Cogn Tech Work (2006) 8:269–282
DOI 10.1007/s10111-006-0044-x

ORIGINAL ARTICLE

Evaluating functional displays for hydropower system: model-based guidance of scenario design

Xilin Li · Penelope Sanderson · Rizah Memisevic · William Wong · Sanjib Choudhury

Received: 1 March 2006 / Accepted: 26 May 2006 / Published online: 5 October 2006

Abstract We discuss the human role in hydropower system control, noting how it is different from other supervisory control environments and noting the typical shortcomings in current displays provided to hydropower system controllers. We describe steps towards evaluating proposed functional displays with industry hydropower controllers whose time with us was limited to a few hours. This involves selecting test scenarios that may maximally demonstrate advantages, if any, of the functional displays. Here we propose a scenario design approach based on a simplified human-control loop model. Starting from a basic human control cycle: detect → analyse → act → evaluate, we identify scenarios in which the functional displays are expected to provide information in a more effective way than current displays. Based on this approach, the prototype of new functional display was tested with industry controllers. The experimental method and subjective results are reported, and the lessons we learned are discussed.

Keywords Human supervisory control · Hydropower system control · Functional displays · Scenario design · Human control model · Evaluation · Situation awareness · Trust

1 Introduction

Designing displays for hydropower system controllers is a particularly challenging undertaking. This is because the hydropower system controller's role encompasses many different kinds of systems that have to work together, and many different timeframes in which activity takes place. However no studies related to this have been done. Most display design work in power systems has been done on transmission network controllers (Overbye et al. 2002; Wiegmann et al. 2005) or on nuclear power plant controllers (O'Hara et al. 2004; Roth and Woods 1988; Vicente et al. 2004; Woods and Roth 1988). Neither of these is relevant for the hydropower plant operator, who must monitor and analyse vast amounts of data from diverse domains and on different time frames to control the energy source, to generate electricity, and to meet economic dispatch targets while keeping an eye on transmission status and constraints, and on changes in the market. Demands on the controller are exacerbated by the fact that hydropower companies often serve peak load, which is highly driven by dynamic market forces.

In this paper we describe the complexity of the hydropower system controller's work and outline

X. Li (✉) · P. Sanderson · R. Memisevic
School of Information Technology and Electrical Engineering, The University of Queensland,
St Lucia, QLD, Australia
e-mail: xilin@itee.uq.edu.au

P. Sanderson
e-mail: psanderson@itee.uq.edu.au

R. Memisevic
e-mail: rizah@itee.uq.edu.au

W. Wong
Interaction Design Centre, Middlesex University,
London, UK
e-mail: w.wong@mdx.ac.uk

S. Choudhury
Snowy Hydro Limited, Cooma, Australia
e-mail: sanjib.choudhury@snowyhydro.com.au

some shortcomings in the current information system with which controllers are typically provided. We briefly present displays we have developed in an attempt to overcome these shortcomings (Memisevic et al. 2005). We outline our steps towards an evaluation of the new displays with industry hydropower controllers whose time with us was limited to a few hours only. A key concern was the design of scenarios for the evaluation. Finally, the initial results of the experiment are reported and some conclusions drawn about the success of our process and lessons learned.

2 Role of the human controller

2.1 Elements of hydropower systems

Figure 1 gives a very simple summary of the different elements of a hydropower system. The energy source for generation is water, a valuable resource that must be used strategically to keep options open. Water can be an uncertain resource, depending upon weather, snow melt, runoff, and so on. Electricity generated is transmitted to the national electricity grid. However the generation of electricity is affected by operation of a deregulated electricity market, especially hydropower companies' wish to serve peak electricity demand which is highly dynamic. Moreover, the generation of electricity is constrained by irrigation requirements—water coming from power stations is used for irrigation and there are agreements and contracts that must be respected for how much water is released, where, and when. Therefore there are multiple constraints, some physical and others intentional (e.g. financial), some in the short term and others in the longer term.

2.2 Human supervisory control in hydropower systems

The role of the hydropower system human controller in a deregulated electricity market is different from the role of his counterpart in regulated or centrally controlled power systems. The differences are explored below, and the role of automation is outlined as well.

Figure 2 is a variant of the familiar Sheridan model of human supervisory control (Sheridan 1987). In the element at right labelled Controller at Company X, the human supervisor interacts with a task through displays and controls that couple him with a central supervisory control and data acquisition (SCADA) computer, which in turn drives remote terminal units (RTU) that interface with sensors and actuators at the physical level of control system. Models of human supervisory control have usually been confined to this element.

We expand the Sheridan (1987) model to illustrate some of the challenges of human supervisory control in a deregulated electricity market. Figure 2 shows that requests to make electricity (bids or rebids) must be sent to an independent market operator—The National Electricity Market Management Company (NEMMCO). Bids may have been set previously by traders at Company X, but sometimes controllers at the same company have chance to update the bids according to market dynamics. NEMMCO receives bids not only from Company X but also from many other companies participating in the market.

All the bids are submitted to an optimisation routine that is designed to match electricity supply with electricity demand at the minimum cost across the market for each dispatch interval (5 min). NEMMCO sends back to each company projections of how much electricity that company will be required to make, or "dispatch", in each dispatch interval for the rest of the day. Throughout the day, the controller at Company X

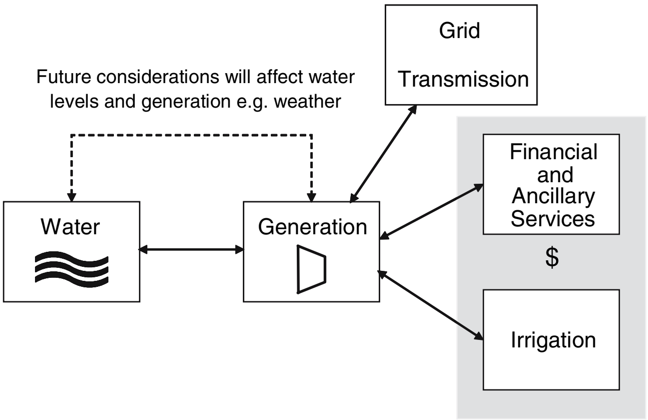


Fig. 1 Diagram of the relation between the different elements of a hydropower system

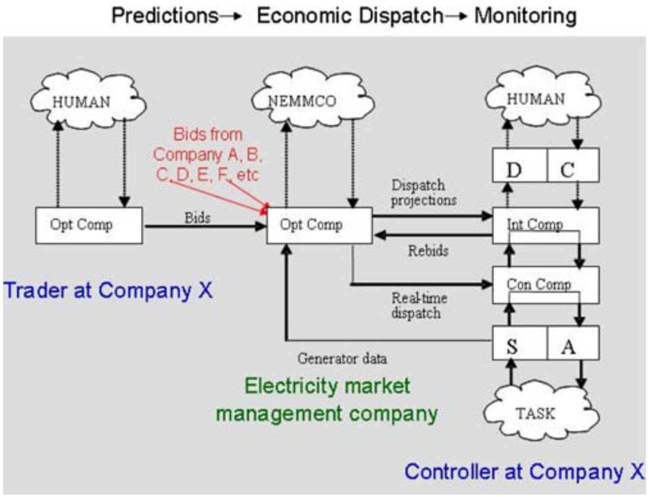


Fig. 2 Role of the human controller in the deregulated electricity market (adapted from Sheridan 1987). *D* display, *C* controls, *S* sensor, *A* actuator

prepares his generators in time to make the electricity requested. The Automatic Generation Control (AGC) driven centrally by NEMMCO controls all generator units inside the market, keeping the frequency of the power system around the nominal value to maintain the stability of the whole electricity network.

The human controller's role is to monitor this process to ensure it is done successfully. The controller intervenes where necessary to ensure an unbroken supply of the electricity the company has been dispatched to make for the market and to avoid penalties. The dynamic nature of the electricity market has led to profoundly different patterns of generation than before, and a need to respond to contingencies and opportunities much faster than before the market existed.

Given this environment, the human controller's work is a mixture of *reactive* and *proactive* control. The process of monitoring for failures and intervening is reactive. However monitoring also involves a proactive search for situations that might become problematic if not handled early. When detected, such problems lead to discussion, decision making and planning of future action, which often takes the form of changing when and from where electricity will be made.

2.3 Relationship between human supervisor and automation

The above description focuses mainly on generation and trading. As indicated in Fig. 1, hydropower system operations also cover the management of water storage and water diversions. In the longer term, there are also irrigation requirements to be met.

The levels of automation that support the above functions are vastly different. Because electricity travels at the speed of light, a high level of automation is needed to maintain power system stability because the human controller simply cannot respond quickly enough. The decision about which generating units will run within each ''aggregate unit'' of power stations is driven by attempts to optimise water use. This decision is also automated because of its complexity and because of the short timeframe in which it sometimes has to be made.

In contrast, water storage management takes place over hours and days in the shorter term and up to months and years in the longer term. Here the human controller makes decisions with the help of off-the-shelf software tools that help him predict future water levels. In particular, water surge is controlled by automation.

Finally, decisions about participation in the electricity market are determined strategically by company traders in consultation with controllers for input relating to plant safety and availability of resources. Controllers have to follow transmission constraints provided by NEMMCO on the basis of overall system stability and security concerns.

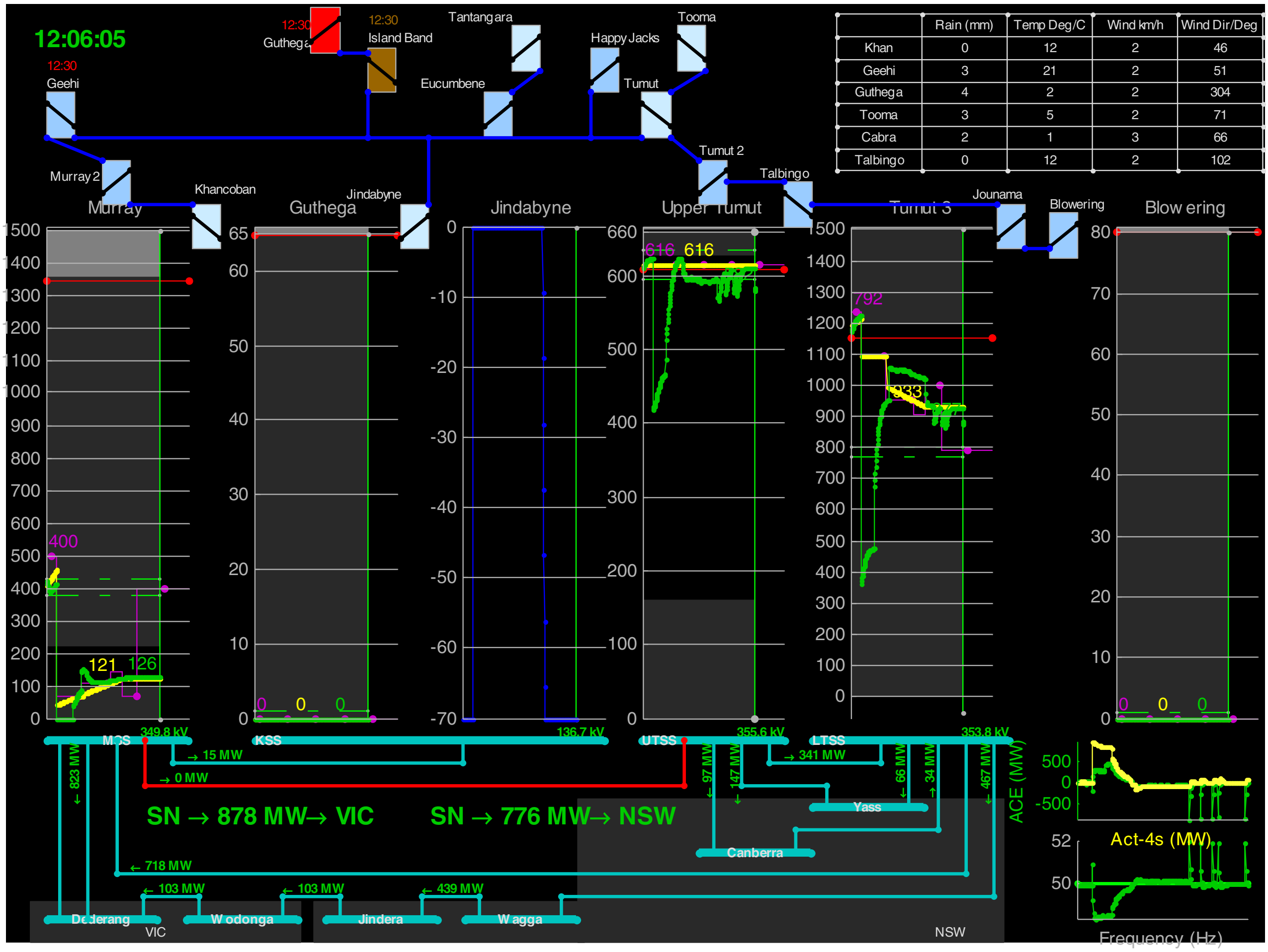
The current displays in the hydropower company control room extract data from SCADA and AGC systems. The displays present scheme status in separate physical systems such as the hydraulic network, generation, the electricity market and the transmission network. For example, to monitor water, controllers use a hydraulic overview display that lets them navigate information related to the water network and drill down to specific components such as gates and valves. However, the current displays do not provide information that integrates past, present, and future views, or that integrates the different subsystems shown in Fig. 1 in a way that allows higher-level properties to be seen. Decision making often depends upon such integrated views. Hollnagel has noted that a control room should be a "room with a view" of past, present, and future states (Hollnagel and Woods 2005). To achieve views that integrate information across the right timeframes, companies either buy further off-the-shelf applications or develop their own tools in house. Even so, it is seldom that the kind of integration is provided that allows controllers to assess consequences across the different subsystems shown in Fig. 1.

3 Functional displays

Figures 3 and 4 show two displays conceived by Memisevic in an attempt to overcome the above problems. The displays are the outcome of an analysis of the work domain and its temporal characteristics that is discussed in Memisevic et al. (2005) and Sanderson (2005). The displays are intended to provide an integrated view of past, present, and future, and in particular to provide a view of the future across timeframes appropriate for different kind of information, and to support controllers' awareness of operational constraints and boundaries. Detailed descriptions and a rationale for the display are forthcoming; here we provide enough detail to understand our description of the scenario design for a first evaluation of the displays.

The energy flow (EF) display in Fig. 3 is intended to support short-term (5–30 min) coordination of real time energy production. In the five general areas that run counter-clockwise from top right to bottom right, the display shows (1) information about weather, (2) information about the energy source (hydraulic network), (3) energy generation against targets (presented across the centre as generation within aggregate units of power stations), (4) energy distribution (transmission network), and (5) power quality.

The hydraulic network provides information about storage levels, whether they are increasing or decreasing, and time at which the level will reach a limit. Generation within each aggregate unit of power



	Rain (mm)	Temp Deg/C	Wind km/h	Wind Dir/Deg
Khan	0	12	2	46
Geehi	3	21	2	51
Guthega	4	2	2	304
Tooma	3	5	2	71
Cabra	2	1	3	66
Talbingo	0	12	2	102

Fig. 3 Energy flow (EF) display (Memisevic et al. 2005)

Fig. 4 Water efficiency (WE) display (Memisevic et al. 2005)

stations for 25 min is shown in each tall rectangular shape in the centre of the display. The functions across each aggregate unit represent 5 min energy targets from NEMMCO for each aggregate unit of the scheme. The actual output (green) of the aggregate unit is plotted across the horizontal dimension. Current constraints on generation are reflected in the different background greyscale values, which let controllers see the range of generation they have available for that aggregate unit, whether it could be immediately available from the generating units currently operating (black) or would require generating units being turned on or off (dark grey). Unavailable capacity is shown at top of an aggregate unit (medium grey). The maximum amount of generation that can come from an aggregate unit is limited by any temporary constraints in the transmission network. This maximum, the transmission constraint, is represented by a horizontal red line with red circles at its ends that sits across the aggregate unit affected. For example, in Fig. 3 there are transmission constraints on the first, fourth and fifth aggregate units. Further problems in the transmission network to the major regions are signalled in the network display at bottom left. At right two graphs indicate how the hydropower scheme's reserve generation is being used to contribute to the control of frequency in the network. Thus an integrated overall snapshot of scheme status is available for the controller.

The water efficiency (WE) display in Fig. 4 is one of several displays that show longer term information about a single storage across a day. The vertical axis represents water storage level (shown via a blue function across the display that is marked off into 30 min market intervals), inflow (a dark purple function), and efficiency (an orange function). The horizontal axis represents 24 h, running from left to right. A vertical green bar down the centre indicates the present moment. Therefore the past, present and future state of the water storage is evident, with future states based on the present and planned use of water through generation and any other water diversions. In addition, the maximum and minimum storage constraints are red horizontal lines at the top and bottom of the display. The target storage level according to the scheme daily operational plan is shown as a green horizontal line across the middle. Similar displays are available for all storages. They can all be viewed singly or can be collected into an integrated overview of smaller versions of Fig. 4.

A further display (not illustrated here) provides financial data over the same time frame as the WE display and in a similar 24 h display format. This display, the Revenue (RE) display, shows all gains and losses of the hydropower company for a whole day. The Revenue display supports a better understanding of the potential economic impact of scheme configuration.

4 Model-based scenario design

The goal of our research was to perform an initial evaluation of the functional displays described above. The evaluation was done at a participating hydropower company with industry controllers whose time with was limited to around 5 h only. For the evaluation, the controllers were to operate the scheme with a subset of the current displays (Current interface) or with the current displays supplemented with the new functional displays (Functional interface). Several scenarios were to be used, each around 30 min. A key concern was the scenarios to present in the experiment.

To achieve a thorough evaluation of novel display concepts, scenarios should capture any potential advantage of new displays as much as possible. Ideally, in our case, this would include most representative financial and operational risks that a generating company might face. The conflict between the ideal requirements for the evaluation and the limited timeframe imposed created a challenge for our design of the evaluation experiment and the scenarios within it.

4.1 Approach

Given the above challenge, we focused on two considerations. The first consideration was to identify emergencies occurring in the hydropower system that would be more effectively represented with the Functional displays than the Current displays. The second consideration was to identify how human controllers might interact with the displays.

Eight fundamental contingencies, listed in the left column of Table 1, were selected to cover the most typical financial or operational risks of a hydropower company as well as to highlight the anticipated advantages of the new displays. A simplified human control model was developed to capture information that the human controller would need for (1) detection, (2) analysis (understanding and projection), (3) determination of possible actions, and (4) evaluation of impact of action.

Most models of human-in-the-loop behaviour have similar phases relating to observation, selection of action, and action execution. For example, Boyd's well-known OODA loop (Observe → Orient → Decide → Act) breaks selection of action into two parts, orient and decide. In contrast, in our proposed model, we divided observation into two parts, observation focused on detecting problems and observation focused on evaluating the results of actions, for the following reason. As noted, hydropower systems are highly automatic, so human controllers only intervene when necessary. In contrast to an automatic control loop that is tightly coupled to changes in the external world, the human control loop is only loosely coupled to such changes. Therefore a key issue for successful human control is when the controller should intervene and when he should leave control with the automation or hand control back to the automation. The human-in-the-loop model used in our case should highlight the above. As a result, detection is general observation that helps controllers notice any contingency and evaluation is a more focused purposive observation to assess whether the system has recovered to normal after an intervention. This leaves us with a Detect → Analyse → Act → Evaluate model. This made it much easier for us to identify the details of the process by which a controller would respond to a contingency.

4.2 Proof of principle study

For each of the contingencies shown in Table 1, a domain expert performed a walkthrough to identify

Table 1 Contingency events proposed for scenarios with some advantages of the Functional displays for handling each contingency that emerged from the walkthrough

Contingency	Advantage of Functional display(s)
1. Failure of a transmission line within the HPS region	EF display shows immediate impact and participation of different aggregate units in the consequences
2. Failure of a generator belonging to HPS	EF display immediately shows whether dispatch target can still be met and supports immediate responding
3. Change in forecast inflow of rain and water runoff	WE display indicates whether a spill or low storage level will result and whether daily water target is met
4. Change in energy demand for some hours ahead	WE shows impact on water use and storage; RE display shows impact on earnings
5. Change of bid from a participant in another region	WE shows impact on water use and storage; RE display shows impact on earnings
6. Change in transmission constraint	EF display shows change immediately and indicates whether generation can be moved elsewhere
7. Change in market target for pumping	EF display will show change for next dispatch interval and how reserve is affected
8. Change in demand (sudden) that changes frequency	EF display will show deviation in frequency and the scheme's recovery process

Table 2 Walkthrough table for evaluating how effectively the Current versus Functional displays support human controllers during the contingency “Change of Transmission Constraint” (contingency 6)

	Detection	Analysis	Possible action	Evaluation
Current displays	1. Detect “price separation” on Energy Dispatch display	1. Identify the cause or type of contingency from market operator’s report and offline software	1. Follow the updated energy target immediately on Unit Auto Control display	1. Assess the deviation between target and current generation on Energy Target display to see whether target is recovered
		2. When in peak demand, might consider shifting generation to recover generation profile, such as browsing Hydraulic overview display	2. Possible rebid to NEMMCO on Merit Order display	2. Watch rebidding result on AUPAB display
Functional displays	1. See the loss of transmission capacity on the affected aggregate unit in Energy Flow display	1. Immediately identify the type of contingency and the affected unit on Energy Flow display	1. Same as above	1. Assess the target deviation and next target on Energy Flow display
		2. Visually determine the possible places for shifting generation, further refer to Water Efficiency display etc. to develop strategy for re-bidding and preparing for reorganising generation	2. Same as above	2. Same as above
				3. Monitor the impact on market income on Revenue display
				4. Examine storage configuration in terms of constraints and daily operational plan

exactly what information the Current display and the Functional display would each provide. The goal was to elicit any potential benefits from the WE and EF displays within each stage of the human control loop: Detect → Analyse → Act → Evaluate.

Table 2 provides results for a walkthrough of a single contingency called “change of transmission line constraint” under the two display conditions. The first row illustrates the Detect → Analyse → Act → Evaluate loop with the Current displays and the second row with the Functional displays. The Current displays referred to, and the order in which they were considered, is illustrated in Fig. 5. When using the Current displays, controllers must look across several displays from different subsystems to gain an overall picture and in addition would use separate software tools to calculate the potential impact or imposed constraints for preparing future actions. However, with the Functional displays there would be far less need to seek information across several displays or to use separate tools to perform integration. In what follows we provide more detailed information about how the displays would be used for the “change of transmission line constraint” example.

Detection The transmission contingency cannot be directly observed in the Current displays. Only the consequence of the contingency can be seen in the current NEMMCO Energy Dispatch tabular display as a so-called “price separation”, where prices increase rapidly and the dispatch target drops in the region affected by the constraint. In contrast, with the new EF display, transmission capacity associated with each aggregate level is represented as a red horizontal line. When the line moves down from the top it indicates the presence of an active constraint, as shown in Fig. 3.

Analysis After detecting the contingency, the controller has to identify the cause from NEMMCO’s Energy Dispatch display and plan possible interventions. If demand for generation is at a peak, then it may be necessary to shift generation to other units so that energy demand can continue to be met. The controller will look for information about storage levels, inflow, weather and the availability of generating units in order to develop a strategy for reorganising energy production across the scheme and re-bidding into the market. In the Current displays such as Hydraulic Overview, Transmission Network and so on the information provided is insufficient to support such planning. For example, the controller has to use a separate software tool to estimate how much capacity was lost with the transmission constraint. In

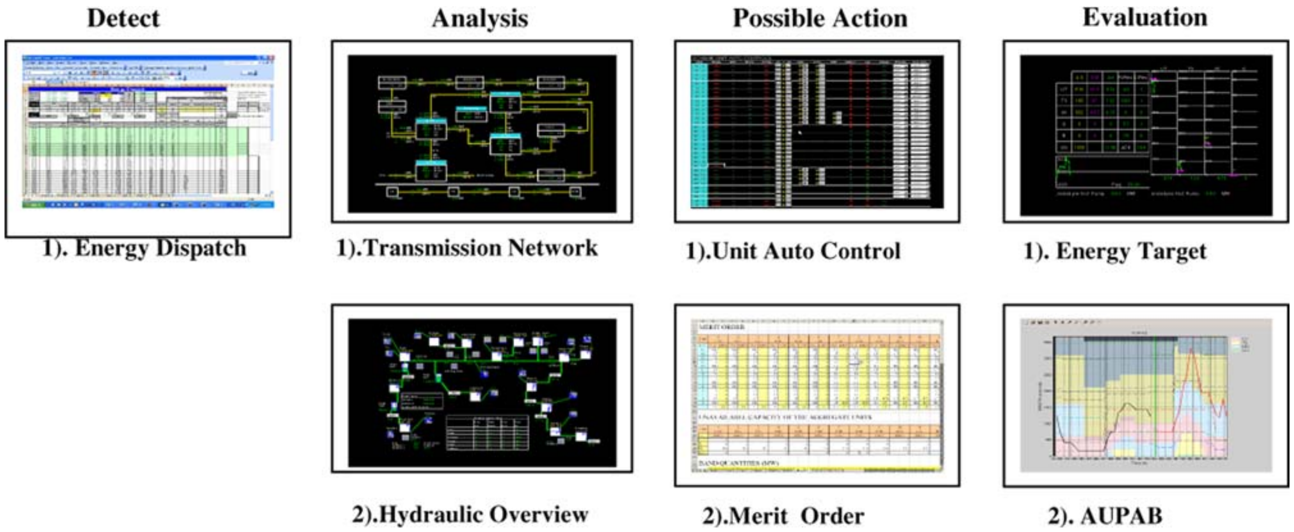


Fig. 5 The Current displays that controllers would walk through for handling contingency ‘‘change of transmission constraint’’

contrast, with the new EF display, with a single glance the controller can see how much capacity was lost and can see whether generation can be moved elsewhere in the scheme. For example, Fig. 3 shows that the leftmost aggregate unit still has a lot of reserve generating capacity. Furthermore, when planning to reorganise generation with the Current display the controller must use separate software to predicate what the storage levels will be at a fixed moment in the future. In contrast, the new WE display lets the controller see not only what the storage levels will be, but also whether future storage levels will fall outside the desired operational boundary and, if so, when. When such information is available for all storages, alternative plans can be drawn up much more easily. For example, Fig. 4 indicates that there will be excess storage levels under present conditions, so it is better to move generation to this power station right now.

Possible action The controller’s first response will be to adjust generation to follow the dropped target immediately to avoid a costly market penalty. Afterwards, he might change the company’s bid to try to restore the initial amount of generation being sold to the market. The displays used in this phase are all panels on which control actions are taken, which we have not redesigned. Therefore there is no difference between testing conditions in how control is exercised.

Evaluation To evaluate whether the generation profile is recovering as desired, with the Current displays the controller reads out the deviation of the target and actual output from an Energy Target display. However, with the Functional displays the EF display includes such information as well. If rebidding is needed, controllers using the Current display can only see the result of the new dispatch schedule for the rest of the day on a so-called AUPAB display, whereas with the new Revenue display (not shown here) controllers can visualise the financial impact of their new configuration for the rest of the day.

4.3 Scenarios

After the walkthrough, a table was created that summarised some of the ways that the Functional display may provide unique assistance to the controller when he handles each contingency (see Table 1). As illustrated above, the assistance involves either providing an integrated view of past, present, and future, allowing the controller to see the impact of the current conditions, or providing an integrated view of the interplay between different subsystems at present, allowing the controller to see immediate possibilities for actions.

The scenarios we used for the evaluation were constructed by combining one or more of the above contingencies into a coherent stream of events (see Table 3). The first criterion for selecting contingencies was that they could be plausibly combined with others. For instance, pumping is hard to integrate with many other contingencies, because a hydropower company does not pump when the price is high—instead, it generates. In contrast, a frequency deviation can occur at any time in the network. Frequency deviation is usually handled by automation so that such a contingency can be easily integrated with others. The second criterion for selecting contingencies was to see whether

Table 3 Profile of the final four scenarios used in the experiment

Scenario	Contingency	Season	Inflow	Peak demand	Description
1	1,4,8	Summer	Low	Yes	20% increase in long term demand in Vic and NSW, ancillary service increase, transmission line 65 failure
3	6,8,3	Winter	High	Yes	Transmission capacity from HPS to NSW dropped 90%, change in short-term demand, change in forecast inflow
4	2,3	Winter	High	Yes	Generation taken off in one aggregate unit, change in forecast inflow
5	7,8	Summer	Low	No	Change of pumping target, frequency deviation

Note For details of each contingency, see Table 1

any two contingencies might have a common or similar human response. For instance, controllers would respond similarly to contingencies such as changes in the bid of other region and changes in energy demand. Therefore we only used one of the above in constructing the final scenario.

Based on the above approach, the four complex scenarios in Table 3 were created. As well as distributing seven of the contingencies across the four scenarios, we had to set up profiles for weather, inflow and demand for each scenario. For example, the first scenario is that during peak demand, the current and forecast demand in two regions keep increasing, which provides a good opportunity to increase generation, but the inflow is so low that only a limited amount of water is available to the controller for generation. The only option is to operate the scheme on the boundary. In contrast, in the third scenario one aggregate unit fails in the peak demand period. However the inflow is high, so that the controller has plenty of water to use but the energy target is low. The controller must try to pick up generation by re-bidding into the market.

In summary, the overall rationale is not only to try to create a good range of scenarios, but also to increase the challenges facing the controller by creating conflicts that might require them to control the scheme on an operational boundary or to rebid into the market to try to recover profitability. The expected benefits of the new displays are most likely to be realised in such challenging situations where the direct perception of operational boundaries or constraints and the integration of high level properties would be very useful.

5 Measures of performance

The goal of measurement was to capture changes due to the new Functional displays in the most sensitive way. Both subjective and objective measures were used in the experiment. In this section we describe the subjective measures which tested how effective the new Functional displays were in the following areas (1) capturing the problem solving time frame as intended, (2) supporting situation awareness, (3) promoting trust and self-confidence.

Whether the intended time frame was accurately captured in each display was tested by asking participants to rate the time frame in which each display would be most helpful for understanding scheme status. Participants could select one or more of the following options: real time, next 5 min, next half hour, time until load changes significantly, rest of the trading day. One of our intentions with the Functional displays was to provide support for thinking effectively within and across timeframes. This is because hydropower systems have constraints operating in different timeframes (Memisevic et al. 2004; Memisevic et al. 2005; Sanderson 2005).

Situation awareness (SA) is a construct that encompasses awareness of system information (SA1), awareness of system state (SA2), and ability to project future states (SA3) (Endsley 2000). Gathering controllers' subjective opinion of how well the Functional displays improved their SA1, SA2 and SA3 allowed us to test how well they thought the displays conveyed scheme status and oriented them to the future, in particular.

Trust and Confidence questions tested controllers' views of how well the displays had helped them in each particular situation (Lee and Moray 1994; Lee and See 2004). Li et al. (2005) noted that to support controllers' trust that they are getting an accurate picture of scheme status, displays should represent information about system performance, underlying processes and high level purposes, which is what we aimed to achieve in the Functional displays. We also asked controllers about their confidence in their ability to control the scheme with the help of the displays.

Besides the above subjective measures, objective measures were also collected to systematically analyse participants' performance. Data logs were collected of scheme status, system events and controller actions and video was collected from freestanding camcorders as well as from a head-mounted camera on the controller. The objective data are still being analysed and will be discussed in future work.

6 Experiment

6.1 Goal

The goal of the experiment was to see whether the new Functional displays would lead to better controller SA, trust and confidence, and whether the intended timeframes represented in the new displays would be successfully conveyed to controllers. The details of the experiment are given below.

6.2 Method

6.2.1 Participants

Consistently with the way roles are allocated in a real hydropower system control room, each run of the experiment included two participants—a controller and a coordinator. There is some overlap between controller and coordinator roles, but, in general, controllers are in charge of hydropower scheme operation whereas coordinators assist controllers and handle market changes. Controllers generally serve as coordinators before becoming controllers. Similarly, coordinators have generally had experience in engineering and technical aspects of scheme operations before becoming coordinators.

The study reported here involved four controller-coordinator pairs. The participants were all experienced professionals who had worked in a large hydropower system control room for a period from around 1 year (the most junior coordinator) to over 10 years (the most senior controller).

6.2.2 Apparatus and simulation

Figure 6 illustrates our experimental configuration. It is relatively unusual for hydropower companies to have full-scale training simulators (unlike the nuclear power industry). Because our industry partner did not have their own simulator we developed a medium fidelity simulator in MATLAB® that simulated hydropower scheme operation, including the water network and storages, generating units, the external market environment and the electricity network (Memisevic et al. 2004).

The simulation ran on two connected laptop computers (3.05 GHz CPU processor, 992 Mb memory, 15″ screen and 1,280 by 1,024 pixel display resolution). As seen in Fig. 6, each laptop drove two large (19″) additional screens through a Digital Tiger video extension card. The laptop screens showed displays used to make control actions, and screens that supported rapid navigation across the different displays. The large screens showed displays used for monitoring only, including the new Functional displays. Participants could rearrange screens as they wished.

The controller used one laptop and the coordinator used the other. The functions in each laptop were allocated according to the principal responsibilities of the controller and the coordinator in the control room. The simulation ran on Windows 2000 and had 36 display pages in total to represent the ‘‘existing’’ interface and all pages mimicked the functionality of the ones in the real control room.

The simulation had an embedded data capture program that captured over 7,000 variables every 4 s about the simulated hydropower system status and about the displays accessed. The data were stored in a MATLAB® matrix file on each laptop and are the source for objective measures of control performance. In addition, two video cameras captured the use of the different displays and comments from participants. As shown in Fig. 6, the controller also wore a small head-mounted camera to track the general direction of gaze during the scenario (Omodei et al. 1997).

6.2.3 Experimental design

A within-subjects design was used in the experiment. Each controller-coordinator pair experienced the four test scenarios shown in Table 3, each around 30 min in length. The order of scenarios was counter-balanced

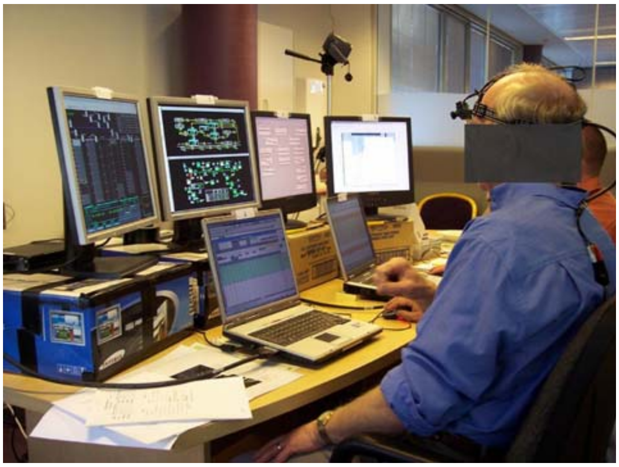


Fig. 6 Experimental setup showing controller (*foreground*) and coordinator (*background* at *right*)

across participant pairs by using a Latin Square. The counterbalancing was arranged so that for each scenario, two participant pairs ran the scenario with the Current displays and the other two participant pairs ran the scenario with the additional Functional displays. The Current display condition was always used in the first two scenarios and the Functional displays in the last two scenarios. The reason for not counterbalancing the order of displays was our limited time with controllers. Our goal was to give participants as much time as possible to become familiar with the simulator's instantiation of the Current displays, and how to make control actions with the control screens, before adding the less-familiar Functional displays. This is a potential confound from the purely experimental point of view but we opted for it in order to acquire as much valid information about performance with, and opinion about, the Functional displays as possible.

6.2.4 *Procedure*

The experiment ran for around 5 h for each participant pair and involved five phases: (1) introduction, (2) familiarisation, (3) scenarios, (4) post-scenario questionnaire, and (5) post-experiment questionnaire. These are described in more detail below.

Introduction In the first 30 min, participants were presented with some sample snapshots of the simulator screens and the new Functional displays. The rationale and a full description for each new display were provided and participants' questions were answered.

Familiarisation This section was about 45 min. With the experimenter's assistance participants worked through a sample scenario to learn some basic operations for running the simulation, such as how to use menus or toolbars to navigate through displays, how to activate a monitor display and move it to a certain position, how to start/stop generation in a generating unit, how to adjust the voltage levels and so on. In almost all cases the procedures were the same as in the real control room, but because there were fewer screens than in the control room participants required some time to review screen navigation and learn idiosyncrasies of the simulator environment.

Scenario The details of the scenarios are given in Table 3. Prior to running each new scenario, participants were given a daily operational plan and an inflow forecast report for reference. These are tools that are provided at the start of a shift in the real control room and they act as orientation to the goals and areas requiring close attention. Participants were also encouraged to communicate as they would in the control room, in a natural task-related conversation.

Post-scenario questionnaire When each scenario was over, there were two phases of questioning. First, the participants worked as a team to answer specific questions about the scenario they had just experienced. The questions focused on whether and when participants recognise the approaching constraints or contingency, how they evaluate the situation and what the consequent control options would be, and what the future scheme status would be like, so exploring their SA status in detail. Second, participants completed a written questionnaire individually and independently of the other participant. They were asked to rate each overview monitor display for which time frames it supported, and to rate their overall trust and self-confidence level and so on. The post-scenario questionnaire took around 25 min.

Post-experiment questionnaires At the end of the experiment, participants individually and independently answered a written questionnaire in which they self-reported their overall SA, trust and self-confidence level on 7-point Likert scales. They were also asked to select their five preferred displays to go on the wallboard in the control room from a set of nine overview displays that included the principal Current and Functional displays. Finally, participants gave open-ended comments on the Functional display as well as on the scenarios and the experiment itself.

6.3 Results

Because of a last-minute shortage of available coordinators, in one experimental session a coordinator who had run in a previous session stood in for a missing coordinator. As a result of his own caution, due to his prior knowledge of the scenario, and a misguided hint from an experimenter intended to prevent the coordinator acting on prior knowledge of the scenarios, the coordinator behaved more passively and did not respond to the controller's clear need for help as effectively as he had done in his previous session. For this reason, we judged that this particular participant pair did not experience the experiment as intended, and that as a result the teamwork between controller and coordinator was not representative of normal teamwork in the control room. Hence we were compelled to drop that participant pair in the data analysis. As a result there are 4 (scenarios) $\times$ 3 (participant pair) = 12 valid trials in the experiment.

As noted, in this paper we focus on participants' subjective responses. Participants' answers on post-scenario as well as on post-experiment questionnaires are analysed as discussed in the sections below.

6.3.1 Post-scenario questionnaire results

Intended time frame. Most participants agreed on the time frames for which the new displays were created (see Table 4 for results). For example, 100 and 83% participants reported using EF display for real-time and short-term control, respectively. Participants reported using the WE display more broadly; with 50% participants using it for planning short-term bidding and 50% for developing a rebid for the rest of day when a contingency occurs. In their free-form comments, two controllers commented that they needed more time to evaluate the WE display, mainly because the WE display was implemented on the coordinator's laptop and so was less directly accessible to the controllers, although quite visible.

Trust/confidence. On a 7-point Likert scale (1 = much less; 7 = much more) in which 4.0 indicated indifference between the two display conditions, participants reported a better overall trust in getting a more complete and accurate picture of the Scheme when the new Functional displays were added (mean rating = 4.83) compared to just with Current displays (mean rating = 4.08); $t(5) = -2.09$, $P = 0.09$ (one-tailed tests are used as we have a directional hypothesis). Participants also had better self-confidence in their ability to control the scheme with the Functional displays (mean rating = 5.25) than with the Current displays (mean rating = 4.33); $t(5) = -2.2$, $P = 0.08$.

6.3.2 Post-experiment questionnaire results

Situation awareness. A 7-point Likert scale was used with the mid-point rating again indicating indifference with respect to each display condition. Ratings significantly higher than 4.0 were not found for SA1, but were found in favour of the Functional displays for SA2 (mean rating = 4.83); $t(5) = 2.19$, $P = 0.08$; and SA3 (mean rating = 5.3); $t(5) = 6.5$, $P = 0.003$. Thus, participants considered that the Functional displays had a greater effect on their comprehension of system state and on their ability to predict future states than on straightforward perception of system information.

Table 4 Percentage of participants allocating each of the Functional displays to the timeframe indicated at left

Time frame	EF (%)	WE (%)
Real time	100[a]	33
Next 5 min	83[a]	50[a]
Next half hour	33	50[a]
Time until load changes significantly	0	33[b]
Rest of the trading day	13	50[a]

Note Multiple choices were allowed

[a] The time frame that each display was intended to support

[b] The timeframe in question may not have been relevant for scenarios selected

However the post-experiment questionnaire did not reveal any differences between displays in overall trust and self-confidence among participants. From participants' comments, this may be due to some anomalies in the simulation data compared with real scheme data, caused by some of the approximation techniques used and our limited access to more commercially sensitive scheme data for use in the simulator.

Wallboard display preferences. Five of the six participants selected the EF display for use in an ideal wallboard, and four of the six participants selected the WE display for use in the ideal wallboard. These ratings were the second and third highest ratings respectively out of the nine possible displays listed. The highest rating was for the existing hydraulic overview (HO) display. These results suggest that (1) the Functional displays successfully complement the Current displays and (2) there may be room for further enhancements of the EF display to provide hydraulic information.

In their comments, most participants agreed that both Functional displays had potential and that they could see their advantages. However, they commented that because the evaluation time was so compressed it was hard for them to ‘‘set up their direction’’ (of control), incorporate the new displays into their routines, and fully test their potential advantages. The participants felt they needed more time and greater realism to fully evaluate the value of new displays. In addition, many participants commented that they also needed to test the displays on the equivalent of a normal day, rather than a highly abnormal day, to see how the displays indicate where they are and where they going.

7 Discussion

Overall, the questionnaire results indicate that the new Functional displays are starting to work as intended. The expected advantages of the Functional displays are evident in the subjective reports. Evident in the participants' subjective reports are comments about the benefits of providing an integrated scheme view of past, present and future, revealing the imposed constraints and the consequent impact of current situations, providing an integrated view of the interplay between different subsystems, and visualising possible action spaces. Participants using the Functional displays reported a better awareness of the significance of current scheme state, a better projection of the future status in terms of planned goals, a greater trust in getting a complete and accurate picture of the scheme, and a greater self-confidence in their ability to control the scheme. Moreover, the intended time frames for which new displays were created were successfully conveyed to controllers.

Given the above, the scenarios that we used in the experiment have clearly gone some way towards achieving the goal of highlighting expected advantages of the new displays. The fact that controllers' opinion of the Functional displays was generally positive is in contrast to many such evaluations, where initial controller reaction to new advanced displays can be negative, even when the new displays reveal performance benefits (Vicente et al. 1996).

However, our experiment had limitations that we hope to correct in future research. First, our scenarios may have ''over-exercised'' the displays. In order to create situations in which the effectiveness of the displays would be tested in a way that created data, we compressed several contingencies into scenarios of 30 min. Therefore, what would normally be considered major contingencies or even major system accidents could occur in relatively close succession, which is certainly not usually experienced. As a result, participants found it difficult to set up their direction. If possible, the follow-on evaluation should have longer scenarios and should include some relatively uneventful scenarios.

Second, although we took pains to introduce participants to the displays and to help them become familiar with navigating the simulation, our time was still too short. Most participants noted that they did not have enough time to become familiar with displays and the navigation, which made the events in the scenarios even more daunting to handle. Although the Functional displays were intended in particular to support controller adaptation to abnormal situations (Rasmussen et al. 1994) a phased evaluation moving from normal situations to abnormal situations would be preferable. In addition, limitations of the simulator in some places had a negative influence on participants' performance and on their trust that they were getting an accurate picture, and therefore on their ability to draw conclusions about the displays.

7.1 Future work

Our remaining work is to complete a systematic analysis of the objective simulation data and the video data. The objective measures to be used are (1) control quality, (2) activity measures, and (3) measures of control adaptation. The results of the objective measures are expected to reveal more about the effectiveness of the Functional displays, despite the noisy data.

Control quality includes proactive measures of how well the controller configures the plant to meet incoming demands. Figures of merit will be revenue, water efficiency and ability to meet daily water goals. Figures of merit will also include how well the target is followed in real time, as reflected in the deviation of actual energy output from the dispatch target, and the application of any market penalties.

Activity measures include simple observables such as the time to detect a problem, time after detection to act, and time for action to take effect.

Control adaptation measures explore patterns of variability in controllers' responses. As discussed elsewhere (Hajdukiewicz and Vicente 2005) small variability within or between participants in achieving high level system goals combined with full use of the discretion available in actual control actions may indicate good human-system coupling; the opposite can be interpreted as poor human-system coupling.

Finally, we hope to run a follow on evaluation using an updated version of the simulation and displays, but with different scenarios. Test scenarios would be much longer and would be phased, starting with scenarios that represent more normal kinds of operation and then proceeding to abnormal situations.

7.2 Summary

Hydropower system controllers work across multiple domains with different properties and constraints, and they monitor and control processes whose time constants range from being instantaneous to running over minutes, hours, and days. If controllers are to maintain effective situation awareness and exercise effective control, their displays must integrate their problem solving across domains and time frames. The Functional displays we have developed and are evaluating are a first step towards this goal.

Evaluation with industry controllers familiar with conventional displays must be carefully designed so as to give novel display concepts the best possible opportunity to reveal their advantages. The proposed model-based steps towards designing scenarios proved to have some effectiveness because participants reported some benefits of using the Functional displays. However the scenarios used may not have tested the displays as sensitively as they might, due to the short exposure time, the density of unusual events, and perceived anomalies in the simulator. The ''ecological validity'' of any evaluation is a function of the match between the expertise of the participants, the realism of the work situations simulated and the fidelity of the work tools, rather than just one of those elements (Rasmussen et al. 1994). Getting the balance right is the major challenge for evaluation of new display concepts.

Acknowledgments We acknowledge the support of the Australian Research Council through ARC SPIRT grant C00107069, the support of Snowy Hydro Limited and the comments from controllers and coordinators at Snowy Hydro Limited.

References

Endsley MR (2000) Theoretical understandings of situation awareness: a critical review. In: Endsley MR, Garland DJ (eds) Situation awareness analysis and measurement. Erlbaum Associates, Mahwah

Hajdukiewicz JR, Vicente KJ (2005) What does computer-mediated control of a thermal-hydraulic system have to do with moving your jaw to speak? Evidence for synergies in process control. Ecol Psychol 16(4):255–285

Hollnagel E, Woods DD (2005) Joint cognitive systems: foundations of cognitive systems engineering. CRC, Boca Raton

Lee JD, Moray N (1994) Trust, self-confidence, and operators adaptation to automation. Int J Human-Comput Stud 40(1):153–184

Lee JD, See KA (2004) Trust in automation: designing for appropriate reliance. Hum Factors 46(1):50–80

Li X, Sanderson P, Memisevic R, Wong W, Choudhury S (2005) Applying the control adaptation method to a real world system: hydropower system example. In: Proceedings of the Annual Conference of the European Association of Cognitive Ergonomics (EACE2005), Crete, Greece, pp 11–18

Memisevic R, Choudhury S, Sanderson P, Wong W (2004) Integrated power scheme simulator for human-system integration studies. In: Proceeding of the Australian Universities Power Engineering Conference (AUPEC04), Brisbane, Australia, 6 p

Memisevic R, Sanderson P, Choudhury S, Wong W (2005) Work domain analysis and ecological interface design for hydropower system monitoring and control. In: Proceeding of the IEEE Conference on Systems, Man, and Cybernetics (IEEE-SMC2005), Hawaii, USA, pp 3580–3587

O'Hara JM, Higgins JC, Persensky JJ, Lewis PM, Bongarra JP (2004) Human factors engineering program review model (NUREG-0711): US Nuclear Regulatory Commission

Omodei MM, Wearing AJ, Mclennan J (1997) Head-mounted video recording: a methodology for studying naturalistic decision making. In: Flin R, Strub M, Salas E, Martin L (eds) Decision making under stress: emerging themes and applications. Ashgate, Aldershot, pp 137–146

Overbye TJ, Yan S, Wiegmann DA, Rich AM (2002) Human factors aspects of power system visualizations: an empirical investigation. Electr Power Compon Syst 30(8):877–888

Rasmussen J, Pejtersen AM, Goodstein LP (1994) Cognitive systems engineering. Wiley, New York

Roth EM, Woods DD (1988) Aiding human performance I: cognitive analysis. Trav Hum 51(1):39–64

Sanderson P (2005) Shapes of human control in time: models and a hydropower system example (invited address). In: Proceedings of the Annual Conference of the Human Factors and Ergonomics Society of Australia [HFESA2005], Canberra, Australia

Sheridan TB (1987) Supervisory control. In: Salvendy G (ed) Handbook of human factors and ergonomics. Wiley, New York

Vicente KJ, Moray N, Lee JD, Rasmussen J, Jones BG, Brock R et al (1996) Evaluation of a Rankine cycle display for nuclear power plant monitoring and diagnosis. Hum Factors 38(3):506–521

Vicente KJ, Mumaw RJ, Roth EM (2004) Operator monitoring in a complex dynamic work environment: a qualitative cognitive model based on field observations. Theor Issues Ergon Sci 5(5):359–384

Wiegmann DA, Essenberg GR, Overbye TJ, Yan S (2005) Human factors aspects of power system flow animation. IEEE Trans Power Syst 20(3):1233–1240

Woods DD, Roth EM (1988) Aiding human performance II: from cognitive analysis to support systems. Trav Hum 51(2):139–172