, consequently, also of organizations) is, by definition, oriented towards the

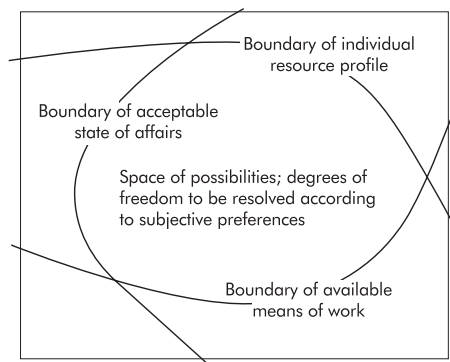


Figure 3 Human behaviour is governed by constraints which must be respected by the actors for the work performance to be successful. Identification of such constraints will specify the “space” in which the human can navigate freely. Violation of the constraints will be considered human error or task violation in the usual sense. For successful performance, humans have to navigate between two boundaries of constraints. One boundary is given by the control requirements posed by the system. The other constraining boundary is given by the human resource profile which depends on individual characteristics such as competence, mental capacity, physical strength, etc. Navigation within the envelope specified by these boundaries will depend on subjective criteria for choice, such as aim to save time, to spare memory load, to have fun, to explore new land, etc.

requirements of the work environment as perceived by the individual. Work requirements, *what* should be done, will normally be perceived in terms of control of the state of affairs in the work environment according to a goal, i.e. *why* it should be done. How these changes are made is, to a certain degree, at the discretion of the agent.

The alternative acceptable work activities, *how* to work, will be shaped by the work environment which defines the boundaries of the range of possibilities, i.e. acceptable work strategies. This range of possibilities will be further bounded by the resource profile of the particular agent in terms of tools available, knowledge (competence), information about state of affairs, and processing capacity. The presence of alternatives for action depends on a many-to-many mapping between means and ends present in the work situation as perceived by the individual; in general, several functions can serve the individual goals and each of the functions can be implemented by different tools and physical processes. If this was not the case, the work environment would be totally predetermined and there would be no need for human choice or decision (figs 3 and 4).

Within the area of acceptable work performance, between the boundaries defined by the work requirements on one side and the individual resource profile on the other, considerable degrees of freedom are still left for the individual to choose among strategies and to implement them in particular sequences of behaviour. These degrees of freedom must be eliminated by the choice of an agent to finally enter a particular course of action. The different ways to accomplish work can be categorized in terms of strategies, defined as types of behavioural sequences which are similar in some well defined aspects, such as the physical process applied in work and the related tools, or, for mental strategies, the underlying kind of mental representation and the level of interpretation of perceived information. In any particular situation-dependent exemplar of actual performance, a token will emerge which is an implementation of the chosen strategy under the influence of the complexity of detail in the environment. The particular token of performance will be unique and impossible to predict, whereas the strategy chosen will, in principle, be predictable. This choice made by individual agents depends on subjective performance criteria related to the process of work,

	Means-end relations
Goals and values, constraints	Work, income, intellectual and esthetic pleasure; social family relations
Priority measures, flow of money, material, information	Expenses, joy probability of success and delays
General functions	Work function: transport Family function: shopping
Physical processes in work and equipment	Train operation, schedules, space for reading; taxi: time spent, traffic jams, price; car: traffic, time, price
Material objects, tools, buildings, etc. topography	Train, taxi, private car routes, distance, scenery; shops and locations, etc.

Figure 4 An example: the activities involved in going to work. The *work given constraints* are related to the location, the time of arrival, and the probability of delays. *Constraints in means* are defined by the transport alternatives, i.e. to take the tube, a taxi, or to drive by yourself. The *subjective process criterion* determining your choice depends on economy, your husband or wife’s request to bring some grocery and, maybe, consideration of the time spent, the likelihood of traffic jams. Given the decision to drive by yourself, the choice of route depends on the secondary task of shopping, of your joy with a particular scenery, and the traffic density. Finally, en route, the speed you choose depends on traffic given constraints, on *formal conditions* such as speed limits or your husband or wife’s anxiety, and ultimately “sporty” criteria related to your driving skill, i.e. to drive fast with smooth gear changes.

such as time spent, cognitive strain, joy, cost of failure, etc. In general, the freedom to choose work strategy is very important as a means to resolve resource-demand conflicts met during performance.

Modelling work activity from this point of view depends on identification of the range of acceptable and possible work strategies (i.e. prototypical sets of behaviour sequences), the human resource profile, and the subjective criteria governing the resolution of the remaining degrees of freedom in different work scenarios. Some work requirements are explicit and discrete, with specified limits of acceptance. Other requirements are formulated as optimizing criteria serving to resolve ambiguity in goal specification, such as the request to reach a solution which is as cheap or as safe as possible. Such product criteria, together with the subjective process criteria, will necessarily lead to an adaptive behaviour seeking to optimize performance according to the criteria, along with evolution of training and expertise (fig 5).

Adaptation, self-organization, and error

It follows directly from this discussion that the structuring of work processes through on-the-job training by an individual will be a self-organizing, evolutionary process, simply because an optimizing search is the only way in which the large number of degrees of freedom in a complex situation can be resolved. The basic synchronization to the work requirements can be based on procedures learned from an instructor or a more experienced colleague, or it can sometimes be planned by the individual in a knowledge based mode of reasoning by means of mental experiments. From here, the smoothness and speed characterizing high professional skill, together with a large repertoire of heuristic know-how rules, will evolve through an adaptation process in which “errors” are unavoidable side effects of the exploration of the boundaries of acceptable performance. During this adaptation, performance

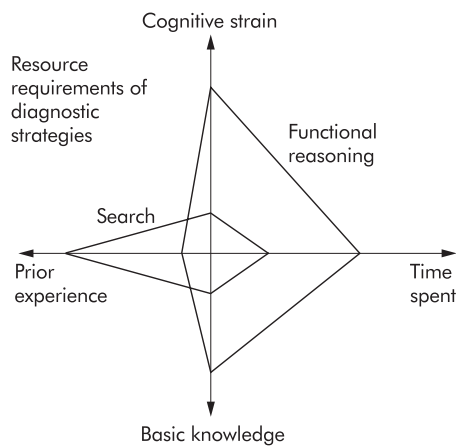


Figure 5 An illustration of the different resource requirements of different mental strategies. This difference makes a shift in strategy when faced with difficulties in a task, an effective way to navigate along the path of least effort, a very popular strategy in skilled performance to adapt behaviour to immediate work situation.

will be optimized according to the individual's subjective process criteria, within the boundary of his individual resources. This complex adaptation of performance to work requirements, eliminating the necessity of continuous choice, will result in stereotype practices depending on the individual performance criteria of the agents. These criteria will be significantly influenced by the social norms and culture of the group and organization.

Conflict will probably be found between global work goals and the effect of local adaptation according to subjective process criteria. Unfortunately, the perception of process quality can be immediate and unconditional, while the effect on product quality of the choice of actor can be considerably delayed, obscure and frequently conditional with respect to many other factors.

In a first encounter, when representation of work constraints is not present in the form of instructions from an experienced colleague or a teacher, and know-how from previous experiences is not available, the constraints of the work have to be explored in a knowledge based mode from explicit consideration of the actual goal and a functional understanding of the relational structure of the work content. For such initial exploration as well as for problem solving during unusual task conditions, opportunities for tests of hypotheses and trial-and-error learning are important. It is typically expected that qualified personnel such as process operators check their diagnostic hypotheses conceptually—by thought experiments—before actual operations if acts are likely to be irreversible and risky. This appears, however, to be an unrealistic assumption, since it may be tempting to test a hypothesis on the physical work environment itself in order to avoid the strain and uncertainty related to unsupported reasoning in a complex causal net. For such a task, a designer is supplied with effective tools such as experimental set-ups, simulation programs, and computational aids, whereas the operator has only his or her head and the plant itself. In the actual situation, no explicit stop rule exists to guide the termination of conceptual analysis and the start of action. This means that the definition of error, as seen from the situation of a decision maker, is very arbitrary. Acts which are quite rational and important during the search for information and tests of hypotheses may appear to be unacceptable mistakes with hindsight, without access to the details of the situation.

Even if a human actor is "synchronized" to the basic requirements of work by effective procedures, there will be ample opportunities for modification of such procedures. Development of expert know-how and rules-of-thumb

depends on adaptation governed by subjective process criteria. Opportunities for experiments are necessary to find shortcuts and to identify convenient and reliable cues for action without analytical diagnosis. In other words, effective, professional performance depends on empirical correlation of cues to successful acts. Humans typically seek the way of least effort. Therefore, it can be expected that no more information will be used than is necessary for discrimination among the perceived alternative for action in any particular situation. This implies that the choice is "under-specified"¹¹ outside that situation. When situations change, e.g. due to disturbances or faults in the system to be controlled, reliance on the usual cues which are no longer valid will cause an error due to inappropriate "expectations". In this way, traps causing systematic mistakes can be designed into the system. Two types of errors are related to this kind of adaptation: firstly, the effect of the test of a hypothesis of salient cues and action which turn out to be negative and, secondly, the effects of acts chosen from familiar and tested cues when a change in system conditions makes the perceived set of alternatives unreliable.

An example in which local adaptation is in conflict with delayed and conditional effects is working instructions which take into consideration the possible presence of abnormal conditions that will make certain orders of actions unacceptable. The instruction to this effect prescribes a certain sequence. If this prescribed order is in conflict with the actor's immediate process criteria, modification of the prescribed procedure is very likely and will have no adverse effect in the daily routine. (If, for instance, an actor has to move back and forth between several distant locations because that sequence is safer under certain infrequent risky conditions, his or her process criterion will rapidly teach him or her to group actions at the same location together, because this change in the procedure will have no visible effect under normal circumstances.)

Even within an established effective sequence of actions, adaptation of the patterns of movements will occur according to subconscious perception of certain process qualities. In a manual skill, fine-tuning depends upon a continuous updating of automated patterns of movement to the temporal and spatial features of the task environment. If the optimization criteria are speed and smoothness, adaptation can only be constrained by the occasional experience gained when crossing the tolerance limits, i.e. by the experience of errors or near errors (speed-accuracy trade off). Some errors therefore have a function in maintaining a skill at its proper level, and they cannot be considered a separable category of events in a causal chain because they are integral parts of a feedback loop. Another effect of increasing skill is the evolution of increasingly long and complex patterns of movements which could run off without conscious control. During such lengthy automated patterns, attention is directed towards reviews of past experience or planning of future needs (fig 6) and performance is sensitive to interference, i.e. capture from very familiar cues.

When delayed or conditional global effects of behaviour are possible, feedback correction and control of the local adaptation is not possible, and adaptation is controlled by an evolutionary "survival of the fittest" work process. In order to compete effectively with the effect of the local process criteria, the perception of fitness of such stored procedures must be maintained in another way (e.g. by artificial reinforcement or, preferably, by rearranging the environment to include the global requirements in the local criteria). Otherwise, simple decay of memory of stored work rules (decay is, in effect, necessary for adaptation to changing requirements from a work environment) will necessarily require a repeated experience of the conflict in order to maintain proper adaptation to characteristics of the environment.

KNOWLEDGE-BASED DOMAIN	Planning in terms of functional reasoning by means of symbolic model: "As Can Be" Achronic	OFF-LINE EVALUATION AND PLANNING	
RULE-BASED DOMAIN	Planning in terms of recall of past and rehearsal of future, predicted scenarios: "As Has Been and May Be" Diachronic	Attention on cue classification and choice of action alternatives Synchronic "As Is"	
SKILL-BASED DOMAIN	Synchronous "As Is"		
	Data driven chaining of sub-routines with interrupt to conscious, rule-based choice in case of ambiguity or deviation from current state of the internal world model		ON-LINE, REAL TIME OPERATION

Figure 6 An illustration of the complex interaction between the different levels of cognitive control. Tasks are frequently analysed in terms of sequences of separate acts. In general, however, control of several acts takes place concurrently. At the level of skilled sensory motor control, activity is like a continuous dynamic interaction with the environment. Attention, on the other hand, is scanning across time and activities in order to analyse past performance, monitor current activity, and plan for foreseen future requirements. In this way, the internal dynamic world model is being prepared for oncoming demands, and the related cues and rules are rehearsed and modified to match predicted requirements, and symbolic reasoning is used to understand responses from the environment and to prepare rules for foreseen but unfamiliar situations. Attention may not always be focused on current activities, and different levels may simultaneously be involved in the control of different tasks, related to different time slots, in a time sharing or in a parallel processing mode.

The structure of cooperative work

So far, the discussion has been focused on the individual adaptation of work strategies to task requirements. In general, however, several people will be active in a work environment, and the allocation or acceptance of the roles of individuals will evolve in a self-organizing mode according to local criteria and within the constraints of externally imposed allocation structures. Such constraints on the evolutionary allocation can have their origin in work requirements as well as in human resource limitations.

Role allocation

Some constraints on work allocation originate in the work domain. Actions can, for instance, be required simultaneously in separate locations; or work can require competence which is dependent on more than one profession. Such conditions will limit the extent to which allocation can be dynamically adapted to the preference of the involved individuals. In some cases, however, constraints are rather lenient and will not be respected strictly during adaptation (e.g. the boundaries between activities which have been assigned to members of different unions by labour market agreements). In other cases, constraints are effectively reinforced, as, for example, when performance is governed by strict quality control standards, as is the case for manufacturing according to machining specifications, or in financial operations with strict legal control. In most cases, however, boundaries among the roles allocated the individual actors are continuously adjusted according to the requirements of the immediate work situation.

As was the case for the choice among alternative work strategies, the dynamic shifting of boundaries among allocated rules will be used to resolve resource demand conflicts and to match performance to individual preferences. The subjective criteria active in this adaptation will be very situation-dependent and directly related to the particular work process, such as perception of differences in workload among colleagues, the amount of communication necessary among agents for coordination, subjective preferences for certain activities, etc. This adaptation of role allocation and coordination to work requirements during normal conditions will endanger functioning during exceptional situations.

Coordination of cooperative work

For concerted work activity, the different processes and functions of work within the various levels of the mean-ends space of a work domain will be allocated among several individuals. Often, coordination will be allocated among other individuals than those directly performing the functions to be coordinated. This is the case in all hierarchic organizations. In effect, boundaries are found between roles at different levels in the hierarchical control structure, as well as among roles within these levels.

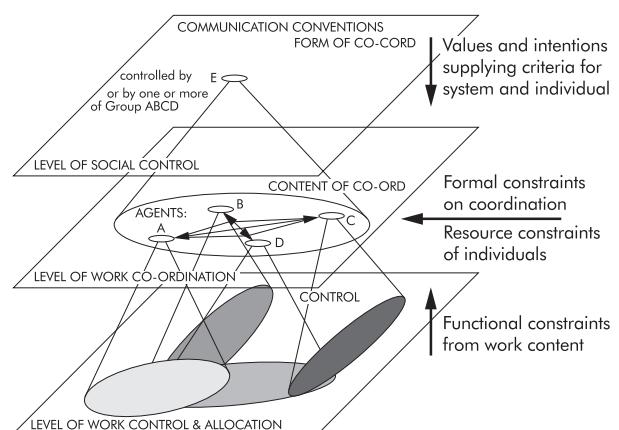


Figure 7 An illustration of the coordination of cooperative work. At the level of work, a dynamically changing allocation to individuals is governed by criteria such as sharing load, minimizing communication, individual interest, etc. At the level of coordination, the content of communication necessary for concerted action is specified by the work content and the actual role allocation. In this way, the *work organization* is dynamically shaped bottom up. Management practice and social values define rules of conduct, i.e. the form of the coordination, and therefore are shaping the *social organization* top down. In addition, formal constraints such as laws, regulations, and union agreements add constraints on allocation and coordination "side in". Within the boundaries defined in this way, there is plenty of room for adaptation guided by subjective criteria. Adapted from Rasmussen.¹²

The basic structure of the allocation depends on the functional requirements of the work content, such as the topographic location of work items, the workload related to certain functions, the timing required between functions in different places, and the time frames to consider in coordination at the various levels. In other words, technology shapes organizations bottom up by imposing strict constraints on allocation of functions to groups and individuals. In many domains, in particular in tightly coupled technical domains like manufacturing, process control, etc. strict control and timing requirements can be explicitly formulated from an analysis of the work requirements (fig 7).

Within the allocation and coordination constraints imposed by the work content, there are many degrees of freedom to arrange the role allocation and to structure the way in which coordination is brought about. Additional formal constraints on allocation can originate in legal requirements (authorization, etc.), agreements (union boundaries), regulations (quality assurance standards) and rules of conduct (military).

System reliability and safety

The dynamic adaptation to immediate work requirements, both of individual performance and of the allocation between individuals, will probably create a very high degree of reliability as long as the interaction is transparent (i.e. critical aspects are visible without excessive delay), and individual process criteria are not in conflict with, or are not overriding, critical product criteria.

Under certain conditions, however, self-organizing and adaptive features will necessarily lead to "catastrophic" system behaviour unless certain organizational criteria are met. Adaptation will normally be governed by local criteria, related to an individual's perception of process qualities in order to resolve the perceived degrees of freedom in the immediate situation. Some critical product criteria (e.g. safety) are conditionally related to higher level combination or coincidence of effects of several activities, allocated among different agents and, probably, in different time slots. The violation of such high level, conditional criteria cannot be monitored and detected at the local criterion level, and monitoring by their ultimate criterion effect will be unacceptably delayed. Catastrophic effects of adaptation can be avoided only if local activities are tightly monitored with reference to a prediction of their role in the ultimate conditional effect, i.e. the boundaries at the local activities are necessarily defined by normal prescriptions, not active functional conditions.

This feature of adaptation to local work requirements probably constitutes the fallacy of the defence-in-depth design principle normally applied in high risk industries.¹⁵ In systems designed according to this principle, an accident is dependent on simultaneous violation of several lines of defence: an operational disturbance (technical fault or operator error) must coincide with a latent faulty maintenance condition in protective systems, with inadequacies in protective barriers, with inadequate control of the location of people close to the installation, etc. The activities threatening the various conditions normally belong to different branches of the organization. The presence of a potentially catastrophic combination of the effects of local adaptation to performance criteria can only be detected at a level in the organization with the proper overview. However, at this level in the control hierarchy (organization), the required understanding of conditionally dangerous relations cannot be maintained in the long term because the required functional and technical knowledge is foreign to the normal management tasks at this level.

The conclusion of this discussion is that catastrophic system breakdown is a normal feature of systems which have self-organizing features and, at the same time, depend on protection against rare combinations of conditions which are individually affected by adaptation. Safety in such systems

depends on the introduction of locally visible boundaries of acceptable adaptation and the introduction of related control mechanisms. What does this mean in terms of organizational structures? What kind of top down influence from "management culture" and bottom up technological constraints can be used to guide and limit adaptation? How can we model and predict the evolution of organizational structure?

CONCLUSION

Work in modern "high tech" societies calls for a reconsideration of the notion of human error: research should be focused on a general understanding of human behaviour and social interaction in cognitive terms in complex, dynamic environments, not on fragments of behaviour called "error". This approach has similarities to the "risk homeostasis" theories of traffic safety, with the reservation that the controlling mechanisms are adaptation in a wider sense than control governed by criteria related to risk.

REFERENCES

- 1 Swain AD, Guttman HE. *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. NUREG/CR 1278. Albuquerque, NM: Sandia Laboratories, 1983.
- 2 Rasmussen J. Human error mechanisms in complex work environments. *Reliability Engineering System Safety* 1983;22:155-69.
- 3 Rasmussen J. Human errors: a taxonomy for describing human malfunction in industrial installations. *J Occup Accidents* 1982;4:311-3.
- 4 Rasmussen J. Coping safely with complex systems. Invited paper for Annual Meeting of the American Association for Advancement of Science, Boston, February 1988.
- 5 Rosch E. Human categorization. In: Warren N, ed. *Advances in cross-cultural psychology*. New York: Halsted Press, 1975.
- 6 Leplat J, Rasmussen J. Analysis of human errors in industrial incidents and accidents for the improvement of work safety. *Accident Anal Prevent* 1984;16:77-88.
- 7 Kuhn T. *The structure of scientific revolutions*. Chicago: University of Chicago Press, 1962.
- 8 Fitzgerald PJ. Voluntary and involuntary acts. In: Guest AC, ed. *Oxford essays in jurisprudence*. Oxford: Clarendon Press; reprinted in White AR, ed. *The philosophy of action*. Oxford: Oxford University Press, 1961:120-44.
- 9 Feinberg F. Action and responsibility. In: Black M, ed. *Philosophy in America*. London: Allen & Unwin, 1965; reprinted in White AR, ed. *The philosophy of action*. Oxford: Oxford University Press, 95-120.
- 10 Rasmussen J, Vicente K. Cognitive control of human activities and errors: implications for ecological interface design. Invited paper for the International Conference on Event Perception and Action, Trieste, Italy, August 1987.
- 11 Reason J. Cognitive under-specification: its varieties and consequences. In: Baars B, ed. *The psychology of error: a window on the mind*. New York: Plenum Press, 1986.
- 12 Rasmussen J. Modelling distributed decision-making. Position paper for International Workshop on New Technology and Distributed Decision-Making, Bad Homburg, May 1988.
- 13 Rasmussen J. Safety control and risk management: topics for cross-disciplinary research and development. Invited keynote presentation for the International Conference on Preventing Major Chemical and Related Accidents, in Institution of Chemical Engineers, Publication Series 110. New York: Hemisphere Publishing Corporation, 1988.

COMMENTARY

HUMAN ERROR

In the 1970s and 1980s there was great interest among applied psychologists and systems reliability engineers in analysing accidents and "near miss" incidents in large scale systems where public safety was a primary concern. Efforts to define and develop taxonomies of human error were motivated by the meltdown at the Three Mile Island power plant near Harrisburg, PA, by the nuclear plant accident at Chernobyl in the Soviet Union, by the poison gas release at Bhopal, India, and by aviation's most deadly crash of two 747 aircraft at Tenerife in the Canary Islands. Key to these efforts were the contributions of Professor Jens Rasmussen of the Riso Energy Laboratory and the University of Copenhagen in Denmark.¹⁻³ Riso

had been assigned the task of evaluating whether Denmark should build nuclear plants (it was eventually decided not to, although neighbours Sweden and Germany had made the decision to go ahead with nuclear power). Human error rather suddenly became a fashionable topic in the human factors or “cognitive engineering” field, and an early series of international meetings was convened by John Senders⁴ of University of Toronto. Subsequently, James Reason⁵ of Manchester University in the UK and Erik Hollnagel⁶ of Linköping University in Denmark wrote well known books on human error. All the while the US medical community, while not ignoring patient safety, seemed reluctant to participate actively in such discussions of human error, the medical culture being oriented to avoiding public scrutiny of medical error for obvious reasons of exposure to litigation. Only in the last two decades has the medical community become open to taking a hard look at medical error. The anesthesiology patient safety movement led by Dr David Gaba of Stanford VA Hospital was one early push in this direction, and there were other patient safety efforts until finally the Institute of Medicine report “*To Err is Human*” pushed the door wide open.⁷

Meanwhile, the human error theoretical community began shifting gears. Rasmussen’s classic paper reproduced here characterises that shift. The nature of the shift was the realisation that human error cannot be studied independently of individuals at work in their institutions, but must be done in the context of such work. As Rasmussen puts it: “errors cannot be studied as a separate category of behavior fragments” but must be viewed within “cognitive control of behavior”.

From my own perspective, as a student of safety and human behavior with training as both a psychologist and a control engineer, “error” is simply a difference between an actual state and a desired state (“state” simply meaning a set of observable well defined variables). There is no control without measurement of state, since controls are applied to reduce that state discrepancy. Efforts to count errors as though there were a consensus on what exactly constitutes “an error” are open to considerable question. What one person considers “an error” another may regard as an acceptable deviation from standard practice because of special circumstances. Some deviations from the norm may be trivial while others may have very grave consequences. Most of what we consider errors in everyday life are compensated for before they have any noteworthy outcome. The threshold of what is error or not error, in other words, is quite arbitrarily dependent on the person making the judgment and the context of the behavior. “Unacceptable” and “acceptable” are words that need to accompany acts that are judged to be in error or not.

CHANGING THE FOCUS

Rasmussen makes the case for moving the focus away from “error” and towards the task and social context in which any “error” may occur. He does this in terms of three types of human–system analysis: (1) traditional task analysis and human reliability estimation, currently the most common way errors are studied in industry; (2) causal analysis of accidents after the fact, currently the most common form of error (or “adverse event”) analysis in medicine; and (3) design of reliable work conditions and sociotechnical systems. His plea is to do more of the latter.

Let us consider each of these in turn and relate the issues that Rasmussen brings up to corresponding issues in health care.

(1) Traditional task analysis works best for fixed procedures where discrete steps can be identified and omission of these steps counted (as errors). The errors with greatest probability and consequence are then focused upon and causal factors are sought. There are situations, however, where traditional task analysis falls apart. One such situation is when the sequence of steps is not fixed so that omission of a step cannot be

identified. Another is when commissions occur: an action is taken that is not part of the appropriate procedure. Further, task analysis is hampered when there is no obvious denominator by which to normalize—for error probability to be estimated one needs to specify that out of x “opportunities” y “error” events occurred, so that y/x is an error probability estimate. But how to define an “opportunity”? Is it a unit of time, or a procedure, or a patient, or a healthcare worker? It is obviously an arbitrary call. While health care is proceduralized to some extent, the procedures are seldom rigid; there are many reasons for variation depending on circumstances, so that the above types of difficulties arise when objective task analysis is contemplated.

(2) Causal analysis of accidents after the fact is common in medicine. However, as Rasmussen points out, the decomposition of causal sequences can be done in a variety of ways; there is no standard. Furthermore, each causal event has precursors going arbitrarily far back in time. What “stop rule” should be used? The tendency has been to focus on immediate or proximal causes such as missteps in a diagnostic or therapeutic procedure, while the more compelling cause may lie much farther upstream with caregiver training, hospital policy, or medical culture. The individual analyst or collection of persons involved, for example, in a morbidity/mortality review meeting may easily jump to conclusions about causes with which they are familiar and ignore other causal explanations which they are not equipped to recognize.

(3) Traditionally, in doing behavioral tracing to the error causation, responsibility for error has been assigned to a person who is immediate in time and space. Rasmussen emphasizes how circumstances may not be under control of that person, and that there may be systematic but unseen traps set by upstream decisions. Short sighted stop rules often do not go far enough upstream. Indeed, the whole idea of responsibility for errors needs to be reconsidered. This problem is very evident in medicine today: many persons assert that the “blame game” is not helpful in getting to the systematic causes of errors. However, others assert that the standards of personal responsibility cannot be relaxed (a position the tort lawyers depend upon to make their livings). In any case, as Rasmussen points out, the aim is to “find conditions sensitive to improvements”.

Rasmussen suggests that the design (or redesign) of systems based on analyses tied to particular paths of accidental events can be dangerous. We all know that no accident ever happened exactly the same way twice. Furthermore, the time course of events is unpredictable. For these reasons, Rasmussen argues, analysis should be carried out at a level higher than a particular accident sequence.

ADAPTATION AND LEARNING

Rasmussen makes a strong case for viewing the human worker—in the current context, the healthcare professional—as being in a continuous state of adaptation and learning. All human work is defined by task goals, task constraints, available tools, and capabilities of individual persons. But there is always some uncertainty with respect to one or more of these factors; circumstances continually change. Rasmussen discusses how necessary adaptation and learning to changing circumstances forces error to occur. This is normal.

Humans seek solutions that take least effort and healthcare workers are no exception. Indeed, no one can blame this tendency for the demands of time and energy can be overwhelming. But the result is that workers see their tasks and perceive process quality in terms of satisfying immediate goals. If the patient makes it through the immediate procedure and there are no adverse events, the tendency is to forget many of the details of what happened. This tendency was obvious to the writer in the course of an observational study of complex surgical procedures at a major Boston hospital. If no serious

adverse event occurred and the patient was not harmed, small safety compromising events were never recorded and were usually forgotten by the members of the surgical team. However, much learning can be gained from reflecting on events that did not go as planned.

Most of us tend to overlook—and not even admit to—uncertainty. Medical students are at their worst in admitting uncertainty. They learn quickly that showing confidence is part of the culture of their education and practice. They are reluctant to admit uncertainty to their mentors and they dare not admit uncertainty to their patients. Their conditioning in personal responsibility is very strong, is reinforced by the Hypocratic oath, and continues through training and right into medical practice. This cannot but reinforce the “blame game”.

Dealing with the familiar is usually preferred over dealing with the unfamiliar because errors tend to be associated with the unfamiliar. Instructions for abnormal conditions conflict with familiar methods and produce stress. Feedback (knowledge of results) is essential to learning, but feedback in patient care is often time delayed by hours or days, making learning difficult. Yet we all know that skill develops as a function of experience with a wide variety of situations, including abnormal situations. How experimental can the healthcare worker afford to be? Ethics demand conservative diagnosis and treatment, yet avoiding admitting to uncertainty and ignoring the occasional need for innovation can lead to robot-like execution of procedures in an unthinking, unreflecting manner which is surely not in the best interest of the patient.

COOPERATION AND ROLE ALLOCATION

Another factor considered in Rasmussen's paper is the problem of cooperation and role allocation. In complex technology based organizations (such as hospitals), roles continuously adjust and boundaries shift. For example, handovers are continually occurring as nurses and physicians go on break or go off shift. Residents and medical and nursing students are continually being mentored, and being asked to perform certain tasks in lieu of the attending physician or regular nurse. Communication and cooperation is therefore essential. Yet the disinterested observer can spot frequent gaps in communication—important information not being recorded or given verbally in a handover; plans and intentions not being communicated by the surgeon to other members of the operating staff or ICU or emergency room team; records not being available at the time or place they are needed. While Rasmussen does not explicitly discuss communication within healthcare systems, the reader can easily translate his ideas into salient pointers for health care.

Hospitals can be viewed as self-organizing systems where workers are simultaneously trying to satisfy high level criteria (such as improving efficiency and reducing costs) and low

level criteria (getting the current task done). Rasmussen discusses how awareness and monitoring of violation of high level criteria is difficult at the local criteria level. He points out that activities which threaten the various conditions normally belong to different levels of the organization, and knowledge on how to fix things is often at a different level from where it needs to be fixed. He asserts that catastrophic system breakdown is a common feature of self-organizing systems. To avoid such catastrophic breakdown there is a need for the boundaries of adaptation to be clear and visible. This includes transparency of role allocation—making clear who is expected to do what. Technology shapes organizations “bottom up” by imposing constraints and, much like the roles of humans, the constraints of the technology must somehow be evident to all who interact with that technology. In the hospital this includes formal postings of assignments and written instructions, but more than that it means continual checking and confirming with one another that goals, plans, roles, and knowledge of how to operate the technology are clear, and that assumptions are shared. It means more emphasis on “safety culture”.⁸

CONCLUDING REMARKS

Although theoretical and not specifically directed towards patient safety, Rasmussen's paper has a lot to say about safety culture. It asserts that past efforts to cope with human error in complex organizations where risk is a significant concern have focused too narrowly on task analysis, after-the-fact causal analysis and error attribution, and not enough on analysis of behavior of individuals and their interactions within the organization.

T B Sheridan

32 Sewall Street, Newton, MA 02465, USA;
t.sheridan@attbi.com

REFERENCES

- 1 **Rasmussen J.** Outlines of a hybrid model of the process plant operator. In: Sheridan T, Johanssen G, eds. *Monitoring behavior and supervisory control*. New York: Plenum, 1976: 371–82.
- 2 **Rasmussen J, Rouse W,** eds. *Human detection and diagnosis of system failures*. New York: Plenum, 1980.
- 3 **Rasmussen J.** Skills, rules, knowledge: signals, signs and symbols and other distinctions in human performance models. *IEEE Trans Systems Man Cybernetics* 1983;**SMC-13**:257–67.
- 4 **Senders J, Moray N.** *Human error*. Hillsdale, NJ: Erlbaum, 1991.
- 5 **Reason J.** *Human error*. New York: Cambridge University Press, 1990.
- 6 **Hollnagel E.** *Human reliability analysis: context and control*. London: Academic Press, 1993.
- 7 **Institute of Medicine.** *To err is human: building a safer health system*. Washington, DC: Institute of Medicine, 2000.
- 8 **Risk Management Foundation.** Teamwork in health care. *Forum* 2003;**23**(3).