LEVEL II

①

Research Memorandum 79-5

AD A077978

ADAPTIVE COMPUTERIZED TRAINING SYSTEM (ACTS): RELATIONSHIPS BETWEEN UTILITY SIMILARITY AND STRATEGY SIMILARITY

Bruce W. Knerr

DDC FILE COPY

EDUCATIONAL TECHNOLOGY AND SIMULATION TECHNICAL AREA

DDC
RECEIVED
DEC 12 1979
D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

U. S. Army
Research Institute for the Behavioral and Social Sciences

April 1979

79 22 5 116

(14) ARI-RM-79-5

(16) Army Project Number 2Q762722A764

Training and Education

(9) Research Memo~~randum~~ 79-5

(6) ADAPTIVE COMPUTERIZED TRAINING SYSTEM (ACTS):
RELATIONSHIPS BETWEEN UTILITY SIMILARITY AND STRATEGY SIMILARITY.

(10) Bruce W./Knerr

(11) Apr 79

(12) 35

Submitted by:
Michael Kaplan, Acting Chief
Educational Technology and Simulation Technical Area

Accession For	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification ____	
By ____	
Distribution/	
Availability Codes	
Dist.	Avail and/or special
A	

Approved by:

Frank J. Harris, Acting Director
Organizations and Systems
Research Laboratory

Joseph Zeidner, Technical Director
U.S. Army Research Institute for
the Behavioral and Social Sciences

Research Memorandums are informal reports on technical research problems. Limited distribution is made, primarily to personnel engaged in research for the Army Research Institute.

408010 Gur

ADAPTIVE COMPUTERIZED TRAINING SYSTEM (ACTS): RELATIONSHIPS BETWEEN UTILITY SIMILARITY AND STRATEGY SIMILARITY

INTRODUCTION

The results of

In 1974 the Army Research Institute initiated an effort to apply Artificial Intelligence (AI) techniques to electronic troubleshooting training. The result of this effort was the Adaptive Computerized Training System (ACTS). While the ACTS has frequently been described elsewhere (see Appendix 1) a description is required here in order to provide the background for the problem addressed in this memorandum.

Simply, the ACTS is an "intelligent" simulator on which a student troubleshoots a complex electronic circuit by making various test measurements, replacing the malfunctioning part, and making final verification measurements. The "intelligent" component is an adaptive computer program which models the student's decision structure, compares this structure to that of an expert, and when complete, will adapt the instructional sequence and provide feedback to eliminate discrepancies between the two. An Expected Utility (EU) model of decision making is the basis of the models of both the student and expert troubleshooters.

P 31

The ACTS consists of four major components: (a) the task model; (b) the expert model; (c) the student model; and (d) the instructional model. The task model is a simulation of the circuit on which the student is to be trained. The circuit currently being used is a modular version of the Heathkit IP-28 Power Supply[1]. A simplified schematic diagram of this circuit is shown in Figure 1. The expert model

[1] Commercial designations are used only for precision of description. Their use does not constitute endorsement by Department of the Army or the U. S. Army Research Institute.

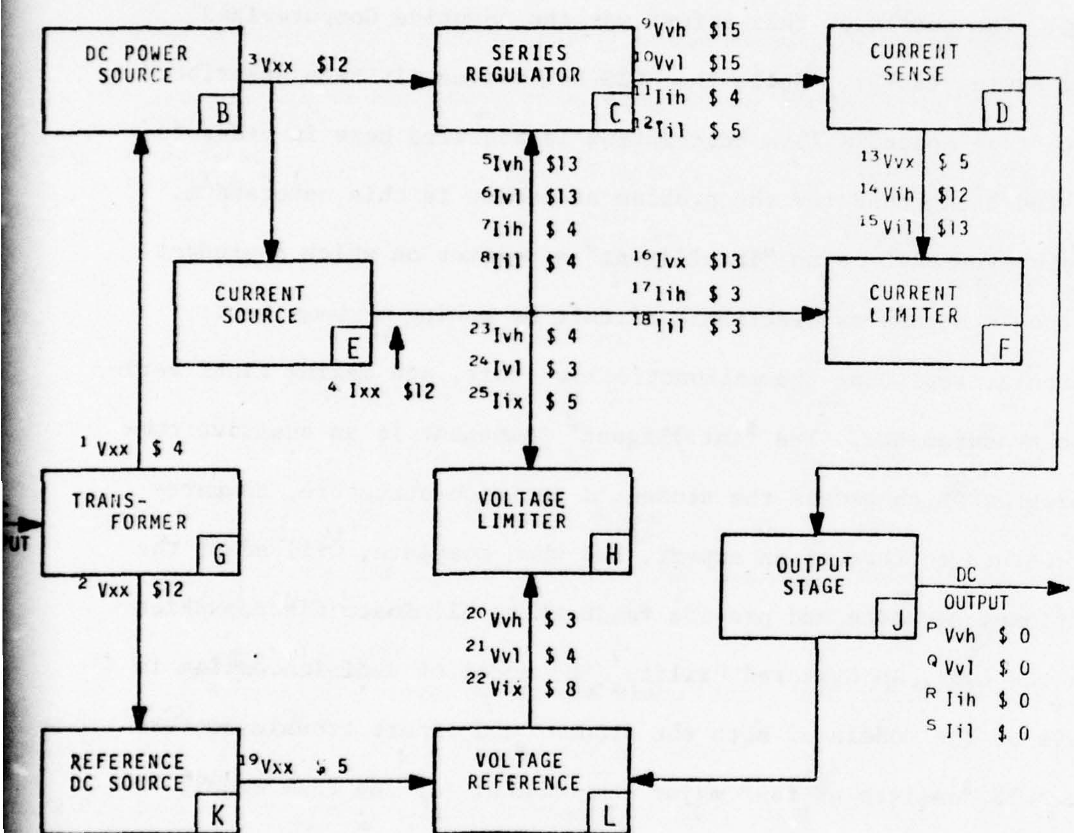

Figure 1. ACTS Circuit Diagram

(a model of an expert troubleshooter) is an EU decision model which predicts the expert's measurement choices as he or she troubleshoots the circuit. It is developed through on-line observation of the expert's troubleshooting behavior. The student model, also an EU model, predicts the student's measurement choices. It is developed through on-line observation of the student's behavior as the student solves troubleshooting problems on the ACTS. The instructional model compares the expert and student models, determines discrepancies between the two and, when complete, will modify the instructional feedback and problem presentation sequence in order to reduce those discrepancies. Currently the instructional model can provide some adaptive feedback, but cannot modify the instructional sequence.

The Expected Utility Model

Troubleshooting a simulated electronic circuit on the ACTS involves selecting a series of actions from a limited set of possible actions. In the case of the Heathkit IP-28 Power Supply, an individual can take any of the 29 possible measurements, or replace any of ten components (modules) for a total of 39 possible outcomes. Each possible IP-28 measurement, for example, has two, three, or four possible outcomes (such as high, low, normal, etc.). Each possible module can be either good or bad. Taking measurements and modules combined, 96 possible outcomes can be obtained. Each possible outcome has three properties. The first is the conditional probability of the occurrence of that outcome given the measurement outcomes previously obtained and given that the appropriate action is selected. The second property is the utility of the outcome to the troubleshooter, i.e., what he gains or loses as a result of that outcome. Utility is subjective, but it should be related to the cost (in time or money) of taking that action. The third property is the amount of information that the outcome provides about the location of the fault. These properties are combined as follows:

$$EU_j = \sum_i^n \alpha_{ij} P_{ij} U_{ij} \qquad (1)$$

where

EU_j = the expected utility of action A_j

P_{ij} = the probability that outcome i, of a set of n outcomes will occur if action A_j is selected

U_{ij} = the utility of outcome i of action A_j

α_{ij} = the information gain resulting from the occurrence of outcome i of action A_j

The P_{ij} and α_{ij} are constant across individuals and can be determined objectively. The EU_j and U_{ij} are determined separately for each individual (expert or student) by "tracking" their behavior as they solve a series of troubleshooting problems. As they do this, they are presented with updated probabilities of measurement outcomes (the P_{ij}) before they select a measurement. The values of P_{ij} and α_{ij} are entered into the model before the individual starts, with the U_{ij} set at some common arbitrary value (usually 100). The model chooses the action which has the highest expected utility. If the individual then chooses the sam action, no changes in the U_{ij} are made. However, if the action selected by the model by the model differs from the action selected by the individual, the model utilities associated with the model choice are punished (decreased) and those associated with the individual choice are rewarded (increased). These processes are represented by the following formulae:

$$U_{ij}^{t+1} = U_{ij}^{t} - P_{ij}\, \alpha_{ij}\, \gamma \text{ (Punish)} \qquad (2)$$

$$U_{ij}^{t+1} = U_{ij}^{t} + P_{ij}\, \alpha_{ij}\, \gamma \text{ (Reward)} \qquad (3)$$

U_{ij}^{t} = the (unadjusted) utility of result i of action j at time t.

U_{ij}^{t+1} = the (adjusted) utility of result i of action j at time t+1.

γ = a constant.

P_{ij} = the probability of occurrence of result i of action j.

α_{ij} = the information gain resulting from result i of action j.

The process contines until the estimated utilities become stable. This will occur when the model is able to predict the choices of the individual accurately.

Adapting Training

Completion of the ACTS requires the development of mechanisms which, ideally, compare the student and expert utilities and, based on this comparison, provide feedback to the student and modify the training sequence to reduce the discrepancies between the student and expert models. The sheer volume of information provided by the utilities makes this a complex task. For the IP-28 power supply, there are 96 utilities. A sample set is shown in Table 1.

Table 1

A Sample Set of Student Utilities

Measurement of Symptom Code	Utility for Normal Outcome	Utility for Non-normal Outcome 1	Utility for Non-normal Outcome 2	Utility for Non-normal Outcome 3	Module Code	Utility for Replacement
P						
Q				100	B	100
R	100	100	100	100	C	100
S	124	113	100		D	100
1	101	100	92		E	100
2	33	60	81		F	100
3	100	100	100		G	100
4	100	100	100		H	100
5	100	100	100		I	100
6	47	34	90		J	100
7	100	100	100		K	100
8	100	100	100		L	100
9	100	86	86			
10	100	86	86			
11	100	100	100			
12	100	100	100			
13	100	100	100			
14	100	100	100			
15	100	100	100			
16	100	100				
17	100	113	113			
18	100	100	100			
19	100	100	100			
20	108	100	100			
21	100	100	100			
22	100	100				
23	100	100	100			
24	86	100	113			
25	100	100				

Existing Mechanisms. Thus far two methods of providing feedback to the student (one of which uses student utilities) have been developed. Neither method alters the problem presentation sequence. The first, and simplest, method presents the expert model's action choice to the student after the student has made a selection and obtained the result.

The second method of providing feedback is based on comparisons among "key" student utilities. The key utilities are those for measurements identified, by an expert, as being of critical importance in the fault isolation process. For the IP-28 power supply, the utilities for the outcomes of six measurements were considered to be key. Based on the relationships among these utilities, a set of six decision rules and feedback statements were developed. Samples are shown in Table 2. The decision rule for the first feedback statement should be read as follows: If any of the utilities for measurement 3 are less than any of the utilities for measurement 19, or if any of the utilities for measurement 3 are less than any of the utilities for measurement 11, present this feedback statement to the student.

The appropriate feedback statements are initially presented to the student after completion of the fifteenth problem, with updated statements presented every fifteen problems thereafter. The student can review them at any time.

Table 2

Sample Decision Rules and Feedback Statements

Rule	Feedback
3<19 or 3<11	Although measurement 3 is located at a good point to isolate the power input modules, it is expensive. Use this measurement after you have eliminated most other possibilities. Measurement 3 should be used when the probability of normal outcome is rather high but not certain (a range of 60% to 80%).
2<11 or 3<11 or 4<11	A good first step in checking the operation of current and voltage feedback loops is to check the output of the series regulator. This should be done with the circuit operating at full output since this fully exercises the circuit functions. Therefore, measurement 9 or 11 should be used even if there is a low probability of a normal outcome. Use measurement 11 since it is much cheaper than 9.

Alternate methods. These approaches, while promising, do not use all of the information available about student performance. The first makes no use of student utilities, and the second uses only a subset of them. Possible alternate methods, which use more of the information provided by the student utilities, have been considered. Both rely on the difference could be based on either outcomes (utilities) or actions, as shown in equations 4 and 5.

$$D_{ij} = U_{sij} - U_{eij} \tag{4}$$

$$D_{.j} = \sum_i (U_{sij} - U_{eij}) \tag{5}$$

where

D_{ij} = difference between student and expert utilities for outcome i of action j

$D_{.j}$ = difference between student and expert utilities for action j

U_{sij} = student utility for outcome i of action j

U_{eij} = expert utility for outcome i of action j

Whatever method is used to calculate this difference, the following procedures would then be used to individualize instruction. The difference having the largest absolute value would be selected. This identifies the measurement or outcome for which the discrepancy in utilities between student and expert is greatest. If the difference is positive, the student has a higher utility for that measurement or outcome than does the expert. A positive sign should result in two actions: (a) presentation of feedback designed to produce decreased use of that measurement or outcome; and (b) presentation of a new problem which does not require the use of that measurement or outcome for proper solution. This process should continue until all differences are less than some specified value.

Implicit in the concept of ACTS operation and, in fact, explicit in the proposed methods of providing feedback, is the assumption that similar troubleshooting strategies result in similar sets of utilities, while different troubleshooting strategies result in different sets of utilities. For example, if a student troubleshoots a circuit by taking a sequence of measurements similar to that taken by the expert, the student utilities derived by the ACTS should be similar to those of the expert. Conversely, if the sequence of measurements taken by the student is different from that taken by the expert, the resultant utilities should also be different.

The purpose of this experiment was to investigate the relationship between similarity of troubleshooting strategies and similarity of sets of utilities. Results indicating that the two were positively related would provide a justification for using discrepancies between student and expert utilities as a basis for adapting ACTS training. Failure to find such a relationship, on the other hand, would indicate that there was no basis for using such techniques.

METHOD

Approach

One possible approach to the problem would have been to treat strategy similarity as an independent variable, and utility similarity as a dependent variable. Using this approach, a series of strategies having specified similarities would be developed, a set of troubleshooting problems would be solved using each of those strategies, and the similarities of the resulting sets of ACTS-derived utilities would be determined. This approach presented two problems, however. The first problem was that the set of utilities derived by the ACTS, for any strategy, is dependent on the amount of "training" (see equations 4 and 5) that the model receives. Since the amount of training required to obtain stable utilities can vary as a function of the particular strategy used, it would be necessary to determine a separate "stopping point" for each strategy.

The second problem was the difficulty of generating strategies of specified similarities. At the start of the study, three alternate operational definitions of strategy similarity were available. There was no rationale for using any particular one of these as a basis for strategy generation. On the other hand, the two alternate definitions of strategy similarity were functionally related. Thus there was a sound basis for generating sets of utilities of known similarity.

For these reasons, the approach taken was to generate sets of utilities having specified similarities. Each set of utilities was then inserted into the ACTS, and the "simulated student" routine activated. The simulated student troubleshoots the circuit according to the utilities it has been given, always selecting the measurement with the highest EU. Since the simulated student is perfectly consistent, i.e., always diagnoses the same fault through the same sequence of measurements, the solutions to a series of problems could easily be converted into a logic tree format. This format defined the strategy used by the simulated student.

Definitions of Utility Similarity

Each set of utilities consists of 96 individual utilities, each utility corresponding to one measurement or module outcome. A standard

standard definition of the difference between two sets of utilities, U_1 and U_2, is Root Mean Square Deviations (RSMD), or:

$$\sqrt{\frac{\sum_{i=1}^{N} (U_{1i} - U_{2i})^2}{N}}$$

where N = 96.

This is equivalent to

$$\sqrt{\sigma_{u1}^2 + \sigma_{u2}^2 - 2\, r_{u1u2}\, \sigma_{u1}\sigma_{u2} + (\bar{U}_1 - \bar{U}_2)^2}$$

where

σ_{u1} = the standard deviation of U_1

σ_{u2} = the standard deviation of U_2

$\bar{U}_1$ = the mean of U_1

r_{u1u2} = the correlation between U_1 and U_2

However, it is the value of a utility (U_{ij}) relative to the other U_{ij} in the same set, and not its absolute value, the determines which action will be selected. Thus the difference between the means of the two sets of utilities can be ignored, and equation 7 simplifies to [2]:

$$\sqrt{\sigma_{u1}^2 + \sigma_{u2}^2 - r_{u1u2}\, \sigma_{u1}\sigma_{u2}} \tag{8}$$

If in addition, it is assumed that all $\sigma_{u}^2.$ are equal to some constant, $\underline{k}$, further simplification is possible:

$$k\sqrt{2(1 - r_{u1u2})} \tag{9}$$

Thus, RMSD reduces to a simple function of the correlation (r_{u1u2}) between the two sets of utilities, and this correlation is the definition of utility similarity used to generate the utility sets. However, since the assumption that the $\sigma_{u}^2.$ were all equal was in retrospect not precisely

[2]This is equivalent to defining Root Mean Square Deviation in terms of deviations of the U_{ij} from $\bar{U}_i$, rather than using the raw score.

correct, the data were also analyzed using equation 8 as the definition of utility similarity.

Definitions of Strategy Similarity

Each strategy can be described as a logic tree, like that shown in Figure 2. The logic tree shows the sequence of measurements taken during the troubleshooting process. The information contained in the logic tree for each strategy formed the basis for the measurement of strategy similarity. Three different measures were used.

The first, and simplest, measure of utility used was termed Similarity of Measurements Used (SMU). SMU is based on the preliminary finding that consistent strategies generally require the use of less than one-half of the available measurements to solve a complete set of problems on the ACTS. For example, the strategy shown in Figure 2 used only 14 of the 29 available measurements (designated P, Q, R, S, and 1 through 25). For each strategy, each measurement is assigned a value of 1 if it is used in that strategy, and 0 if it is not. SMU is calculated as a $\emptyset$ coefficient.

Clearly, SMU ignores all information about when, or how frequently, a measurement is used within a strategy. Measurement 7, which is taken in every path through the tree (see Figure 2), seems intuitively more important than measurement 4, which is taken in only two such paths. In order to take this into consideration, a second measure of similarity, Similarity of Order of Use (SOU), was devised. In calculating SOU, measurements are first ranked according to the order in which they were

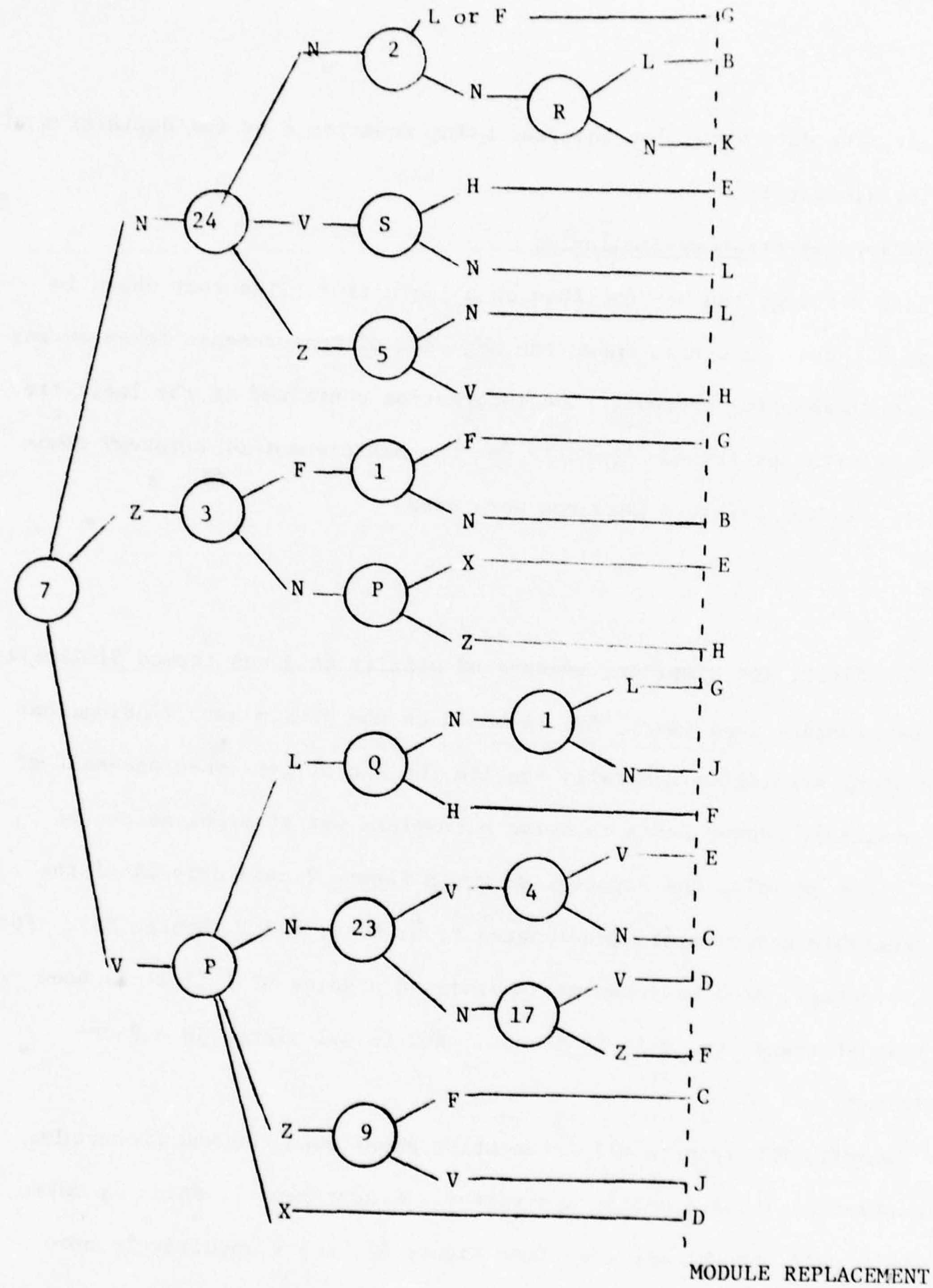

Figure 2. Logic tree for troubleshooting strategy.

○ indicates measurement taken.

used in the strategy. For example, in Figure 2 measurement 7 was assigned a rank of one, measurements 24 and 3 a rank of 25, etc. Measurements used in different orders in different paths were assigned a rank on basis of their mean order (P, for example, was assigned a rank of 4). All measurements not used in a particular strategy were assigned to the lowest rank. SOU is calculated as a rank-order correlation.

A third measure, Similarity of Frequency of Use (SFU) was calculated by assigning a rank to each measurement, based on the frequency with which it was used in solving a series of problems. The measurement used most frequently was assigned a rank of one, and all measurements not used were assigned the lowest rank. SFU is also calculated as a rank-order correlation.

Experimental Design

The experimental design is shown in Table 3. Five sets of utilities (called standard sets and designated U_{10} through U_{50}) were generated. For each standard set, five additional sets ($U_{.1}$ through $U_{.5}$) were generated, having correlations (utility similarity r_{u1u2}) of +.95, +.50, .00, -.50, and -.95, respectively, with the standard set.

Procedures for Generating Utility Sets

The formulae used to generate utility sets were:

$$U_{ijk} = \sigma_{uij} \left(1 - r^2_{uiouij} A_{ijk} + r_{uiouij} A_{iok} \right) + \bar{U}_{ij} \qquad (11)$$

for sets U_{i1}, U_{i2}, U_{i3}, U_{i4}, and U_{i5}

and

$$U_{iok} = \sigma_{uio} A_{iok} + \bar{U}_{io}$$

for set U_{i0}

Table 3

DESIGNATIONS OF UTILITY SETS

STANDARD SETS	UTILITY SIMILARITY (r_{u1u2})				
	+0.95	+0.50	0.00	-0.50	-0.95
U_{10}	U_{11}	U_{12}	U_{13}	U_{14}	U_{15}
U_{20}	U_{21}	U_{22}	U_{23}	U_{24}	U_{25}
U_{30}	U_{31}	U_{32}	U_{33}	U_{34}	U_{35}
U_{40}	U_{41}	U_{42}	U_{43}	U_{44}	U_{45}
U_{50}	U_{51}	U_{52}	U_{53}	U_{54}	U_{55}

Note - Utility similarity is defined as the correlation with the corresponding standard set.

where

i designates standard set (1-5)

j designates level of similarity (1-5)

k designates the individual utility (1-96)

A_{ijk} and A_{iok} are sets of 96 normally distributed random numbers having a mean of zero and a standard deviation of one

$\sigma_{uij} = 15$

$\bar{U}_{ij} = U_{io} = 100$

Equation 10 assumes that A_{ijk} and A_{iok} are independent, i.e., have zero correlation. If this assumption is met, the U_{ijk} will have a mean of $\bar{U}_{ij}$, a standard deviation of σ_{uij}, and a correlation with the standard set (U_{io}) of r_{uiouij}. However, variations in the correlation between A_{ijk} and A_{iok} could be expected by chance. In order to minimize the impact of this chance fluctuation, a group of utility sets (U_{io}, U_{i1}, U_{i2}, U_{i3}, U_{i4}, and U_{i5}) was rejected if more than one of the 17 parameters (5 correlations, 6 means, and 6 standard deviations) differed from the desired value at the .05 level.

Problem Set

The complete set of 27 troubleshooting problems, each of which contained a unique fault, was used. A list of faults is shown in Table 4.

Computer Hardware/Software

The standard ACTS software (Kuppin, 1976) was used. This software operates on an Interdata Model 70 minicomputer under the Real-Time Operating System.

Table 4

Circuit Faults

FAULT NO.	FAULTED MODULE	FAULT	EFFECT
0	None	No Fault	Normal Outcomes
1	G. Transformer	Break in Primary	No Output at Main or Reference Secondary
2		Short in Secondary	Reduced Output at Main and Reference Secondary
3		Break/Short in Main Secondary	Reduced Output at Main Secondary
4		Short in Reference Secondary	Reduced Output at Reference Secondary
5		Break in Reference Secondary	No Output at Reference Secondary
6	B. DC Power Source	Capacitor Short	No Source Output
7		Open Diode or Leaky Capacitor	Reduced Source Output
8	E. Current Source	Open Collector	No Output Current
9		Base-Collector Short	Excess Output Current
10		Base-Collector-Emitter Short	No Current
11		Bad Value, Bias Resistor	Reduced Output Current
12	C. Series Regulator	Base-Collector-Emitter Open	No Output
13		Low Beta Value	Low Output Current
14		Base-Collector-Emitter Short	Excess Output Current
15	D. Current Sense	Open Series Resistor	No Output Current
16		Potentiometer Short	Reduced Control Voltage
17		Improper Resistor/Potentiometer Value	Reduced Current in Current High State
18	F. Current Limiter	Transistor Short	Excess Control Current
19		Transistor Open	Reduced Control Current
20	K. Reference DC Source	Bad Value, Match Resistor	Low Reference Voltage
21		Zener or Capacitor Short	No Reference Voltage
22	L. Voltage Reference	Open Potentiometer Wiper	No Current to Voltage Limiter
23		Short Potentiometer Wiper	Excess Current to Voltage Limiter
24	H. Voltage Limiter	Transistor Short	Excess Control Current
25		Transistor Open	Reduced Control Current
26	J. Output Stage	Open Resistor or Capacitor Short	No Output Current
27		Leaky Capacitor	Reduced Output Current

RESULTS

Utility Sets

Utility similarities (r_{u1u2}), means, and standard deviations for each utility set are presented in Table 5. None of the utility similarities and only one mean (U_{32}, $\bar{U} = 95.97$) and one standard deviation (U_{12}, $\sigma_u = 17.75$) differed significantly from the desired values ($p<.05$). Considering all three parameters, two of the 85 values (2.4%) are significantly different from the desired values at the .05 level. Since this is fewer than would be expected by chance, it can safely be concluded that, in terms of individual parameters, the utility sets have the desired characteristics. However, there was the possibility that the combination of small fluctuations in correlation and standard deviation could cause substantial fluctuations in RMSD, as defined in Equation 8. In order to check this, RMSD for each utility set was calculated, with the results shown in Table 6. The significance of the difference of each value from the desired value was tested as an F ratio, $RMSD_{(largest)}^2/RMSD_{(smallest)}^2$, with 95 degrees of freedom in both the numerator and denominator. None of these tests revealed a significant difference at $p <.05$.

Correlational Analyses of Utility Similarity and Strategy Similarity

Correlations among the two measures of utility similarity (RMSD and r_{u1u2}) and the three measures of strategy similarity (SMU, SOU, & SFU) are shown in Table 7. The two measures of utility similarity are highly negatively correlated (-.97), as would be expected from their functional

Table 5

Characteristics of the Utility Sets

	U.0	U.1	U.2	U.3	U.4	U.5
Utility Similarity (r_{u1u2})						
Desired Value	______	+.95	+.50	.00	-.50	-.95
Standard Set						
U_{10}	______	+.97	+.53	+.04	-.58	-.96
U_{20}	______	+.95	+.47	-.02	-.41	-.96
U_{30}	______	+.95	+.47	-.15	-.46	-.94
U_{40}	______	+.94	+.57	-.04	-.49	-.96
U_{50}	______	+.94	+.36	-.09	-.52	-.94
Mean Value[a]	______	+.947	+.485	-.038	-.495	-.955
Mean						
Desired Value	100.00	100.00	100.00	100.00	100.00	100.00
Standard Set						
U_{10}	102.65	102.57	101.09	101.78	99.63	96.93
U_{20}	101.88	101.14	101.79	99.81	99.51	98.41
U_{30}	100.51	100.93	95.97[b]	101.94	100.92	98.91
U_{40}	99.60	98.91	99.14	100.71	97.31	101.17
U_{50}	100.38	99.34	101.97	97.90	97.57	99.51
Mean Value[c]	101.01	100.58	99.99	100.43	98.98	98.99
Standard Deviation						
Desired Value	15.00	15.00	15.00	15.00	15.00	15.00
Standard Set						
U_{10}	16.24	16.56	17.75[b]	16.61	14.47	16.73
U_{20}	14.37	14.14	15.61	14.20	13.83	14.72
U_{30}	14.54	14.38	14.24	15.36	13.84	15.11
U_{40}	14.55	14.62	14.83	14.63	15.56	15.26
U_{50}	14.56	13.98	13.57	15.97	14.36	14.47
Mean Value[d]	14.85	14.73	15.20	15.35	14.41	15.26

[a] Calculated using Fisher's z.

[b] Significantly different from desired value ($p < .05$).

[c] Grand mean = 99.996

[d] Grand mean = 14.968

Table 6

Actual Utility Similarity
(Root Mean Square Deviations of Each Utility Set From Its Standard Set)

	Utility Set				
	$U_{\cdot 1}$	$U_{\cdot 2}$	$U_{\cdot 3}$	$U_{\cdot 4}$	$U_{\cdot 5}$
Desired Value	4.74	15.00	21.21	25.98	29.62
Standard Set					
U_{10}	4.24	16.48	22.78	27.28	32.67
U_{20}	4.93	15.51	20.38	23.72	28.79
U_{30}	4.62	14.77	22.75	24.28	29.23
U_{40}	5.20	13.69	20.26	26.01	29.49
U_{50}	4.96	15.92	22.58	25.23	28.58
Mean Value	4.79	15.27	21.75	25.30	29.75

relationship. The three measures of strategy similarity are highly positively correlated, with values ranging from +.97 (SMU and SFU) to +.99 (SOU and SFU).

As would be expected, correlations between utility similarity and strategy similarity are relatively invariant as a function of which measure of each is used. These correlations range from +.78 to +.81 when r_{u1u2} is used as the measure of utility similarity, and from -.81 to -.85 when RMSD is used as the measure of utility similarity. Thus r_{u1u2} accounts for between 61% and 66% of the variance in strategy similarity, and RMSD, between 66% and 72%. While the correlations suggest a tendency of RMSD to be more highly correlated with strategy similarity than r_{u1u2} this apparent effect is not statistically significant.

SMU, SFU and SOU as a function of RMSD and r_{u1u2} are shown in Figure 3.

Additional Analyses

Examination of Figure 3 raised two issues. First, it was apparent that there was considerable variation among utility sets. Second, there appeared to be some tendency toward a curvilinear relationship between utility similarity and strategy similarity, particularly when r_{u1u2} was used as the measure of utility similarity. In order to determine whether a significant curvilinear component was present, and to determine the extent to which the utility sets varied in terms of strategy similarity,

Table 7

Intercorrelations Among Measures of Utility and Strategy Similarity

	Utility Similarity		Strategy Similarity		
	RMSD	r_{u1u2}	SMU	SOU	SFU
RMSD		-.97	-.85	-.83	-.81
r_{u1u2}			.81	.80	.78
SMU				.98	.97
SOU					.99
SFU					

Note - N = 25. All correlations are significant at $p < .001$

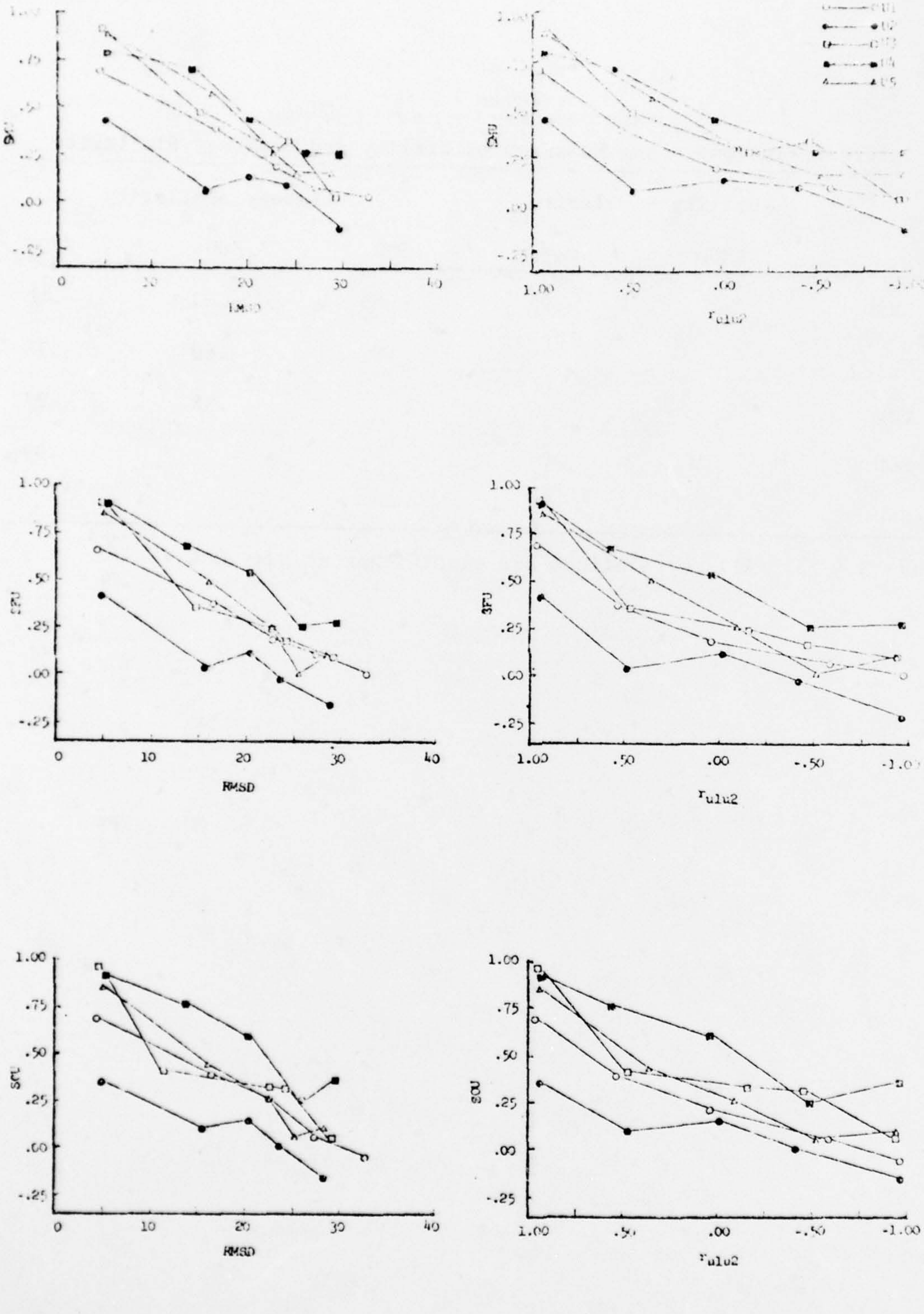

Figure 3. SMU, SFU, and SOU as a function of RMSD and r_{ulu2}

data were re-analyzed using analysis of variance. The particular analysis of variance model used was a one factor (utility similarity) repeated measurements design, with additivity assumed (Myers, 1966, pp. 153-156). Utility sets (U_1. through U_5.) were treated as subjects. Five levels of utility similarity, corresponding to the desired r_{u1u2}'s of +.95, +.50, .00, -.50, and -.95, were used. Significant utility similarity effects were further analyzed to determine the significance of linear and higher-order trends (Myers, 1966, pp. 348-357). Weights for these comparisons were derived on the basis of both mean r_{u1u2} and mean RMSD for each level. The estimated proportion of variance accounted for, ω^2, was calculated for each significant effect (Hayes, 1963, pp. 406-408).

The analysis of variance summary tables for SMU, SOU, and SFU are presented in Tables 8, 9, and 10, respectively. The effect of utility sets is significant for all three measures of strategy similarity, accounting for between 20 and 27% of the total experimental variance. This confirms that there is substantial variation among utility sets, as suggested in Figures 3 through 8.

The overall effect of utility similarity is also signficant for each measure of strategy similarity. Utility similarity accounts for between 63 and 68% of the total experimental variance, approximately what would be expected on the basis of the correlations between utility similarity and response similarity presented in Table 7.

When trend analyses of the utility similarity effects are performed using weights derived from mean RMSD, the linear component of the trend

Table 8

SMU Analysis of Variance Summary Table

	df	SS	MS	F	ω^2
Utility Sets (S)	4	0.4358	0.1089	11.75*	0.20
Utility Similarity (U)	4	1.3940	0.3485	37.59*	0.68
RMSD					
Linear	1	1.3919	1.3919	150.12*	0.68
Dev. from Linear	3	0.0021	0.0007	0.08	
r_{u1u2}					
Linear	1	1.3043	1.3043	140.67*	0.64
Dev. from Linear	3	0.0897	0.0299	3.22	
SU (error)	16	0.1484	0.0093		

*p < .001

Table 9

SOU Analysis of Variance Summary Table

Source	df	SS	MS	F	ω^2
Utility Sets (S)	4	0.5742	0.1435	19.20*	0.24
Utility Similarity (U)	4	1.5736	0.3934	52.61*	0.68
RMSD					
Linear	1	1.5446	1.5446	206.58*	0.67
Dev. from Linear	3	0.0290	0.0097	1.29	
r_{u1u2}					
Linear	1	1.4361	1.4361	192.06**	0.62
Dev. from Linear	3	0.1375	0.0458	6.13*	0.05
SU (error)	16	0.1196	0.0075		

* $p < .01$

** $p < .001$

Table 10

SFU Analysis of Variance Summary Table

Source	df	SS	MS	F	ω^2
Utility Sets (S)	4	0.6546	0.1637	16.82*	0.27
Utility Similarity (U)	4	1.4756	0.3689	37.92*	0.63
RMSD					
Linear	1	1.4546	1.4546	149.50*	0.62
Dev. from Linear	3	0.0210	0.0070	0.7191	
r_{u1u2}					
Linear	1	1.3769	1.3769	141.52*	0.58
Dev. from Linear	3	0.0777	0.0259	2.66	
SU (error)	16	0.1557	0.0097		

*p < .001

is always significant, while deviations from the linear component (quadratic, cubic, and quartic) are never significant. Thus, there is no evidence that the relationship between RMSD and strategy similarity is non-linear.

Trend analysis using weights derived from mean r_{u1u2} produces somewhat different results. Again, the linear component is always significant. However, deviations from the linear component are significant for SOU, and approach significance($.05 > p > .10$)for SFU and SMU. This effect accounts for 5% of the total variance in SOU. Thus, it cannot be concluded that the relationship between r_{u1u2} and strategy similarity contains only a linear component.

DISCUSSION

For the values of utility similarity considered, a strong relationship was found between utility similarity and strategy similarity, as both were defined in this study. Even in the worst case, more than 61% of the variance in strategy similarity (SFU) was accounted for by utility similarity (r_{u1u2}). Furthermore, the relationship between utility similarity and strategy similarity was predominately linear, although a slight curvilinear component was evident when utility similarity was measured by r_{u1u2}.

Under the conditions of this study, then, the greater the similarity between a set of utilities and a criterion (or standard) set, the more similar their strategies will be. If one set of utilities is that of a student, and the criterion set is that of an expert, the similarity of their strategies will increase as the student utility set becomes more similar to that of the expert. However, the conditions of this study were different from those that would be found in a "real world" training situation, and the overall effect of those differences would be to produce stronger relationships between utility and strategy similarity in the experimental than in the training setting.

The rationale for investigating the relationships between strategy similarity and utility similarity in an artificial situation was twofold: (a) only limited resources were available; and (b) if a relatively strong relationship between utility and strategy similarity could not be discovered in the artificial situation, it certainly would not be discovered in the training situation. Thus the experimental results are a necessary, but not sufficient, prerequisite to obtaining similar results in the training situation.

It is appropriate to consider why a stronger relationship between utility and strategy similarity would be expected in the aritifical situation than in the training situation. First, human troubleshooting behavior is not perfectly reliable. Although each individual may have a "true" strategy which can be represented by a logic tree, deviations from the true strategy can be expected. This inconsistency in human performance precludes obtaining a perfect relationship between utilities and strategies: hence, utilities will not predict human performance perfectly (May, Crooks, and Freedy, 1978).

Second, as previously noted, this study was run "in reverse." Utilities generated strategies, which is the opposite of what would occur in the training situation. In actual applications the student (or expert) model will necessarily produce identical utilities for identical (and "error-free") strategies only if the same value of γ is used for both, and behavior is observed over the same sequence of problems. Otherwise, differences between the utilities <u>may</u> occur.

Finally, little is known about the extent of variation in the skewness, kurtosis, and variance of sets of utilities, and their impact on measures of utility similarity. Based on previous data a variance of 225 seemed reasonable, but substantial deviations may occur from individual to individual.

A note of caution concerning the variation among utility sets is also in order. This variability could be of practical importance in situations in which multiple sets of expert utilities have been developed. This would occur when there are several equally effective ways to troubleshoot the same circuit. If a student's utilities are more similar to those of Expert A than to those of Expert B, it does no necessarily follow that the student's strategy will be more similar to that of Expert A than to that of Expert B.

Summary

A strong positive relationship between utility similarity and strategy similarity was obtained. This is a prerequisite to using discrepancies between student and expert utilities as a basis for adapting training.

Because the experimental conditions were somewhat artificial, the strength of the relationship between utility similarity and strategy similarity could be expected to be attenuated in a "real world" training situation.

References

Hayes, W. L. Statistics for psychologists. New York: Holt, Rinehart, & Winston, 1963.

Kuppin, M. A. Adaptive Computerized Training System (ACTS): User's Manual (Perceptronics System Document 1028-76-1). Woodland Hills, CA: Perceptronics, Inc., December 1976.

May, D. M., Crooks, W. H., and Freedy, A. Application of adaptive decision aiding systems to computer-assisted instruction: Experimental studies ARI Technical Report TR-78-A4). Alexandria, VA: US Army Research Institute, February 1978. (NTIS No. AD-A050887)

Mayers, J. L. Fundamentals of experimental design. Boston: Allyn & Bacon, 1966.

APPENDIX: BIBLIOGRAPHY OF ACTS PUBLICATIONS

Crooks, W. H., Kuppin, M. A., & Freedy, A. Application of adaptive decision aiding systems to computer-assisted instruction: Adaptive Computerized Training System (ACTS) (Perceptronics Technical Report PATR-1028-77-1). Woodland Hills, CA: Perceptronics, Inc., January 1977.

Freedy, A., and Crooks, W. H. Use of an adaptive decision model in computer-assisted training of electronic troubleshooting. In Proceedings of the conference on new concepts in maintenance training research. Orlando, Florida: April 1975.

Knerr, B. W. Artificial intelligence techniques in electronic troubleshooting training: Development and evaluation. In Training: Technology to policy, proceedings of the 1976 symposium of the Potomac Chapter, Human Factors Society (ARI Special Publication P-77-6). Alexandria, VA: US Army Research Institute, December 1977. (NTIS No. AD-A049629)

Knerr, B. W. Problems of performance measurement in computer-based simulation. Paper presented at the Computer Adaptive Testing '77 Conference, Minneapolis, Minnesota, July 1977. (Proceedings in press)

Knerr, B. W., and Nawrocki, L. H. Development and evaluation of an Adaptive Computerized Training System (ARI Product Utilization Report). Alexandria, VA: US Army Research Institute, in press.

Kuppin, M. A. Adaptive Computerized Training System (ACTS): User's Manual (Perceptronics System Document 1028-76-1). Woodland Hills, CA: Perceptronics, Inc., December 1976.

May, D. M., Crooks, W. H., and Freedy, A. Application of adaptive decision aiding systems to computer-assisted instruction: Experimental studies (ARI Technical Report TR-78-A4). Alexandria, VA: US Army Research Institute, February 1978. (NTIS No. AD-A050887)

May, D. M., Crooks, W. H., Purcell, D. D., Lucaccini, L. F., Freedy, A., & Weltman, G. Application of adaptive decision aiding systems to computer-assisted instruction (ARI Technical Report TR-77-A26). Alexandria, VA: US Army Research Institute, October 1977.