HEWLETT PACKARD

Real-Time Clinical Decision Making

Enrico Coiera, Vanessa Tombs,
Graham Higgins, T.H. Clutton-Brock
Intelligent Networked Computing Laboratory
HP Laboratories Bristol
HPL-94-79
August, 1994

decision making, real-time, anesthesia, diagnosis, monitoring, control, planning, attention, data overload, cognitive overload

Objectives: To study the intraoperative behavior of anesthetists in order to characterize clinical decision making processes under real-time conditions.

Design: Non-participatory observation and retrospective subject interview. Clinical behavior was captured on video, and annotated by voice commentary provided by shadow anesthetist not participating in patient care. Subjects were interviewed pre and post operatively, and approximately one week later, when selected events were reviewed.

Settings: Cardiac operating theaters at 2 UK teaching hospitals.

Subjects: 12 anesthetists, of which 7 were consultants and 5 were senior registrars in training.

Results: Analysis of the data resulted in the development of a behavioral model which yielded several key observations: Anesthetic decision making is primarily focused on physiological control decisions rather than diagnosis. Most clinical activity followed structured plans. When physiological systems were not directly measurable, control strategies either chose an indirect but observable parameter as a substitute, or attempted to estimate the unobserved parameter from multiple pieces of evidence. Cognitive loading may swamp attentional resource and result in perceptual errors while monitoring, and may be caused by plan reformulation or multi-parameter estimation. Data overloading was not observed in the study. Information gathering is structured around anesthetic tasks rather than physiological systems.

Conclusions: Clinical decision making is normally described in terms of the diagnostic process, but in real-time domains like anesthesia decisions seem to be control oriented and plan based.

1 Introduction

Real-time problems require clinical decisions to be made about the moment to moment state of a patient. They span the management of acute clinical emergencies, through to the management of critically ill patients. The classic studies of medical decision making have focused on the traditional medical consultation, usually in an idealised experimental setting [1]. Here the focus is on the careful elicitation of facts through questioning and examining the patient, and the evolution of diagnostic hypotheses which drive subsequent management. In real-time situations, one seeks to make the most optimal decision given the time and resources to hand, with the goal typically being the attainment of a stable state for the patient.

Relatively little is known about real-time clinical decision processes. Yet these represent some of the most difficult tasks facing clinicians because of their time pressure and complexity. However, it is difficult to anticipate the onset of clinical emergencies, making them problematic for behavioural study. Further, they represent a heterogeneous group of events, making it difficult to control for event type. Fortunately the delivery of anaesthesia offers a similar set of decision pressures, but allows for relative control of event type and onset. Whilst the study of decision making in anaesthesia has many specific characteristics, it presents us with a model for the general study of clinical decision making under real-time constraints.

In the past, analyses of anaesthesia practice have aimed at improving anaesthetic outcome by focusing on the causes of error [2][3]. An alternate strategy is to examine routine behaviours, which constitute the bulk of clinical practice [4]. This approach has two advantages. Firstly, by focusing on the elements of normal behaviour that are responsible for sub-optimal performance, we may optimise behaviours that clinicians spend the majority of their time engaged in, rather than supporting critical but rare failures. Secondly, major errors usually arise in the context of a sequence of minor errors [5][2]. By analysing the dynamic context within which minor errors occur, we may obtain a clearer idea of how to prevent them. An important consequence may be that potentially critical errors can be prevented by optimising "normal" clinical practice. An analysis of normal behaviour should also reveal ways in which anaesthetists routinely successfully recover from incidents, since in the main, most significant events are successfully dealt with and do not result in negative outcome [5].

2 Method

In order to characterise real-time clinical decision making, observations of anaesthetists at work with follow-up interviews post-operatively were undertaken. Observations were enhanced by commentary on clinical performance obtained at the time of surgery provided by an expert observer. This commentary provided a link between the observation and interview components of the study.

Two UK sites were used for the study. Six participating clinicians with cardiac anaesthetic experience were selected from each site. Of these twelve, five were senior registrars in anaesthesia, who were able to take entire responsibility for the delivery of the anaesthetic, and seven were consultant anaesthetists.

2.1 Study process.

2.1.1 Stage 1 - Observation

The observation technique consisted of passive recording of intraoperative behaviour, with no questioning of subjects. The subject anaesthetist was the one primarily responsible for delivery of the anaesthetic. Subjects were informed that the purpose of the data collection was to assist in the design of monitoring equipment.

All monitored signals available to the subject were recorded by a secondary monitoring device which operated in parallel with the one used by subjects. Other data included a record of drugs given by the subject, as well as video and audio recordings of the subject during the operation. The visual field of view for the video recording centred on the subject and the monitor, but also included the anaesthetic machine, ventilator, drip stands, bypass pump, and the surgeon. Audio recordings were made of pre- and post-operative case assessments by the subjects. Commentary from the observing clinician recorded anaesthetic events, surgical phase, etc. The commentary was made via a close-range pick up radio microphone so that comments could be made *sotto voce*.

2.1.2 Stage 2 - Expert review

Because a typical cardiac bypass operation takes several hours to complete, it would not have been practical to review the whole operation with each subject in the follow-up interviews. Instead, two (occasionally three) clinically significant events were selected for review.

26 events were culled from the data, typically being 15-30 minutes long. No major errors resulting in significant morbidity were observed. 10 of the 26 events demonstrated good performance by the subject, as judged by the expert anaesthetist (Table 1). The remainder of the events demonstrated suboptimal performance. This

Table 1: Distribution of performance during clinical events (N=26) in the study by stage in subject training

	Good Performance	Suboptimal Performance
Senior Registrar	5	5
Consultant	5	11
Total	10	16

was primarily manifested by a patient parameter moving to clinically unacceptable levels for periods of several minutes or more (Table 2).

2.1.3 Stage 3 - Subject review

Post-observation interviews with subjects were arranged for the week following the observation sessions and lasted about one and a half hours. To cue subjects' recall,

Table 2: Primary abnormality in monitored parameter that characterised selected events (N=26)

	ABP	SaO2	ETCO2	Temperature	Cardiac Rhythm
Senior Registrar	7	1		1	1
Consultant	13	2	1		
Total	20	3	1	1	1

they were allowed access to the data set including the audio-visual record, but excluding the observing clinician's commentary and event interpretation.

Subjects was asked about the selected events. The emphasis was on building an understanding of the information used in the subject's decision making and subjects were specifically questioned about what information was used and what role the information played in the decision. They were questioned about their goals during the selected events, and the rationale for particular decisions. Questions were posed in a neutral fashion in an attempt to avoid the subjects feeling obliged to defend themselves by providing interpretations that justifying their actions.

2.1.4 Stage 4 - Event analysis

The objective of the event analysis was to characterise clinical decisions. Close attention was paid to identifying what information was available to the clinician, what information could be positively identified as having been used and what information could have been used. Analysis involved the detailed study of the observing clinician's commentary, the electronic data and the transcript of the interviews. The audio-visual records were used to resolve any ambiguities or contradictions appearing in the transcripts.

In order to identify decisions we produced a chronological map detailing subject actions for each event. Each event map was next interpreted in relation to the inferred goals and reasoning of the subject, as revealed through the interview transcripts. Finally, an analysis was made of the subject's relative performance over the event period. In particular, the focus was on deciding whether, given the information available, alternative decisions could have been made which might have resulted in a more favourable outcome. An example summary of a typical event analysis is presented in the appendix.

Finally, the event maps were pooled, and common behaviour patterns and decision processes were abstracted. These abstract descriptions were then assembled into a general model of the real-time decision making processes observed.

2.2 Methodological Basis

There are a number of alternative methods available for the study of decision-making processes, and these can be roughly grouped as follows:

Interviews: Asking subjects to introspect on their own decision processes suffers from a number of methodological difficulties. The first is attitudinal bias - people's actions differ from their verbalised reports [6]. Secondly, these methods require self-reporting. When asked why they made a certain judgement people are capable of providing apparently plausible reports of their mental events [7]. However this is a contentious investigative method because there is no way of independently checking the validity of self reports [8][9].

Non-participatory observation: A relatively more objective approach is to observe individuals in the field. There are many approaches that one could take, including techniques derived from ethnography [10]. Here one tries to understand the actions of individuals through detailed observations of them carrying out tasks.

Formal psychological experiments: At the most rigorous level, hypotheses about specific aspects of clinical behaviour, may be tested by controlled studies. Some of the classic work in clinical decision making has been of this type [11].

The need to characterise medical decision processes as much as possible in the way they really occur, as opposed to the way they *should* occur strongly favours initially pursuing non-participatory methods in the field. These are sufficiently formal to allow robust statements to be made about results, and have the advantage of being grounded in the realities of the clinical workplace. Since formal psychological studies are conducted under laboratory conditions, they factor out the interruptions, short cuts and pressures of the clinical workplace. Formal studies can thus be best used to test specific hypotheses that arise out of non-participatory studies.

In the absence of strong hypotheses about the nature of real-time decision processes, this study combined non-participatory observation with interviews. After pilot studies, it was clear that without both the observing clinician commentary and subject interviews, the large and complex data set would be opaque to detailed analysis. The rich data record validated subject recollections, and their reports on decision processes were checked against our expert clinician's assessment. Thus the data and interpretations collected in the first two stages acted as controls on our subject's verbal reports in the third stage.

In the absence of clear recall, the richness of the recorded data allowed subjects to reconstruct a hypothetical analogue of events when giving answers to questions. In this situation we were likely to get reports which contained clinicians' *a priori* theories [9] about their likely responses as well as partial recall of their actual responses.

The approach described in stage 4 of the study is a permutation of a qualitative research technique known as 'Grounded Theory' ([12], [13]) which emphasises the emergence of concepts from the data rather than the imposition of an a priori theory by the researcher.

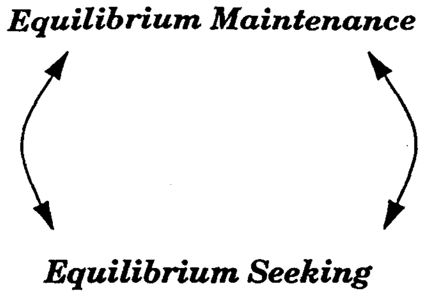

Figure 1. Study subjects exhibited two major control behaviours depending upon whether the patient was in a stable state, or required active intervention following the development of an instability.

3 Results

3.1 Modelling Real-Time Decision Making

The primary anaesthetic task is to control patient state. A patient is placed into a state in which surgery can be carried out safely, and then retrieved from that state. We identified two basic control behaviours that appeared across the study events (Figure 1.):

Equilibrium maintenance: This behaviour is observed when a clinician is comfortable with the patient's state, and carries out actions that are intended to maintain that state. Typically adjustments to control are fine grained e.g. minor variations in the percent of anaesthetic gas delivered.

Equilibrium Seeking: When a disturbance of patient state occurs, the clinician is forced to change control strategy. The clinician may undertake diagnostic strategies to obtain an explanation for the change, and then institute appropriate therapy to push the patient into a new stable state.

Switching between maintenance and seeking: A trigger causes the switch between these two behaviour modes. The clinician may observe a clinical event, or discover that a significant change has occurred in patient state. Both result in a need to understand the implications of that change. Delays in switching between control modes will delay initiation of appropriate therapy, and may prolong inappropriate therapy. In some of the study events we observed inappropriate equilibrium maintenance behaviour becoming quite exaggerated, as the clinician tried to maintain control - drugs were given repeatedly in increasing doses in the expectation that the drug was appropriate but simply that more of it was needed. Finally, at some stage during the equilibrium seeking behaviour, the clinician must decide that the situation is under control and that a switch back to maintenance is possible.

3.2 Characterising Behaviours

Equilibrium maintenance and seeking behaviours are respectively *strategic* and *reactive*. Strategic behaviours are responsible for selecting and implementing treatment plans, while reactive behaviours cope with unexpected events. Clinical events can similarly can be thought of as *intended* or *unintended*. Intended events are part of a predefined strategy, and unintended ones must be reacted to.

These behaviours can be further broken down into task types. While control is the major goal, there are actually three generic tasks that are repeatedly performed by clinicians - *monitoring*, *diagnosis*, and *control*. Control describes the institution of a therapeutic loop, from administration of a therapy to the modification of subsequent doses based upon observed patient response. Diagnosis is a directed action to isolate the cause of unexplained observations. Monitoring can be seen as a link between diagnosis and control - the clinician observes a variety of parameters, and institutes diagnosis and control behaviours based upon these observations. Anaesthetists also engage in *administrative* tasks, which include record keeping, manipulation or setting up equipment, and drawing up drugs. Table 3 summarises the distribution of clinical behaviours observed in the events selected for the study.

Table 3: Primary clinical task that characterised selected events (N= 26). Some events were characterised by multiple tasks.

	Monitor	Diagnosis	Control	Planning	Total Events
Senior Registrar	4	0	3	3	9
Consultant	5	0	8	6	17
Totals	9	0	11	9	26

The notion of strategic intent, and generic task type can be used to categorise individual elements in the observed behaviour of clinicians. Six possible combinations arise:

	strategic	reactive
diagnosis	SD	RD
monitoring	SM	RM
control	SC	RC

3.2.1 Contrasting the Behavioural Categories

3.2.1.1 Reactive vs Strategic Monitoring

These two behaviours represent opposite ends of the spectrum in vigilance. Strategic monitoring is focused and goal oriented, while reactive monitoring is unfocussed and event driven.

Strategic monitoring relies on the clinician having a defined set of parameters to look at which directly assist in the completion of a specific task. For example, one may check the ECG to specifically look for arrhythmias after having started treatment for rhythm disturbance. Associated with a predefined set of observations will probably come some notion of how frequently these observations need to be made. Some clinical contexts will trigger strategic observations - for example checking oxygen saturation after bypass for the development of pulmonary oedema.

Reactive monitoring describes an event driven examination of a patient's clinical state. For example, the incidental detection of an unforeseen abnormality on a monitor screen, or having a set of laboratory results come in at an unplanned time.

3.2.1.2 Reactive vs Strategic Diagnosis

Studies of clinicians undertaking medical consultations reveal general strategies for the collection of information to formulate and refine a differential diagnosis [9].

In real-time situations, a reactive diagnosis is a response to an unplanned event. A diagnosis is often drawn from a very short list of alternatives, where each alternative is associated with a control strategy. There is little value apparently in selecting an over-refined diagnosis that does not confer any additional advantage when instituting control.

A strategic diagnosis on the other hand, represents a *preplanned* choice between predefined alternatives. For example, when a patient comes off cardiac bypass, the anaesthetist should check to see that the lungs have re-inflated. Pulmonary oedema may also develop at this time if cardiac performance is poor. Both conditions will manifest as a decrease in the arterial oxygen saturation. Aware of both possibilities, the clinician knows exactly where to look to isolate the cause of the low saturation. Thus the differential diagnosis is strategically precompiled for specific types of operative stage, and particular types of patient.

3.2.1.3 Reactive vs Strategic Control

With reactive control, the clinician is faced with the need to *regain* control of a disturbed physiology. The cause of the disturbance may be obvious from the context, or have been determined by a reactive diagnostic step.

If the clinician has not previously experienced the disturbance in a particular context, there may be several possible treatment options that need to be considered. If it is not clear which treatment will work, a particular therapy may be trialled, and perhaps discarded in favour of another. A sequence of reactive diagnosis and control actions may thus be repeated until a sufficiently accurate diagnosis has been made to allow the institution of the most appropriate control strategy.

The situation is less uncertain when commencing a strategic control action, such as at the induction phase of an anaesthetic. The clinician has preplanned manoeuvres to execute, which will shift the patient's physiology in a specified direction. The medications and dosages are already largely selected prior to commencing the action. Refinements in the doses to obtain optimal control are usually within bounds that are predefined by the clinician.

3.3 Planning Behaviours

A variety of complex behaviours were observed across the studied events. The combination of event specific circumstances with a generic set of behaviours seems to account for much of the variety observed. Faced with a set of problems and a number of ways to deal with them, clinicians seemed to assemble behavioural sequences to accomplish their goals. It was found that such behaviours could be described by assembling sequences of the behavioural components already identified. For example, consider the general situation in which an unheralded arrhythmia develops, and is treated by a clinician. The following sequence of behaviours is likely:

$$RM \rightarrow RD \rightarrow RC \rightarrow (SM;SC)$$

RM represents the clinician detecting the unexpected arrhythmia from a chance observation. RD represents the clinician focusing down in an attempt to diagnose the arrhythmia type. The next component RC represents an attempt to regain control of cardiac rhythm through administration of appropriate therapy. Having re-established some control, the final component of the behaviour is instituted - (SM;SC). This represents a switch from reactive to strategic mode, with the situation now under control and a plan for control maintenance having been formulated. The maintenance action has two concurrently active components - regular planned observations of the ECG (SM), and the institution of maintenance therapy (SC).

Each transition in such a sequence of behaviours may be associated with a satisfaction condition or trigger - one moves on to the next action in the sequence once that condition has been met. The condition may be a particular set of clinical observations, or the completion of a physical task like the administration of a medication.

3.3.1 Goals and Plans

Central to understanding the way in which behavioural sequences are constructed is the notion of a *plan*, where a plan corresponds to a prespecified sequence of tasks.

We begin by noting that the clinician can be characterised as having a set of beliefs, goals, and intentions. *Beliefs* correspond to everything the clinician knows - knowledge of medical practice, the patient and the clinical context. *Goals* are states that the clinician wishes to attain or maintain, such as "maintain coronary perfusion" or "maintain cerebral oxygenation". *Intentions* represent goals that the clinician actually chooses to commit to, in preference to others [14].

From an examination of the study events it is clear that many goals can be committed to concurrently - the clinician seeks to satisfy all of them over time. To achieve this, there is associated with each goal, a plan of action. These plans will often be composed of actions which are instances of diagnosis, monitoring or control. Thus there may be multiple plans of action in effect at any one time, each one aimed at achieving a different goal.

Since most people carry out a single action at a time, one can think of the clinician as maintaining a stack of actions to perform next, with actions from different plans being interleaved. For example, a clinician may wish to achieve two goals G_1 and G_2. Associated with each of these goals is a plan:

$$G_1: P_1^a \rightarrow P_1^b \rightarrow P_1^c$$

$$G_2: P_2^a \rightarrow P_2^b \rightarrow P_2^c$$

The plans are interleaved into a sequence of actions that allow both plans to progress over time e.g.

$$P_1^a \rightarrow P_1^b \rightarrow P_2^a \rightarrow P_2^b \rightarrow P_1^c \rightarrow P_2^c$$

3.3.2 Scheduling Actions

While plans may be predefined, selecting which action to do next is not. This is a dynamic activity because new goals may need to be satisfied in response to unexpected events, or because an existing plan cannot be executed because a resource is unavailable. For example, one may plan to give a medication, but the drug is temporarily unavailable. While waiting for it to arrive in theatre, other actions can profitably be performed to satisfy other goals. Thus while strategic behaviours can be formulated in advance, their enaction is strongly influenced by events imposed by the external world.

Current plans may be interrupted when an event establishes a new goal that demands attention. High priority actions are placed at the head of the queue, and only when they are satisfied will previous plans be returned to. For example, a clinician may be engaged in determining the status of a patient's fluid balance when the patient has a cardiac arrest. The fluid balance plan is immediately suspended, and a new goal of returning normal cardiac function is established. A plan is put into place to achieve it, and enacted. Only when this new goal has been satisfied may the clinician return to the task of assessing fluid balance.

It should be noted that an interrupted plan may be returned to with the patient in a different state to that prior to the interruption. In the above example, the patient may have had a significant amount of IV fluid during the resuscitation effort following asystole. Thus, after a plan interrupt, the world may have changed, and the motivation for continuing with previous plans may need to be updated. Failure to do so may result in a now inappropriate plan being enacted. Another consequence of an interrupt is that a clinician may *forget* the previous task and so leave it uncompleted.

3.3.3 Goal interactions

While a clinician may have several goals, he may commit to only some of them. This is because goals may conflict and a decision needs to be made about which is the most important one. Sometimes a trade-off is possible between such goals. For example, if a patient is bleeding during cardiac surgery, one goal is to keep the blood pressure reasonably low to minimise blood loss. The conflicting goal might be the need to maintain blood pressure to keep coronary perfusion adequate. The actual blood pressure selected will depend on the trade-off the clinician considers acceptable between these two goals. Errors may occur in the selection of the appropriate trade-off point in such situations.

In contrast, some actions may satisfy more than one goal simultaneously. Administering the anaesthetic agent Enflurane, for example, can satisfy the goal of maintaining depth of anaesthesia as well as lowering the blood pressure (BP), as it has both these effects upon the patient.

3.3.4 Side-Effects

An additional complexity arises when actions do not just result in the satisfaction of a goal, but also have side-effects. For example, glyceryl trinitrate (GTN) is given for its effect on coronary perfusion, but as a side-effect drops the systemic BP. A clinician may plan to deal with that side-effect by reducing the inhalational anaesthetic agent and perhaps increase the rate of intravenous fluid administration. Side-effects, if not consciously considered and planned for, may lead to control difficulties.

Intended actions then may have unintended, but expected side-effects. The sense here is that the clinician knows that a side-effect will probably happen, but would rather that it didn't, and indeed would be happy if it did not. There is no commitment to the side-effect happening i.e. the clinician would not seek other means of causing it to occur if it did not appear. In contrast, if the primary and intended result of an action does not eventuate, the clinician will seek an alternative way of achieving it.

3.3.5 Plan Structure

The notion of a plan is familiar to anaesthetists, who talk of formulating an *anaesthetic plan* for each case. This plan is formed by making a pre-operative assessment of the patient's condition, considering the type of surgery proposed, the surgeon who will perform the operation, and so on. Plans can be quite detailed, with predetermined actions planned across several stages [15]. Anaesthetists often make the analogy that the plan structure of an anaesthetic is like flying a plane. Take-off, flight and landing are similar to the induction of anaesthesia, its maintenance, and reversal. Coronary bypass surgery has even more clearly defined stages. Subject clinicians were readily able to articulate details of plans if asked to. For example, heparin administration will always be planned for the period immediately prior to going on to bypass.

Plans are only as detailed as practicable, with fine details being fleshed out as the plan is enacted. Associated with each plan is a set of expectations based upon past experience that determines the likely starting doses of drugs, and the likely response of patients to therapy. Hence one may plan to control pre-bypass instabilities in BP with a GTN infusion, but the dose used, and any changes to that dosage will depend on the clinician reacting to the actual blood pressure. Thus plan reactivity does not just manifest itself as the creation of new goals, but also in the fine tuning of prespecified plans. This is necessary because it is not possible to plan for the way a patient will respond to treatment (although with past experience one can build up an approximate model and base the initial dose upon that experience).

The interaction of predefined plans with clinical events may be characterised as a combination of macro and micro planning strategies. Macroplanning is responsible for creating global strategies based upon skeletal plans. Microplanning involves the instantiation of skeletal elements with the particular conditions existing at the time. It may involve refining detail, such as choosing doses based upon patient response. It can be more complex, choosing among several more specific subplans at a choice point. Thus a clinician may expect that a patient's BP will dip during atrial cannulation prior to bypass. Depending upon whether the BP returns quickly or stays depressed, the clinician will enact different plans to deal with the situation.

3.3.6 Reactive Planning

Reactive plans, like strategic plans are composed of predefined elements and exist as a set of contingencies. For example, plans exist for dealing with asystole, arrythmias, blood loss etc. Limitations on the time available to deal with these events means that clinicians have to have clear plans ready to deal with them should they arise.

If events become sufficiently complex, or are poorly understood there may be no alternative but to modify existing plans or form new ones. Preformed plans exist because they are a cognitively economic way of working. Novel plan formulation may place large cognitive demands on a clinician. Subjects demonstrated that under only minimally novel circumstances demanding the plan reformulation, potentially significant errors could occur (see Appendix).

3.4 Monitoring and control strategies

For a control loop to be set up, measurements must taken of the system being controlled, and coupled to control actions. Many clinical control problems are difficult because direct measurements of the system being controlled are not possible. For example, there is no routine measurement of depth of anaesthesia available at present, yet it is routine to estimate anaesthesia depth and try to control it. Coronary perfusion and myocardial oxygenation are estimated through ST segment changes on the ECG.

3.4.1 Single parameter estimates

The simplest strategy observed for dealing with unmeasurable states was to identify a parameter that was indirectly related to the system of interest, and to treat it as if it were a direct measure. Depth of anaesthesia was estimated by all subjects by gauging the rise in arterial blood pressure in response to surgical stimulus. This strategy is problematic for a number of reasons. Since blood pressure is an indirect measure of anaesthesia depth, the relationship is complex. For blood pressure to meaningfully rise in response to pain, one must assume a normal pressure response. Further, blood pressure may rise for other reasons.

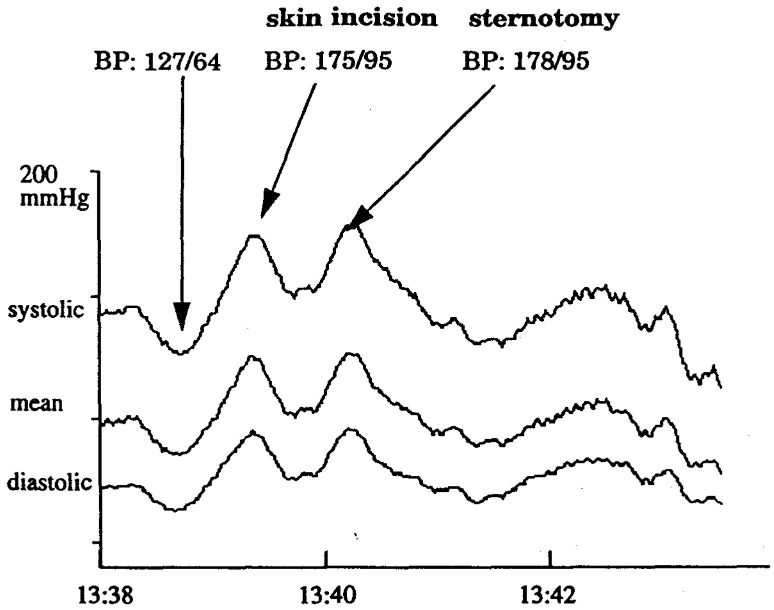

Figure 2. Arterial pressure rises due to lightness of anaesthesia were often detectable prior to sternotomy.

The main strategy used by subjects to cope with the many possible influences on pressure was to schedule monitoring for a period when it was most likely that a rise was directly related to surgical stimulation. Most subjects chose to look at pressure rises during sternotomy, believing that this was the best indicator of lightness of anaesthesia.

Invariably this meant that subjects missed earlier data that would have indicated that anaesthesia was light. In most cases the pressure response at the time of skin incision was sufficient to indicate that more analgesia was necessary (Figure 2.). Indeed, it was often the case that a significant pressure response was evident even earlier, when the skin was prepared for surgery. When the data was reviewed with each subject postoperatively, they were surprised to see the earlier pressure responses on skin preparation and incision. Thus while selective monitoring of indirect measurements at periods of peak influence seems to be an efficient way of dealing with multiple influences, it does not guarantee a timely response.

A similar set of behaviours were seen for control of fluid balance. Here the indirect measure chosen to substitute for intravascular volume was usually central venous pressure (CVP). As an intermediate variable, CVP also suffers from not being a unique indicator since it is also influenced greatly vessel capacitance. In addition to reflecting fluid balance, it can for example be affected by vasodilators.

The use of a fluid challenge to estimate whether a patient is hypovolaemic was observed in a number of cases, and this can be seen as another attempt at maximising the likelihood that the intermediate measure being observed is actually reflecting the system of interest. By giving a fluid bolus, the likelihood is that an immediate change in blood pressure is causally related to the bolus, and thus reflects volume status. In a similar way, sternotomy was used by subjects as 'pain bolus' to test their control of depth of anaesthesia.

3.4.2 Multiple parameter estimates

Some subjects chose to use multiple indirect measurements, and combined the evidence to estimate the state of unobserved systems. For example, when estimating fluid balance, the heart rate, CVP and blood pressures and the temporal relationship of changes in each of these could be examined. When estimating depth of anaesthesia, some subjects factored in estimates of drug half lives in an effort to direct their monitoring strategies - they scheduled observations around the period they expected drug effects to be wearing off.

One interesting error typical of the multiparameter strategy was a failure to factor out all confounding influences. For example the presence of a rising heart rate in the presence of a falling arterial pressure may signify hypovolaemia. Two subjects missed an evolving hypovolaemia because they assessed the rising heart rate to be within acceptable limits. In both cases they failed to consider that the patient was being treated with Beta Blockers which obtunded the heart rate response. So, although the rise in heart rate was small, in both cases it represented a significant increase.

3.4.3 Single vs Multiple parameter strategies

While the data from this study are insufficient to allow firm conclusions to be drawn, it seemed that subjects who employed a multiparameter strategy outperformed subjects who chose single parameter strategies. Difficulties arise here because of the low number of subjects, and the lack of precise performance estimates.

As might be expected, multiparameter strategies are much more complex, and require a significantly more sophisticated understanding of physiology. They offer apparently more accurate estimates of the behaviour of unobservable systems, but at the cost of increased effort on the part of the clinician. Being simpler, single parameter strategies demand less effort from the observer, but probably are less accurate than the multiparameter approach.

Although expert-novice differences in performance were not examined in this study, one might speculate that novices would favour single parameter estimates, and that experts would favour multiple parameter methods. This is because although they are probably more accurate, multiple parameter methods require a greater degree of confidence and knowledge on the part of the observer. Inexperienced observers may feel that the trade-off between complexity and accuracy that they make when selecting a strategy favours the single - while they may not perform as well as a consultant with a single parameter strategy, they would be worse if they chose the multi-parameter strategy.

4 Discussion

The model of real-time decision making processes developed for anaesthesia has a number of interesting implications, both for the way clinicians should approach decision making, and also for the way decision support systems like patient monitors should be developed in the future.

4.1 Attentional Load

It is often suggested that clinicians in real-time situations are prone to data overload. Data overload arises when more data is physically displayed than can reasonably be comprehended by an individual [16]. A typical nuclear power plant supervisor for example, may have up to 3000 controls and displays to deal with [17]. Typically, clinicians in our study had on the order of 40 parameters available, and choose to display about a quarter of these. Thus several orders of magnitude difference exist in the quantity of data requiring display between industrial process control and anaesthesia.

Current understanding of human sampling behaviour suggests that in principle, anaesthetists should be able to handle this amount of data reasonably well. Since sampling behaviour is influenced by the observer's models of the dynamic and stochastic properties of signals, robust mechanisms are available to deal with multiple data streams [18] [19].

However, several of subjects performed suboptimally because they failed to detect significant data events, *despite the fact that they apparently were looking directly at the monitor* for extended periods. This may be due to attentional fixation on a different subset of the displayed data. One may conjecture that the subjects focused exclusively on a few of the displayed parameters, and ignored other information. Were they subject to genuine data overload conditions, one might explain this by suggesting a fundamental perceptual limit had been reached. For example, it may not have been physically possible to examine all the relevant parameters within a given time. However, no such physical limits were apparent during the study.

A more likely explanation is the notion of attentional loading, where attention is equated with the amount of cognitive resource available to carry out a particular task [20]. As an individual scans a set of signals, several activities are occurring in parallel. Each signal is filtered based upon its perceived relevance - individuals only look at those things that are regarded as interesting. These signals are then interpreted. For example, the numeric display of BP is identified as *being* the BP and then a meaning is attached to its value.

There is a trade-off that occurs as reasoning and perception compete for attention. Allocating attention to one means that less is available for the other. The attentional fixation observed may be explained by hypothesising that these subjects devoted most of their resource to the process of reasoning, leaving little to successfully filter signals. The end result is that the filtering is particularly savage, pruning away all but those signals that are directly relevant to the task at hand. There is no attention left to detect events slightly outside of the narrowed task focus.

One source of attentional loading is plan reformulation - having to change an existing plan can consume large amounts of attention. Signal interpretation for control tasks, in particular multiparameter strategies for estimating unobservable variables, may be another source. Because the appropriate data are not available, we can say that anaesthetists are *data underloaded* for some tasks. The cost is processing overload.

Another hypothesis that may explain the apparent task fixation by some clinicians may be that they are reluctant to abandon the mental models they have constructed for one task to complete another. This is because they may be carrying a number of intermediate computations that cannot be easily committed to memory. Thus while the nature of anaesthetic care demands that clinicians have multiple tasks active at one time, the high cost of setting up a model to deal with one task may make it difficult for that task to be temporarily suspended, since the clinician will "lose" intermediate calculations.

Thus, clinicians may miss significant events in the signals displayed to them because of high cognitive processing load, rather than the quantity of data they need to process. This load may either steal resources from perceptual filtering, or produce a reluctance to unload complex mental models. In the case study presented in the appendix, the study subject was overloaded both by plan reformulation and variable estimation from multiple parameters, resulting in task fixation and aggressive perceptual filtering.

4.2 Supporting Clinical Decisions

4.2.1 Diagnosis and Control

A common characteristic of proposals for anaesthetic decision support is that they focus on the need for diagnostic support for clinicians [21]. The assumption is that diagnosis is difficult. An examination of the behaviours of the subjects in our study does not support the need for diagnostic assistance. Most, if not all, of the diagnostic process has been completed by the time a patient presents for anaesthesia. There may be some requirement for diagnosis to take place in the pre-operative evaluation. Having made such an initial assessment, usually in the light of the patient's existing diagnosis and reasons for coming to surgery, the requirement for intraoperative diagnosis is minimal. Such diagnoses, when they are made, are taken from a reduced set e.g. cardiac arrhythmias, anaphylaxis. Some exceptionally rare conditions may not be as well handled - malignant hyperthermia is often cited as an example.

No cases were seen in which an error occurred because a clinician was unable to diagnose a clinical state. More importantly, when diagnosis occurred, it was usually of a simple form posing no great problems for the clinician. Usually the diagnostic behaviour we witnessed was shallow, and did not involve complex reasoning about physiological processes, but relied upon fairly simple rules. In contrast, quite complex inferences were often needed to manage control loops. For example, making an estimate of cardiac performance requires the clinician to estimate fluid balance, myocardial contractility, cardiac perfusion, vascular resistance and so on.

The reason diagnosis has been traditionally seen as a key task was possibly motivated by early research in decision analysis which suggested that human judgements were prone to decision biases [13], and the corollary which stated that such flawed judgements could be normalised by relying on sounder, more formally based diagnostic systems. There is now growing evidence that questions such an assessment. Based upon studies of individuals in the field rather than in controlled laboratory situations, evidence now suggests that the primary effort for decision makers is not at the moment of choice, but rather in *situation assessment* [22]. In other words, establishing a clear picture of the state of the world, and clarifying their goals and assumptions, prior to attempting to make a diagnosis.

If it is the case that situation assessment is the key bottleneck in decision making, this has implications for the design of decision support tools for monitoring and control. Systems which assist clinicians in making an assessment of monitored data may be of more utility than systems that attempt to manufacture a diagnosis. Supporting diagnosis of complex or rare conditions may be useful, but does not address the most cognitively demanding aspects of anaesthetic management. Supporting complex inferences needed to perform some control tasks may be more appropriate.

4.2.2 Planning

The behavioural model that has arisen from our analysis suggests that real-time decision making is a highly structured activity. Much of the observed variability in behaviour results from clinicians scheduling actions based upon shifting resource and task priority decisions.

While some goals clearly are of obvious priority and deserve to interrupt current actions (like attending to an asystole), the prioritisation and scheduling of actions from different goals may not always be obvious. Errors in deciding which task to do next may be a source of clinical error. There were clear examples of suboptimal clinical behaviour arising from difficulties in plan selection and following. Supporting the procedural aspects of clinical behaviour, particularly through treatment protocols is both feasible, and potentially beneficial. Further, since planned actions seem to constitute the bulk of observed clinical activity, decision support can have a far more pervasive role than previously expected.

4.2.3 Data Presentation

At present, monitoring systems present patient data in a relatively unstructured way. Many decision support system proposals suggest that grouping information based upon physiological systems would be beneficial.

We observed many anaesthetic tasks which cut across physiological groupings. For example, maintaining blood pressure requires inferences about the state of analgesia, volume loading, vasoactivity and myocardial contractility. Grouping data by physiological system could require a clinician to look at several different systems in the course of a single task. Some tasks, such as maintaining oxygenation, could be satisfied by a single set of data grouped around the respiratory system. Thus it is not that physiological systems are never an appropriate view of clinical data, but rather that they are appropriate when they coincide with a given clinical task. If we group information to reduce attentional demands, then the more appropriate grouping is one based on the task at hand.

5 Conclusion

The study of clinical decision making processes has fundamental implications for the way in which clinicians approach clinical practice, for clinical training, and for the design of computer based decisions support systems.

While the results of the present study present a clear model for real-time behaviour, significant work needs to be done, both in validating the model, and in expanding it. In particular there is a need for a more detailed exploration of the role of attentional loading on clinical error, and on novice-expert differences in control strategy.

References

[1] A. S. Elstein, L. S. Shulman, S. A. Sprafka, *Medical Problem Solving - An analysis of Clinical Reasoning*, Harvard University Press, Cambridge Ma., (1978).

[2] J. B. Cooper, C. D. Long, R. S. Newbower, Human Error in Anesthesia Management, in B. I. Grundy, J. S. Gravenstein et al. (eds), Quality of Care in Anesthesia, (1982) 114-130.

[3] J.B. Cooper, R.S. Newbower, R.J. Kitz, An Analysis of Major Errors and Equipment Failures in Anesthesia Management: Considerations for Prevention and Detection, *Anesthesiology*, 60, 1, (1984), 43-42.

[4] J. S. McDonald, R.R. Dzwonczyk, A time and motion study of the Anaesthetist's intraoperative time, *British Journal of Anaesth*esia, 61, (1988), 738-742.

[5] D. M. Gaba, M. Maxwell, A. DeAnda, Anesthetic Mishaps: Breaking the Chain of Accident Evolution, *Anesthesiology*, 66, 5, (1987), 670-676.

[6] A. W. Wicker, Attitudes v. Actions: the Relationship of Verbal and Overt Responses to Attitude Objects, in *Attitudes*, N. Warren, M. Jahoda (eds), 2nd. Ed. Penguin Books Ltd., Middlesex, (1976).

[7] R.T. Kellog. When can we introspect accurately about mental processes? *Memory and Cognition*, 10:141–144, 1982.

[8] K.A. Ericsson and H.A. Simon. Verbal reports as data. *Psychological Review*, 87:215–251, 1980.

[9] R.E. Nisbett and T. D. Wilson. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84:231–259, 1977.

[10] D. Fafchamps, C.Y. Young, and P.C. Tang. Modelling work practices: Input to the design of a physician's workstation. In *Proceedings of the 15th. Annual Symposium on Computer Applications in Medical Care*, Washington, D.C., 1991.

[11] D. Kahneman, P. Slovic, A. Tversky (eds), Judgment under uncertainty: heuristics and biases, Cambridge; New York, Cambridge University Press, (1982).

[12] N.F. Pidgeon, B.A. Turner, and D.I. Blockley. The use of grounded theory for conceptual analysis in knowledge elicitation. *International Journal of Man-Machine Systems*, 35:151–173, 1991.

[13] K.L. Henwood and N.F. Pidgeon. Qualitative research and psychological theorizing. *British Journal of Psychology*, 83:97–111, 1992.

[14] P. Cohen, H. Levesque, Intention is Choice with Commitment, *Artificial Intelligence,* 42, (1990), 213-261.

[15] L. L. Bready, R. B. Smith, *Decision Making in Anesthesiology*, B.C. Decker Inc, Toronto, (1987).

[16] J. Lovesey. The instrument explosion - a study of aircraft cockpit instruments. *Applied Ergonomics*, 8,(1), 1977.

[17] C. D. Wickens, *Engineering Psychology and Human Performance* (2nd Ed), Harper Collins, New York, (1992).

[18] H. G. M. Bohnen, M.A.M. Leermakers, Sampling Behaviour in a Four-Instrument Monitoring Task, *IEEE Trans. Sys. Man Cyb.*, 21,4,893-897, (1991).

[19] J. D. Gould, A. Schaffer, The Effets of Divided Attention on Visual Monitoring of Multi-Channel Displays, *Human Factors*, 9 (3), 191-202, (1967).

[20] C. D. Wickens, Processing Resources in Attention, in R. Parasuraman, D. R. Davies (eds), *Varieties of Attention*, Academic Press, London, (1984).

[21] E. Coiera, Editorial: Intelligent Monitoring and Control of Dynamic Physiological Systems, *Artificial Intelligence in Medicine*, 5, (1993), 1-8.

[22] G. A. Klein, R. Calderwood, Decision models: some lessons from the field, *IEEE Trans. Sys. Man Cyb*, 21, 5, (1991), 1018-1026.

Appendix - An Example Event - Missed Hypoxia.

The clinician

A consultant anaesthetist with substantial experience of cardiovascular surgery.

The patient

A 47 year old gentleman who had a large myocardial infarction several months prior to surgery. Following the infarction he had recurrent episodes of left ventricular failure and subsequently developed a left ventricular aneurism. The patient was admitted to have the aneurism excised.

The problems noted by the subject anaesthetist prior to surgery were recurrent bouts of ventricular failure with episodes of pulmonary oedema. The patient's relatively poor left ventricular contractility with dyskinesia from the aneurism was considered a problem and the subject suggested that the patient could require fairly intensive haemodynamic support when coming off cardiac bypass.

The event

The event demonstrates a gradually declining SaO_2 level post bypass, probably caused by pulmonary oedema (Figure 3.). The downward trend started off with an SaO_2 at 100 immediately after bypass, and dropped down to 87. The clinician did not detect the trend until he took a routine blood gas sample 30 minutes after the drop had commenced. The alarm capabilities of the monitoring device had been suspended (many clinicians perceive that patient monitors have a high rate of false alarms). Consequently, no alarm sounded. However the downward trend would have been visible for quite some time before the SaO2 dropped into the alarm region.

Chronology of clinical actions

The subject normally had his patients coming off cardiac bypass with an infusion of an inotropic agent to support cardiac contractility. In this case, the surgeon preferred instead to have the patient on an infusion of GTN. Thus, counter to the

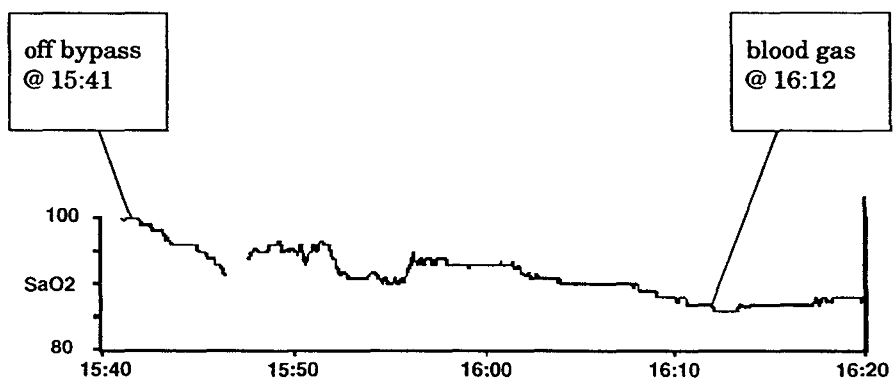

Figure 3. A decline in patient's SaO2 level in the immediate post-bypass phase remained undetected for over 40 minutes, despite the clinician repeatedly examining the patient monitor screen.

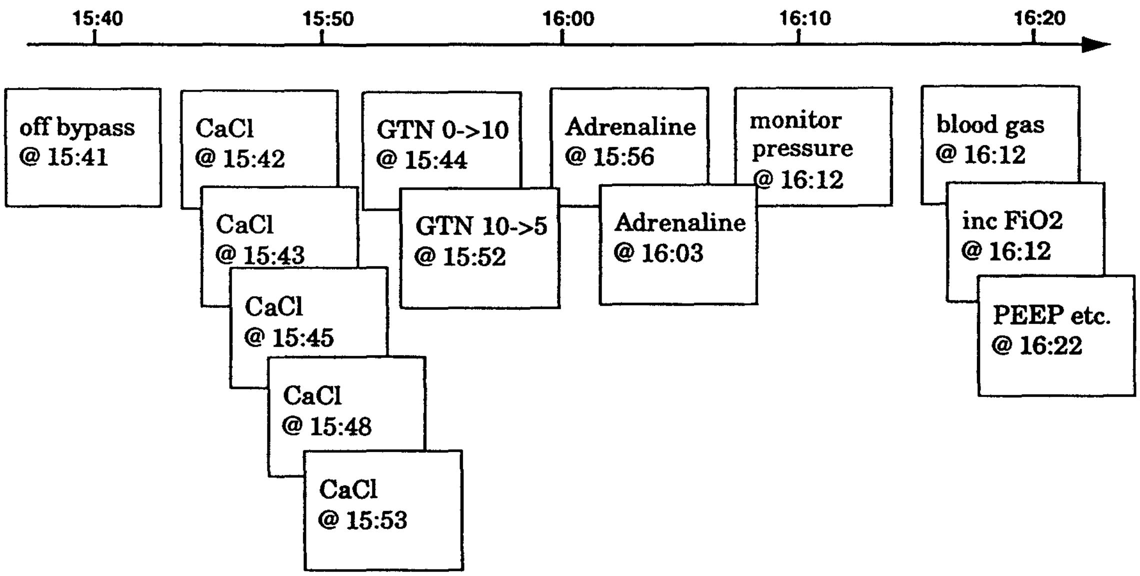

Figure 4. Clinical actions taken during the period of evolving hypoxia

subject's normal plan, an intravenous infusion of GTN was running at the end of the cardiac bypass.

Concerned about the patient's cardiac performance, and in the absence of his preferred strategy, the subject gave the patient repeated doses of Calcium Chloride (CaCl) (Figure 4.) This was given for its inotropic action. The clinician had a rule that by 20 minutes after bypass the various pressures should have returned to normal range. Since this had not happened, he finally felt justified to depart from the surgeon's practice and commence treatment with standard inotropes, and administered adrenaline at 15:56.

At 16:12, he took a routine blood gas. In so doing noted the low SaO_2 value, and immediately increased the patient's inspired oxygen (FiO_2). He commenced hand ventilation, and increasing the end-expiratory pressure on the ventilator (PEEP). The routine blood gas value taken at 16:12 showed an arterial PO_2 of 7.2.

An analysis

There were several possible explanations for the prolonged drop in saturation. Firstly, the lungs may have been partially collapsed after bypass. In view of the patient's previous history, and the high filling pressures, pulmonary oedema was a more likely explanation. The addition of PEEP and increase in FiO2 were commenced in response to this possibility.

Much of the clinical activity throughout the drop in saturation was seen to be reactive (Figure 4.). The subject was forced to deal with a combination of events for

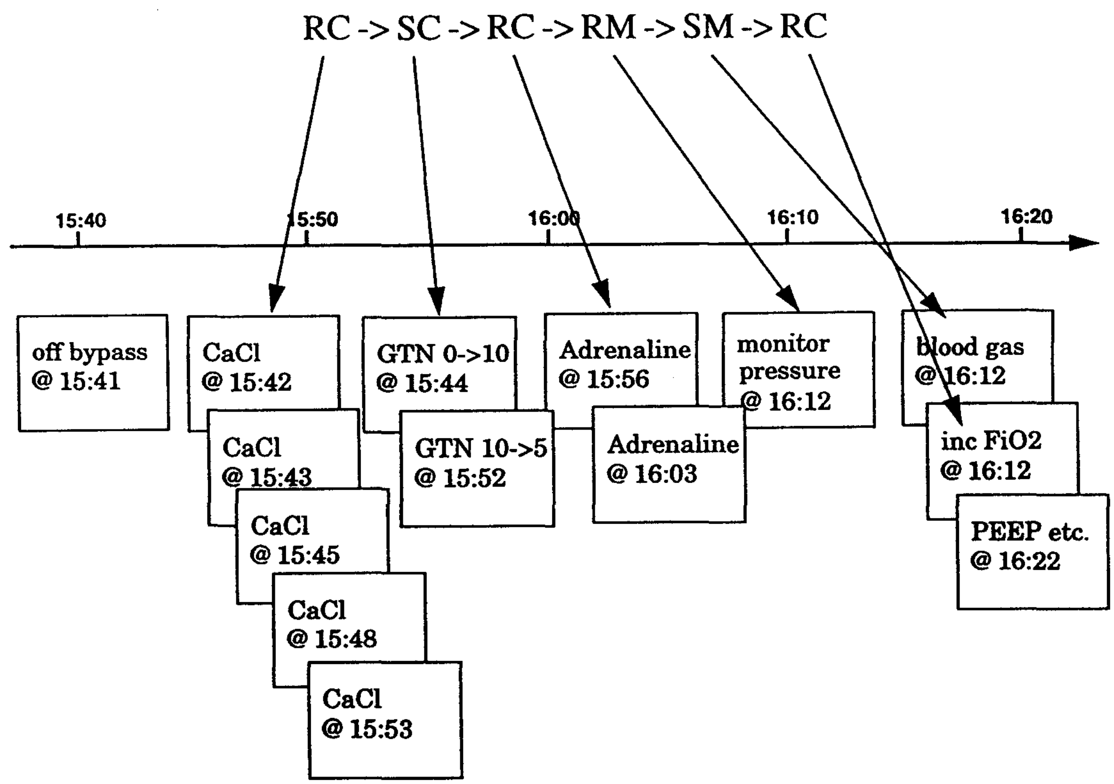

Figure 5. Strategic(S) and Reactive (R) monitoring and control actions taken during the period of evolving hypoxia

which he had not developed a preplanned response. In particular, he was faced with a poorly performing heart, and pulmonary oedema. However, as an experienced clinician, these factors on their own would not be expected to trouble him. The third confounding element of the scenario was that the subject was forced to adopt the therapeutic plan favoured by the surgeon.

It appears that this combination of circumstances led to an period of intense preoccupation with the management of the patient's cardiovascular performance. One consequence of this may have been that, despite looking repeatedly at the monitor, the subject did not detect the saturation trend.

The video and physiological data provide evidence to substantiate this hypothesis. He switches the CVP transducer to measure LA pressure repeatedly during the period between 16:05 and 16:10 (Figure 6.). The intraoperative video demonstrates the subject spending long periods in front of the monitor, dispelling any doubts that the SaO2 value was missed because there had been any failure to use the monitoring equipment.

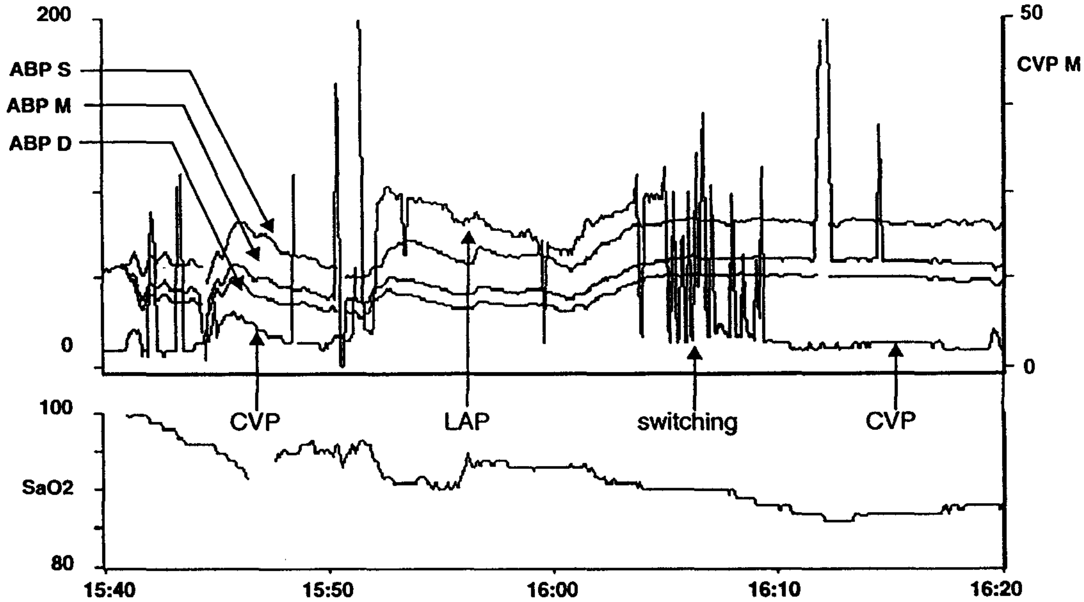

Figure 6. Blood pressure changes during the dropping SaO_2 trend revealed relatively poor cardiac performance immediately post-bypass. By 16:00 the ABP systolic value is below 80 mmHG. The lower curve in the blood pressure set shows central venous pressure (CVP). For the middle third of this curve, the clinician uses the CVP transducer to measure left atrial pressure (LAP), and towards the end of this middle third, he switches repeatedly between the two (as evidenced by the spikes in the trace). The switching behaviour indicates a period of intense concentration on these values.

5.1 A possible explanation

It seems reasonable to suggest that the subject was trying to develop a mental picture of cardiac function over the period in question. This involved the integration of several pieces of indirect evidence (the cardiac pressures) and therefore involved the constructing of some kind of mental model of the evolving cardiac haemodynamic status. At particular points, there is clear evidence of focus on this activity - especially the switching behaviour around 16:05-16:10.

Cognitive psychological theories of attentional resource allocation suggests that the subject's difficulties in detecting the SaO_2 trend may have been due to this reasoning process swamping his available mental resources. So intense was the central reasoning activity, that other activities like monitor observation were heavily filtered - the result being that changes in numeric values on their own were not recognised. A change in the physical properties of the numeric values such as colour may have escaped perceptual filtering and come to the subject's attention.