Int. J. Human–Computer Studies (1994) **40**, 153–184

Trust, self-confidence, and operators' adaptation to automation

JOHN D. LEE
Battelle Seattle Research Center, Seattle, WA 98105, USA

NEVILLE MORAY
Engineering Psychology Research Laboratory, Department of Mechanical and Industrial Engineering, University of Illinois, Urbana, IL 61801, USA

(Received 14 December 1992 and accepted in revised form 13 September 1993)

The increasing use of automation to supplant human intervention in controlling complex systems changes the operators' role from active controllers (directly involved with the system) to supervisory controllers (managing the use of different degrees of automatic and manual control). This paper examines the relationship between trust in automatic controllers, self-confidence in manual control abilities, and the use of automatic controllers in operating a simulated semi-automatic pasteurization plant. Trust, combined with self-confidence, predicted the operators' allocation strategy. A Multitrait-multimethod matrix and logit functions showed how trust and self-confidence relate to the use of automation. An ARMAV time series model of the dynamic interaction of trust and self-confidence, combined with individual biases, accounted for 60.9–86.5% of the variance in the use of the three automatic controllers. In general, automation is used when trust exceeds self-confidence, and manual control when the opposite is true. Since trust and self-confidence are two factors that guide operators' interactions with automation, the design of supervisory control systems should include provisions to ensure that operators' trust reflects the capabilities of the automation and operators' self-confidence reflects their abilities to control the system manually.

1. Introduction

1.1. SUPERVISORY CONTROL, AUTOMATION, AND TRUST

As the technological sophistication and complexity of systems increase, the role of the human operator changes from direct manual control to management of different levels of computer control. In many systems the magnitude of the forces, the physical danger and the required precision, together with the time constants of the systems, combine to make direct physical control inappropriate (Sheridan, 1987; Woods, O'Brien & Hanes, 1987). Therefore, humans assume the role of a supervisory controller, interacting with the system through different levels of manual and automatic control (Sheridan & Johannsen, 1976). Just as a traditional supervisor interacts with subordinate human associates, the supervisory controller of a modern system interacts with computerized subordinates. These subordinates translate humans' directives into actions, and report on the consequences (Sheridan, 1987).

Very little work has addressed how operators of complex systems adopt strategies for allocating automatic and manual control. However, several authors (Halpin,

1071-5819/94/010153 + 32$08.00/0

Johnson & Thornberry, 1973; Sheridan & Hennessy, 1984; Muir, 1988) have suggested that operators' trust in the automation may play a major role in guiding their allocation strategy. Just as the relationships between humans depend on trust (Barber, 1983; Rempel, Holmes & Zanna, 1985), the relationships between supervisory controllers and subordinate machines may also depend on trust. In addition to trust, the choice of automatic or manual control probably depends on a number of factors such as risk, responsibility, management policy, and mental workload (Boettcher, North & Riley, 1989), but trust seems to play a prominent role. In documenting accounts of the interaction between automation and operators, Halpin, Johnson and Thornberry (1973), Zuboff (1988) and Muir (1988) all suggest that operators' subjective feelings of trust towards automation play an important role in the effective use of automation. When operators mistrust automation they may use it inappropriately (Muir, 1988).

In contrast to the case-study based approach of both Zuboff and Halpin *et al.*, Moray and Muir developed the first laboratory-based study to investigate the role of trust in mediating human–machine relationships in a supervisory control situation (Muir, 1989). Muir's research demonstrated a strong correspondence between changes in operators' trust in automation and operators' reliance on automation. Lee and Moray (1992) and Lee (1992) provide a summary of her results.

Contrary to Muir's findings, Lee and Moray (1992) showed that under certain conditions changes in the operators' reliance on the automation failed to correspond to changes in the operators' trust. Lee and Moray speculated that reliance on automation may depend not only on operators' trust in the automation, but also on operators' self-confidence in their own abilities. This paper presents the results of a study that investigated the relationship between trust, self-confidence, and operators' reliance on automation.

1.2. SELF-CONFIDENCE

Just as operators' trust in automation may influence their reliance on automation, so too may their self-confidence influence their reliance on manual control. Operators' self-confidence (anticipated performance during manual control) may interact with their trust in the automatic system to guide the allocation policy. If operators' self-confidence fails to correspond to their actual abilities then they may allocate automation inappropriately, just as mistrust may lead to an inappropriate allocation strategy.

Much research has addressed self-confidence and the results suggest that people are often overconfident in their abilities, both in forecasting future events (Fischhoff & MacGregor, 1982), and in their general knowledge (Fischhoff, Slovic & Lichtenstein, 1977). In those studies bankers, executives, civil engineers, clinical psychologists, and psychology graduate students, among others, all displayed overconfidence in their abilities. See Lichtenstein, Fischhoff, and Phillips (1982) and Yates (1990) for reviews. Furthermore, in many cases subjects' overconfidence increased as they received more information during a judgment task. For example, in the context of clinical diagnosis, Oskamp (1982) showed that additional information only slightly improved subjects' judgment accuracy, but it led to a large increase in their confidence in their judgment. Fischhoff (1982) suggests that the general bias towards overconfidence represents a robust effect, and is more than an artefact of contrived laboratory conditions.

Interpreting these results in terms of operators' interaction with automation, it seems that when operators engage manual control self-confidence should rise markedly, even without accompanying increases in performance. However, this increase will be more pronounced when the operator perceives he or she is successful in controlling the system. At the same time, operators would become less and less likely to change allocation strategies and delegate control to automation. Because overconfidence may be more than just an artefact of laboratory experiments, miscalibrations in operators' self-confidence may prove to be a substantial impediment to the effective use of automation. If operators consistently overestimate their capabilities then they are likely to maintain manual control, failing to benefit from the capabilities of the automation, even though the system design will usually have been optimized under the assumption that the plant will normally be run under automatic control.

Moray and Lee (1992) showed that it is possible to develop a quantitative model of operators' trust. The latter varied as a function of the reliability of the plant and the productivity as perceived by the operator, and showed some inertia from trial to trial. In this paper we extend their approach to the interaction of trust and self-confidence as determinants of manual intervention. The concepts of self-confidence and trust may provide a basis for understanding the factors that guide operators' allocation strategy when acting as supervisory controllers of a complex system. Previous research on self-confidence has shown that humans often develop miscalibrations; in particular, they often tend towards overconfidence. This, combined with the possible miscalibrations of trust, points towards potential problems in the human operators' relationship with automation. Although other factors may be important in guiding the choice of automatic or manual control (Boettcher, North & Riley, 1989; Kirlik, 1993), this research focuses on the relationship between trust, self-confidence, and the use of automation. Moreover, we examine this relationship in the context of operators adapting their allocation strategies to accommodate faults affecting automatic and manual control.

2. Methodology

The following paragraphs describe both the simulated system and the operators' task. The description of the simulated system includes a general discussion of the characteristics of the simulation and their implications for the generalizability of the experimental results. In addition, it includes a description of the relationship between the control mechanisms, as well as an explanation of the control algorithm that governs the automatic feedstock pump. The feedstock pump develops a fault and degrades plant performance. A description of the operators' task follows. This description explains how operators could interact with the system, the experimental design, and the subjective questions to which the operators responded.

2.1. PROCESS CONTROL TASK

A simulation of a semi-automatic process control plant was adopted with the expectation that findings from experiments using this system may generalize to other semi-automatic systems. The term "semi-automatic system", as used in this paper, describes a system whose subsystems can be operated under automatic or manual control. In other words, these systems are supervisory control systems, where operators need not use manual control, but normally allow the plant to run automatically, intervening only if they judge it necessary. While this simulation is orders of magnitude less complex than actual systems, we hope that it captures some of the essential elements of the supervisory control task.

The orange juice pasteurization plant adopted for this experiment was derived from the plant used by Muir (1989) and Huey (1989). Unless otherwise noted, the experimental procedure and apparatus used in this experiment matched the procedure and apparatus used in the experiment reported by Lee and Moray (1992). The only differences were the type of fault and its occurrence during the experiment, the control algorithm of the automatic feedstock pump, and the addition of subjective ratings of self-confidence. While this plant incorporated many of the complexities of an actual process control plant, and was based on a description of a real plant, it represented a somewhat simplified and abstracted version of actual process control plants. As such, it was not a replication of any real process control plant, but instead it was a "microworld" (Brehmer, 1990, 1992), designed to capture some of the complexities of the actual work domain, but in a way that allowed a greater degree of experimental control.

The simulation included sufficient complexity so that operators could achieve their goals in a variety of ways. For example, a variety of combinations of automatic and manual control could have been equally effective in controlling the plant. In addition, performance of the system as a whole depended on the complex interrelation of subsystems in a way that replicated the complexity of real systems. Because it included some of the complexity of actual work situations, the experiment imposed cognitive and motivational demands on subjects similar to those experienced in actual work situations. For this reason the results of these experiments should be more generalizable to actual work domains than simple laboratory tasks. At the same time, the lack of complete experimental control may limit the interpretation of the results. As an investigation using a microworld, that provided the operators with a wide variety of possible strategies, this research represents an attempt to bridge the gap between highly controlled but unrepresentative laboratory research and observational field studies.

2.2. CHARACTERISTICS OF THE SYSTEM

Figure 1 depicts the display that the operators saw when they controlled the simulated pasteurization plant. This mimic diagram also shows the relationships between the different subsystems of the plant. For example, the flow of mass and energy through the system, the level of the input vat and the successful pasteurization of the juice depended only on the feedstock pump rate, the steam pump rate, and the steam heater setting. The feedstock pump rate specified the juice flow completely and the steam pump rate specified the flow of steam completely. Manipulating the steam heater setting did not specify a directly observable variable. Instead, it specified the amount of heat energy transferred to the steam from the steam heater. Each of these controls operated either under manual control or by engaging an automatic controller. Whether controlled manually or automatically, the controls responded with a lag, changing slowly to reach the target value. This lag was most severe with the feedstock pump, which required about 10 iterations (18 s) to reach a new setpoint. The steam pump, on the

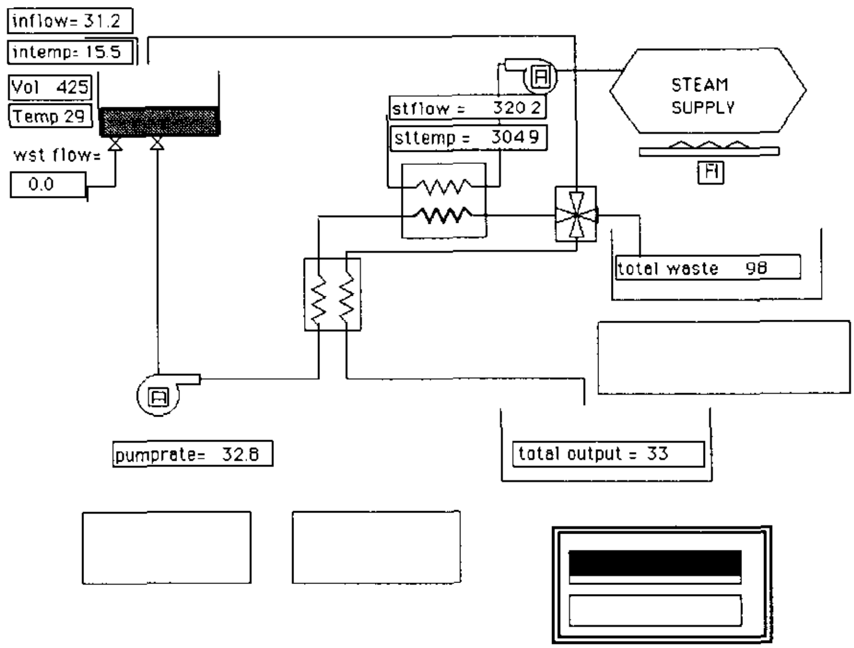


FIGURE 1. The mimic diagram of the pasteurization plant.

other hand, reached a target setpoint in about 2 iterations (3.6 s), while the steam heater reached target setpoint in 5 iterations (9 s). These lags became important as the operator tried to maintain a constant level in the input vat and constant temperature of the juice leaving the heaters. For example, if the input vat level began to drop because the feedstock pump was running too quickly, the lag in the feedstock pump might (if it was set at a very high rate) have made it impossible to reduce the pump rate quickly enough to avoid depleting the input vat, causing the plant to shut down. Likewise, the lags with the steam heater and steam pump led operators to waste juice because their attempts to raise or lower juice temperature were thwarted by a gradual response of the steam heater and the steam pump to operator commands.

Each automatic controller operated under its own control algorithm, acting independently of the other automatic controllers. For example, the automatic controller of the feedstock pump adjusted the rate of the feedstock pump independently of the settings of the other controllers. This algorithm depended solely on the level of the input vat and did not consider other aspects of the system, such as the final state of the juice flowing out of the heating subsystem. Likewise, the control algorithms for the steam heater and steam pump depended solely on the temperature of the juice flowing out of the active heat exchanger, and not upon the settings of the feedstock pump, steam heater or steam pump.

Because of the interactions between the heating and storage subsystems, the algorithms of the automatic controllers could have easily produced an unstable system. With some time constants the automatic controllers in the heating subsystem might not have been able to compensate for the increases in the feedstock pump rate. This would lead to juice being returned to the input vat, and further increases in the feedstock pump rate. Eventually, a perpetually overflowing input vat would result, as the feedstock pump increased its rate to bring the level to an acceptable range, and the heat transfer required to match the mass flow would fall further and further behind. In fact, the time constant of the feedstock pump was substantially larger than the time constant of the steam heater or steam pump, so the automatic controllers maintained a stable input vat level.

The automatic controllers of the steam heater and the steam pump, on the other hand, generated a limit cycle when they operated together, never producing a stable temperature setting. The source of the problem stemmed from the fact that both controllers relied on the same error signal; the temperature of the juice at the three-way valve. Both subsystems compensated for the error and the combined compensation was too large. This interaction of the two automatic controllers resulted in an unstable system and some waste of juice, as it was sometimes overheated and burnt. At other moments, the heat transfer was insufficient and the juice was recycled. Even with these inefficiencies, the automatic systems working together maintained the input vat level near the middle of its range, seldom reaching the danger limits.

2.3. AUTOMATIC FEEDSTOCK PUMP CONTROL ALGORITHM

Table 1 illustrates the control algorithms used in the experiment reported by Lee and Moray (1992) and in the current experiment. Table 1 shows that, in the original experiment, automatic control of the target feedstock pump rate (*Ptar*) continually changed in response to an error signal, based on deviations of the input vat level and the flow into the input vat (*Vatin*). These fluctuations led to alternately burning and recycling the juice. To maintain a constant output temperature, the amount of energy imparted to the juice by the two heat exchangers fluctuated continuously. The lags inherent in the process made compensation for the fluctuations in flow rate, using either automatic or manual control of the heaters, impossible. Therefore, efficient control of the system (efficiencies above 90%) was very difficult to achieve using the automatic controller of the feedstock pump.

TABLE 1
The control algorithms of the feedstock pump used in Lee and Moray (1992) and the current experiment

Used in Lee and Moray (1992)		Current experiment	
Input vat range	Control relation	Input vat range	Control relation
>750	$Ptar = Ptar + 17$	>750	$Prate = Prate + (setpoint + 10 - Prate)/3$
600–750	$Ptar = Vatin + 8$	600–750	$Ptar = Ptar + 2$
400–600	$Ptar = Vatin + 3$		
350–400	$Ptar = Ptar - 3$	300–600	$Prate = Prate + (setpoint - Prate)/2$
150–350	$Ptar = Vatin - 8$	150–300	$Ptar = Ptar - 2$
<150	$Ptar = Ptar + 16$	<150	$Prate = Prate - 8.23$

Ptar = target pump rate, *Prate* = current pump rate, *Vatin* = inflow rate to the input vat, *setpoint* = mean inflow rate to the input vat.

The control algorithm used in the current experiment remedies this situation. While the two algorithms share similarities, modifications incorporated in this experiment provided a constant feedstock pump rate (*Prate*) for a range of vat levels. With an input vat level of between 300 and 600 the automatic pump maintained a constant flow rate. Levels of the input vat above and below these levels triggered a change in the feedstock pump rate, which brought the input vat level back to the acceptable range. The more consistent flow through the heat exchangers led to less wasted juice and a generally more efficient plant, compared to the algorithm used by Lee and Moray (1992). Besides incorporating a range of constant flow rate, the algorithms for the pump controller in this experiment adjusted the flow more efficiently when the input vat level drifted high or low. With these changes, efficiencies above 90% were possible using the automatic controller of the feedstock pump and manual control of the other variables. However, because of interactions among the three subsystems, the use of all three automatic controllers would have led to efficiencies between 75% and 80%. Although a better tuned set of automatic controllers may have enticed more operators to rely on automatic control, the inefficiencies of the automation used in this experiment helped to reveal how operators might intervene and manage imperfect automation.

2.4. THE OPERATORS' TASK

Before operating the system, operators received an extensive written description of their objectives in controlling the plant, the possibility of faults, and the thermo-hydraulic processes involved in the control of the plant (see the Appendix). During the first hour of the first day each operator spent 10 trials learning to control the plant. During these trials each operator controlled the plant on alternate trials using only manual control or only automatic control of the feedstock pump, the steam pump, and the steam heater. After the training trials, operators controlled the plant as they liked, switching between automatic or manual control of these three subsystems whenever desired. Each of the first 10 trials lasted for 3 min. After the training trials operators controlled the plant for a further series of trials, each of which lasted 6 min until 20 trials had been completed. This required about 2 h. During a second session the operators returned and completed a series of 18 6-min trials. This required 2 h.

Like actual operators, operators in this experiment balanced the completing goals of safety and performance, using automatic control, manual control, or any combination of the two. Performance was measured by the amount of successfully pasteurized input divided by the total input. Operators received bonuses (10¢/trial) for achieving a performance above 90%. Safety, on the other hand, depended on operators maintaining sufficient volume in the input vat. If the vat emptied (volume = 0.0) the plant shut down, and operators lost all the accrued rewards. If the vat overflowed, then the juice that overflowed was classified as waste. This wasted juice was included when calculating the operators' payments: any wasted juice reduced the efficiency of the plant, reducing the operators' chance of getting a bonus. At the end of each trial, the computer calculated the net productivity of the pasteurization system and displayed it to the operators.

After the training trials, operators could control the three subsystems using manual control, automatic control, or any combination of automatic or manual control, to manipulate the flow rates and temperatures of the juice and steam to maximize plant performance. Operators could request automatic control from the keyboard for the whole trial or any part of a trial, switching between automatic and manual control as they wished. With complete reliance on the automatic controllers of all the subsystems (feedstock pump, steam pump and steam heater), the system produced juice at an efficiency of 75–80%. Under manual control, operators entered changes to the pump and heater settings from the keyboard, and the subsystems would then try to reach these "target" values. Because the three subsystems could be controlled either by using automatic or manual control, a wide variety of control strategies were available to the operators. (In Muir's (1989) experiment only the feedstock pump could be manually controlled.)

The display in Figure 1 shows all the information viewed by the operators. This display showed all the information operators received about the plant. This information included quantitative information about the temperatures and flow rates, and the levels of the vats. In addition, the display included qualitative information, such as whether the juice was in the proper temperature range, and the mode of the controllers (automatic or manual). This display indicated the level of the input vat graphically as the height of the shaded area of the input vat on the mimic diagram. In addition to illustrating the volume of the input vat graphically, changes in the flows through pipes were displayed as changes in the color of the pipe. Normally the pipes were black; when juice flowed through them they changed color to indicate flow. The graphical representation of the plant linked the system state variables to a mimic diagram of the plant to facilitate an unambiguous perception of the plant state. The display also included a simple annunciator, located in the lower left of the screen. This panel contained a flashing message whenever the input vat level entered a danger zone. An auditory signal accompanied the annunciator message.

While this display contained all the information needed to control the plant, the information presented to the operators was not exhaustive. For example, the display presented only qualitative information (too hot, too cold, pasteurized) concerning the juice temperature at the three-way valve. In addition, the display did not contain the steam heater setting or the temperature of the juice between the two heat exchangers. The additional information might have aided operators in their task, especially the information about the heater setting. Withholding this, and other information, increased the complexity of the task, forcing operators to infer the behavior of the subsystems.

Besides controlling the simulation, operators were responsible for logging data about the process. The data logging task required operators to record three system variables every 15 s. The purpose of this task was to replicate some of the other responsibilities of a process control task. Muir (1989) reported a tendency for operators to adopt complete manual control, so this task increased operators' workload to encourage them to use the automatic controllers. The anticipation of future plant states, combined with the task of integrating information across the system, made the additional task of data logging substantial, especially when the system failed to conform to operators' expectations and became unstable. As will be seen, data logging was not entirely successful in forcing the operators to use automatic control.

A total of 12 (mixed male and female) undergraduate University of Illinois students participated in this study. Each operator completed two 2-h sessions and was paid \$3.50 an hour with an additional bonus on performance. By using naive operators it was possible to track the development of trust, self-confidence and control strategies as operators were trained with new equipment. Although the operators were initially naive, the stable control strategies they developed by the second day indicated an understanding and a pattern of controlling the system that might be comparable to trained operators of more complex systems.

2.5. INTRODUCTION OF FAULTS

On the second day, faults were introduced into the operation of the plant. The first six trials on the second day were normal, containing no faults. The following 12 trials contained two types of faults. One type affected only the manual control of the feedstock pump, and the other affected only the automatic controller of the feedstock pump. The 12 subjects were divided into two groups, with one group experiencing six trials with a fault in the manual controller, followed by six trials with a fault in the automatic controller. The other group experienced six trials with a fault in the automatic controller, followed by a fault in the manual controller. Figure 2 shows the locations of faults in the plant, while Figure 3 shows the experimental design.

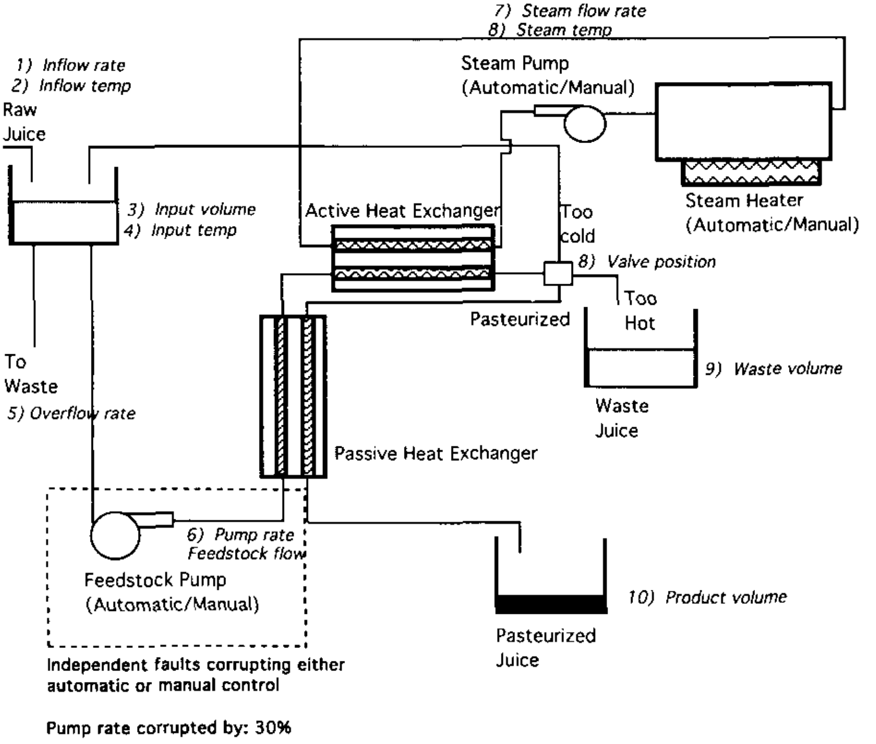


FIGURE 2. An illustration of the plant with the location of the faulty subsystem encircled by a dotted line.

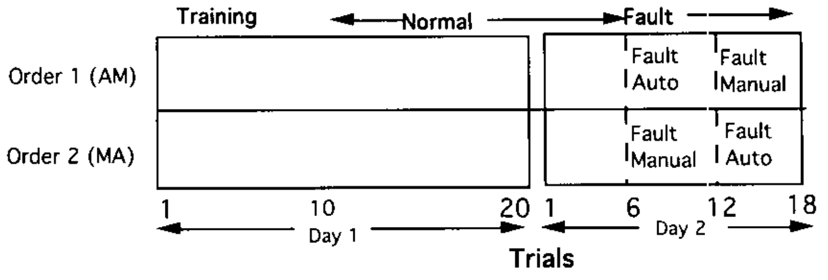


FIGURE 3. The two experimental groups and the occurrence and type of faults.

The fault in the automatic controller corrupted the pump rate by randomly adding or subtracting 30% of the pump rate. This caused the feedstock pump rate to fluctuate widely while under automatic control. The fault in the manual control corrupted the operators' specified pump rate by randomly adding or subtracting 30% of the specified pump rate. While these faults were designed to disrupt the control of the system, they were easily circumvented by the operators if they changed their allocation strategy. For example, if a fault occurred with the manual controller the operator could allocate automatic control and avoid its consequences.

2.6. SUBJECTIVE QUESTIONS

The subjective questions differed in two ways from those in Lee and Moray (1992). First, the questions not only addressed operators' trust in the automation, but required them to rate their self-confidence. Second, the subjective questions were more specific, referring to operators' trust and self-confidence in the feedstock pump, steam pump and steam heater, as opposed to the overall system. Because these questions included both trust and self-confidence, a new set of instructions and example scales were developed. The Appendix includes a copy of these instructions. This experiment excluded the ratings of Predictability, Dependability and Faith used in an earlier study (Lee & Moray, 1992), and added questions concerning self-confidence.

3. Results

3.1. A GENERAL OVERVIEW AND A QUALITATIVE ANALYSIS OF ALLOCATION STRATEGIES

Figures 4–7 summarize operators' performance, subjective ratings and reliance on automation. Figure 4 illustrates the efficiency of the plant, and the number of times the plant inadvertently shut down due to the operator emptying the input vat. Figures 5 and 6 show the change in trust and self-confidence during the experiment. Figure 7 shows the average use of the automatic controllers for each subsystem. The

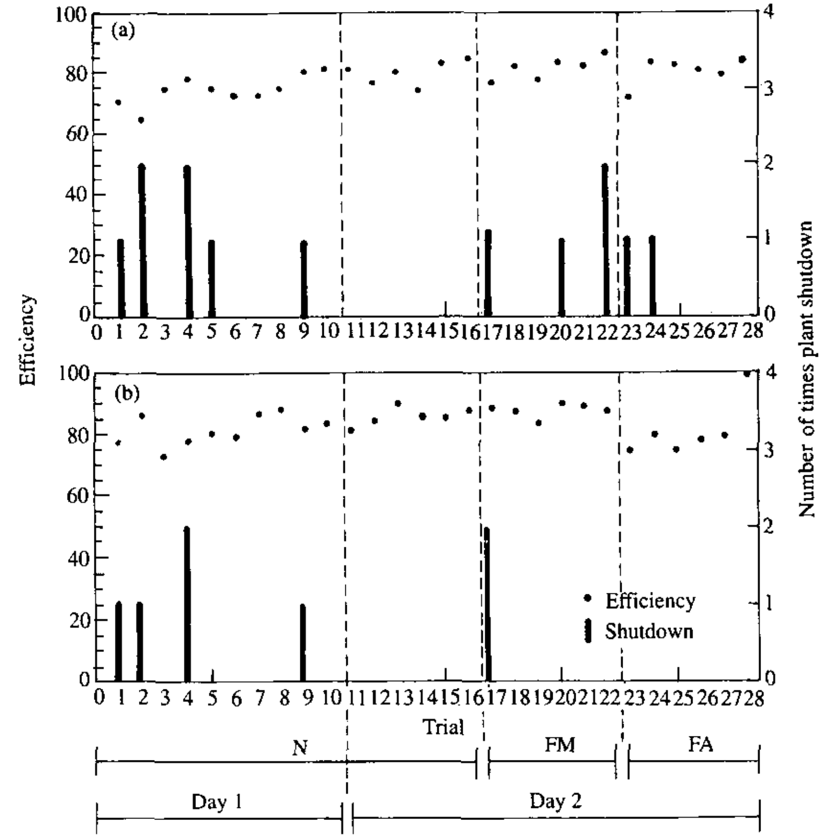


FIGURE 4. The fluctuation of plant efficiency and the occurrence of plant shutdowns, for the two experimental groups: (*a*) order 1: FA followed by FM; (*b*) order 2: FM followed by FA. Each point represents data from six operators. (N = normal operations, FM = fault affecting manual control, FA = fault affecting automation.)

data on all of these graphs start with the first trial following the training, and represent the average of six operators. Each set of graphs represents data from two groups. One group experienced a fault with the automation followed by the fault with the manual control, while the other experienced a fault with the manual control followed by the fault with the automation. Figures 4(*a*), 5(*a*), 6(*a*), and 7(*a*) all represent data from the group experiencing a fault with the automation, followed by a fault with the manual control. Figures 4(*b*), 5(*b*), 6(*b*), and 7(*b*) all represent data from the group experiencing a fault with manual control, followed by a fault with the automation. The faults first appeared on the seventh trial of the second day, trial 17 on the graphs. This paper contains only a qualitative analysis of the data, but Lee (1992) contains a detailed discussion of a repeated measures analysis of variance performed on this data.

Figure 4 shows that operators quickly learned how to control the system, and by the end of the training trials they were able to produce an average efficiency of 83.7%. These graphs show that performance generally increased as the experiment progressed, but fell slightly during the occurrence of faults. More specifically, in

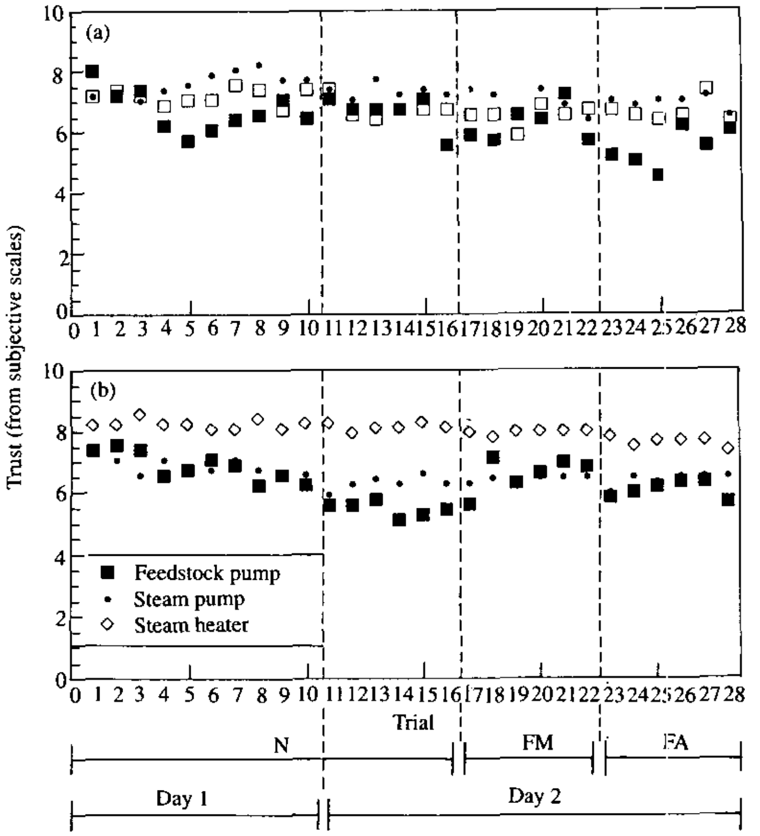


FIGURE 5. The fluctuation of trust in the automatic controllers, for the two experimental groups: (*a*) order 1: FA followed by FM; (*b*) order 2: FM followed by FA. Each point represents data from six operators. (N = normal operations, FM = fault affecting manual control, FA = fault affecting automation.)

Figure 4(*a*) efficiency remained nearly constant, while the only shutdowns during the second day occurred during the two faults. Figure 4(*b*) shows a slightly different pattern, with efficiency dropping slightly when the fault occurred with the automation. These graphs illustrate that the faults had only a minor effect on the performance of the system, reducing efficiency slightly and increasing the number of shutdowns slightly. Trust, on the other hand, seems to have decreased slightly as the experiment progressed, while self-confidence rose. The faults did not affect operator trust dramatically, at least when the data were pooled in this way. Figure 7(*a*) and Figure 7(*b*) show quite different allocation strategies for the two groups. Figure 7(*a*) shows a generally greater reliance on manual control, particularly for the steam heater. This allocation strategy is reflected in Figures 5(*b*) and 6(*b*), where trust greatly exceeded self-confidence for the steam heater. Because these graphs represent data pooled over six operators, it is not clear how they relate to the actual strategies of operators.

Table 2 illustrates the average amount of time spent using the automatic controllers by each of the 12 operators during the first six trials of the second day.

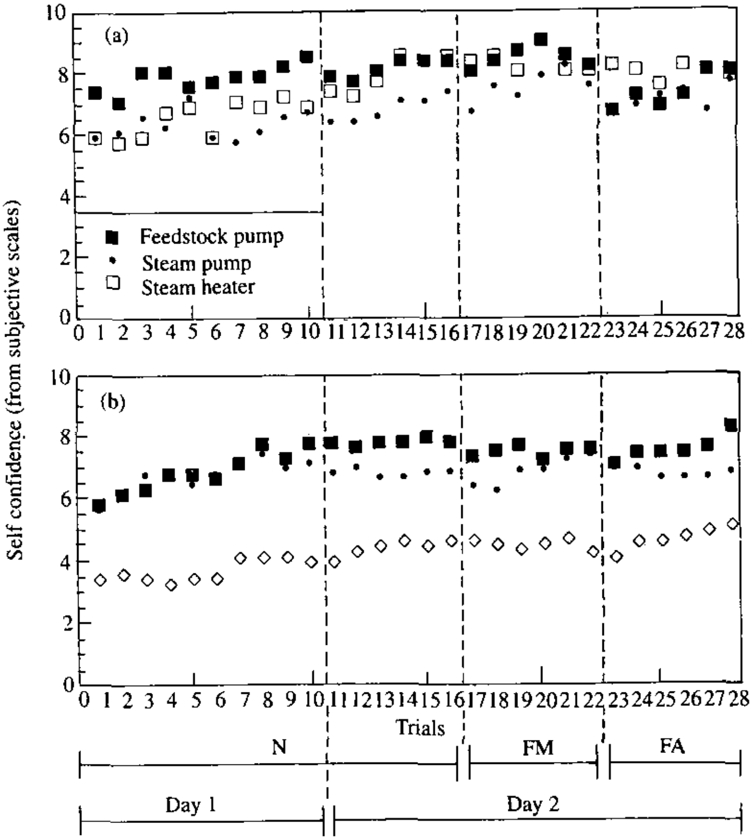


FIGURE 6. The fluctuation of self-confidence in the ability to control the subsystems manually, for the two experimental groups: (*a*) order 1: FA followed by FM; (*b*) order 2: FM followed by FA. Each point represents data from six operators. (N = normal operations, FM = fault affecting manual control, FA = fault affecting automation.)

While one of the goals of this experiment was to encourage a greater reliance on the automatic controllers, Table 2 shows that this experiment was only partially successful in meeting this goal. Only three operators adopted automatic control of the feedstock pump early in the experiment. Instead, nearly all operators depended heavily upon either the automatic controller of the steam heater, or the automatic controller of the steam pump. Only two operators chose to control the system entirely with manual control. This table shows that operators seldom engaged the automation for only part of a trial. Before the faults they either relied on the automation nearly 100% of the time, or they assumed complete manual control. Only four operators failed to follow this pattern.

Table 3 provides a more detailed description of the allocation strategies of each operator (1–12). This table shows how the 12 operators allocated automatic and manual control over the course of the experiment, during normal trials and those with faults. The first column describes how operators engaged the automation. The letter "A" denotes greater than 85% reliance on the automation, and the letter "M" denotes primarily manual control, where automation is used less than 15% of the

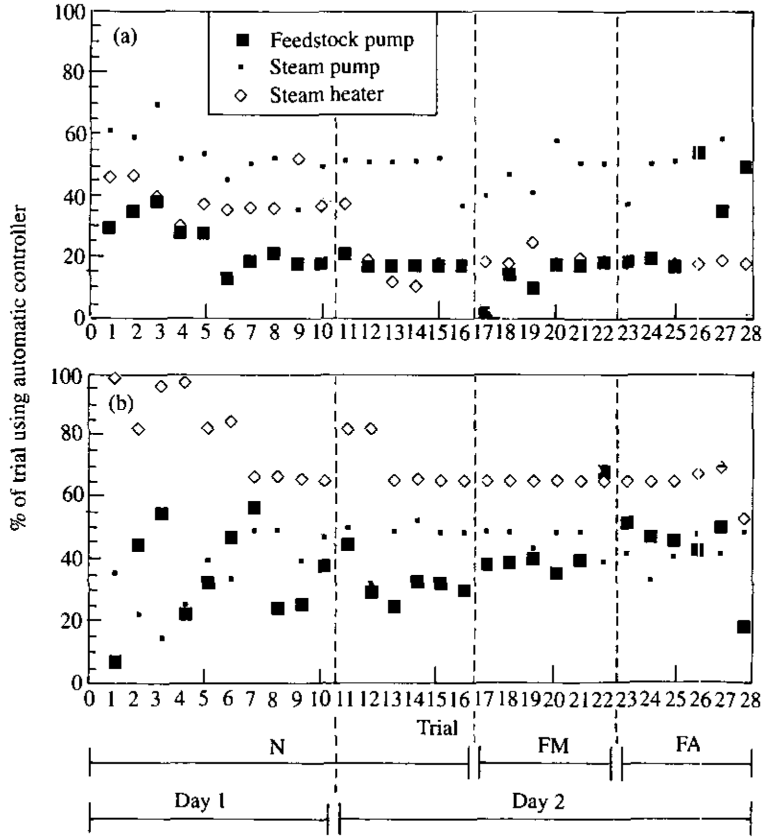


FIGURE 7. The fluctuation of the use of the automatic controllers, for the two experimental groups: (*a*) order 1: FA followed by FM; (*b*) order 2: FM followed by FA. Each point represents data from six operators. (N = normal operations, FM = fault affecting manual control, FA = fault affecting automation.)

time. The letters "MX" mean that the operator used some combination of automatic and manual control. For example, the first line of the table shows that operator 1 delegated control of the feedstock pump and the steam pump to the automation, and controlled the steam heater manually during the normal trials and the trials where a fault affected manual control. When a fault plagued the automatic controller of the feedstock pump, operator 1 controlled the feedstock pump and the steam pump with a mixture of automatic and manual control, while leaving the steam heater under complete manual control. Interestingly, the majority of operators adopted a dichotomous allocation pattern, with nine out of 12 operators selecting some combination of all manual or all automatic control of each subsystem during normal operation. When faults affected automatic control, operators who had adopted nearly complete automatic control of the feedstock pump switched to a mixture of manual and automatic control. However, when the fault affected manual control most operators maintained manual control with only 3 operators (operators 5, 6 and 7) shifting to automatic control to circumvent the fault. While Tables 2 and 3 and Figures 4–7 provide a general understanding of how the operators interacted

TABLE 2
The average use (in percent) of the automatic controllers by each of the 12 operators during the first six trials of the second day, with the standard deviation in parentheses

Operator	Feedstock pump (S.D.)	Steam pump (S.D.)	Steam heater (S.D.)
Group AM			
1	99.75	86.16	0.75
	(0.61)	(33.85)	(0.88)
2	25.41	100.00	18.08
	(36.85)	(0.00)	(40.14)
3	0.83	1.83	85.25
	(0.81)	(1.12)	(20.45)
4	91.50	5.25	100.00
	(9.57)	(8.72)	(0.00)
5	4.41	100.00	2.00
	(7.93)	(0.00)	(4.89)
6	8.58	1.58	100.00
	(7.53)	(1.16)	(0.00)
Group MA			
7	0.75	2.75	3.47
	(0.41)	(1.29)	(1.68)
8	15.75	100.00	100.00
	(9.77)	(0.00)	(0.00)
9	0.75	3.58	4.75
	(0.41)	(2.31)	(1.69)
10	3.41	0.75	100.00
	(3.78)	(1.83)	(0.00)
11	0.75	100.00	18.83
	(0.52)	(0.00)	(39.96)
12	59.83	83.33	16.66
	(26.27)	(40.82)	(40.82)

with the system, they do not provide a good basis for investigating how the subjective ratings related to the use of automation, or how faults affected the operator's trust, self-confidence, and control strategies. These issues are addressed in the following sections.

3.2. TRUST, SELF-CONFIDENCE, AND THE RELIANCE UPON AUTOMATIC CONTROLLERS

This section investigates the relationship between trust, self-confidence and the use of automatic controllers in four different ways. First, a qualitative description of the reliance upon the automatic feedstock pump by a single operator provides an intuitive basis for understanding the factors affecting the subjective measures and their relationship to the use of the automatic controllers. Second, the correspondence between the use of automation and the difference between operators' subjective ratings of trust and self-confidence is considered in a Multitrait-multimethod approach. Third, the functional relationship between the difference between trust and self-confidence and the use of automation is addressed. Fourth,

TABLE 3

The allocation strategies of operators 1–12, where each row represents a different combination of automatic and manual control and the numbers represent the operators who adopted those allocation schemes

Feedstock pump	Steam pump	Steam heater	Normal	Fault with manual	Fault with automatic
A	A	M	1	1	
A	MX	MX			4
A	M	A	4	4	
MX	A	A	8	8	8
MX	A	MX	2		
MX	A	M		2, 5, 12	
MX	MX	M			1, 2, 12
MX	M	A		6	6
MX	MX	MX	12		
MX	M	M		7	
M	A	M	5	11	5
M	M	A	3, 6, 10, 11	3, 10	3, 10, 11
M	M	M	7, 9	9	7, 9

M represents an average of less than 15% reliance on the automation, A represents a greater than 85% reliance on the automation, and MX represents a mixture of automatic and manual control (greater than 15%, but less than 85%).

an analysis of instances that do not correspond to this general pattern illustrates the possibility of exploratory behavior. Fifth, a time series analysis illustrates the dynamics that guide the use of automation over time. Each of these approaches provides a different perspective on the relationship between the operators' subjective feelings and the reliance upon automation.

3.2.1. *A qualitative analysis of the use of automation*

Visual inspection of Figure 8 suggests that the allocation of function changed as the difference between trust and self-confidence changed. Before the faults occurred, the operator relied primarily on manual control, using the automatic controller for less than 20% of the trials. During these trials, trust in the feedstock pump declined slightly and the operator's self-confidence increased, surpassing trust by trial 5. From trial 18 to trial 24, a fault affected manual control. This fault corrupted commands to the feedstock pump. With the fault affecting manual control, self-confidence declined, while trust in the automatic controller increased, surpassing self-confidence on trial 18. Following this shift in trust and self-confidence, the allocation of functions changed from nearly complete manual control to nearly complete automatic control. During the last six trials of the experiment, the automatic controller of the feedstock pump failed. This failure of the automation resulted in wide fluctuations in the feedstock pump rate. After a lag of several trials, trust declined, self-confidence increased, and the operator switched back to manual control of the feedstock pump.

Like several operators in this experiment, operator 6 switched between nearly complete manual and complete automatic control. Other operators used a mixture

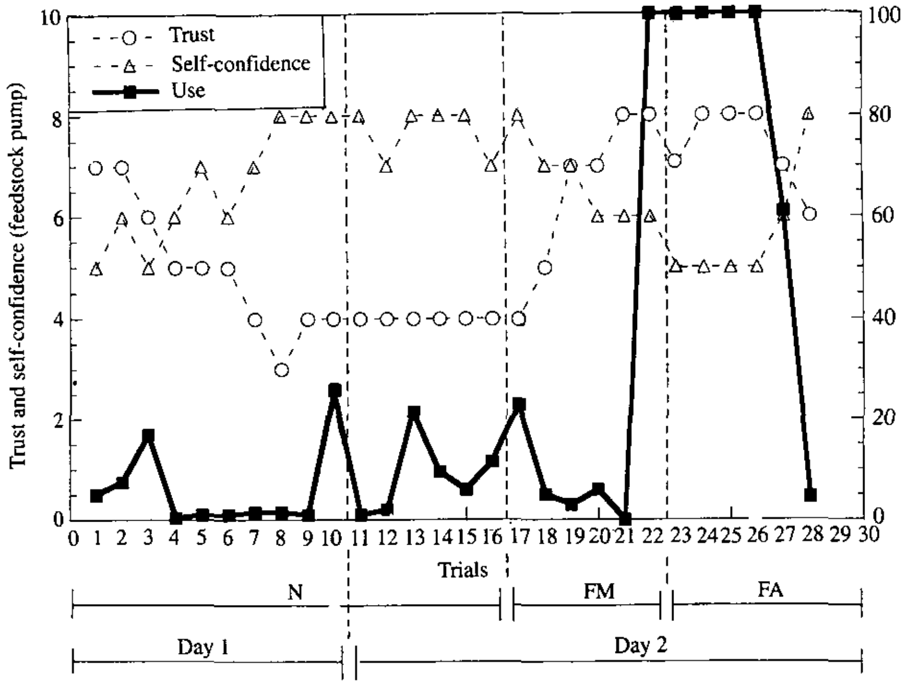


FIGURE 8. Trust, self-confidence, and the use of the automatic feedstock pump for Operator 6. (N = normal operations, FM = fault affecting manual control, FA = fault affecting automation.)

of automatic and manual control, never completely relying upon either automatic or manual control, but, in general, operators behaved as operator 6 and adopted a dichotomous allocation scheme, full automation or full manual control for the feedstock pump. While operator 6 switched from automatic to manual control, others adopted either complete manual or complete automatic control, never switching during the duration of the experiment. When operators adopted either automatic or manual control exclusively, trust and self-confidence sometimes fluctuated, but for those using automatic control trust generally exceeded self-confidence, and for those adopting manual control self-confidence generally exceeded trust. This pattern of allocation strategies was seen for the steam pump and the steam heater as well. Although Figure 8 supports a qualitative understanding of operators' changes in strategies, the following sections present a more quantitative analysis of the relationship between trust, self-confidence, and the use of automatic controllers. In particular, the following sections apply quantitative analyses to investigate the specificity, sensitivity, and the dynamics of the relationship.

3.2.2. Multitrait-multimethod analysis

The primary purpose of this experiment was to investigate how the subjective rating scales, and the reliance upon automatic controllers, reflect changes in the operators' levels of trust and self-confidence. Our hypothesis is that operators will rely on automatic controllers when trust exceeds self-confidence and they will rely on

TABLE 4

A Multitrait-multimethod matrix showing the mean of z-transformed correlation coefficients of the difference between trust and self-confidence and the use of the automatic controllers for 12 operators, $n = 28$ for each operator. $T - SC$ is the difference between operators' subjective ratings of trust and self-confidence

	Use of automatic feedstock pump	Use of automatic steam pump	Use of automatic steam heater
T − SC feedstock pump	0.32	−0.01	0.10
	($p < 0.005$)	N.S.	N.S.
T − SC steam pump	0.08	0.27	0.00
	N.S.	($p < 0.005$)	N.S.
T − SC steam heater	0.10	−0.05	0.53
	N.S.	N.S.	($p < 0.001$)

manual control when trust falls below self-confidence. The subjective ratings should also reflect these changes in trust and self-confidence. Two measures of validity might be used to test this hypothesis. First, there should be a strong correlation between use of the automatic controllers and the difference between trust and self-confidence (T − SC) in those controllers (convergent validity). Second, the use of the automatic controller of one subsystem and changes in T − SC of the other subsystem should be independent (discriminant validity). Campbell and Fiske (1959), and later Hammond, Hamm and Grassia (1986), have developed the ideas of convergent and discriminant validity in the context of the Multitrait-multimethod matrix, which supports an evaluation of both convergent and discriminant validity. If trust and self-confidence are to be useful constructs in understanding operators' behavior, then all the different measurement methods should reflect the underlying changes (convergent validity). Likewise, these methods should discriminate between trust and self-confidence in different aspects of the system (discriminant validity).

Table 4 illustrates a portion of a Multitrait-multimethod matrix, and shows the relationship between the difference between trust and self-confidence (T − SC) and the time spent using the automatic controllers. The values in this table are the average of the z-transformed correlations of the difference between trust and self-confidence for each subsystem and operators' reliance on the subsystem automation. Each entry in the table represents the data from 12 operators. Correlations on the diagonal represent the association of the difference between trust and self-confidence as reflected by subjective scales and reliance upon the automatic controllers. As expected, the use of the three automatic controllers is strongly related to the difference between trust and self-confidence in the corresponding subsystem.

While the correlation between the difference in trust and self-confidence in each of the three subsystems provides convergent validity, demonstrating the influence of trust and self-confidence on allocation strategies, the data should exhibit discriminant validity as well. The changes in the difference between trust and self-confidence in one subsystem (the feedstock pump) should not correspond to changes in the use of another subsystem (the steam heater or the steam pump). The off-diagonal values in Table 4 indicate that the changes in $T - SC$ in one subsystem are not related to the use of the automatic controllers of other subsystems. This shows that operators can partition the system into subsystems and identify trustworthy and untrustworthy subsystems based on information derived from the display. In addition, Table 4 shows that changes in trust in one subsystem can be independent of changes in trust for other subsystems. Thus, the changes in the subjective ratings and in the use of the automatic controllers demonstrate both convergent and discriminant validity.

A detailed analysis of the operators' allocation strategies suggests that the correlation coefficients reported in Table 4 may underestimate the actual relationship between the difference in trust and self-confidence and the use of automatic controllers. In many instances, operators adopted an allocation strategy that involved very little or no variance in the use of automatic controllers, leading to a very low correlation. For instance, three operators allocated complete automatic control to the steam heater for the duration of the experiment, making it impossible to calculate a correlation. In these cases, the correlation coefficient was assumed to be zero. In other cases, operators used the automatic control only as long as it took them to switch from automatic to manual control at the start of the trial. These cases show that although the operators' data were entirely consistent with the theoretical predictions (use of automatic control when trust exceeds self-confidence), the correlation coefficients in Table 4 are artificially low. When the data for operators who used entirely manual or automatic control were removed from the analysis, the average correlation between $T - SC$ for the feedstock pump and the use of the automatic feedstock pump increased from 0.32 to 0.43. Likewise, the correlation for the steam pump increased from 0.27 to 0.34, and the correlation for the steam heater increased from 0.53 to 0.78.

3.2.3. Logit functions

Figures 9, 10 and 11 suggest that the use of automatic controllers as a function of the difference between trust and self-confidence may not follow a linear relationship. Instead of a linear relationship, a logit function might better represent the relationship; with complete use of automatic control when trust is much greater than self-confidence, and complete manual control when trust is much less than self-confidence. Such a relationship is conservatively estimated by a linear model, such as the correlation coefficient. Moreover, to interpret the data as a logit function would relate the data to the standard form of the psychophysical functions found in many psychological experiments (Finney, 1971). In addition, models of a wide variety of dynamic systems show that the logit function accurately accounts for transitions between equilibrium states (Morrison, 1991). In this experiment this might correspond to transitions between stable allocation strategies.

The logit functions, shown in Figures 9, 10 and 11, represent the pooled data of all the operators. Fitting a logit function to the data requires two parameters, both of which have a meaningful interpretation in the context of this experiment. The logit function can be written as $y = 100/(1 + e^{s(x-b)})$ with "x" as the independent variable (the difference between trust and self-confidence) and "y" as the dependent variable (the percentage of time spent using the automation). In this equation, "b"

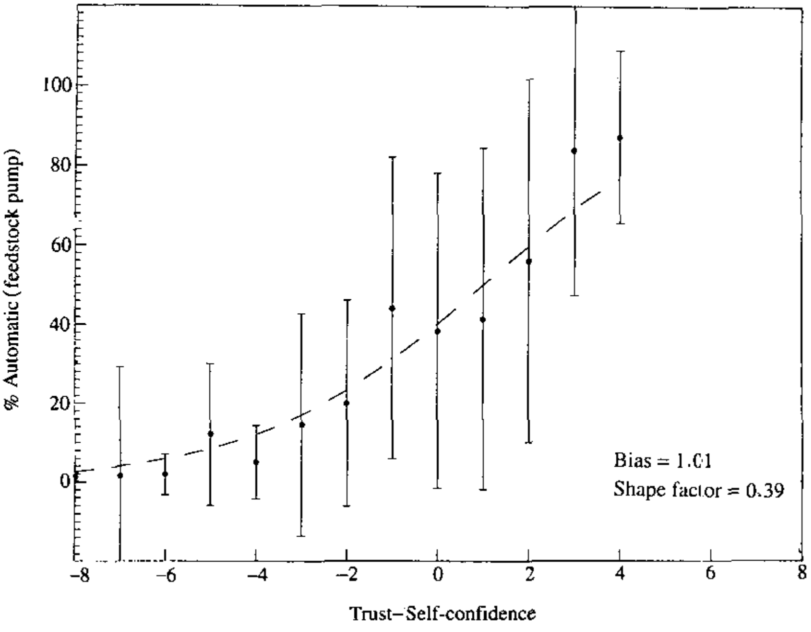


FIGURE 9. The use of the automatic controller of the feedstock pump as a function of the difference between trust and self-confidence (a single outlying data point at T − SC = 6 and % Automatic = 0 has been deleted from the graph). This graph represents the pooled data of all the operators, fitted by a logit function.

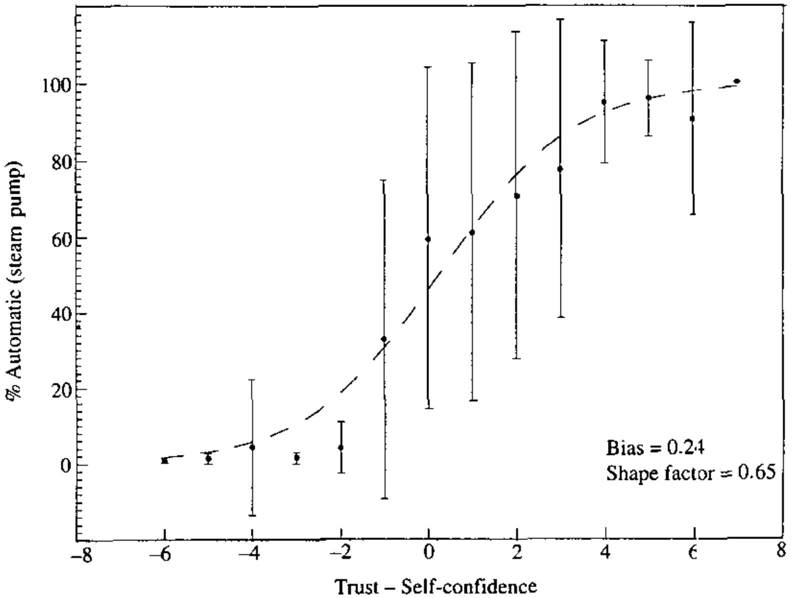


FIGURE 10. The use of the automatic controller of the steam pump as a function of the difference between trust and self-confidence. This graph represents the pooled data of all the operators, fitted by a logit function.

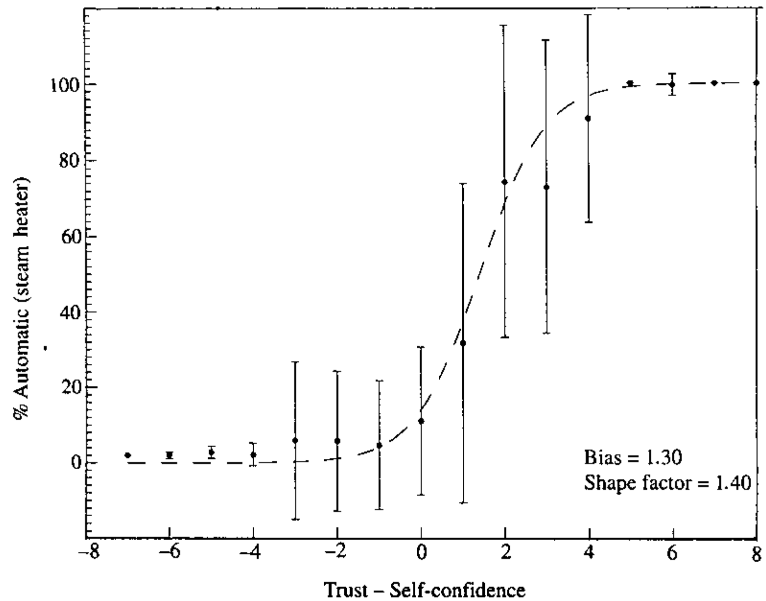


FIGURE 11. The use of the automatic controller of the steam heater as a function of the difference between trust and self-confidence. This graph represents the pooled data of all the operators, fitted by a logit function.

determines the bias, the distance that the curve is shifted along the "x" axis. As "b" increases the curve shifts to the right. The parameter "s" governs the shape of the function. A large "s" results in a steeper curve, while a small "s" results in a more gradual rise in the value of the dependent variable. In the context of this experiment the bias parameter represents the proclivity for manual control. That is, for a given difference in trust and self-confidence, the larger bias will correspond to a greater use of manual control; and for a given percentage of automatic control a larger value of T − SC is required. On the other hand, the shape parameter "s" relates the incremental change in the use of automatic controllers to an incremental increase in the difference between trust and self-confidence. The smaller the shape parameter, the smaller the change in the allocation of automatic control for a change in trust and self-confidence. In the extreme, a very small shape parameter would lead to a somewhat linear relationship between changes in the use of the automatic controller as a function of the difference between trust and self-confidence. A large shape parameter would lead to a more dichotomous relationship, with a small change in the difference between trust and self-confidence triggering swings between complete manual and complete automatic control.

The logit functions fitted to the data illustrate the implications of different shape and bias parameters. In all three cases the bias parameter is slightly greater than zero. The slight positive bias suggests that trust must be slightly greater than self-confidence before operators engage the automatic controllers. This result is consistent with previous research (Muir, 1989; Lee & Moray, 1992), as operators seem to have a tendency to adopt manual control in preference to automatic control. The bias parameter might also reflect operators' exploratory behavior, where they engage manual control even though they question their abilities. The next section includes a complete discussion of exploratory behavior. Another potential explanation of the slight positive bias in these curves is the interaction between automatic controllers. If operators perceive an interaction between the controllers they might trust each controller independently, but not engage it because of its poor performance in combination with other controllers. This experiment cannot differentiate between a tendency towards manual control and exploratory behavior, or the operators' perception of interactions between the subsystems. A subjective rating of trust and self-confidence in the control of all subsystems, operating as a group, is needed to resolve these results.

In the case of the automatic steam heater, bias is more pronounced, with the value of "*b*" equal to 1.30, compared to 1.01 and 0.24 for the automatic feedstock pump and the automatic steam pump respectively. This bias suggests that operators are more reluctant to transfer control to the automatic controller of the steam heater, even though they trust it more than they have self-confidence in themselves. The data support two interpretations. First, the operation of the steam heater is more opaque than the operation of either the steam pump or the feedstock pump. The display does not contain information about the temperature setting of the heater, but only the amount of power supplied to it. Therefore, it is difficult for operators to understand the operation of the automatic controller. Another possible explanation for this result comes from the casual comments of some operators who reported that they trust both the steam heater and the steam pump to work individually, but they did not trust them to operate together. (This may be due to the limit cycle mentioned earlier.) If this is the case an operator might assume manual control, even with a highly trusted automatic controller. As with the interpretation of the bias parameters for the other subsystems, this experiment cannot differentiate between these two possibilities.

The shape parameter also differs slightly among the three curves, with the steam heater having the highest value. The shape parameter is 1.40 for the steam heater, but only 0.65 and 0.40 for the steam pump and the feedstock pump respectively. This reflects the dramatic change in the use of the automatic steam heater as the value of (T − SC) changes. At the other end of the extreme, the use of the automatic feedstock pump changes relatively slowly as a function of the increasing difference in trust and self-confidence. These shape parameters suggest that operators will engage complete manual or automatic control of the steam heater, while the feedstock pump might be used for parts of the trial under manual or automatic control. Table 3 supports this claim, showing only two instances where the steam heater was not used in a dichotomous manner. With the feedstock pump, on the other hand, there were twelve instances where it was used partly in manual, and partly in automatic mode.

3.2.4. Exploratory behavior and interacting subsystems

Perhaps more interesting than the correspondence between the difference between trust and self-confidence and the use of the automatic controllers, are the cases where the use of automatic controllers disassociates from the difference between trust and self-confidence. Table 4 and Figures 9, 10 and 11 show that when trust

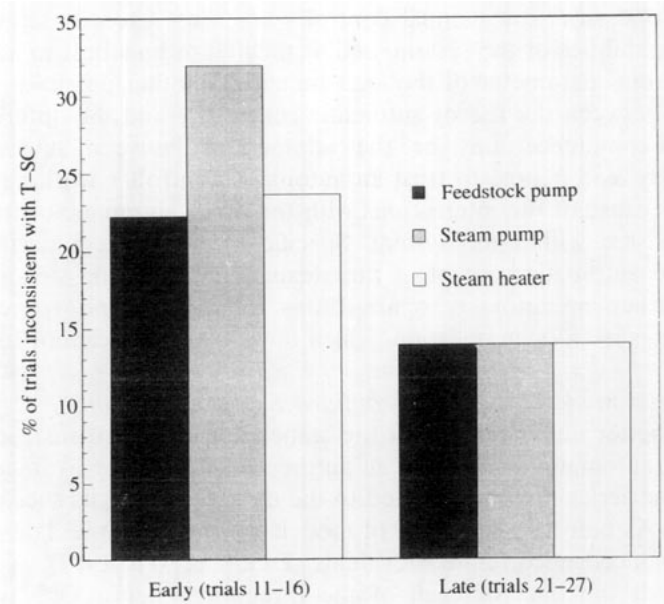


FIGURE 12. The percentage of trials with predominantly manual control, when trust exceeded self-confidence (exploratory behavior). The data in this figure has been pooled over all 12 operators.

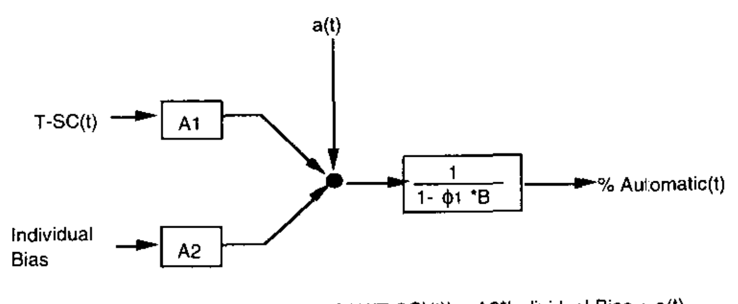
exceeded self-confidence, operators tended to engage the automatic controllers, whereas when trust was less than self-confidence, operators tended to use manual control. Although a predominant tendency, it was not completely universal. For example, occasionally operators controlled the system manually when trust exceeded self-confidence. Figure 12 shows the frequency with which the use of automation failed to correspond with operators' trust and self-confidence. The use of manual control was considered "exploratory" when trust exceeded self-confidence and when operators adopted manual control for more than 50% of the trial. Figure 12 suggests a greater disassociation early in the experiment as compared to late in the experiment. For example, in nearly 25% of the early trials, operators controlled the pump manually, even though their trust exceeded their self-confidence. This compares to 13% of the trials during the first part of the second day. Likewise, a greater amount of disassociation occurred during the trials that contained a fault compared to trials during the first part of the second day.

One possible interpretation of these results is that they reflect exploratory behavior. Early in the experiment operators seem to explore the capabilities of manual control, as they try different strategies to increase productivity. Likewise, during faults operators might experiment with different manual control strategies to circumvent the effects of the fault. This result seems similar to the interpretation of the remnant in manual control research. For example, in discussing how drivers adapt to changing road conditions, Sheridan and Ferrell (1974) suggested that drivers inject "evaluation steps" into the system and evaluate the resulting system response. Like the data in this experiment, such exploratory behavior manifests itself as a remnant, not accounted for in the linear models of manual control. Thus, it seems that operators use manual control when trust exceeds self-confidence to explore the capabilities of the system, and as they learn how best to adapt to faults.

As with the bias parameter of the logit function, another possible reason for the disassociation between the use of automatic controllers and the difference between trust and self-confidence may be the interactions between subsystems. These interactions may lead people to trust an automatic controller working by itself, but not to use it because of the interactions with the other automatic controllers. These interactions existed within the system. Specifically, the system performed poorly when both the automatic control of the steam pump and the steam heater were engaged. Whether operators recognized this interaction and reflect it in their subjective ratings is an open question, which this experiment cannot resolve.

3.2.5. Time series model of trust, self-confidence and automation

To develop a better understanding of the association between trust, self-confidence and the use of automatic controllers, an autoregressive moving average vector form (ARMAV) time series model was fitted to the data, following its success in Lee and Moray (1992). As before, a sequence of models was fitted to the data until one was found that produced uncorrelated residuals (Pandit & Wu, 1987). A single model was found to fit the data of each of the three subsystems. Only variables that significantly decreased the error variance were retained in this model. In addition, the validities of the three models were verified by showing that the residuals were independent, with Durbin–Watson statistics of approximately 2, and autocorrelations of less than 0.07. The model common to all three subsystems indicates that the use of automatic controllers depends not only upon the difference between trust and self-confidence, but also upon the previous use of the automatic controllers and operators' individual biases. The parameter that represents individual operation bias is calculated through a least squares estimation procedure, and is a different constant for each of the 12 operators. The bias parameter represents the tendency for operators to adopt either manual or automatic control regardless of the difference between trust and self-confidence. Figure 13 shows a block diagram representation



$$\%\ \text{Automatic}(t) = \phi 1 * \%\ \text{Automatic}(t-1) + A1*((T\text{-}SC)(t)) + A2*\text{Individual Bias} + a(t)$$

FIGURE 13. A block diagram representation and discrete differential equation showing the factors that influence the use of the automatic controllers. A1 and A2 represent the weights of T − SC and the individual bias towards manual control, a(t) represents normally distributed independent fluctuations, and ϕ represents how strongly the current use of automation depends on the past use of automation.

of the ARMAV model. This model accounts for a large part of the variance in the use of automatic controllers of the three subsystems: 60.9% for the use of the feedstock pump, 80.5% for the steam pump, and 86.5% for the steam heater.

This model has two interesting characteristics. First, the use of automatic controllers depends not only upon the difference between trust and self-confidence, but also on the past use of the automatic controllers. This implies that reliance on an automatic controller in one trial will generally lead to reliance on that automatic controller in the next trial, independent of the change in trust and self-confidence. This result is consistent with Edwards (1968), Tversky and Kahneman (1974), and Dorner (1987), who describe a reluctance of humans to change their hypotheses or decisions, even as contradictory information accumulates.

The second interesting characteristic of the model concerns the individual differences of operators. The data show that not only does the use of automatic controllers depend upon trust and self-confidence and the past use of controllers, but that it also depends upon individual differences between operators. That is, some operators were consistently prone to using automatic controllers, regardless of their ratings of trust and self-confidence. Other operators consistently preferred manual control. Observations of operators' comments support this result. Some operators took the task as a challenge, and tried to master the system, using manual control even though their self-confidence was relatively low. Other operators seemed to prefer automatic control, even using distrusted automatic controllers, which they monitored closely.

4. Conclusion

This experiment sought to investigate the importance of self-confidence in the operators' allocation strategy, as well as the ability of the operators to change their allocation strategy to accommodate faults. The results provide convincing evidence that operators' allocation policy depends on the difference between trust and self-confidence. In other words, changes in self-confidence and trust correspond to changes in the operators' allocation policy. A Multitrait-multimethod approach also demonstrated this result, showing that changes in the subjective ratings of trust and self-confidence were related to changes in the use of automation, and were specific to individual subsystems. Additional analyses showed that a logit function related the difference in trust and self-confidence to the use of automation, suggesting that, although operators could use automation for whatever length of time they wished, they tended to engage the automation continuously or rely completely on manual control. Rarely did operators switch between automatic and manual control during a trial. A time series analysis provided further support for this result, showing that the allocation strategy possessed an inertia, with the current allocation policy depending on the recent use of automation. The results of the logit analysis, combined with the results of the time series analysis, show that operators prefer complete automatic or manual control, and that they were reluctant to change their allocation strategy, even when trust and self-confidence changed. Besides trust, self-confidence, and the past use of the automation, the time series analysis revealed that individual biases influenced the choice of automatic or manual control. As well as the individual biases identified in the time series analysis, the logit functions, which represent the pooled data of all operators, show that operators were generally predisposed to manual control.

These results support four general conclusions that have implications for the design and development of supervisory control systems. First, the experiment showed a strong relationship between the difference in operators' trust and self-confidence and their reliance on automation. This conclusion suggests that system designers should consider how characteristics of the system affect operators' subjective feelings of trust and self-confidence. Specifically, it stresses the need to provide operators with information regarding their performance, and that of the automation, so that operators' self-confidence and trust reflect true capabilities, and promote the appropriate use of automation. Second, reliance on automation displayed an inertia, where the current use of automation depended on the previous use of automation. This finding suggests that operators may fail to switch allocation at the appropriate times, indicating the need for designers to minimize the obstacles involved in reallocating control. Eliminating complex reconfigurations of the system needed to invoke automatic and manual control should encourage operators to switch between manual and automatic control as they judge necessary. Third, the data showed that operators were biased towards selecting manual control. Fourth, this bias towards manual control was particularly prevalent during the initial interaction with the system. These two conclusions suggest that designers should not assume that operators will engage automation, even when they feel their capabilities are less than those of the automation. If designers want users to make the best use of automation, special provisions need to be made to encourage users to engage automation. Overall, this paper shows the need to consider more than traditional interface design guidelines; successful implementation of semi-automatic systems will require attention to issues such as operators' subjective feelings of trust and self-confidence.

These conclusions are based on the results of a single experiment and further investigation is needed to confirm them and establish the limits of their application. Using a "microworld" (Brehmer, 1992) enabled this experiment to capture many of the demands of actual process control plants (multiple conflicting goals, dynamic decision problems, and complex interactions between system components) that more traditional laboratory tasks ignore. However, the system used for these experiments is much more simple than any commercial system, and the findings of this experiment may not scale to more complex systems where a broader variety of factors may also influence operator reliance on automation (Roth, Bennett & Woods, 1987; Boettcher *et al.*, 1989; Kirlik, 1993). In addition, the automation examined in this experiment was quite simple. Real systems have multiple levels of automation and operators adjust parameters to optimize system performance rather than simply intervening or delegating tasks to the automation. Transitions between levels of automation and adjustments to automation blur the distinction between "manual" and "automatic" control, and may limit the generalization of this experiment. Thus, this experiment represents an initial examination of how trust and self-confidence influence operator performance in supervisory control systems. Future research will reveal the importance of these factors in moderating operators' use of automation.

We are thankful for the helpful comments of David D. Woods and his anonymous reviewers who helped make this a more readable and useful paper. We would also like to thank Penelope Sanderson and Kim Vincente for their contributions to this research. This work was supported by a grant from the University of Illinois Research Board and the University of Illinois Beckman Fund.

References

BARBER, B. (1983). *Logic and the Limits of Trust.* New Brunswick, NJ: Rutgers University Press.

BOETTCHER, K., NORTH, R. & RILEY, V. (1989). On developing theory-based functions to moderate human performance models in the context of systems analysis. *Proceedings of the Human Factors Society 33rd Annual Meeting,* pp. 105–109, Denver, Colorado, USA, 16–20 October.

BREHMER, B. (1990). Towards a taxonomy of microworlds. In J. RASMUSSEN, B. BREHMER, M. DE MONTMOLLIN & J. LEPLAT, Eds. *Taxonomy for Analysis of Work Domains. Proceedings of the first Mohawc Workshop.* Roskilde: Risö National Laboratory.

BREHMER, B. (1992). Dynamic decision making: human control of complex systems. *Acta Psychologica,* **81,** 211–241.

CAMPBELL, D. T. & FISKE, D. W. (1959). Convergent and discriminant validation by the Multitrait-multimethod matrix. *Psychological Bulletin,* **56,** 81–105.

DORNER, D. (1987). On the difficulties people have in dealing with complexity. In J. RASMUSSEN, K. DUNCAN & J. LEPLAT, Eds. *New Technology and Human Error.* New York: Wiley.

EDWARDS, W. (1968). Conservatism in human information processing. In B. KLEINMUNTZ, Ed. *Formal Representation of Human Judgment.* New York: Wiley.

FINNEY, D. J. (1971). *Probit Analysis.* New York: Cambridge University Press.

FISCHHOFF, B. (1982). Debiasing. In D. KAHNEMAN, P. SLOVIC & A. TVERSKY, Eds. *Judgment Under Uncertainty: Heuristics and Biases,* pp. 421–444. New York: Cambridge University Press.

FISCHHOFF, B. & MACGREGOR, D. (1982). Subjective confidence in forecasts. *Journal of Forecasting,* **1,** 155–172.

FISCHHOFF, B., SLOVIC, P. & LICHTENSTEIN, S. (1977). Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance,* **3**(4), 552–564.

HALPIN, S., JOHNSON, E. & THORNBERRY, J. (1973). Cognitive reliability in manned systems. *IEEE Transactions on Reliability,* **R-22**(3), 165–169.

HAMMOND, K. R., HAMM, R. & GRASSIA, J. (1986). Generalizing over conditions by combining the Multitrait-multimethod matrix and representative design of experiments. *Psychological Bulletin,* **100,** 257–269.

HUEY, B. M. (1989). *The effect of function allocation schemes on operator performance in supervisory control systems.* Unpublished Doctoral thesis. George Mason University, Washington, DC.

KIRLIK, A. (1993). Modeling strategic behavior in human–automation interaction: why an "aid" can (and should) go unused. *Human Factors,* **35,** 243–263.

LEE, J. D. (1992). *Trust, self-confidence, and adaptation to automation.* Unpublished Doctoral thesis. University of Illinois, Urbana, IL.

LEE, J. & MORAY, N. (1992). Trust and the allocation of function in the control of automatic systems. *Ergonomics,* **35,** 1243–1270.

LICHTENSTEIN, S., FISCHHOFF, B. & PHILLIPS, L. D. (1982). Calibration of probabilities: the state of the art to 1980. In D. KAHNEMAN, P. SLOVIC & A. TVERSKY, Eds. *Judgment Under Uncertainty: Heuristics and Biases,* pp. 306–334. New York: Cambridge University Press.

MORRISON, F. (1991). *The Art of Modeling Dynamic Systems: Forecasting for Chaos, Randomness, and Determinism.* New York: Wiley.

MUIR, B. M. (1988). Trust between humans and machines, and the design of decision aids. In

E. HOLLNAGEL, G. MANCINI & D. D. WOODS, Eds. *Cognitive Engineering in Complex Dynamic Worlds*, pp. 71–84. London: Academic Press.
MUIR, B. M. (1989). *Operators' trust in and use of automatic controllers in a supervisory process control task.* Unpublished Doctoral thesis. University of Toronto, Ontario, Canada.
OSKAMP, S. (1982). Overconfidence in case-study judgments. In D. KAHNEMAN, P. SLOVIC & A. TVERSKY, Eds. *Judgment Under Uncertainty: Heuristics and Biases*, pp. 287–293. New York: Cambridge University Press.
PANDIT, S. M. & WU, S. M. (1987). *Time Series Analysis with Applications.* New York: Wiley.
REMPEL, J. K., HOLMES, J. G. & ZANNA, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, **49**, 95–112.
ROTH, E. M., BENNETT, K. B. & WOODS, D. D. (1987). Human interaction with an "intelligent" machine. *International Journal of Man–Machine Studies*, **27**, 479–525.
TVERSKY, A. & KAHNEMAN, D. (1974). Judgement under uncertainty: heuristics and biases. *Science*, **185**, 1124–1131.
SHERIDAN, T. B. (1987). Supervisory control. In G. SALVENDY, Ed. *Handbook of Human Factors.* New York: Wiley.
SHERIDAN, T. B. & FERRELL, L. (1974). *Man–Machine Systems.* Cambridge, MA: MIT Press.
SHERIDAN, T. B. & HENNESSY, R. T. (1984). *Research and Modeling of Supervisory Control Behavior.* Washington, DC: National Academy Press.
SHERIDAN, T. B. & JOHANNSEN, G. (1976). *Monitoring Behavior and Supervisory Control.* New York: Plenum.
WOODS, D. D., O'BRIEN, J. F. & HANES, L. F. (1987). Human factors challenges in process control: the case of nuclear power plants. In G. SALVENDY, Ed. *Handbook of Human Factors*, pp. 1724–1770. New York: Wiley.
YATES, J. F. (1990). *Judgment and Decision Making.* Englewood Cliffs, NJ: Prentice Hall.
ZUBOFF, S. (1988). *In the Age of the Smart Machine: The Future of Work and Power.* New York: Basic Books.

Paper accepted for publication by Associate Editor Professor D. Woods.

Appendix

GENERAL INSTRUCTIONS

Assume you are a controller of an Orange Juice processing plant. Your job is to heat the Orange Juice to a temperature between 75–85°C to destroy any parasites. Since the oranges arrive sporadically the flow of the juice into the plant varies.

The plant is fully automated, and can be run automatically. As an operator you can assume manual control of the plant, or assign any combination of manual and automatic controls to the plant.

As in many plants, an automatic data logging system records system parameters. To evaluate the accuracy of the plant's data logging system part of your task controlling the Orange Juice processing plant involves recording several of the system parameters. To monitor the plant's data logging system we would like you to record three system variables (Inflow to input vat (VI), Inflow temperature (VT) and Waste vat volume (WV)). You should write the values of these parameters on the sheet provided whenever the computer gives the double beep signal.

Try to produce as much orange juice as possible. You will be rewarded with a $0.10 bonus for every trial in which 90% or more the orange juice input is processed. Your percentage score will be shown at the end of each trial. In order to receive the bonus you must log the three parameters for all double beep signals from the computer.

Like any real life system, the plant is subject to wear and tear, and so occasionally faults may occur. You will work with the plant for a total of 2 days (4 hours). Make yourself familiar with the system. Learn what happens when the plant runs automatically and what happens when you take over manual control.

We are interested in your attitudes and feelings about the plant. Determine how trustworthy the automatic controllers are and how much self-confidence you have in controlling the system manually. After each trial you will rate your estimates of self-confidence and trust. You may find that there is some variability in the plant's behavior; work out the best way to control it.

DESCRIPTION OF PLANT

The plant is a simulated pasteurizer. Orange juice enters the system through a pipe at the top left of the screen, and collects in the vat, which starts with 400 liters in it. The orange juice enters the vat at a variable rate. If the volume of the vat exceeds 950 liters the orange juice overflows into the waste vat. If the vat level drops to 0 liters the plant shuts down.

The pump at the bottom left pumps the juice through two heat exchangers, in the second of which steam is used to heat it between 70–85°C. An automatic pump controls the rate at which the steam is pumped through the exchanger, and an automatic controller controls the temperature of the steam.

If the temperature of the juice when it leaves the second heat exchanger is too low, then it is sent back to the vat along the G pipe at the top of the screen, to be processed. Flow through this pipe changes the pipe's color to pale blue.

If the temperature of the juice is too high, it is spoilt, and it is dumped into the waste vat. Flow through this pipe changes the pipe's color to red.

If it is within the correct range, it goes on through a pipe which turns green, and gives up some of its heat to the incoming cold juice in the first heat exchanger. It then collects in the final output vat. This forms the basis of your final output score.

The goals of the system consist of the following three items:

1. Pasteurize as much as possible of the incoming supply of juice. The percent achieved is shown at the end of the run under the output vat.
2. Minimize the wasted juice.
3. Try to keep the input vat containing around 500 liters, and neither overfill it nor allow it to become empty.

Each trial lasts about 5 min. There may be slight differences from trial to trial due to equipment problems in the plant, making manual or automatic control less effective. Watch the different parts of the system. In each trial there may be differences in how well the various components carry out their designed task. You are to rate each component of the system on their ability to satisfy the three goals.

At the end of each run we want you to perform six ratings:

1. **How high was your self-confidence in controlling the feedstock pump?**
2. ***How much did you trust the automatic controller of the feedstock pump?***
3. **How high was your self-confidence in controlling the steam pump?**
4. **How much did you trust the automatic controller of the steam pump?**

5. How high was your self-confidence in controlling the steam heater?
6. How much did you trust the automatic controller of the steam heater?

Remember—plants vary somewhat from time to time, so try not to let your rating of one trial influence your rating of the next.

INTRODUCTION TO SUBJECTIVE RATING

In this study we are interested in your judgments of trust in the automatic control. In addition, we are interested in how much self-confidence you have in your ability to control the system yourself.

First, think about your trust in people. We all trust some people more than others. If you think about people you know, you can probably think of some whom you trust very much and others whom you trust much less. We do not trust all people equally, and we can express how much we trust a particular person.

We also think about trusting things, such as products. For example, I trust my Chrysler to start in the morning because it has never failed to do so. I trust my wife's Chevrolet much less because of a history of trouble. I trust one of my computers because I have never had trouble with it, while another is constantly giving me trouble when I try to log on, and I trust it much less.

Similar to trust we can also consider the self-confidence in our own abilities. For example, you might have a great deal of self-confidence in your ability to walk to school because you have been doing it every day for the last year. Your self-confidence might be substantially less in composing a paper on a new word processing system. If you are unsure of your abilities to complete a task your self-confidence will most likely be lower.

If we think about it for a moment, we could rate our degree of trust and self-confidence in many of the things we use on a scale like that on the attached sheet. You'll be using a scale like this in the experiment, so I'd like to give you a bit of practice in using it. Let's rate a few of the things you may use often.

Now please rate your trust and self-confidence in each of the following:

1. Trust in the local bus service to get you to the store on time./Self-confidence in your ability to get to the store in time.
2. Trust in your calculator or computer to produce the right answer./Your self-confidence in your ability to arrive at the correct answer doing the calculations manually.
3. Trust in the heating system where you live to keep you comfortable./Your self-confidence in your ability to turn the heater on and off manually to keep you comfortable.
4. Trust in your watch to tell the correct time./Your self-confidence in your ability to estimate the correct time.

In this experiment, you will be asked to rate an industrial plant, either as you operate it or as you watch it perform its task automatically. In each trial you will be asked to assess the plant's performance based on two criteria: the amount of trust that you place in the automatic controllers, and the amount of self-confidence you have in your ability to manually control the plant.

At the end of each trial the computer displays six screens, one for self-confidence ratings and one for trust ratings for each of three system components. These screens each contain a rating scale similar to the four scales shown on the following pages. To select a rating simply place the mouse cursor on the desired rating and "click" it.

Any questions? If you have any questions about the scales, don't hesitate to ask.

There are no "right" answers. We are interested in how you feel about the operation of the plant. Your answers will help us in our research on how to improve the relation between humans and machines.

CONTROLS

As an operator you will be able to control three aspects of the plant: the Juice pump, Heater, and Steam pump. These three components operate either manually or automatically:

Juice pump

You will be able to control the speed that this pump operates. This pump drains the juice from the vat. The pump has a range of 0–100 L/min.

control example: P A (sets the pump to run on automatic)
P M 100 (sets the pump to 100)

Heater

You will be able to control the amount of power going to the heater. The heater has a range of 0–100% power. **Remember!** you are not setting a temperature, you are setting the amount of power going to the heater. The effect of your choice of setting shows in the "steamtemp" display on the screen.

control example: H A (sets the heater on auto)
H M 50 (sets the heater to 50% power)

Steam pump

You will also be able to control the steam pump. The steam pump moves the steam that is produced by the heater. The steam pump has a range of 0–100 L/min.

control example: S A (sets the steam pump to auto)
S M 675 (sets the steam pump to 675)

**Note: There must be blank spaces between the object (P, H, or S) and control (A or M) and number you choose (in manual), but you *do not* need to use capital letters.

ALARMS

There is also an alarm panel in the bottom left corner of the display. Warnings pertaining to the state of the initial vat are displayed there. Such warnings are "Vat Low", "Vat High", "Emergency Vat Low", and "Emergency Vat High".

STRATEGIES

1. Keep the vat level around 500 L. NEVER let the vat become completely empty. If you do, the run comes to an end, and you lose any bonuses you have earned on that day.
2. Stay on top of the system, pay attention to what is going on at all times.
3. Enter correct commands (it saves time).
4. Stay relaxed!