UNIVERSITY OF ILLINOIS

MAY 9, 19 90

THIS IS TO CERTIFY THAT THE THESIS PREPARED UNDER MY SUPERVISION BY

JOHN HOWARD MENDENHALL, JR.

ENTITLED MULTIPLE FAULT DIAGNOSIS AND HUMAN BEHAVIOR:

DIAGNOSING A TWO-BIT BINARY ADDER

IS APPROVED BY ME AS FULFILLING THIS PART OF THE REQUIREMENTS FOR THE

DEGREE OF BACHELOR OF SCIENCE IN LIBERAL ARTS AND SCIENCES

Penelope [signature]

Instructor in Charge

APPROVED: [signature]

HEAD OF DEPARTMENT OF PSYCHOLOGY

O-1364

Multiple Fault Diagnosis and Human Behavior: Diagnosing a Two-Bit Binary Adder

By

John Howard Mendenhall, Jr.

Thesis

for the
Degree of Bachelor of Science
in
Liberal Arts and Sciences

College of Liberal Arts and Sciences
University of Illinois
Urbana, Illinois

1990

Abstract

A 2-bit binary adder was simulated to examine the effects of multiple faults on troubleshooting behavior. The multiple faults were two broken connections between components of the adder. An a priori analysis of the symptoms displayed by the dual faults revealed different types of interactions associated with the various dual faults. Interaction types were classified according to the predicted levels of difficulty experienced by subjects diagnosing the dual faults. Sixteen electrical engineering and computer science students were recruited as subjects. Each subject diagnosed four dual faults, one each from four different interaction classifications. The results did not support the predicted difficulty levels assigned to the interaction classifications. Subjects were able to diagnose the dual faults using the symptom states that displayed symptoms of one broken connection or the other, but not both and thus avoid any confusion associated with a high level of interaction. Further experiments with the adder will analyze subjects' performance using an idealized strategist for comparison and all subjects will experience various conditions of single and multiple faults.

Table of Contents

Introduction

The complexity of our technology will continue to increase as the engineering tools of the designers improve and the use of these tools expands. The use of complex, informationally rich technology during the design and development phase of our systems has resulted in improved ergonomics and reduced maintenance requirements. However, its use has also resulted in complex systems that are much more difficult for troubleshooters to diagnose and repair. Wohl (1982,1983) warned that the increased use of computer-aided design may eventually create problems that are insurmountable for the troubleshooter. As the interconnections between a system's components rise above seven, troubleshooters begin experiencing extreme difficulties as measured by the time it takes them to repair a system. The theoretical limits associated with processing in short term memory may place real limits on troubleshooters' diagnostic abilities so that the time needed to isolate and repair malfunctioning components could hypothetically approach infinity.

The trend toward ever more elegant refinement and precision in our systems is part of an overall paradigm that drives most of our technological development. Since it is unlikely that such a trend is reversible, troubleshooters will have to adapt and adjust to the increasing complexity of our systems. Adapting to the diagnostic environments created by the current generation of complex systems already places extreme demands on the troubleshooter's cognitive processes, and this is especially true when diagnosing novel, intermittent, and multiple faults (Bereiter & Miller, 1988; Rasmussen & Jensen, 1974; Wohl, 1983; Woods, Roth, & Bennett, 1987; Yoon & Hammer, 1988). Not only does the increased complexity make it difficult for troubleshooters to understand the structure and functioning of a system, but there are many more components that can malfunction and disable the system.

In addition, it is reasonable to assume that a sufficiently complex system composed of components with a given reliability has a greater probability that two or more of its components will malfunction simultaneously rather than in temporal isolation (Sanderson, 1990). With the likelihood that multiple fault diagnosis will typify the troubleshooting process, there is a tremendous need to understand the phenomenon in terms of its effects on the human troubleshooter. Are there parallels between the impediments to timely diagnosis of single faults and the problems faced by troubleshooters diagnosing multiple faults? Because in any environment the "human factor" behaves in accordance with a loosely defined set of innate cognitive strengths and weaknesses, we must assume that these are salient issues in any fault diagnosis domain. Therefore, it would be helpful to examine some of the research that focuses primarily on single fault diagnosis.

Rasmussen and Jensen (1974) identified different search strategies employed by troubleshooters to isolate faults. A topographic search involves a series of good/bad checks along a structural pathway or within the component under consideration Good/bad judgements are made for each test according to the rules governing the component. Symptomatic search involves matching an observed symptom to a symptom retained in records of symptom patterns analogous to a library or database. If the library does not contain the observed symptom, then hypotheses must be generated to explain the symptom in a process of search by hypothesis and test. The troubleshooter must employ a mental representation of the system to imagine system states that could explain the observed symptoms or invoke the symptoms in an identical functioning system. In either case this is the most cognitively demanding strategy because the troubleshooter must strive for a deep understanding of the system while considering the abnormal states that could cause the symptoms. While troubleshooters show a preference for the topographical search, complex systems with pathways or components that contain redundancies and feedback loops often force troubleshooters to use the more cognitively demanding strategies (Rasmussen & Jensen, 1973, 1974; Rouse, 1981).

Other researchers have focused on the role of troubleshooters' mental models of systems and how these mental representations affect troubleshooting abilities in general (deKleer & Brown, 1983; Gitomer, 1988; Sanderson & Murtagh, 1990). Gitomer (1988) studied a group of radar technicians who had about the same amount of job experience. He found a significant disparity between those technicians rated as more and less skilled by their supervisors when he tested them on a radar system's functional structure. Gitomer concluded that the technicians rated as less skilled were unable to use their knowledge of specific systems they had repaired in the shop and apply it to the functioning of radar systems in general.

Sanderson and Murtagh (1990) identified some of the impediments to successful fault diagnosis when the diagnosis is based on faulty and incomplete mental models of a system. Their subjects were allowed to explore a simulated logic network until they were confident they understood its structure and functioning. Faults were then inserted in those areas of the network where the subject's mental model was incomplete or inaccurate. While many of the subjects were able to detect an abnormality, none were able to diagnose it if it challenged a false belief about that part of the network's structure. It seems that the subjects were either unable or unwilling to adjust their mental models when the evidence called for it.

In their discussion of "mechanistic mental models" (p. 155), deKleer and Brown (1983) make a distinction between the quality of a mental model needed to "envision" the structure of a device and that of the model needed to describe the "running" or functioning of the device. The assumptions made and the ambiguities between these two models may cause the difficulties experienced by the troubleshooter. In fact the troubleshooter does the opposite of the envisioning process in that the needs of fault diagnosis require a shift from the model of the device's functioning to a new model of the structure created by the fault in order to explain the malfunctioning. Even a perfect mental model of the functioning of a device may not be enough to assure diagnosis if "the fault so perturbs the function that the function is not even part of the intrinsic mechanism." (p. 181)

Yoon and Hammer (1988) make a distinction between shallow and deep reasoning and, in the case of the latter, discuss the cognitively taxing exercise of reasoning about the functioning of complex systems. Ideal performance in a fault diagnosis task would require optimal rationale to deduce with certainty the cause of a malfunction in a complex system. However, cognitive limitations often preclude such optimal behavior and the troubleshooter must depend on inductive reasoning to infer the fault and compensate for a limited processing ability.

Multiple Faults

Given the cognitive processing needed to diagnose faults in general, how and in what new ways do multiple faults exacerbate the troubleshooter's limitations? Before we can prescribe effective aiding for troubleshooters, we must first answer these and other more specific questions related to multiple fault diagnosis.

The field observations reported in Bereiter and Miller (1988) would indicate that multiple faults do create new problems when troubleshooting complex systems. Multiple faults often confuse troubleshooters because the latter tend to focus on single fault hypotheses which subsequently do not explain all of the observed symptoms. They will then abandon the hypothesis and continue to search for a single fault to explain the symptoms, even when they have actually isolated one of the faults causing the abnormal system states.

Despite a clear need for laboratory research that examines the effects of multiple faults on humans' diagnostic abilities, very little has been conducted to date. Most of the multiple fault research that is reported in the literature pertains to the artificial intelligence domain and focuses on automatic process and quality control (Cox, Ivanov, Agarwal, & Rajski, 1988; deKleer &Williams, 1987; Pazzani, 1987).

However, Moray and Rotenberg (1989) have studied human management of multiple faults in a simulated process control task. Two sequential faults were introduced into the process and subjects' eye movements and control actions were analyzed. Even though subjects spent time observing the second fault in the sequence, they took no action to remedy its effects until they had finished managing the effects of the first fault. This does not bode well for real world process control in that the accumulation of faults across time could threaten a system, especially if the later faults carry more severe consequences for the systems processes.

The current study examines multiple faults in a static context that might represent the kind of diagnosis that occurs in the repair shop or with a system that has been taken off line or shut down specifically for repair. While timely and accurate diagnosis is certainly the goal, the time issues associated with fault management in a continuing operation are not present in this environment. In addition, faults do not occur sequentially but are present in the system when troubleshooters begin the diagnostic task.

In this context the effects of multiple faults on troubleshooters' diagnostic abilities are examined in a simulated 2-bit binary adder. The adder environment is complex enough to explore some of the salient issues described by researchers in the field (e.g., Bereiter & Miller, 1988) but simple enough for a detailed analysis of the problem space and subsequent development of specific hypotheses predicting troubleshooters' behavior.

The current study is the first in series of experiments with the adder and our primary goal was to establish a set of objective criteria that could be used to define the level of difficulty associated with various multiple faults in the adder. In this first experiment, multiple faults were limited to two broken connections between the components of the adder (see Figure 1). The symptoms of these dual faults were analyzed in accordance with any interactions they displayed in the adder's output states. Different types of interactions were identified and these interactions were classified according to the predicted difficulty they would cause our troubleshooters. As we gain a deeper understanding of troubleshooting behavior in this context-specific domain, we hope to be able to generalize our conclusions to the broader issues of training and processing aids as they relate to multiple fault diagnosis in a variety of complex systems.

INSERT FIGURE 1 ABOUT HERE

Analysis of the Diagnostic Task

The multiple faults used in this experiment were restricted to two broken connections for each trial. There were 16 unique combinations of binary inputs, and their corresponding output states were used to evaluate the symptoms displayed by each broken connection as a single fault. There were 210 possible pairs of these faults (see Figure 1). For each of the 16 possible input states, the symptoms displayed by these dual fault combinations were compared to the symptoms displayed by their single fault components. It was this comparison that defined how the 2 single faults interacted to produce the symptoms of the dual fault compositions. Based on the degree of interaction between these symptoms, the dual faults classified into unique categories. The categories were subsequently ordered to reflect the predicted difficulty of diagnosing the dual fault interactions. These levels of interaction were used to define the four types of faults used in the experiment. We predicted that the knowledgeable subjects would experienced more difficulty while diagnosing the dual faults would be least for Type 1 and greatest for

Type 4. It should be kept in mind that all dual faults were type cast into the highest level of interaction possible based on at least one output state displaying symptoms that met the criteria as described above. Based on this a priori analysis of the dual faults' symptoms, four levels of interaction were used in the experiment (see Figures 2 to 5). To illustrate the predicted difficulty of these four fault types, it is helpful to first consider the highest level of interaction (and subsequently the highest level of difficulty) then work back to the lowest level.

Type 4 Interactions. The highest level of interaction (Type 4) displayed at least one output that mimicked normal output where the single fault components had instead displayed fault symptoms. Figure 2 is an example of this type of interaction. The interaction of the two faults eliminated the symptoms displayed by each singly. Two faults that produce a Type 4 dual fault display symptoms at the same output state when they are analyzed as single faults. When they occur together, this output no longer displays symptoms of either fault X or Y. In fact, the output is indicative of a normally functioning adder. The two faults interact to so perturb this output state that all symptomatic information is lost and the circuit appears normal. Since the evidence that appears for the two faults when they occur singly is lost due to their interaction, it was hypothesized that dual faults displaying Type 4 interactions would be the most difficult to diagnose.

INSERT FIGURE 2 ABOUT HERE

Type 3 Interactions. Figure 3 is an example of the second highest level of interaction (Type 3) which displayed at least one state whose symptoms appeared in neither fault X nor fault Y when analyzed as single faults (i.e., they were "unique"). The symptoms were unique due to the interaction of fault X and fault Y occurring together. The second highest level of interaction (Type 3) is similar to the Type 4 interaction. However, instead of the two faults interacting to eliminate symptomatic information, their interaction produces symptoms that are unique. That is, the symptoms are different for the given output state when compared to the symptoms of the two contributing single faults. It was hypothesized that the symptoms caused by the single faults' interaction would make the resulting dual fault somewhat difficult to diagnosis, but not as difficult as dual faults with Type 4 states where all symptomatic information is lost.

INSERT FIGURE 3 ABOUT HERE

Type 2 Interactions. The next highest level of interaction (Type 2) displayed an "intermediate interaction" (see Figure 4). In this case, at least one output state displayed symptoms of both fault X and fault Y. For an intermediate level of interaction, the symptoms of the dual fault display the symptoms of both component faults in the given output state. Their occurrence together does not perturb the symptomatic evidence of the output, in that the symptoms are indicative of both single faults simultaneously. Dual faults with Type 2 states were hypothesized to be less difficult to diagnose than dual faults with Type 3 or Type 4 interactions.

Type 1 Interactions. Figure 5 is an example of the lowest level of interaction (Type 1) defined as having symptoms displaying "no interaction." None of the 16 output states of the no interaction level (Type 1) displayed symptoms of both single faults simultaneously. Each state displayed either a normal output, a symptom for fault X, or a symptom for fault Y. Since any given output state could display symptoms that were only attributable to one component or the other, it was hypothesized that Type 1 dual faults would be the easiest to diagnose. Subjects would be able to quickly diagnose either fault using symptomatic evidence that was free of interactive interference from the symptoms of its counterpart.

In summary, these dual fault types were defined during the initial analysis of the 2-bit binary adder and their definitional characteristics were incorporated into the experimental design. It was hypothesized that the highest level of interaction seen for a particular dual fault would be correlated to the level of difficulty knowledgeable subjects would experience while diagnosing the fault. Therefore, objective measures of difficulty would be greatest for the Type 4 dual faults and least for the Type 1 dual faults.

INSERT FIGURE 4 ABOUT HERE

Method

Subjects

Sixteen subjects were recruited from Electrical Engineering classes at the University of Illinois. Subjects were required to have completed introductory courses in digital logic and computer design. All subjects were paid for their participation.

Materials

A 2-bit binary adder was simulated on a Macintosh IIx computer (see Figure 1.). There were four input displays and four output displays. The display for the carry value of the lower bit (CARRY0) was included for diagnostic purposes. The logical structure of the adder was constructed using nine logic gates; five AND gates, two OR gates, two exclusive OR gates (XOR), and one inverter. The broken connections subjects were to diagnose were programmed to always float low, so that the input to the succeeding circuit component was always zero. A start button and a finish button also were displayed and were used to initiate and end each trial. In addition, a diagnosis window in which a complete record of the subject's diagnoses was displayed. The program for the interface and diagnostic task was written in Pascal.

Subjects were encouraged to verbalize their problem solving processes and these protocols were recorded with an audio cassette recorder. A pencil and paper workload rating was also used to compare the subjects' assessment of the task with their actual performance as recorded by the computer. The scale used was the NASA TLX workload rating scale (Hart & Staveland, 1989).

Design

A 4 x 4 x 4 Latin square design was used in this experiment. All subjects diagnosed one dual fault from each of the four types. Four groups of four subjects each were tested with a different fault from each type except for Type 4 faults. Only two Type 4 faults had unique diagnostic solutions, so two groups saw one of these and two groups saw the other. Therefore, the total number of dual faults with unique diagnostic solutions used in this experiment was 14 and not 16. To control for practice effects, a Latin square design was used in each of the four groups (see Table 1). Each of the four subjects within a group experienced a different order of presentation for the same four faults.

INSERT TABLE 1 ABOUT HERE

Procedure

Subjects were instructed to identify and register the two broken connections in each trial. To do this they had to change the input to the circuit and evaluate its corresponding output, then record their diagnoses.

The subjects were given verbal instructions about how to test the circuit and diagnose the faults (see Appendix). They were to use the mouse to both change the values of the input and to diagnose broken connections. When subjects clicked in the input boxes, the value flipped from 1 to 0 or from 0 to 1. The inputs determined the corresponding output state that appeared in the output boxes. These in turn reflected the effects of the dual fault for any given trial. Any deviations from the expected normal sum of the binary addition were interpreted as symptomatic evidence of one, the other, or both broken connections in the trial. This was the only manner in which subjects could test the circuit for broken connections. Subjects were informed that broken connections were simulated as floating "low." That is, they always simulated a value of zero at that point in the circuit.

Once subjects had hypothesized a broken connection, they registered their diagnosis by clicking on the two components directly linked by the connection. For example, in Figure 1 if they believe the connection between XOR2 and XOR8 was broken, they would register that diagnosis by clicking on the XOR2 and XOR8 components. They could do this in any order. The components would temporarily appear with their background and foreground reversed, then return to normal and the solid black line connecting them would become lightly dotted. Subjects were allowed to make two kinds of diagnoses. This distinction was programmed into the interface because subjects in a pilot study were using the diagnose function to visually explore their fault hypotheses. Its intended use was as a register of their final diagnoses and we were unable to distinguish between exploratory versus definite diagnoses.

The first distinct diagnosis was an "exploratory" diagnosis. Subjects were encouraged to register all hypotheses in the set they were currently considering. This served two purposes: 1) to capture some of the subjects' diagnostic strategies in their files, and 2) to provide the subjects with a visual aid to track their current hypotheses. Exploratory diagnoses were made by simply clicking on the components, as described above, and the connection would become light gray and dotted. This in turn would be registered in the "Diagnosis" window in its chronological order with a light gray line joining the names of the components clicked.

The second kind of diagnosis was a "definite" diagnosis. Subjects were instructed to make these diagnoses only when they were fairly certain that they had narrowed their hypothesis set to connections that would cause the symptoms they observed, if in fact they were broken. They registered definite diagnoses in the same manner as for the exploratory diagnoses except that they pressed the command key while clicking on the components. In this case, the connection became a light red dotted line and the diagnosis was registered in the "Diagnosis" window with the word "red" in parentheses following the diagnosis.

Both types of diagnoses could be removed at any time prior to the end of a trial. Subjects would simply click on the components linked by the diagnosed connection and the connection would return to a solid black line. This "Undiagnosis" would also be registered in the "Diagnosis" window with a solid black line between the component names.

Subjects could test and make or remove diagnoses at any time during the trial. No particular order was imposed. When subjects were satisfied with their diagnoses, they informed the experimenter, then clicked on the "finished" button to end that part of the trial. At this point, a dialogue box appeared and listed any of the 16 input states the subjects had not entered during this part of the trial.

Since subjects were allowed to actively survey the circuit's function by choosing the inputs to test, it was possible that they would make their final definite diagnosis without testing all 16 inputs and examining their corresponding output states. Because the dual fault types were defined according to given symptoms produced by specific inputs, subjects had to see the effects of all 16 inputs in order for us to objectively evaluate difficulty levels based on the data from subjects' records. Therefore, if the dialogue box listed any "untested" states, the experimenter made a note of these, directed the subjects back to the active testing interface, and had the subjects enter the untested inputs. As subjects entered each of these, they were asked to re-evaluate their diagnoses based on the corresponding output states. They were allowed to make other tests or change their diagnoses until they were satisfied that they had correctly identified the two faults. This process continued until they had seen all 16 input states, at which point the trial was ended.

At the beginning of the session, subjects performed a pair comparison of all TLX workload dimensions, mental demand, temporal demand, physical demand, effort, performance, and frustration. This was to determine the relative weights of each dimension so that a weighted workload scale could be calculated using the later trial-by-trial workload ratings. After the second single fault and each subsequent dual fault, subjects rated their subjective workload on the six dimensions of the TLX.

Results and Discussion

The dependent variables used for measuring the levels of difficulty incorporated various dimensions of the time spent diagnosing the circuit, the frequency of diagnoses and tests, the success or failure to diagnose for each dual fault trial, and the subjects' workload ratings. Most measures were taken from the period subjects spent till their first definite solution, prior to their viewing the input states they had not tested. However, the time till subjects finished each trial also was analyzed. Time was measured in seconds from the beginning of each trial, and this was broken down further into dwell times for viewing various output states and times to discovery and diagnosis of the two fault components. The amount of time spent testing and diagnosing the circuit was also analyzed. The diagnoses themselves were measured as; a) neither fault component correctly diagnosed, b) one component but not the other, or c) both components correctly diagnosed.

Time and Correctness of Diagnoses

None of the results were significant in terms of the predicted difficulty for the types of interactions displayed by the dual faults (see Table 2 for fault type means). The time till the first solution, the time till the finish of each trial, and the time spent testing and diagnosing did not vary with the four types of interactions: all F values were less than 1.0. The only significant results were practice effects reflected by the decrease in time to the subjects' first solutions across trials, $F(3, 15) = 4.275$, $MSe = 69662.709$, $p < .01$, time spent testing the circuit until subjects' first solution, $F(3,15) = 2.916$, $MSe = 62198.593$, $p < .05$, time till subjects finished the trial, $F(3,15) = 4.469$, $MSe = 68612.065$, $p < .01$, and time spent in diagnosis activities until the subjects' first solution, $F(3,15) = 3.09$, $MSe = 5829.179$, $p < .05$ (see Table 3 for practice effect means). Figure 6 shows the mean times "till the first solution" for the

14 dual faults arranged according to the type of interactions they displayed. Figure 7 shows the same means sorted in ascending order.

INSERT TABLES 2 AND 3 ABOUT HERE

INSERT FIGURES 6 AND 7 ABOUT HERE

The number of exploratory diagnoses and the number of tests made also were unrelated to the difficulty of the dual faults as defined by the a priori analysis. A chi-squared was performed on the relation between which of the two broken connections' symptoms was viewed first and which one was subsequently diagnosed first. The null hypothesis in this case was that no tendency would emerge and equal probabilities were assigned to matches and mismatches between the first symptoms seen and the first broken connection diagnosed. Dual faults displaying Type 1 interactions revealed a significant tendency for subjects to diagnose first those faults whose symptoms were seen first, $\chi^2(1, \underline{N} = 16) = 4.0, \underline{p} < .05$. The results from the other three categories of dual faults showed no significant tendency toward an ordered diagnostic strategy.

Subjects' overall workload ratings and weighted workload ratings were compared to the type of interaction associated with each dual fault. There was no relation between the subjects' ratings and their performance on the dual faults.

There were very few incorrect diagnoses overall and these occurred in only 3 of the 14 dual faults during the time to first solutions (dual faults BF, DF, and NU). One of the four subjects who saw the Type 1 dual fault "BF" misdiagnosed "F" while three of the four subjects who saw the Type 1 dual fault "DF" also misdiagnosed "F." In all four of these cases, "F" was incorrectly diagnosed as the broken connection labeled "H" in Figure 1. The Type 3 dual fault "NU" was incorrectly diagnosed by three of the four subjects who saw it, and in each instance "N" was misdiagnosed as the broken connection labeled "J."

Examination of Incorrect Diagnoses

All subjects who got "F" wrong saw disconfirming evidence that should have dissuaded them from their incorrect diagnosis. Figure 8 and Figure 9 show the amount of time spent viewing the symptoms of "F" for each of the eight subjects who saw the two dual faults involving "F." The time spent viewing the symptoms and the order of appearance for these dual faults were not functions of the correct versus incorrect performance measure. It may be that the difficulty experienced by subjects diagnosing "F" was instead a function of the type of logic gate it immediately precedes, which is the only inverter in the circuit. Each subject was informed at the beginning of their sessions that broken connections changed the circuit as if a zero value proceeded through the circuit from that point forward. However, it is also a common industry practice to design digital circuitry that "floats high" (i.e., passes a value of one) in corresponding real world instances of our fault simulations. When this is coupled with the function of the inverting logic gate, the incorrect diagnoses may be a reflection of the subjects' confusion caused by a conflict between their knowledge of how a real world circuit should function and the symptoms displayed by our simulation.

INSERT FIGURES 8 AND 9 ABOUT HERE

In the case of the incorrect diagnoses for the Type 3 dual fault "NU," subjects had to see at least one of the two unique symptoms produced by the interaction of "N" and "U" in order to correctly diagnose "NU." In addition, a correct diagnosis required they see the only output state that could disconfirm any but the correct hypothesis for "N." If subjects saw either or both of the unique interaction symptoms without also seeing the disconfirming output state, the symptoms were the same as those for "JU," and this is in fact how "NU" was misdiagnosed in all three cases. Only one of the three subjects who misdiagnosed "NU" saw the disconfirming output state, and that subject had NU first in order of appearance (that is, the subject had had no prior practice diagnosing the dual faults).

Conclusion

The results show that the types of interactions displayed by the dual faults was not an effective criterion for predicting the amount of difficulty knowledgeable subjects would experience diagnosing the dual faults. None of our dependent variables showed a relation to the predicted levels of difficulty. The only significant result showed a practice effect, which would be expected as subjects become more familiar with the computer interface, the requirements of the task, and the problem space in general. Therefore, our initial hypothesis was wrong in that the types of interactions we found within the various dual faults did not translate into clear-cut levels of difficulty actually experienced by our subjects as they diagnosed the 2-bit adder.

The reason for this is that subjects were able to avoid any confusion associated with the higher levels of interaction by focusing on symptom states that were indicative of one broken connection or the other. In some cases, only one of the 16 output states displayed the symptoms that classified the given dual fault as having a particular type of interaction (e.g., KN and KP). For the same dual fault, at least one, and often several of the other 15 output states displayed symptoms of one broken connection or the other, but not both. The remaining states displayed normal output. In fact, subjects often correctly diagnosed a dual fault without seeing the symptom state(s) that classified its interaction type. The only exception to this was the dual fault NU which did not have any output states that were symptomatic of the broken connection N by itself. In fact, three of the four subjects who saw NU misdiagnosed it as JU.

A post-analysis examination of the subjects' data files reveals that most of our subjects spent very little time viewing the symptom states used to classify the higher types of interaction, if they viewed them at all. Clearly, our a priori analysis was inadequate in that the dual fault symptoms were evaluated in isolation as to their effect on our subjects' abilities to diagnose the broken connections. Our hope was to establish objective measures of difficulty for these dual faults within the context of a 2-bit binary adder. However, without a consideration of the symptomatic evidence present in the other output states, accurately predicting the difficulty of a dual fault was not possible.

Summary

Based on our results, we are unable to answer how multiple faults may exacerbate the cognitive limitations associated with reasoning during fault diagnosis in general. However, we can draw some parallels to the problems experienced by troubleshooters diagnosing single faults.

The difficulties experienced by subjects diagnosing dual faults involving the broken connection F is best explained by a poor mental representation of the circuit's functioning at the inverter. Similar to the effec - seen by Sanderson and Murtagh (1990) a poor mental model of our simulation's structure at this component will increase the likelihood of a misdiagnosis for F.

In the case of the dual fault NU, two of the incorrect diagnoses are best explained in terms of a bias for confirmation of hypotheses. These two subjects never sought the evidence to disconfirm their errant hypotheses (Klayman & Ha, 1987). However, when they were forced to see the output from the untested input state that provided the disconfirmation, they changed their diagnoses to the correct one.

Subjects' strategies were limited to topographical searches and searching by hypothesis and test as described by Rasmussen and Jensen (1974). Although directly testing the components within the circuit was not possible (i.e., subjects were not able to probe the individual connections) subjects tested individual components by mentally tracing the values of the inputs through pathways and evaluating the integrity of these pathways in terms of the resulting outputs. When a particular input produced a bad output, subjects developed a set of possible hypotheses to explain the bad output and proceeded to test the pertinent components with different inputs. The efficiency of the subjects' strategies is reflected in the cumulative, chronological histories of the tests they made and the hypotheses they entertained as exploratory diagnoses.

This is in contrast to our initial hypothesis of predicted difficulty that was based on the interaction of the dual faults as reflected by symptom states taken in isolation. Therefore, we are developing an "idealized" automated strategist that will evaluate subjects' performance in terms of the efficiency of their tests and the corresponding validity of their exploratory diagnoses. This will allow us to shift our focus away from quantifying the difficulty associated with this particular diagnostic environment. Instead, the performance of subjects in future experiments with the 2-bit adder will be analyzed in terms of the quality of their reasoning under different fault conditions.

Several questions remain about how multiple faults affect troubleshooters' abilities to efficiently diagnose a complex system. We told our subjects that their task was to diagnose two broken connections in the circuit. While this initial experiment 's purpose was to quantify the innate difficulty of the dual faults, this is not how the task is assigned to troubleshooters in the real world. Troubleshooters are not told ahead of time that they are diagnosing a given number of malfunctions. As Bereiter and Miller (1988) observed, troubleshooters tend to become fixed on sets of single fault explanations and find it difficult to begin considering multiple causes for the observed symptoms.

In order to understand how and under what conditions troubleshooters shift their focus from single to multiple fault hypotheses, the next set of experiments will vary the conditions of single faults and multiple faults for each subject. Their performance will be compared to the idealized strategist. Hopefully, we will be able to pinpoint where in the diagnostic sequence the evidence triggers the cognitive shift from single to multiple hypotheses, if it does trigger a shift, and whether the shift is warranted based on the cumulative history of the subjects' tests.

In conclusion, there is an urgent need for basic research in multiple fault diagnosis. We need to understand the cognitive processing needs of the people we call upon to maintain and operate our complex systems before we can develop effective aids to facilitate the process. In addition, this knowledge will add to the tools available to human factors practitioners involved in the design of our systems. It is hoped that this knowledge will contribute to realistic design strategies that consider all facets of human-machine interaction, and not just ease of use.

References

Bereiter, S. J., & Miller, S.M., (1988). Sources of difficulty in troubleshooting automated manufacturing systems. In W. Karwowski, H. R. Parsaei, & M. R. Wilhelm (Eds.), Ergonomics of hybrid automated systems I (pp. 37-50). Amsterdam: Elsevier.

Cox, H., Ivanov, A., Agarwal, V. K., & Rajski, J., (1988). On multiple fault coverage and aliasing probability measures. International Test Conference, 314-321.

deKleer, J., & Brown, J. S., (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner & A. L. Stevens (Eds.), Mental Models (pp.155-190). Hillsdale, NJ: Erlbaum.

deKleer, J. & Williams, B. C., (1986). Reasoning about multiple faults. Artificial Intelligence Conference, 132-139.

Gitomer, D. H., (1988). Individual differences in technical troubleshooting. Human Performance (pp. 111-131). Hillsdale,NJ : Erlbaum.

Hart,S. G. & Staveland, L. E., (1989). Development of multi-dimensional workload rating scale: results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human Mental Workload. Amsterdam: North Holland Press.

Klayman, J. & Ha, Y. W., (1987). Confirmation, disconfirmation, and information in hypothesis testing. Psychological Review, 94, 211-228.

Moray, N. & Rotenberg, I., (1989). Fault management in process control: eye movements and action. Ergonomics, 32, 1319 -1342.

Pazzani, M. J., (1987). Failure-driven learning of fault diagnosis heuristics. IEEE Transactions on Systems, Man, and Cybernetics, 17, 380-394.

Rasmussen, J. & Jensen, A., (1973). A study of mental procedures in electronic trouble shooting (Riso-M-1582). Riso: Danish Atomic Energy Commission.

Rasmussen, J. & Jensen, A., (1974). Mental procedures in real-life tasks: A study of electronic trouble shooting. Ergonomics, 17, 293-307.

Sanderson, P. M. & Murtagh, J. M., (in press). Predicting fault diagnosis performance: Why are some bugs hard to find? IEEE Transactions on Systems, Man, and Cybernetics.

Wohl, J. G., (1982). Maintainability prediction revisited: Diagnostic behavior, system complexity, and repair time. IEEE Transactions on Systems, Man, and Cybernetics, 12, 241-250.

Wohl, J. G., (1983). Connectivity as a measure of problem complexity in failure diagnosis. (Available from Alphatech Inc., 3 New England Executive Park, Burlington, Massachusetts.

Woods, D. D., Roth, E. M. & Bennett, K., (1987). Explorations in joint human-machine cognitive systems. (Available from Westinghouse Research & Development Center, Pittsburgh, Pennsylvania, 15235.)

Yoon, W. C. & Hammer, J. M., (1988). Deep-reasoning fault diagnosis: An aid and a model. IEEE Transactions on Systems, Man, and Cybernetics, 18, 659-676.

Author Notes

I would like to thank Jeff Holden for his helpful comments and his programming expertise. Thank you to Teresa Appeddu for all our discussions and her assistance running the experiment. Finally, I would like to thank Mike Coles as the honors program director for his advice and encouragement, and Penny Sanderson, my advisor, for her guidance and support during my participation in the honors program.

Table 1

Latin Square Design of Experiment

Interaction Type and Order of Appearance

Group		Type 1	Type 2	Type 3	Type 4
Group A	Dual fault	1 = DE	2 = BS	3 = PV	4 = KN
Subject 1		1	2	4	3
5		2	3	1	4
9		3	4	2	1
13		4	1	3	2
Group B	Dual fault	5 = BE	6 = DN	7 = UV	8 = KP
Subject 14		5	6	8	7
2		6	7	5	8
6		7	8	6	5
10		8	5	7	6
Group C	Dual fault	9 = DF	10 = BU	11 = SV	12 = KN
Subject 11		9	10	12	11
15		10	11	9	12
3		11	12	10	9
7		12	9	11	10
Group D	Dual fault	13 = BF	14 = DV	15 = NU	16 = KP
Subject 8		13	14	16	15
12		14	15	13	16
16		15	16	14	13
4		16	13	15	14

Table 2

Means Across Fault Types

Dependent variables	Means in minutes for the interaction types			
	Type 1	Type 2	Type 3	Type 4
Time till the first solution	9.25	8.00	8.37	7.70
Time to finish trial	11.42	9.55	10.40	9.22
Time spent testing till first solution	7.35	7.25	6.57	6.22
Time spent diagnosing till first solution	2.02	1.62	1.97	1.50

Table 3

Means Across Positions

Dependent variables	Means in minutes for ordered position			
	Position 1	Position 2	Position 3	Position 4
Time till the first solution	10.98	8.15	8.72	5.45
Time to finish trial	13.00	9.85	10.38	7.37
Time spent testing till first solution	8.65	7.12	7.20	4.40
Time spent diagnosing till first solution	2.50	1.85	1.62	1.17