Title
Fast and Frugal Operators Sub-Optimally Adapt To Machine Failure

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN
1069-7977

Authors
Ohlsson, Stellan
Youmans, Robert J.

Publication Date
2006

Peer reviewed

Fast and Frugal Operators Sub-Optimally Adapt To Machine Failure

Robert J. Youmans (ryouma1@uic.edu)
Stellan Ohlsson (stellan@uic.edu)

Department of Psychology, 1007 W. Harrison St.
University of Illinois, Chicago
Chicago, IL 60607 USA

Abstract

One safeguard against instrument malfunction is to provide backup instruments for machine operators. In previous studies, prior training caused operators of a simulated machine to adapt to instrument malfunction by adopting a suboptimal decision rule rather than by reallocating attention to backup instruments. One hypothesis for these findings is that operators do not notice when their main instruments malfunction. Here we examine warning systems that force operators to notice instrument problems. Our results indicate that warnings did not help operators to reallocate attention to backup instruments. Instead, operators fail the simulation and make sub-optimal adaptations afterward that lead to further failures.

Introduction

The operation of complicated machines often requires a flexible human operator that can adjust the operations of a machine to meet task demands. This is especially so because of the inherent fallibility of complicated machines; people must react to normal feedback from a machine in order to operate them as intended, and must react to abnormal feedback from machines when a corrective action is required. In prior work, we documented how operator performance suffered when instruments that an operator had been trained to use suddenly begin to provide inaccurate information, even when a second, valid instrument was available that could correct the error (Youmans & Ohlsson, 2005). The finding suggested that machine operators have difficulty switching from their usual instruments to a secondary or backup source of information.

Why do operators fail to utilize seemingly obvious secondary instrumentation when the primary instruments that they have been using malfunction? One explanation is that training produces biases and automaticity that might interfere with rapid adaptation to changing task demands. Although quickly switching from one task set to another might subjectively seem to progress smoothly and effortlessly, evidence strongly suggests that switching between even simple task sets can be quite difficult (e.g., Allport, A., Styles, E. A., & Hsieh, S., 1994). To what extent are people limited by prior experience when faced with the need to adapt to changing task conditions?

We investigate these questions with the help of a simulated human-machine interface in which the degrading of one set of instruments poses a need to re-allocate attention.

A Simulated Machine Interface

In our simulation, participants assume the role of the operator of a juice factory. Two *Holding Tanks*, tank A and tank B, were shown on the upper left side of a computer screen, connected with pipes to a *Mixing Tank* shown to the right; see Figure 1.

On the lower half of the screen was the gauge equivalent of the color information. Here, three realistic looking temperature gauges representing tanks A, B, and the Mixing Tank were displayed; see Figure 1.

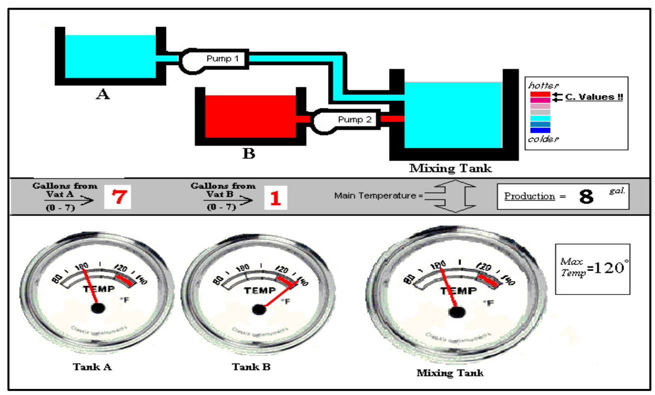


Figure 1: Example of factory interface. *Note.* Factory interface was in color.

Each Holding Tank contains liquid at a certain temperature. The factory is operated by adding some amount of liquid from tank A and some amount from tank B into the Mixing Tank. The amount and temperature of the juice is determined by the amount and temperature of the previous content of the Mixing Tank, the added input from Tank A and the added input from Tank B. Once a participant entered these amounts, the simulation was animated; the colored liquid was shown flowing through pipes into the Mixing Tank, and the Mixing Tank's color and gauge responded appropriately to the new input. Once the two inputs were added, the resulting state of the Mixing Tank was computed and displayed [1], and the operator could make the next decision about how much liquid to add from either Holding Tank. The task of the operator was to maximize the production of juice without overheating the facility, a type of trade-off situation.

As shown in Figure 1, the display was divided into two sections by a thick gray bar. Above the bar is the section that

[1] TEMPcurrent = [(14 * TEMPprior.) + (7 * TEMPa) + (7 * TEMPb)] / 3, rounded to the closest whole value.

we will refer to as the *color instrument* portion of the screen. Beneath the bar is the section that we will refer to as the *gauge instrument* portion of the screen. In the beginning of each experiment, both the color and gauge instruments provide the necessary information to successfully run the simulation. For example, in the color portion, the temperatures in all tanks are represented by the liquids' shades of blue or red. These colors represent common indicators of heat (e.g., many bath fixtures represent temperature in shades of blue and red). To help ensure that participants understood these color values, a color-to-temperature guide was on screen at all times. In the gauge instrument portion of the screen, a red 'needle' in each gauge indicated those same temperatures.

It was the temperature of the Mixing Tank that was most important in the simulation. Although the Holding Tanks' liquids could be hot or cold, the Mixing Tank could not accommodate extreme heat. If temperatures in the Mixing Tank ever rose above a certain point, the juice was ruined. This point was represented by the two deepest shades of red (colors), and the two highest needle readings (gauges), and is referred to as the *critical temperature*. If the content of the Mixing Tank reached this temperature, then the pasteurization process was spoiled. Tanks A and B could safely hold liquids across the entire range of possible temperatures; only the Mixing Tank had this temperature restriction.

In both sections, the temperatures of the liquids in tank A, tank B, and the Mixing Tank are indicated. In normal operation mode, the factory may be operated on the basis of either the color or gauge instruments; these two information sources are redundant. Because the colors and gauges present identical information, the task can be solved equally well on the basis of either.

The simulated instruments were implemented so that they could be made to malfunction at a determined point in the simulation. When malfunctioning, both sets of instruments still displayed temperature values for the three tanks, but either the gauges or colors in the Holding Tanks became inaccurate. Only one source of information became inaccurate, so the operator always had the option of discontinuing use of the malfunctioning source, and switching to the other.

Simulation Settings. The juice factory simulation may be set to remain reliable, or to malfunction, and may be set to provide a variety of practice with the simulation before the experimental conditions. All participants begin the simulation with a *practice session* that is intended to teach participants how to operate the simulation successfully. The simulation may be set to provide instructions for using either the color or gauge instruments, but not both. In the experiment reported here, all conditions were set to receive color instrument training.

During practice, participants are told how tanks A and B relate to the Mixing Tank, and are told repeatedly that their goal is to use the factory to produce as much juice as possible, without overheating the Mixing Tank. Following these instructions, participants operate a partial version of the factory simulation that displays only that portion of the screen that they have received instructions about. For example, if a participant is in a color training condition, then in the practice session they will not see the gauge instruments. The result is that participants practice with only one of the sources of information, never both.

During practice, participants learn to produce the maximum volume of juice per *trial*. Each trial consists of two judgments about how much liquid to input, one amount from tank A (judgment 1) and one from tank B (judgment 2), into the Mixing Tank. Participants indicate how much juice, from zero to seven gallons, should be entered into the Mixing Tank from each Holding Tank by typing the appropriate digit on the keyboard. Juice production accumulates across these trials to a preset level in the simulation, usually 300 gallons. In this paper, a series of such trials is referred to as a *round*.

If the Mixing Tank heats into the critical temperature range during practice, a warning appears, and the system pauses for 6 seconds while displaying this warning. Otherwise, practice ends when participants have produced the preset level of juice. After this, the participants are instructed with respect to (but do not receive practice with) the backup instrument that had been absent, and are told that the backups convey the same information as that with which they have just practiced with, albeit in different form. For example, if a participant trains with color instruments, then they would receive instructions about the function and duplicitous nature of the gauge instruments, but would not actually practice with those gauges. At the end of practice, participants were told *that if they have trouble using one of the two types of instruments in any part of the simulation, they should switch to the other instrument.*

Following the practice session, participants operated the factory for three rounds. These rounds constitute the experimental conditions of the simulation, generally defined by either no-malfunction or malfunctioning instrumentation. In no-malfunction conditions, both color and gauge portions of the screen are presented simultaneously and participants are asked to produce 150 gallons of juice as quickly as possible without overheating the Mixing Tank. If the Mixing Tank is heated past the critical temperature, the round ends and a failure display is presented indicating that the participant had overheated the system.

In malfunction conditions, either the color or gauge instrument representing the Holding Tanks becomes unreliable roughly halfway through each round. As an example of this type of malfunction, color instruments might indicate that there are two tanks of cool liquid waiting in the Holding Tanks, while the gauges indicate the true temperature of the tanks correctly as hot. The moment at which the mismatch between the instruments occurs will be referred to as the *malfunction point* in this paper.

The malfunction point is the moment during a round at which a successful participant that is basing his or her decisions on the malfunctioning source of information should recognize that the instrument has malfunctioned, and follow instructions by switching to the other instrument. In malfunction conditions, the malfunction point always occurs when a participant has produced between 70 and 90 gallons of juice *and* the Mixing Tank is a neutral or less temperature. If these conditions do not occur, then the malfunction point automatically occurs at 90 gallons of production. These criteria help to ensure that, after the malfunction point, there is still leeway for the Mixing Tank's temperature to increase.

Feedback. The simulation provides three types of feedback to the operator. (a) *Success/failure*. When the operator succeeded in producing the target amount of juice or overheated the factory, he or she is informed of this fact. We refer to this as *between-rounds feedback*. (b) *Instrument mismatch*. When one set of instruments malfunctions, that display will conflict with the reliable instrument. Although this mismatch is not feedback per se, it is feedback in the sense that this information could inform an operator that they had advanced to a point in a round where malfunctions were taking place. (c) *Mixing Tank outcomes*. If a set of Holding-Tank instruments malfunctions and an operator uses those instruments, then the effect of that liquid on the temperature of the Mixing Tank will not make sense. For example, transferring hot liquid that was falsely displayed as cool would increase the temperature of the Mixing Tank, although the operator would expect the opposite. We refer to this information *within-trial feedback* (see Figure 2).

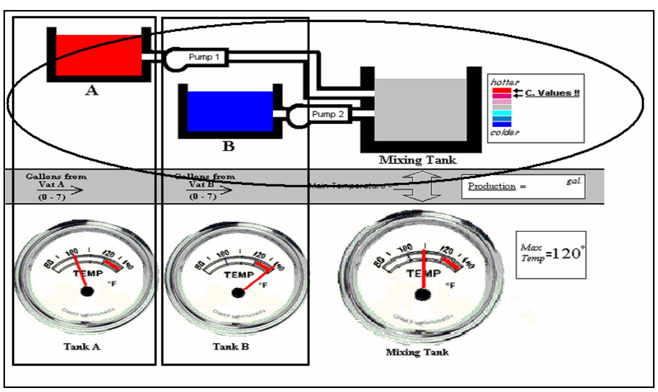


Figure 2: The two rectangles highlight the first form of within-round feedback, the mismatch between the color and gauge instruments. The oval highlights the second form of within-round feedback, the mismatch between what a participant would expect to happen to the Mixing Tank when the color instrument is used, and the actual outcome whereby the Mixing Tank's temperature increases rather than decreases. *Note.* Factory interface was in color.

Previous Findings

The findings reported here build on four prior experiments that utilized the basic juice-factory simulation. In two previous experiments, we demonstrated that practice causes operators of this simulation to adapt to instrument malfunction with a suboptimal decision rule rather than by reallocating attention to a reliable backup instrument (Youmans & Ohlsson, 2005). In Experiment 1, operators became fixed on the instrument they practiced with, regardless of whether it was the color or gauge instruments. Fixation occurred despite task instructions that the secondary instruments would be helpful, the within round feedback that indicated primary instrument malfunction, and repeated failures across three rounds of the simulation.

Table 1: Simulation Success by Condition in Experiment 1.

Round	Condition	Color Practice	Gauge Practice
1	Colors Malfunction	2 (6%)***	21 (70%)
	Gauges Malfunction	18 (60%)	7 (23%)***
	No Malfunction	19 (66%)	19 (66%)
2	Colors Malfunction	5 (16%)***	20 (67%)
	Gauges Malfunction	22 (73%)	12 (40%)*
	No Malfunction	22 (76%)	18 (62%)
3	Colors Malfunction	11 (36%)**	21 (70%)
	Gauges Malfunction	21 (70%)	15 (50%) *ns*
	No Malfunction	24 (83%)	20 (69%)

Note. Significant differences reflect comparisons between groups who practiced with and without a malfunctioning instrument. * $p < .05$; ** $p < .01$; *** $p < .001$

In Experiment 2, we showed evidence that operators *failed to notice* the within-round feedback that was being provided. Specifically we showed that reaction times taken prior to and after the primary instrument malfunction remained constant, suggesting that operators took no pause to even consider the within-round feedback indicating that their actions were driving the simulation to unacceptable temperatures.

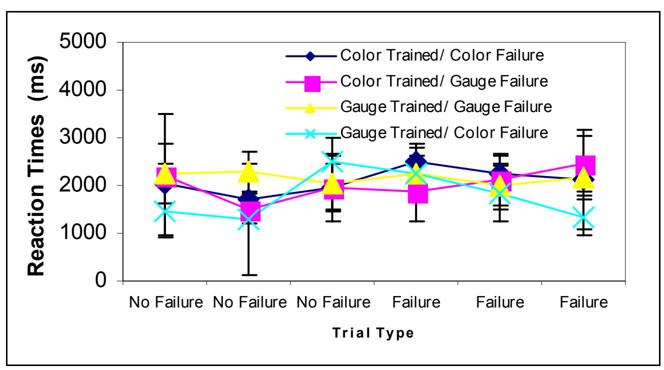


Figure 3: Participant's average reaction times in Experiment 2 for the three trials before and after instrument malfunctions by condition in Round 1. Due to space constraints, highly similar null results, found in Rounds 2 and 3, are not here shown.

Two follow-up experiments replicated these prior findings, even when our operators received specialized training regimens that we had predicted would help reduce functional fix by increasing awareness of within-round feedback (Youmans & Ohlsson, In Review). In Experiment 3, our operators experienced a greatly enhanced training regimen that highlighted the relationships between the Holding Tanks and the Mixing Tank. Specifically, during training our operators were required to note the cause-and-effect relationships between the Holding Tanks and the Mixing Tank on a worksheet, received hands-on training with an experimenter, and were given more time with the simulation before advancing to the experimental conditions. The result of the expanded training was a main effect on performance with little effect on fix; participants were slightly better at the overall task in all conditions, but rarely switched away from their primary instrument set even when not doing so led to failure. Reaction time analysis confirmed that our operators gave little thought to the within-round feedback that preceded task failure.

In Experiment 4, we provided our operators with a different form of enhanced training also designed to increase awareness of within-round feedback. Operators were taught to use the simulation, and then asked to *troubleshoot* a malfunctioning simulation to determine which of the two instrument clusters was malfunctioning. Thus, our operators had already experienced unreliable instruments before advancing to the experimental rounds. Unfortunately, this type of training also produced a main effect of training on all conditions, but little effect on operators switching between instrument sets. As in Experiment 3, RT analysis of our operators' decisions showed no pause when within-round feedback indicated problems with the primary instruments.

Finally, in Experiments 1-4, we reported evidence for a type of sub-optimal adaptation to the task involving our operator's decisions about the *amount of juice* to use from trial to trial in the experiment. In particular, Experiments 1-4 reported that the average amount of juice that was sent from an Holding Tank either remained constant or increased across experimental rounds when participants were successful, but *decreased* when participants failed. This effect was revealed to occur even before the malfunction points in a round, suggesting that participants made these strategy choices as a result of outcome failure between rounds, rather than as a reaction to the malfunction points. This strategy adaptation resulted in an increase in the average number of trials per round, but did not lead to successful adaptation to the within-trial feedback. Given that failing participants' reaction times did not increase across malfunction points in any round of any experiment, these results raised the possibility that the decision to utilize less juice was made between rounds in response to task failures, rather than within a round in response to within-round indicators that a current strategy was leading towards failure.

Why did participants choose to reduce the average input into the Mixing Tank? One possible reason would be that participants wanted to increase the number of trials they had in which to hypothesis test about why they had failed in previous rounds. This would be consistent with the idea that participants suspected that they had failed to notice some key aspect of the task, and by increasing the number of trials they were exposed to, also increase their chances of detecting the source of error in subsequent rounds. However, this theory presents something of a paradox: if participants had intended to increase their chances of success by increasing the number of trials they were exposed to, then why did they not take *more time per trial*, especially once the malfunction point had occurred?

The current experiment was designed to test two possible theories that might account for the inconclusive results of Experiments 1-4. The first, the *diligent hypothesis tester theory*, proposes that operators adopted a reduced input strategy in order to increase their potential for hypothesis testing during a round, but still failed to notice the malfunction points in the experiment despite the increased number of trials. This theory would be consistent with the notion that operators attempt to discover why they fail, but are unable to discern when within-round feedback indicates a problem with their strategy, and so they never find reason to pause within a round or abandon primary instruments.

In contrast to this theory is the *sub-optimal adaptor theory*, which proposes that an operator's decision to reduce Holding Tank input is, in and of itself, their adaptation to malfunctions. According to this theory, operators simply are unable to abandon mental set, and instead come up with somewhat plausible, but ultimately failing iterations of their original strategy. Operators that are sub-optimal adaptors do not attempt to problem solve by engaging in the details of their error, but rather, are carrying out something of an attenuated version of their original strategy.

Method

Participants

Participants in this study were 102 undergraduate psychology students from the University of Illinois at Chicago. The participants were randomly assigned to each of three groups of 34.

Procedure

Informed consent and debriefing were done off-line, but in the experiment participants interacted with a computer.

Experimental conditions. This study utilized three conditions to test between the diligent hypothesis tester theory and sub-optimal adaptor theory. In one condition, the *control condition*, no instrument malfunctions occurred. In the *color-malfunction condition,* the color instrument stopped matching those temperatures displayed by the gauges, and instead indicated that the liquid in Tanks A and B were shades of cool blue. In the *warning condition* color instruments also became unreliable, but when this occurred a highly salient warning was given onscreen that highlighted the mismatch between the color and instrument clusters that reappeared on every subsequent malfunction screen. The warning paused the action in the program, drew arrows between the color and gauge instruments, and presented the text 'Warning! Potential Color/Gauge Mismatch' in two places onscreen. Thus, the participants were reminded each time within-round feedback became relevant to the task.

Results

Our analysis focused on the success or failure of participants in the simulation, those participants' reaction times, and on the average input that those participants utilized prior to the malfunction points in each round.

Success / Failure

As shown in Figure 3, participants in the control condition did well at successfully operating the juice factory simulation, while participants in the color-malfunction and warning conditions did not. In all rounds, participants in the no malfunction condition reliably succeeded more frequently than either the malfunction or warning conditions. In Round 1, participants experiencing no malfunction succeeded in the simulation 88.2% of the time, compared with 38.2% when a malfunction occurred, and 32.4% when a malfunction occurred with a warning, χ^2 (2, N = 102) = 25.72, $p < .001$. In Round 2, participants experiencing no malfunction succeeded in the simulation 100% of the time, compared with 29.4% when a malfunction occurred, and 38.2% when a malfunction occurred with a warning, χ^2 (2, N = 102) = 40.80, $p < .001$. In Round 3, participants experiencing no malfunction succeeded in the simulation 88.2% of the time, compared with 52.9% when a malfunction occurred, and 32.4% when a malfunction occurred with a warning, χ^2 (2, N = 102) = 22.27, $p < .001$.

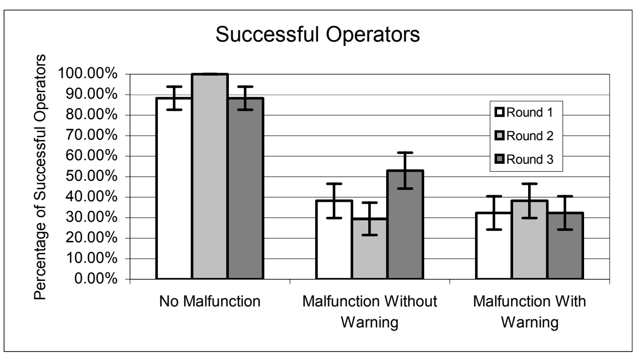
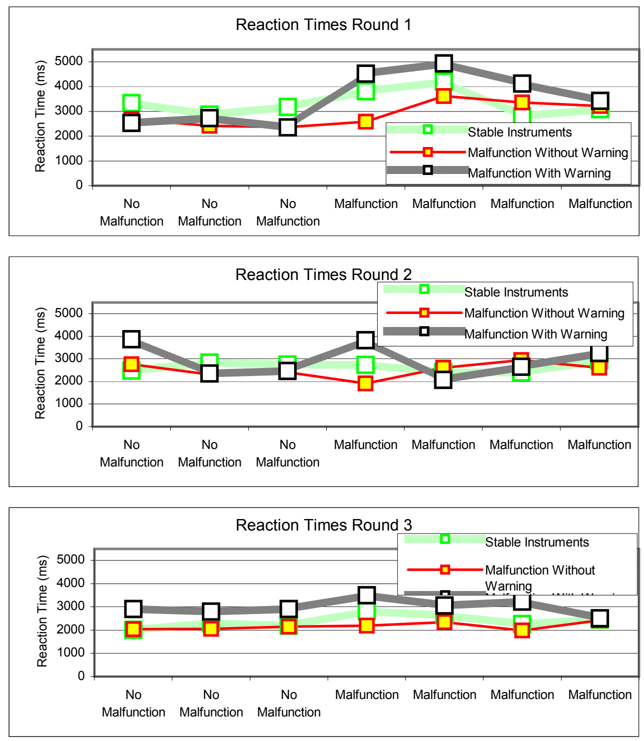


Figure 4: Percentage of successful participants across rounds in the simulation by experimental condition.

Reaction Times

As shown in Figures 5-7, reaction times in the three trials before the malfunction point, and the four trials after, were averaged within each condition in each of the three rounds. We conducted a 3 (condition: no malfunction, malfunction without warning, malfunction with warning) by 7 (trial: pre-malfunction 3, pre-malfunction 2, pre-malfunction 1, malfunction 1, malfunction 2, malfunction 3,malfunction 4) repeated-measures ANOVA on participants' reaction times for the three trials prior to malfunction and the four trials following malfunction for each of the three experimental rounds. Our analysis of Round 1 revealed no main effect of condition, $F(2, 111) = 1.01$, *ns*, a main effect of trial, $F(2, 111) = 5.30$, $p < .0001$, and no interaction, $F(2, 111) = 1.39$, *ns*. Our analysis of Round 2 revealed no main effect of condition, $F(2, 111) = .65$, *ns*, a marginal main effect of trial, $F(2, 111) = 1.82$, $p = .09$, and a significant interaction, $F(2, 111) = 2.13$, $p = .014$. Our analysis of Round 3 revealed no main effect of condition, $F(2, 111) = 1.75$, *ns*, no main effect of trial, $F(2, 111) = .37$, *ns*, and no interaction, $F(2, 111) = .58$, *ns*. In sum, the warning condition appeared to notice the warnings in Round 1, and slow their decision times per round because of them. However, by Round 2 these effects became marginal, and by Round 3, no condition appeared to notice malfunctions.



Figures 5-7: Participants' average reaction times across the malfunction point for rounds 1-3, split by condition.

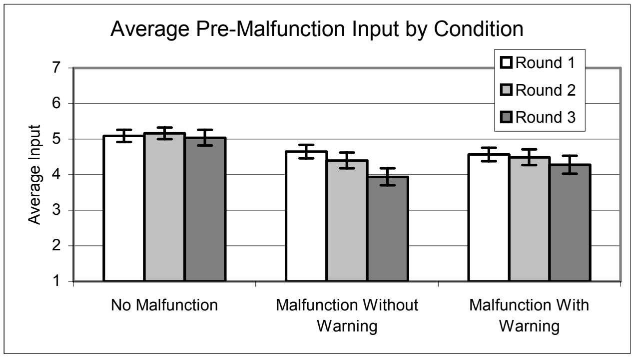
Average Input

We conducted a 3 (round: 1, 2, 3) by 3 (condition: no malfunction, malfunction, warning) repeated-measures ANOVA on the average input of participants prior to the malfunction point. This analysis revealed a main effect of round, $F(2, 196) = 4.58$, $p < .05$, a main effect of condition, $F(1, 98) = 2206.14$, $p < .05$, and no interaction, $F(1, 196) = 1.29$, *ns*.

Follow-up planned contrasts of pre-malfunction input data collapsed across conditions revealed that Round 1 pre-malfunction input was significantly greater than Rounds 2 and 3, $F(2, 196) = 4.12$, $p < .05$, and Round 2 was significantly greater than Round 3, $F(2, 196) = 4.46$, $p < .05$.

Follow-up planned contrasts of pre-malfunction input data collapsed across rounds revealed that the malfunction and warning conditions were significantly different than the control condition, $F(2, 196) = 42.60$, $p < .05$, and no significant difference between the malfunction condition and warning condition, $F(2, 196) = .37$, *ns*.

This pattern of results indicates that malfunctions, both with and without warnings, produced reliable drops in average pre-malfunction input, and that average pre-malfunction input dropped across the experimental rounds.



General Discussion

When two instruments are available and one of them malfunctions it seems as if a fully rational operator would switch to the other one. In prior studies, we observed a reduction in success rates when the information source a participant had practiced with malfunctioned, and a slow recovery with additional experience. Even in the face of repeatedly failed rounds, participants' *within* round reaction time data did not support the notion that they became more sensitive to *noticing* malfunction points. Instead, participants in past studies responded to malfunction by decreasing the amount of juice they sent through the factory simulation, possibility to increase the number of experimental trials before failing.

In this study, noticing malfunction points was controlled for by directly pointing out mismatches between color and gauge instruments. Our findings support the notion that our operators made sub-optimal strategy decisions between rounds, and this is a counterintuitive finding. It seems reasonable to assume that when someone is performing a task and fails, that they would then attempt to determine the cause of their errors through hypothesis testing in subsequent procedures. However, here operators generated hypotheses about the problems in their behavior only when they were not engaged in the task, between rounds.

Theoretically, this finding impacts theories of how people deal with errors during the performance of a complex task. In past work, one of us (see Ohlsson, 1996) proposed a detailed hypothesis about how people might interpret and unlearn an error in a cognitive skill. The finding reported here suggests that people might react to an error in a different way than by engaging the details of that error. Our participants apparently conducted no analysis while controlling the simulation about what was going wrong. Instead, they seemed to follow a very general disposition that we perhaps can formulate as *if our action causes trouble, try an attenuated version of that action*. This heuristic for how to deal with an undesirable outcome is reminiscent of the fast and frugal heuristics that Gigerenzer, Todd and the ABC research group (1999) have observed with respect to decision making and judgment in other domains. From a practical point of view, the lesson is interface designers should not assume that operators engage in extensive cognitive processing of the information available in a given machine interface.

If this form of adaptation turns out to generalize outside our juice factory simulation, there are several practical implications for designing safeguards against instrument malfunction. Operators may or may not engage in deep causal reasoning about the device they interact with, and these and other factors may influence their willingness to abandon their current strategy. The need to keep cognitive load within reasonable limits may inhibit causal accounts of an anomaly and encourage simple heuristics instead. It is important to know which fast and frugal heuristics device operators are likely to utilize, something that might not be possible without empirical studies. Knowledge of the heuristics that operators fall back on when an anomaly occurs might save system builders from implementing costly safeguards which are entirely reasonable, but likely to be overlooked or bypassed by frugal operators.

Acknowledgements

The authors would like to thank Dr. Alex Kirlik for his helpful comments on machine interaction in general, and heuristic approaches to HMI in particular. We also thank Heidi Beebe, Brendan Green, and Michael Holman for their help with data collection and analysis. Finally, we thank Ben Jee for many helpful discussions regarding the interpretation of our data.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting attentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P., and the ABC Research Group (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, 103, pp. 241-262.

Youmans, R., Ohlsson, S. (in review). Suboptimal adaptations to redundant instrument failure: The role of noticing behavior in feedback effectiveness. *Ergonomics*

Youmans, R., Ohlsson, S. (*2005*). Training produces suboptimal adaptation to redundant instrument malfunction in a simulated human-machine interface. *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society*.