# Keeping the patient asleep and alive: Towards a computational cognitive model of disturbance management in anaesthesia

K.Keogh [a],* E.A.Sonenberg [b]

[a]*School of Information Technology and Mathematical Sciences, University of Ballarat, Mt. Helen VIC. 3353 Australia*

[b]*Department of Information Systems, University of Melbourne, Parkville, VIC 3052 Australia*

## Abstract

We have analysed rich, dynamic data about the behaviour of anaesthetists during the management of a simulated critical incident in the operating theatre. We use a paper based analysis and a partial implementation to further the development of a computational cognitive model for disturbance management in anaesthesia. We suggest that our data analysis pattern may be used for the analysis of behavioral data describing cognitive and observable events in other complex dynamic domains.

*Key words:* behavioral analysis, computational cognitive modelling, disturbance management

## 1 Introduction

*Cognitive models* are attempts to describe the cognitive processes which may be performed by a human whilst engaged in a task [43]. To build a cognitive model, it is important to gain insights from real-world data as real-world examples expose the complexity of the domain [44]. The task setting we report on comes from the complex and dynamic environment of an operating theatre -a rich domain from which to gather complex behavioral data. Specifically, we studied the decision making of an anaesthetist when a critical incident

* Corresponding author.
  *Email addresses:* `k.keogh@ballarat.edu.au` (K.Keogh),
`l.sonenberg@unimelb.edu.au` (E.A.Sonenberg).

occurred – with a *critical incident* being an event under anaesthetic care which has the potential to lead to an undesirable outcome if left to progress [38]. Interspersed with complex reasoning about possible causes of the incident, bearing in mind that the disturbance may be due to multiple causes, actions must be performed to keep the patient alive [11–13]. This situation has been classified as *disturbance management* [47].

Descriptive, paper-based models of hypothesised cognitive processes are a first step in building systematic explanations of observed behaviours, but are limited in their power to capture the complexities of cognitive activity. Researchers have explored more detailed models, implemented as computer programs, where the program represents a runnable, inspectable model that can be observed 'in action' e.g. [15,17,30,45].. The availability of mechanistically reproducible activity streams that relate to the mental phenomena being explained in turn provides a means for testing the sufficiency of the explanation [35,44]. Testing according to the *behavioral outcome modelling* approach [44] commences with a gathering of behavioral data from human subjects, and involves analysis of that data at appropriate levels of abstraction to determine the extent to which the computational model can produce the same kinds of decisions as do human subjects, under comparable conditions. This has also been termed *cognitive simulation* [36]. Cognitive simulation can show the interaction between data and processes and incorporate the potential to predict the ordering of cognitive tasks that are involved.

Of course, computational models that attempt to explain invisible mental processes of human cognitive activity can only do so in a highly hypothetical way - rarely, if ever, will we have cognitive data that leads unambiguously to a computational model. In other words, models are generally *underconstrained* from data [44], even with all the methods of cognitive task analysis and verbal protocol analysis available. Nevertheless, it is accepted that close study of particular domain problems in the form of computational cognitive models can aid progress towards the extremely challenging goal of understanding human cognitive activity. For example, simulations have been shown to be useful as tools to reveal the underlying cognitive functions involved in disturbance management [48] and other problem solving tasks [15,30].

To facilitate the development of computational cognitive models, a variety of cognitive architectures have been designed and used for model building [2,31,42]. Indeed the process of model building from data can also shed light on the design requirements for cognitive architectures, e.g., [33]. More recently there has been interest in developing higher level abstractions for modelling cognitive processes that in turn can be mapped to specific architectures [16]. Endeavours to produce an abstract language for cognitive modelling have demonstrated that there are important differences between the most prominent cognitive architectures, but that much of the work involved in building specific cognitive

models (e.g. structured task analysis to elicit human behavioral data) is the same no matter which architecture is targeted [18].

The data on which our modelling is based came from a study where clinical anaesthetists were observed in an exercise in which a software simulated patient was experiencing difficulty under anaesthetic. The problems simulated needed to be solved within 2-3 minutes to avoid catastrophic results. Verbal protocol data collected during these experiments, aligned with information on the actions conducted by the clinician, provided us with the behaviour of clinicians during the disturbance management episode and provided insight into their management of decision making. Data collection and early aspects of the analysis are briefly explained in section 2 and in detail elsewhere [21,20,22,39].

A feature of the anaesthesia domain is that many patient state changes are not able to be anticipated by the anaesthetist. Expert performance in anaesthesia involves parallel processing, multitasking and iteration through a cycle of observation, decision, action and re-evaluation. The anaesthetist must perform dynamic decision making, relying not only on medical domain knowledge but also on expert management of the cognitive demands presented by the task [11–13].

It has been noted that cognitive demand and performance issues such as selective attention and management of multiple goals are aspects of cognitive processing that are not easily revealed by analysis of capturable human performance data. Further, control knowledge which influences *how and when* the subtasks in a model might be selected is not easily represented in a paper-based analysis. It has been argued that these can be analysed through the evolutionary process of attempting to build and refine a computational model [32].

In this paper, we present a computational cognitive model STAM (Select and Test Model with Action and Meta Skills) as a partial model of the tasks involved in disturbance management in anaesthesia. This model has been derived from our analysis of the domain specific behavioral data and our attempts to initially align that data to the generic Select and Test (ST) model of diagnostic reasoning proposed by Ramoni and colleagues [28].

Ramoni's ST model provides an iterative cycle for tasks at the skill and rule based level without directly showing more abstract knowledge level tasks. We explain our analysis pattern and the techniques we have applied successfully as part of this analysis. We propose that this analysis pattern is suited to similar complex dynamic domains and we argue that there was significant benefit to the development of our model by programming a partial computational model to complement the paper based analysis.

As we demonstrate below, attempting to design and implement a computational model of complex decision mechanism in a dynamically changing con-

3

text forces the designer to make explicit the knowledge and processing mechanisms in the model, particularly with respect to modelling management and coordination. The importance of explicit modelling control structures has been highlighted since the very earliest days of cognitive modelling [1] and has been reinforced in both pragmatic and philosophical analyses, e.g. [27,33,34,49]). For example:

> "To exert executive control, an agent needs to be able to timely initiate, suspend, cease and co-ordinate the execution of adequate intentions, attentively monitor the proceeding execution of the proximal intentions as well as the environmental changes, detect errors and conflicts, actively maintain and sustain the execution of appropriate intentions, and inhibit intervening or distracting processes." [49]

Indeed, it has been argued that the most significant difference between AI planning and cognitive modelling is in their treatment of control knowledge, in particular control knowledge specific to a given domain [2]. Goal driven architectures, such as Soar and ACT-R 4, are insensitive to supporting stimulus driven or bottom-up changes [33], whereas many cognitive tasks of interest (including the tasks essential for disturbance management in the domain of anaesthesia) require the actor to exhibit sophisticated executive control. In the evolution of the STAM model we describe below, it is perhaps unsurprising that it was necessary to insert features in the model to provide explicit meta-level knowledge processing, but we believe that additional insight can be obtained by understanding the general 'control' problem in the setting of a particular, challenging, domain.

The remainder of this paper is structured as follows.

Section 2 provides some background to our work. The complexity of the chosen domain is described. The method we adopted as part of our analysis pattern is explained. We describe the process used to analyse, describe and visualise our data (Event Maps, Event Matrices and Hypothesis Trees). Some findings based on our data analysis are provided, more detailed examples and findings can be found elsewhere [20,21].

In section 3 we outline a computational model, a partial implementation of the STA cognitive model: Select and Test Model with Action incorporated. We discuss the limitations in the STA cognitive model which were clearly evident in the performance of the computational model. Based on the analysis of our data and drawing from theory, we explain the limitations of STA model and justify these changes by proposing the STAM model(Select and Test Model with Action and Meta Skills).

In section 4, we discuss computational modelling as a tool for analysis toward 'realistic' cognitive models. We comment on our suggested pattern for

the analysis of complex dynamic, behavioral data and suggest it could be applied to other similar data. We mention the limits of the STAM model and the difficulty in evaluating such models, however, we argue that our analysis and computational modelling progressed us toward a cognitive model of disturbance management in anaesthesia with a 'good fit'.

## 2  Representing complexity

### 2.1  Introduction

In disturbance management in anaesthesia, complex decision-making interspersed with actions for therapy and diagnosis provides us with rich temporal data. In this section, we describe the complexity in our data, how the data was collected and our novel methods for representation and initial analysis of the data. We present the results of our analysis as summaries of strategies and behaviour in the subjects.

### 2.2  Complexity in Anaesthesia

In the domain of anaesthesia, the patient's state is constantly changing, not only as a result of actions performed by and drugs introduced by the anaesthetist, but due to physiological changes in the patient's condition. The patient is *not* a static system! Anaesthetic management is event driven, with severe time constraints and involves risk and uncertainty [11–13]. The decision maker must address patient disturbances that are short term consequences of the problem whilst simultaneously diagnosing underlying 'faults'. Features such as: incomplete data availability; uncertainty; dynamically changing state; time constraints; and multiple interleaved sub-tasks make management of such a complex task a cognitively demanding process [5].

As the amount of data available about the unfolding situation is greater than can be immediately managed, the anaesthetist must be selective in what data to take notice of, and how often this monitoring is required [12]. Careful juggling of multiple goals: administering therapy to keep the patient alive; and performing diagnostic reasoning and diagnostic tests; is necessary in the successful management and resolution of the problem. Meta cognitive skills in supervisory control and resource management are important in the expertise developed by the successful anaesthetist [12,13]. The strategies employed by the anaesthetist to dynamically adapt their thought processes are significant to their successful management of critical incidents which may occur [12].

5

Indeed, it has been found that the meta-cognitive elements of disturbance management are more significant for safe intraoperative patient management than deep medical knowledge [11].

Gaining insight into the meta-knowledge used by decision makers was a focus of the study reported here. This strategic knowledge was needed to help describe *how* the model could be applied as a predictive model. A limitation of this study is that we did not gather any specific data representing the domain knowledge of each anaesthetist. We have focused on the skills of coordination and on management processes.

## 2.3   The experimental data

Earlier a number of patient simulation experiments were conducted and preliminary analysis reported [39,40]. A Macintosh based patient simulator, (the Poor Ventilation Simulator, PVS), was used to produce patient data representative of ventilation problem conditions which might occur soon after a patient has been anaesthetised and a procedure known as endotracheal intubation has occurred. The PVS simulator was created based on simple heuristics which managed to capture the patterns of patient degeneration in these situations quite credibly. Using the PVS simulator enabled the same incident to be presented to multiple subjects and it automatically recorded the anaesthetists' actions whilst managing the problem. The incidents being simulated by PVS involve difficulties with patient ventilation and present symptoms which may have multiple explanations. The anaesthetist must simultaneously consider multiple hypotheses and perform therapeutic actions to keep the patient alive.

Three converging data sources were used to obtain evidence of the subjects' knowledge structures and decision steps whilst managing the problem: a computer generated (PVS) record of actions taken and the clinical context in which they occur; a verbalised record of the problem solving session; and a verbalised record of a stimulated recall session immediately following the simulation exercise [39]. Using multiple sources of evidence provides a more complete framework of knowledge and processes used during the problem solving session [39]. In each run of the simulation, the anaesthetist was asked to manage 'the patient' whilst the simulation was driven by the experimental team. The anaesthetist sat in front of the monitor giving directions for patient 'manipulation' (actions to be performed on the patient, for therapeutic reasons and/or the purpose of generating further information e.g. listen to chest). Patient data that would ordinarily be available in the operating theatre by direct observation of the patient or a display monitor, such as Blood Pressure, Oxygen Saturation, Endtidal $CO_2$, Ventilation, Skin Colour and ECG data was continuously displayed on the PVS monitor. Data that was not continuously

visible was only displayed on the PVS monitor in response to the anaesthetist's requests or instructions to act.

Data transcribed from the audio tape was organised into time segments aligning the think-aloud protocol with the retrospective protocol and the values for visible patient data available from a log file. Further interpretations of the data were produced identifying information referred to and any reasoning or action which took place within each segment. The process of integrating each data source is described in more detail elsewhere [39].

The three problem scenarios simulated were: Oesophageal Intubation (tube holding gases was intubated into oesophagus instead of trachea.); Kinked Tube (tube is kinked disrupting flow of gases into patient.); and Anaphylaxis (a massive allergic response, in anaesthesia this is typically due to an intravenously administered drug, which results in cardiovascular collapse, respiratory difficulty and possibly cardiac arrest).

Seven subjects were given the simulation of the problem: Oesophageal Intubation. The simulation experiment was based on the problem scenario of the tube containing gases being accidentally inserted into the oesophagus instead of the trachea. The data from the Oesophageal Intubation simulation was analysed in detail and is the focus for the analysis described in this report. The subjects' experience as an Anaesthetist varied from 4 to 19 years.

*2.4 An Analysis Pattern*

The following major steps were involved in our analysis, in an iterative cycle toward the development of our computational cognitive model.

*Cognitive Decision and Action Categorisation* Based on the verbal protocol data, we categorised the actions and cognitive processes (eg. decisions, hypotheses) that may have occurred. We chose tasks based on the cycle of reasoning proposed in Ramoni, Stefanelli, Magnani and Barosi's [28] Select and Test Model [22]. Each event identified in the data was categorised and tabularised to show the trace over time of each event type occurring. The synchronised raw data, was analysed in terms of potential cognitive tasks involved from the Select and Test Model.

As part of this analysis, hypotheses identified as under consideration were listed, along with actions performed and clinical data which was given attention. An example of this analysis is given elsewhere [20,21].

*Hypothesis Trees: Visual Tools to highlight individual's strategic behaviour.* Whilst attempting to define the process in more detail, toward the imple-

7

mentation of a computational model, we have examined strategic patterns of behaviour used by individuals. In separate publications [20,21] we have provided a detailed example showing how we built hypothesis trees representing each subjects' hypothesis generation during the scenario.

*Event Matrices: Visual Tools to highlight reasoning versus action* Time–Event Matrices and Event Maps were constructed as a schematic representation of the reasoning and events involved in each scenario. These were used to aid the objective interpretation and analysis of the data. Time–Event Matrices as tools for describing and presenting complex data representing thoughts and activities, have been shown [41] to be visually helpful to aid comprehension and to aid in the calculation of numeric scoring measures. Such metrics are helpful when comparing the behaviour of two or more different individuals. Streufert and Nogami use Time–Event Matrices to analyse planning behaviour [41]. Such planning information is not the focus of this study, however a modified form of Event Matrices were a useful tool for describing and presenting the data involving reasoning and actions [21].

*Event Maps: Visual Tools to facilitate comparison between subjects* Time–Event Maps were developed as a valuable picture of each subject's performance. These enabled the researchers 'at a glance' to see the interaction between cognitive and actual activity that may have been involved for each subject [21]. They were also useful as a comparative tool between subjects. Using this analysis tool, common strategies were identified in all subjects' behaviour. Also, differences between subjects, particularly with respect to control of the selective attention to data was apparent.

*Computational Model Development and Refinement* The process of the development of a model is an iterative cycle. Informed by the literature, and all our previous analysis including the findings drawn from each of the previous stages in our analysis pattern, our earlier model (STA model)[22] was refined and extended. The STAM model explained later is our current proposal for a cognitive model which may match the anaesthetists' behaviour during disturbance management.

*2.5 Describing the data*

The raw data was dense and difficult to conceptualise. Event Matrices and Event Maps were introduced to represent the reasoning and events involved in each scenario. The data was easier to analyse and compare using these visual methods for presentation. We have discussed the creation of Event Maps and Event Matrices in more detail elsewhere [21,20] . Based on our experience with data in the domain of Anaesthesia, we found a natural connection be-

tween the generic categorization of actions as initiative or respondent adopted by Streufert and Nogami [41], and the concepts of reasoning-driven actions and data-driven actions identified separately in our domain [39]. These categorizations gave rise to insights in highlighting how our subjects managed to interleave reasoning with action.

Time event matrices and maps allow a visual representation connecting times of observation with decision points and actions performed as a consequence of the data attended to. Each unique action was represented at a different height on the graph, as was each unique hypothesis. Actions are plotted below the X-axis and Hypotheses above the X-axis. The visual plot showing action and hypothesis consideration with time allows an immediate impression of the relationship between actions and reasoning and how active each subject was both mentally and 'physically'.

For example, Subject 1, a registrar, with 4 years experience, engaged in a problem scenario over 2 minutes and 52 seconds. During this time, 6 unique diagnostic and 2 unique therapeutic actions were made and 7 hypotheses were considered. Each of the actions were represented visually aligned with the time on the Time Event matrix and Time Event Map. The hypothesis 'Endotracheal Tube problems - not in trachea' was considered 47 seconds into the scenario and reconsidered 1 minute later and then refined to the final (correct) hypothesis 'Oesophageal intubation' at 2minutes 25seconds. The final action as a consequence was to extubate the patient. The Event Matrix highlights actions that follow as a consequence of or as if planned based on previous information, by connecting lines. An example of such a relationship is the mental task of abstraction following from the action of attending to new data.

The event maps depict clearly how long hypotheses were considered and visually show how active the subject was. Comparing the visual representations of data from each subject was helpful toward the analysis and complemented metrics collected. Hypotheses considered and the period of time they remained under consideration was easily identified using unique colours for each hypothesis in the event map plots. When hypotheses were maintained over a period of time, the Event Map shows constant plateaus rather than sharp peaks above the x-axis.

The number of actions performed and whether they were initiative driven reasoning-based actions or actions in response to other information were used to suggest more about each subject's directionality of reasoning. The behavioral trends and strategies of successful subjects differs from those of the unsuccessful subjects.

The successful subjects performed a smaller range of possible actions during the fault disturbance episode than successful subjects. Successful subjects

9

performed between 5 and 10 unique actions, a behaviour similar to that of experts: high breadth of functioning (number of unique actions) is not a strategy associated with expertise. It is claimed that experts selectively perform fewer actions than novices, choosing actions which will help to differentiate between hypotheses under consideration [3]. Unsuccessful subjects, performed comparatively more actions: between 10 and 13 unique actions.

The two unsuccessful subjects (subjects 2 and 3) performed more initiative (reasoning driven) actions than the successful subjects. By definition, these actions must be based on their reasoning rather than in response to information provided (one could suggest these might be (backward) reasoning based on current hypotheses, guess work or theorising from first principles). It could be suggested that the successful subjects, by comparison, were more efficient. They performed fewer actions to provided the necessary information to help isolate/eliminate hypotheses. This is consistent with findings in other studies that accurate diagnoses are associated with pure forward reasoning backed up by backward reasoning only to explain unrelated data in terms of the hypothesis [14,24–26].

In summary, the Event Maps and the Time–Event Matrices served to provide a useful picture of the behaviour of subjects which allowed for a quick visual comparison between subjects. Interestingly, the Event Maps that represent behavioral data for the two subjects who were unsuccessful at diagnosing the problem appear to have different characteristics to Event Maps describing the other subjects. The Event Maps show that these subjects performed more initiative actions (quite probably due to backward reasoning) which increased the workload for these subjects.

## 2.6  Patterns of hypothesis generation

Table 1 provides information regarding hypotheses considered by each subject. These were generated based on both the think-aloud and retrospective protocols. If a hypothesis was considered (referred to) on more than one occasion, it was counted on both occasions for the second column, (HR: total number of Hypotheses (Re)considered) but only counted once in column one (UH: count of Unique Hypotheses).

Unsuccessful subjects did not generate as many initial hypotheses. They continued to generate new hypotheses to test rather than attempting to eliminate from a larger initial set. Subject 2 failed to generate a number of initial hypotheses from which to eliminate and refine. This subject kept eliminating hypotheses and generating alternate (previously unconsidered, though related) hypotheses, similar to the behaviour of intermediate novices in other

| Subject | UH | HR |
|---------|-----|-----|
| subject 1 | 7 | 9 |
| subject 4 | 4 | 6 |
| subject 5 | 5 | 5 |
| subject 6 | 6 | 10 |
| subject 7 | 1 | 1 |
| subject 2 | 6 | 9 |
| subject 3 | 11 | 15 |

Table 1
Number of unique hypotheses(UH) and total hypotheses(HR) considered by each subject

studies [3,26]. Subject 3 suffered from a lack of new data on which to base reasoning and new hypotheses. This subject engaged in more hypothesis driven (backward) reasoning rather than reasoning based on data.

Hypotheses considered by successful subjects as the scenario unfolds are generally more connected to previous hypotheses, as refinements or reconsidered hypotheses, than hypotheses generated by unsuccessful subjects. The successful subjects tended to initially identify more appropriate general hypotheses which were subsequently refined or eliminated. These general hypotheses could be compared with the notion of facets described as capturing a group of findings as one concept that is recalled [4,9]. This is consistent with the possibility that these subjects were using long term memory retrieval structures to enhance their processing of information[10]. The unsuccessful subjects seemed to backtrack and reconsider previously eliminated hypotheses more.

Our observations are consistent with the findings of Arocha with respect to the directionality of reasoning used by practitioners at different levels of expertise [3]. In our study, successful subjects seemed to be following strategy patterns associated with expertise. For example, the successful subjects attempted to constrain the problem space rather than expand it which is a strategy used by experts [26]. Successful subjects engaged in forward reasoning and generated more initial hypotheses which were narrowed, refined or eliminated. The successful subjects did not seem to generate new totally unrelated hypotheses later in the scenario, though they did reconsider previously eliminated hypotheses in the light of new data.
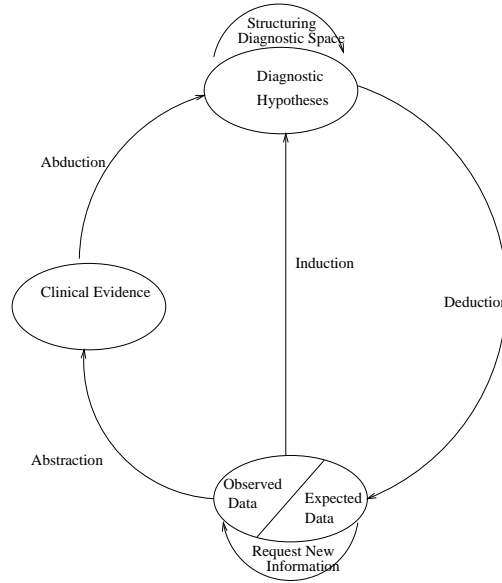
Fig. 1. STmodel [28]

## 3 Modelling disturbance management in anaesthesia

### 3.1 Introduction

In this section, we describe our analysis and modelling in detail. Our initial model, STA Model was an adaptation of ST Model [28] to include action but it did not directly incorporate strategic control knowledge and was not capable of progress. We discuss the performance of our computational model of the STA model and describe our new proposed model, Select and Test with Action and Meta skills (STAM).

### 3.2 A Computational Model : STA Model

Ramoni et. al. [28] suggested that their Select and Test Model (STModel) can be applied to describe the medical *reasoning* tasks of diagnosis, therapeutic planning and patient monitoring. Patel, Groen, Ramoni and Kaufman [26] have exercised the STModel with data. Data from subjects with varying levels of expertise have been examined in terms of the STModel [3] indicating that the cycle was not followed in the same way by subjects at different levels of expertise, but that experts' data matched it well. Joseph also found that expert behaviour during diagnostic reasoning seemed to be consistent with the behaviour suggested by the STModel [19].

Ramoni et. al [28] claimed that the STModel cycle described not only diagnosis
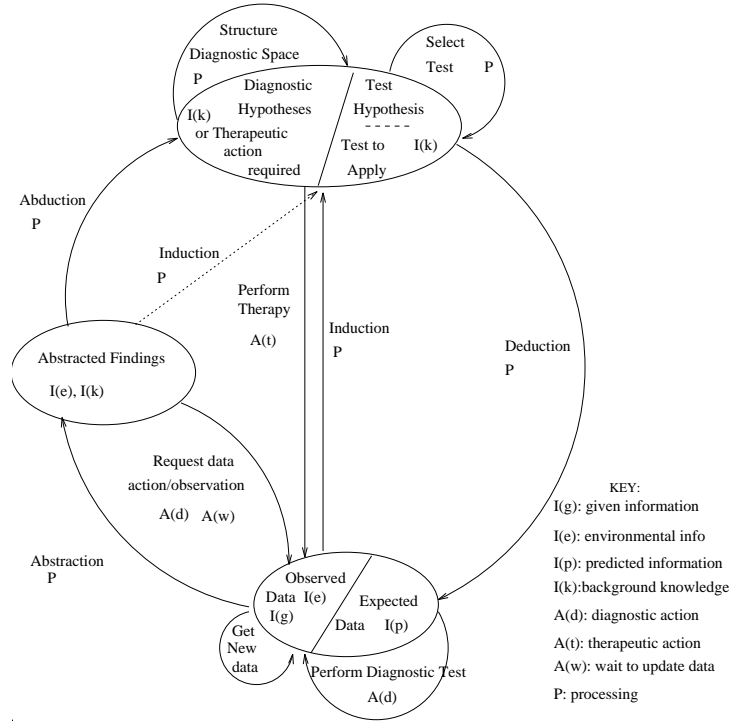
Fig. 2. The Select and Test model with Action (STA Model)

but also cycles of therapy planning and monitoring. They defined diagnostic reasoning as providing 'an explanation of the current situation in the patient.' Therapy planning and monitoring were executed following a previous understanding of the patient's situation (i.e. *following* a diagnosis). The therapeutic planning task they described was that of choosing an appropriate therapy, however, it did not involve performing that therapy.

The STModel does not incorporate action into its cycle of selection and testing. It focuses on reasoning separate from any action, implying that any action occurs only as a result of a final diagnosis/therapy plan. The patient monitoring task described involves continuous cycling between deduction and induction steps given a prior diagnosis or therapy plan and as a result, previous diagnosis and therapy planning tends to be confirmed or questioned. A new, separate cycle of diagnosis or therapy planning can then follow this cycle.

Based on our initial paper-based data analysis, we modified STModel to produce STA Model *STModel with action* [22]. Figure 2 shows our STA Model, the Select and Test model with Action. STA Model included *action* for therapy and diagnostic evaluation. It included an explicit loop for requesting particular data as part of the diagnostic selection phase (the left section of the cycle leading from observed data at the bottom to diagnostic hypotheses at the top). STAM model, shown in Figure 3a, is built upon STA Model with extensions to show meta level procedural knowledge.

13

On paper, the behaviour of subjects in our study matched the STA Model well. Each subject's behaviour for a time segment was tabulated into categories: observed data, abstracted data, hypotheses, diagnostic test and expected data. The mental tasks of abstraction and abduction were indicated in keeping with the model cycle and where the data suggested that they occur. The analysis of each subject's behaviour in terms of the STA Model was examined particularly noting deviations from the STA Model. In general, based on this paper based analysis, each subject performed similar tasks and in a similar order to those suggested by the STA Model. The synchronised raw data, was further analysed in terms of potential cognitive tasks involved from the Select and Test Model. Further details regarding this analysis are provided elsewhere [20].

### 3.2.1 An implementation of STA

As a descriptive tool to model our data, based on a general examination of data, STA Model appeared to provide a good basis for the cognitive tasks involved in diagnostic reasoning in this domain. To explore further the 'fit' of the model we built a (partial) computational model using Soar [31] We have outlined elsewhere a task level description of a Soar computational model of the diagnostic reasoning of anaesthetists based on STA Model [22]. We implemented two communicating Soar agents – the *patient* and the *diagnoser*. To partially implement the diagnoser agent, 120 production rules were written, using 11 problem spaces (diagnosis, do-abstraction, monitor-patient, abduction, induction, anaesthetic-abduction, deduction, perform-therapy, structure-diagnostic-space, select-diagnostic-test, perform-diagnostic-test) corresponding to the top down tasks in STA model.

The *diagnoser* agent is described in more detail in [20], and the *patient* was built using the same design as the simulated patient used in data collection [39]. This state machine was designed by an expert anaesthetist to represent a realistic picture of the patient's state given the problem condition being modeled and the actions of the diagnoser. The patient agent regularly sends updates to the diagnoser agent describing immediately visible patient data such as pulse, blood pressure, oxygen saturation, ventilation and skin colour. Additional patient information may be sent to the diagnoser agent in response to a particular request (e.g. Auscultate the left lung, Increase oxygen). Each agent contains a separate knowledge base. The patient model comprises simple heuristic knowledge and attributes that mimic a real patient.

The diagnoser agent has its own internal model of the patient which is updated to represent its current knowledge of the patient's status. Communication between agents is achieved with programmed links in the C programming language enabling a loop of dynamic feedback to the diagnoser agent.

14

Communication links from the patient to the diagnoser are set up so that the information is made available immediately to the diagnoser on the diagnoser's input link, but until the diagnoser chooses to accept such data it is not reflected in the diagnoser's own patient model. This design was chosen to mirror the deliberate processing of an anaesthetist in deciding to watch for particular data, ignoring some data, and sometimes noticing other data. The bottom up observation tasks are represented in our design decision to make the information immediately accessible on the diagnoser's input link. This approach of combining bottom-up processing with a top down design has been successfully adopted in similar work [33].

The implemented computational model did not support resource management or supervisory control. When the system was faced with a choice, i.e. when more than one task was proposed as acceptable, it didn't have built in domain knowledge to suggest how to proceed. There were very limited process rules that allowed some progress (outlined below) but not sufficient knowledge to successfully juggle and choose between simultaneous competing tasks. Our system was able to read in the initial patient input values and perform some basic abstraction of this data in readiness for the abduction task to generate hypotheses. The new input values which then became available fired competing tasks to attend to the new input, rather than proceed with the cognitive task of abduction. Typically, the system kept attending to new data and didn't progress further to other cognitive tasks such as abstraction and abduction based on the patient state. Limited knowledge was incorporated with a rule that suggested the diagnoser prefer abstraction over monitoring tasks when some monitoring had already been performed. When some data had been attended to, this meta-knowledge was needed to aid decisions to revise the data as well as how to decide when to perform diagnostic reasoning tasks.

The system was provided with some domain specific meta-knowledge. When new input was available, it could be ignored if it was within provided threshold values for acceptability. For example, for the clinical cue, Blood Pressure (BP), a delta value 5 was provided to say that when data fluctuations occur that were less than or equal to 5, it was not necessary to perform further abstraction to re-classify the new value. In addition, there were search control rules which provided a limited ability to choose between different monitoring tasks and among abstraction tasks. Without these rules, the system reached an impasse due to tied operators almost immediately, as it didn't know which monitoring task to perform first. Also, priority settings were explicitly incorporated into the system to enable limited decision-making between competing attention tasks based on these priority settings. These values were hard-coded and no meta-knowledge was incorporated to enable dynamic changes to these priority settings. The priority settings were based on our observations of the data collected and how often our anaesthetists referred to the different data available. For example, BP deserves more attention than ETC02 levels. This knowledge

was provided by giving BP priority 2 and ETC02 priority 1. A search control rule provided meta-knowledge: prefer operators with higher priority. For example, in the do-abstraction problem space, when proposals existed to think about two different clinical cues, priority was given to cues which were identified as having higher priority, otherwise both proposed operators were treated with equal preference.

The implicit control provided in the STA Model cycle limits the ordering and selection of tasks to some extent – e.g. abstraction cannot be proposed until (some) patient monitoring has been performed; abduction cannot be performed until abstraction on (some) data has been performed etc. This control was implemented using the patient-condition attribute as a trigger. However, ordering of the tasks according to the top down STA Model cycle was not sufficient to use as a generator of appropriate behaviour sequences. Executive control of goals is needed for deliberate action [27]. As noted during the analysis, there were some acceptable deviations from the STA model which occurred in our human performance data. Heuristics to describe such strategic knowledge were needed before it could be claimed that the model could be used to predict performance. The performance of our computational model confirmed that implementing more domain specific meta-knowledge in a separate meta-level problem space was necessary to enable the system to progress efficiently and intelligently. Drawing from our data and from literature we make some suggestions toward how this meta-knowledge could be structured in the next section.

*3.3   STAM: Incorporating Action and Meta level skills*

Based on our analysis and existing literature regarding dynamic disturbance management and decision making in complex domains, we proposed STAM, *the Select and Test model with Action and Meta level skills*, shown in figures 3a and 3b and explained in more detail in [20].

STAM includes some of the major tasks that could be involved during the management of a critical incident in a complex domain such as anaesthesia. The STAM model we propose is based on the top-down goal driven tasks in the STA model (Figure 2): abstraction, abduction, induction, deduction, diagnostic testing, therapeutic action, patient observation and patient model updating. Additional cognitive management tasks in STAM are proposed as: priority management, supervisory control, multi-tasking resource management, and automatic tasks for data perception. The shaded areas and annotations in Figure 3a indicate the components distinguishing the STA model from the STAM model.
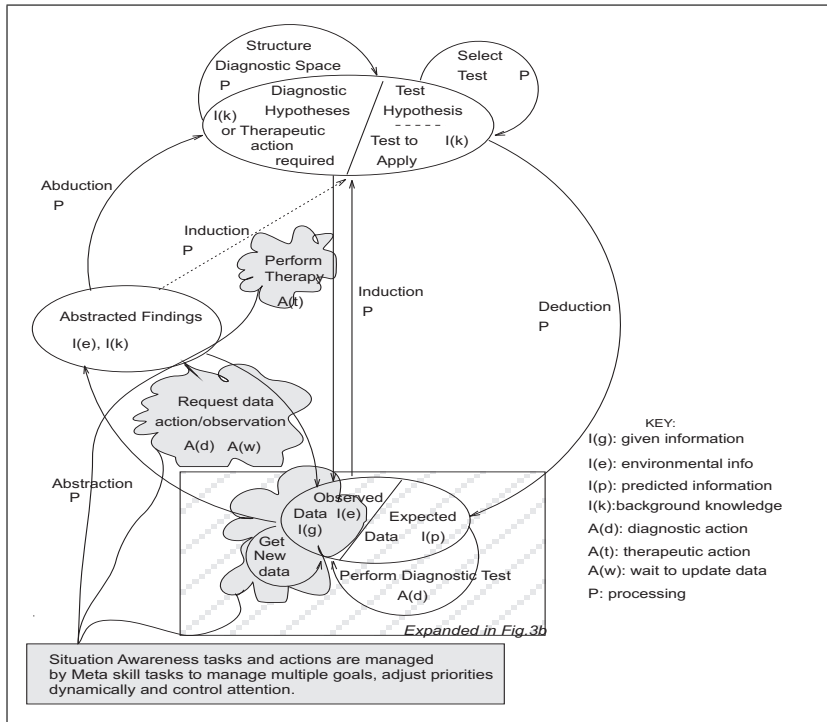
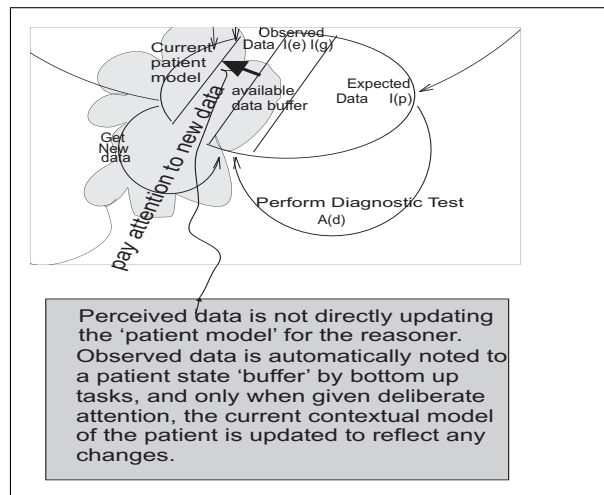Fig. 3a. STAM: Select and Test Model with Action and Meta skills



Fig. 3b. STAM: Select and Test Model with Action and Meta skills - zooming in on data observation

The distinction between perceived data and data that has been attended to as shown in Figure 3b was made to remain realistic to human behaviour. Automatic environmental perception data is available in a buffer, then when this data has been attended to, it changes 'state' to flag that it has been attended to. This approach has also been used by Schoelles and Gray [33].

The STAM model includes meta-level tasks such as supervisory control and resource management that have been shown to be crucial in the anaesthesia domain [11,12]. The STAM model also includes action integrated into the

cycle of decision making.

Figure 3a shows the cycle of top down diagnostic tasks beginning at the bottom with new data, through abstraction of data to form more general classifications and findings, abduction of possible hypotheses and or therapies that may be needed, then selection of tests chosen to eliminate hypotheses from those under consideration followed by deduction to generate an expected outcome for these tests. The expected result is compared with the actual outcome, leading to induction task which may lead to elimination or confirmation of a hypothesis. This cycle is inherited from both the STModel and STA Model.

The STAM Model differs from the earlier models on which it is based. It includes the meta control tasks: Multitasking resource management, priorities management, and supervisory control. Based on Rasmussen's taxonomy of decision making tasks [29], these would be the highest level abstract knowledge cognitive skills. Rule based tasks which are based on previous experience and relationships between situations and procedural responses to that state, include diagnostic reasoning tasks and monitoring tasks. We have identified conscious situation awareness tasks and actions as deliberate tasks based on decisions to attend to or search for available data. Automatic, low level cognitive tasks in this domain relate to the automatic observation tasks which could be described as bottom-up tasks. The STAM model distinguishes between the patient state and the reasoner's current contextual model of the patient's state. The reasoner's model of the patient is only updated following a deliberate decision to attend to new patient data as highlighted in Figure 3b.

The difference between STAM model and STA model is the addition of the higher level cognitive tasks responsible for management and coordination. The flow of control between new data, abstraction and abduction is not prescribed as merely following these rule based tasks in a cyclic manner, but is now managed by these meta-control tasks. Following the process of updating the current patient model, instead of the flow of control moving automatically to abstraction, the meta-level Supervisory Control knowledge decides what to do next. The Multitasking resource management task can interrupt reasoning when needed. The changing priorities are managed in an explicit abstract knowledge level task.

Strategic meta-level knowledge is very significant in this domain [11,12]. We have found our data to be consistent with other studies which suggest that success is related to pure forward reasoning strategies [14,24–26]. Based on our analysis, complemented by theory, we have identified a number of strategies which could be associated with success. The process of building a (partial) computational model to test our cognitive model proved a powerful method of analysis.

As previously outlined, the preliminary implementation of STA Model was based on problem spaces derived to correspond to each of the main cognitive tasks in the STA Model. The implementation had no explicit supervisory control level. Meta-level tasks for resource management and allocation of attention were not explicitly incorporated into the design. The cyclic nature of the STA Model allows for some implicit control in terms of the repeated iteration cycling through the main cognitive tasks: abstraction, abduction, deduction and induction. This control was implemented based on trigger conditions that must be present in order to propose a task in the diagnosis problem space.

The data analysis conducted suggested that STA model is feasible as a descriptive model. In other words, the data analysis performed shows that behaviour can be categorised to fit this model. However, the program based on this model was not capable of successful progression past the observation and monitoring tasks. Due to the continuous presence of new data describing the changing patient state, the program would focus on new data as it arrived without progressing on to problem recognition and decision making tasks. This supports claims by other researchers regarding the behaviour and management of complex and dynamic tasks [5,12]. Additional explicit meta-control tasks are essential in order to decide enough monitoring has been done and to move on to 'thinking' about this data. Such strategic meta reasoning knowledge may bring STAM closer to being a predictive model of performance in this domain. Preliminary work in this direction can be found in [20].

## 4 Conclusion

Computational modelling of complex human decision making is a step towards the goal of building systems that can provide intelligent decision support to human operators. This paper has reported a descriptive and computational study drawing on rich data on human decision making in a complex and demanding setting. Whilst domain knowledge is clearly required and domain expertise has been shown to affect the efficiency of a practitioner during diagnostic reasoning [19], considerable management expertise is also required. It has been shown that meta-level cognitive tasks that enable a person to be adaptive and flexible in prioritising their actions and to coordinate juggling attention and resources between competing tasks are a major component of the 'expertise' required for success e.g. [6,12]. In our domain the behaviour of an expert managing a disturbance or crisis of complexity, involves therapeutic actions interspersed with diagnostic reasoning and action over a short time period.

Modelling the decision making in such situations involved first analysing the behaviour of experts engaged in successful disturbance management. Then

suggesting a descriptive model of this behaviour and ultimately a predictive dynamic model that could simulate some of the expert behaviour. Such computational modelling provides us with an analysis process for mapping people's cognitive behaviour whilst performing complex tasks. We have not attempted to model the anaesthetist's knowledge (see, e.g., [46] for attempts to model knowledge in this domain). The cognitive activity engaged to successfully manage and perform decision making in complex environments is not easily determined. Building a computational model, which can be run, 'watched in action' and then compared with real data provides us with valuable insights into the processes which *may* occur. The analysis pattern we adopted can arguably be applied to other forms of complex management tasks in domains where diagnostic reasoning must be interspersed/competing with actions. We have attempted to computationally model some of the processing  [48]. This has progressed the development of a cognitive model of processing in this domain.

We have developed the STAM model and suggested some domain specific strategies that would apply to an implementation of STAM model in the anaesthesia domain. The STAM model is less reliant on an implicit predictive cycle of tasks than Ramoni's ST model, and relies on meta-level knowledge to manage and control which tasks are performed according to a contextual internal model of situated awareness. Although we do not attempt to provide a complete detailed architecture we gained insight as to the types of components needed to generate appropriate behaviours.

Our work has illustrated how on-paper analysis can be inadequate to test a model as a predictive performance model. The search for strategic knowledge can be aided with paper analysis, but the testing of such knowledge can be done best via implementation and running of a computational model.

Evaluation of computational models is far from straightforward. It is necessary to keep separate, implementation details and the theoretic cognitive model [7,8,37] and there are many confounding factors in the search for what Sun and Ling have termed the *source of power* in a cognitive model [44]. It is suggested to test different versions of a computation model by varying parameters to see their effect [44]. In our case, significant further work is needed to implement the proposed control knowledge in our computational model. Following this it would be possible to experiment with changing difference strategic knowledge and to compare the model performance with individual human performance. Of course, even at that point the question would remain about the robustness of the explanation supported by such data alignment. As Krebs frames it: "are results obtained from experiments that are essentially performed on data structures, equivalent to results from 'real' experiments?" [23]. As Sun and Ling observe, "there is always a many-to-many mapping between computational models and the cognitive phenomena to be modelled, so the development and application of fundamental, abstract cri-

teria in analysing and comparing models are essential [44]". In terms of the success criteria they propose, our work is limited to basic goodness-of-fit. But we see a contribution being made by our close analysis of human behavioral data towards a computational cognitive model, in a domain that is highly complex and dynamic.

## 5    Acknowledgements

## References

[1] Newell. A. You can't play twenty questions with nature and win. In W. G. Chase, editor, *Visual information processing*, pages 283–308. New York: Academic Press, 1973.

[2] R.S. Amant, S.P. McBride, and F.E. Ritter. An AI planning perspective on abstraction in ACT-R modeling: Toward an hlbr language manifesto. In *ACT-R Workshop proceedings 2006*. 2006.

[3] J F Arocha, V L Patel, and Y C Patel. Hypothesis generation and the coordination of theory and evidence in novice diagnostic reasoning. *Medical Decision Making*, 13(3):198–211, 1993.

[4] Jose F. Arocha, Dongwen Wang, and Vimla L. Patel. Identifying reasoning strategies in medical decision making:a methodological guide. *Journal of Biomedical Informatics*, 38:154171, 2005.

[5] L Bainbridge. The change in concepts needed to account for human behavior in complex dynamic tasks. *IEEE Transactions on Systems, Man and Cybernetics*, 27:351–359, 1997.

[6] L Bainbridge, T M J Lenior, and R W van der Schaaf. Cognitive processes in complex tasks: introduction and discussion. *Ergonomics*, 36:1273–1279, 1993.

[7] G Barosi, L Magnani, and M Stefanelli. Medical diagnostic reasoning: epistemological modeling as a strategy for design of computer-based consultation programs. *Theoretical Medicine*, 14:43–55, 1993.

[8] R Cooper, J Fox, J Farringdon, and T Shallice. A systematic methodology for cognitive modelling. *Artificial Intelligence*, 85:3–44, 1996.

[9] Evans DA and Gadd CS. Managing coherence and context in medical problem-solving discourse. In Evans DA and Patel VL, editors, *Cognitive science in medicine: biomedical modeling*, page 21155. Cambridge, MA: MIT Press, 1989.

[10] K A Ericsson and P G Polson. An experimental analysis of the mechanisms of a memory skill. *Journal of Experimental Psychology:Learning, Memory, and Cognition*, 14(2):305–316, 1988.

[11] D M Gaba. Dynamic decision making in anesthesiology: Cognitive models and training approaches. In D A Evans and V L Patel, editors, *Advanced Models of Cognition for Medical Training and Practice*, pages 123–148. Springer Verlag, 1992.

[12] D M Gaba, K J Fish, and S K Howard. *Crisis Management in Anesthesiology*. Churchill Livingston, 1994.

[13] D M Gaba and S K Howard. Situation awareness in anaesthesiology. *Human Factors*, 37:20–31, 1995.

[14] G J Groen and V L Patel. The relationship between comprehension and reasoning in medical expertise. In M Chi, R Glaser, and M Farr, editors, *The nature of expertise*, pages 287–310. Lawrence Erlbaum Associates, 1988.

[15] T R Johnson, J Krems, and N K Amra. A computational model of human abductive skill and its acquisition. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates, 1994.

[16] Crossman J.A. Lebiere C Jones, R.M. and B.J. Best. An abstract language for cognitive modeling. In *Proceedings of 7th ICCM, Italy 2006*. 2006.

[17] G Jones, F E Ritter, and D J Wood. Using a cognitive architecture to examine what develops. *Psychological Science*, 11:1–8, 2000.

[18] R Jones and R Wray. Comparative analysis of frameworks for knowledge-intensive agents. *AI Magazine*, 27(2), 2006.

[19] G-M Joseph and V L Patel. Domain knowledge and hypothesis generation in diagnostic reasoning. *Medical Decision Making*, 10(1):31–46, 1990.

[20] K Keogh. A computational model of disturbance management in anaesthesia. Masters Thesis, Department of Computer Science, The University of Melbourne, 2002.

[21] K Keogh and E A Sonenberg. Analysis and visualisation of complex bahavioural data: A case study of disturbance management in anaesthesia. In *Proceedings of the HF2002 Human Factors Conference*. Swinburne University of Technology, 2002.

[22] K Keogh, E A Sonenberg, and J A Lawrence. Towards a computational model of disturbance management in anaesthesia. In *Proceedings of the Eighth Australian Joint Conference of Artificial Intelligence*. World Scientific Publishing, 1995.

[23] P R Krebs. Smoke without fire: What do virtual experiments in cognitive science really tell us? In *International Conference on Cognitive Science*, 2004.

[24] V L Patel and G J Groen. Solution strategies in medical reasoning. *Cognitive Science*, 10:91–116, 1986.

[25] V L Patel, G J Groen, and J F Arocha. Medical expertise as a function of task difficulty. *Memory and Cognition*, 18(4):394–406, 1990.

[26] V L Patel, G J Groen, M F Ramoni, and D R Kaufman. Machine depth versus psychological depth: A lack of equivalence. In E Keravnou, editor, *Deep models for Medical Knowlege Engineering*, pages 249–272. Elsevier, 1992.

[27] A Qusaibaty and N Howard. Intentional cognitive models with volition. Technical report, Center for Advanced Defense Studies, Directorate of Research, Center for Advanced Defense Studies, Washington DC, 2006.

[28] M Ramoni, M Stefanelli, L Magnani, and G Barosi. An epistemological framework for medical knowledge based systems. *IEEE transactions on systems, man and cybernetics*, 22(6):1361–1375, 1992.

[29] J Rassmussen. Skills, rules, and knowledge; signals, signs, and symbols, and other distinction in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13:257–266, 1983.

[30] M M Recker. Explorations in the parameter space of a model fit to individual subjects' strategies. Technical Report ER-MR 93-03, Georgia Institute of Technology, Atlanta, Georgia, 1993.

[31] P S Rosenbloom, J E Laird, and A Newell. *The Soar Papers: Research on Integrated Intelligence*, volume 1 & 2. MIT Press, Cambridge, Massachusetts, 1993.

[32] E M Roth, D D Woods, and Jr H E Pople. Cognitive simulation as a tool for cognitive task analysis. *Ergonomics*, 35:1163–1198, 1992.

[33] M J Schoelles and W D Gray. Top-down versus bottom-up control of cognition in a task switching paradigm. In D. Doerner F. Detje and H. Schaub, editors, *Proceedings of the Fifth International Conference on Cognitive Modeling*, pages 295–296. Bamberg, Germany: Universitats-Verlag Bamberg, 2003.

[34] M. J. Schoelles, H. Neth, C. W. Myers, and W. D. Gray. Steps towards integrated models of cognitive systems: A levels-of-analysis approach to comparing human performance to model predictions in a complex task environment. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pages 756–761, 2006.

[35] C.D. Schunn and W.D. Gray. Editorial. introduction to the special issue on computational cognitive modeling. *Cognitive Systems Research*, 3, 2002.

[36] H A Simon. *The Sciences of the Artificial*. MIT Press, 1969.

[37] T Simon and G Halford. Computational models and cognitive change. In T Simon and G Halford, editors, *Developing Cognitive Competence: New Approaches to Process Modelling*, chapter 1. Lawrence Erlbaum Associates, 1995.

[38] Howard SK, Gaba DM, Fish KJ, and Sarnquist FH Yang G. Anesthesia crisis resource management training: teaching anesthesiologists to handle critical incidents. *Aviation, Space, and Environmental Medicine*, 63(9):763–770, 1992.

[39] E A Sonenberg, J A Lawrence, and J Zelcer. Modelling disturbance management in anaesthesia: a preliminary report. *Artificial Intelligence in Medicine*, 4:447–61, 1992.

[40] E A Sonenberg, J A Lawrence, and J Zelcer. Keeping the patient asleep and alive: Investigating specialists' disturbance management in anaesthesia. In *Proceedings of the Second Australian Cognitive Science Conference*, pages 136–138, The University of Melbourne, February 1993.

[41] S Streufert and G Y Nogami. Analysis and assessment of planning: The view from complexity theory. In S L Friedman and E K Scholnick, editors, *The Developmental Psychology of Planning: Why, How, and When Do We Plan?*, pages 157–182. Lawrence Erlbaum Associates, 1997.

[42] R Sun. The clarion cognitive architecture: Extending cognitive modeling to social simulation. In R Sun, editor, *Cognition and Multi-agent Interactions: from Cognitive Modeling to Social Simulation*. Cambridge University Press, 2005.

[43] R Sun, L A Coward, and M J Zenzen. On levels of cognitive modeling. *Philosophical Psychology*, 2005.

[44] R Sun and C Ling. Computational cognitive modeling, the source of power, and other related issues. *AI Magazine*, 19(2):113–120, 1997.

[45] R Sun, X Zhang, and R Matthews. Modeling meta-cognition in a cognitive architecture. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, 2005.

[46] A ten Teije and E Rotterdam. Diagnostic reasoning with anaesthesia knowledge. *Systems analysis, modelling, simulation : Journal of mathematical modelling and simulation in systems analysis*, 33:327–345, 1998.

[47] D D Woods, L Johannsen, and S S Potter. Human interaction with intelligent systems: An overview and bibliography. *ACM SIGART Bulletin*, 2(5):39–50, 1991.

[48] D D Woods and E M Roth. Symbolic AI computer simulations as tools for investigating the dynamics of joint cognitive systems. In J-M Hoc, P C Cacciabue, and E Hollnagel, editors, *Expertise and Technology Cognition and Human-Computer Cooperation*, pages 75–92. Lawrence Erlbaum Associates, 1995.

[49] J Zhu. Understanding volition. *Philosophical Psychology*, 17(2):247–273, 2004.