Proceedings of the 1985 Winter Simulation Conference
D. Gantz, G. Blais, S. Solomon (eds.)

COMPUTERIZED EVALUATION SIMULATOR BASED ON THE CIPP MODEL

Dr. Takeshi Ohara
Dr. Kenneth Pickard
School of Business Administration
California State University
Long Beach, CA 90840

ABSTRACT

An attempt was made to build a computer simulation model based on an evaluation model. The model (CIPP) systematically examined each stage of evaluation activities (context, input, process, and product of a project). The administrators, funds, and program effectiveness were considered as major factors in a simulation model. An optimal solution approach was chosen to evaluate the simulation results. Jackson's time independent solution was used to calculate the solution.

The purpose of this project was to construct an evaluation simulation model based on the CIPP model, so that the evaluator would be able to envisage the best-fit evaluation design using a given educational setting. The major factors were the characteristics of the students and the school, the preferential attitude of the administrators to a specific project and program, and the availability and capability of programs and funding agencies.

This paper presents the methodology used in building the CIPP simulation. The study is composed of four sections: 1. A theoretical framework in which a systems analytical view of the CIPP model is presented. This includes descriptions and interactions of the functions within and between the components of CIPP model. 2. A system synthesis of the CIPP model in which the concept of probabilistic operations is implemented. 3. A theoretical presentation of the optimal solution of a simulation model which gives an estimation of what would be a good strategy an evaluator may adopt. 4. A summary of the results from the simulation experiment.

THEORETICAL FRAMEWORK

An evaluation system is the framework in which evaluation activities are planned, implemented and evaluated. The evaluation system exists as long as there are behavioral objectives to be met. Once the system achieves the objectives and the degree of achievement is assessed, the system becomes inactive. The evaluation system of this study is composed of four components--Context, Input, Process, and Product evaluation, which are derivatives of the CIPP model. Each component is subject to a time dimension. In other words, the component exists until it achieves the given objectives. Context evaluation assesses the conditions of the educational environment by sampling unmet needs and prioritizing them. Because of this, context evaluation receives input from the educational environment. Then, based on the prioritized needs, the behavioral objectives are formulated and determined.

In input evaluation, strategies to achieve the adopted objectives are considered and designed. Every available alternative is sought to maximize the successful attainment of final student achievements.

Once evaluation strategies and procedures have been determined and funds are available, process evaluation is the next stage. Process evaluation constantly monitors the daily activities of a project and keeps record of them in terms of efficacy of the implemented program and detected invalidities in strategies. If the degree of the strategical invalidities overshadows the degree of efficacy of a program, the current strategies are reviewed and the adjusted strategies are implemented and put into the evaluation procedure. This adjustment mechanism includes process evaluation and product evaluation, together called the evaluation cycle (4).

Product evaluation assesses the degree of the achievements of programs compared to the objectives formulated at context evaluation. Records of daily activities are carefully reviewed and provide the necessary information to judge the nature of the project achievements. By reviewing the records, sources of invalidities in the implemented design are detected thus assisting an evaluator to interpret correctly the student's achievement.

The evaluation systems are limited by two constraints--the preferences of an administrative staff and the availability of funds. The administrative staff plays the role of decision-maker in each stage of the evaluation activities. In context evaluation, the administrative staff exercises influence over both the ranking of process needs and the formation of project objectives. The staff decides whether to do an input evaluation when they accept the project objectives or repeat a context evaluation to make sure that the objectives reflect their preferences. In input evaluation, the administrative staff frequently joins the decision making process to decide on the strategies to be taken, the programs to be chosen, and the funds available to project. If they find that the chosen strategies do not meet their subjective criteria, then the evaluation activities go back to the stage of context evaluation. The administrative staff's influence in process evaluation is as strong as in input evaluation. It is essential that the administrative staff adjusts strategies which do not provide expected products. If the staff strongly feels that the strategies should be adjusted, then the current programs are implemented along with the adjusted strategies. In product evaluation, the role of the administrator is to make decisions based on the information provided by an evaluator. Such decisions include whether or not the project has been successful and what should be done with the final results. The scope of the CIPP evaluation system under consideration provides useful information to the administrative staff however it does not provide them with a decision making process. In other words, the administrators must decide if a project is successful or not. Decision making is not the evaluator's concern.

An important constraint in evaluation systems is the availability of funds. Without funds, no evaluation can begin. In input evaluation, the amount of available funds determines the kind of implementing strategies, available programs, and personnel. Availability of funds also determines the duration of an evaluation project. In process evaluation, a record of how much money is spent each evaluation day is kept and reviewed by the administrative staff. If the funds are used up, the project must be terminated.

The evaluation system is also distinguished by interactions. Interactions are defined as the dynamic communications between project personnel across time, the interactions occur between and within components. The within interactions indicate specific evaluation activities of one component to achieve its objectives. For example, in context evaluation, the within interactions are realized in a formation of objectives based on the unmet needs.

The between interactions are indicative of a transition step from one component to another. If an evaluation moves from context evaluation to input evaluation; the information obtained, such as educational objectives, is transmitted to input evaluation. The between interactions also occur when each of the components is under the influence of the constraints. For instance, in context evaluation, the preference of administrative staff intervenes during the process of objective formation and controls a system flow.

SYSTEM SYNTHESIS

The objective of this section is to demonstrate how to build a simulator which encompasses the components described under the theoretical framework. The objective of a context evaluation is to assess the educational needs and to establish the objectives based on the ranked needs. In a simulator (3), we assume that the unmet educational needs such as the improvement of spelling and reading are functionally distributed in the educational environment, and that the needs assessed are simulated by random sampling from the unmet needs distribution function. In the educational environment, more than one thousand unmet needs may be categorized under common attributes such as basic skills, affective domains, and social skills. Assuming that we have five different categories among the unmet needs in our hypothetical school environment, we can give as parameters a hypothetical probability weight of .30 to category I, .20 to both category II and III, and .15 to category IV and V. By assigning a probability weight of .30 to category I, we imply that 300 out of 1000 unmet needs are categorized as I. These weights are the parameters to our simulator and should be manipulated by a user when the realistic educational environment has been assessed. After a random sample of 1000 unmet needs from the unmet needs distribution, the needs can be ranked in order of priorities. The three top-ranked needs are adopted and used as the system objectives. At this time, the administrative staff makes a decision about the system objectives. We assume that the preference of the administrative staff in regard to the three objectives ranked at the top is a discrete uniform distribution, in which we hypothesize that in seven cases out of ten the administrative staff agrees with the objectives, then the evaluation activities go back to context evaluation and the staff approaches the problem in a different manner. In our simulator, the different approach is symbolized by using a new random initiator to sample from the distribution functions. Table 1 shows the simulation flow of context evaluation. The unmet needs are generated 1000 times by the normal distribution called GAUSS. Then these unmet needs are grouped in five major categories and given a rank.

If the administrative staff agrees with the three objectives, then an input evaluation is done. The input evaluation provides information concerning the possible programs and their costs, and the possible funding agencies. In our simulator, two possible programs for each objective are chosen from a discrete uniform distribution. Hypothetically, the probability that program I or II is chosen is .50. If program I is sampled, then the base to calculate would be three dollars per person per day. The base for program II would be two dollars per person per day. (These figures are based on the past funding of Title I.) With 100 repetitions of the same sampling procedure, we will obtain the frequency of programs I and II. One hundred samples are taken from the budget distribution in order to simulate the adjustment procedure of the budget. Without the consent of the funding agencies, there is no evaluation project. In order to gain their consent, an adjustment of evaluation strategies must be done. We assume in the simulation model that the budget distribution is a unit normal distribution which has the probability for each adjustment. Again, the hypothetical parameters are as follows; the adjustment of 50 cents per student has the probability of .6, one dollar and 50 cents has a probability of .15, and two dollars has a probability of .10. These dollar and probability amount are based on the observations obtained from the evaluation practices. The adjustment is done in the model by subtracting the adjustment amount from the originally planned amount. Again, the administrative staff intervenes in the adjustment procedure and it has the right to determine whether the project will proceed or return to the planning board. We assume that the probability that the administrative staff will reject the allocated amount is .3. This probability is another manipulable parameter that a user must provide to the simulator. Table 2 shows that when the programs are chosen, they are checked by the administrative staff. The availability of funds is also determined in the flow. If the funds are not available, then the process of input evaluation is repeated.

The function of process evaluation is to constantly monitor and to keep a record of the daily activities, the program effectiveness, the sources of internal invalidity in the implemented strategies, and the budget spending behaviors. In the simulator, the efficacy of a program for a student is randomly sampled from the program efficacy normal distribution. The distribution has five probability areas, each of which is weighted from 1 (very low effectiveness) to 5 (very high effectiveness). Hypothetically, weight 1 has the probability of .10, weight II .20, weight III .40, weight IV .20 and weight V .10. The probability interval is designed to reflect the natural settings of the program effectiveness.

While the efficacy of a program per student per day is being recorded, the sources of internal invalidity are being constantly monitored. According to the reviews of 20 Title IV evaluation reports in Illinois (3), the history regression effect, testing, maturation and selection bias are the major sources of the internal invalidity encountered in the evaluation projects. The probabilities calculated are as follow: for history (.21), regression (.21), selection bias (.23), maturating (.12) and testing (.23). Table 3 presents the empirical correlation coefficients among the invalidities, in which most coefficients are found nonsignificant. Thus, it is assumed in this simulator that these invalidities are statistically independent. Here we assume that each of these sources has one unit of diminishing effect on the observed effect of the implemented programs. This assumption is based on the concept that an observed score is composed of a true score plus error, which is one unit of diminishing effect in an arbitrarily chosen unit in a simulator. It may be more or less than one unit, but that awaits further empirical research. The manner used to monitor these sources is also per student per day per program.

In process evaluation, if the administrative staff detects the necessity of changing the ongoing strategy, then this would be reflected in the simulator. A different random number initiator is given and the effect of the new strategies will be added to the old strategies. The mode of detecting the deterioration of the old strategies is operationalized in a manner that if the total effects of the program are overcome by the total amount of the internal invalidity, the deterioration of the strategies begins and the project is doomed to fail.

Another important function of process evaluation is the manner in which time or funds are used. Once the evaluation project starts, time and money are spend daily. Thus, the administrative staff should watch carefully how much money they have left to spend, and if they see that they are running out of funds then the winding up operation of the project should be done. In a simulator, the lack of either time or money for the project indicates that the process evaluation should be terminated. Table 4 shows the evaluation activities taking place in process evaluation. The efficacy of each implemented program is recorded daily together with the invalidity sources of the strategies. If the invalid effect overcome the valid effects of a program, a new strategy is implemented with the consent of the administrative staff.

Product evaluation provides useful information concerning the total project effectiveness along with procedureal information. Our simulator provides information about the accumulated effectiveness of each program, the accumulated invalidity of each strategy in which the program is implemented, and the true effects of each program. The real effects are the results of the subtraction of the accumulated invalidity of a program from the accumulated effectiveness of a program. Whether the project is a failure or a success depends on the criteria with which the administrative staff gleans from their experiences and the comparisons from other similar projects. Table 5 depicts the simulation flow of product evaluation.

OPTIMAL SOLUTION OF A SIMULATOR

An optimal solution is the equilibrium state probability distribution, which implies that the waiting time (queue) to be served would be constant. Actually, the constant value is equal to the average of the waiting times from each state of the models.

Before discussing the applications of an optimal solution, an illustration of a solution is given below for the purpose of clarification. Suppose there are two stores, one of them sells bananas and the other one sells cucumbers. The customers are instructed to buy bananas first and then cucumbers. The customers arrive at the first store and get served, if no other customers are waiting in line. Then they proceed to the second store and get served, again assuming no other customers are waiting in line. For our model we will assume that the banana store needs n1 seconds to sell its goods and that the cucumber store requires n2 seconds to sell theirs, and that the average waiting time for the banana store is n1 seconds and for the cucumber store n2 seconds.

The model is defined as the two dimensional vector N=(n1,n2). The total number of jobs in the system is $N=\sum_i n$. Let the model be in state k=(k1,k2) at the time t. We now compute the conditional probabilities p (N, t + h/k, t) of the model being in state N = (n1, n2) at time t + h, where h is so short that at most one event may take place in the model during the time interval (t, t + h). Let the mean arrival time of a customer be λ. Then in this short time interval (t, t+h), the probability of an arrival is $h\lambda$, and the probability that service center i completes the processing of a job is $h\lambda u_i$. The following cases will be considered (1, p.199):

a. N=k (which means no event occurs), thus
$$p\ (N,\ t + h\ /\ k,\ t) = 1 - h\lambda - h\sum_{i=1}^{2} u; \tag{1}$$

b. N=k except n1 = k1 + 1 (one arrival), thus,
$$p\ (N,\ t + h\ /\ k,\ t) = h\lambda \tag{2}$$

c. N=k except n2 = k2 - 1 (one departure), thus
$$p\ (N,\ t + n\ /\ k,\ t) = h\mu_2\ (1 - b_{21}), \tag{3}$$

where b21 is the probability that a customer will go to the cucumber store and get served then go to the banana store and get served.

d. N = k except n1 = k2 - 1 or n2 = k1 + 1 (one service completion at a banana store followed by the job's immediate transfer to a cucumber store), thus,

$$p\ (N,\ t + h\ /\ k,\ t) = h\mu 1 * b21 \tag{4}$$

note that b21=0.0. Thus, (3) will be

$$p\ (N,\ t + h\ /\ k,\ t) = h\mu 2 \tag{5}$$

From probability theory, we know that

$$p\ (N,\ t + h) = \sum_k p\ (k,\ p). \tag{6}$$

$$p\ (N,\ t + h\ /\ k,\ t).$$

Substituting (1), (2), (3), and (4) into (6), we have

$$p\ (N,\ t + h) = p\ (N,\ t) * (1 - h\lambda - h\sum_{i=1}^{2}\mu_i) +$$
$$h*p\ ((n1 - 1,\ n2),\ t) + h\mu 2*p\ ((n1,\ n2 - 1),t) \tag{7}$$

We now subtract p(N, t) from both sides of (7), divide both sides by h and then let h approach zero. Thus, we obtain a difference-differential equation. According to Jackson (2), a time-independent solution to the equation (7) exists and is unique.

$$p\ (N) = \lim_{t \to \infty} (p(N,t)) \tag{8}$$

is given by,

$$p\ (n1,\ n2) = p1(n1) * p2(n2) \tag{9}$$

where

$$p_i(n_i) = (1 - e_i)\ e_i^{n_i}\ (i = 1,\ 2) \tag{10}$$

with

$$e_i = \frac{\lambda_i}{\mu_i} \quad (11)$$

and

$$\lambda_1 = \lambda \quad (12)$$

$$\lambda_2 = b12\lambda_1$$

Thus, the solution is:

$$\lambda_1 = \lambda$$

$$2 = b12\lambda \quad (13)$$

This shows that the model of these two stores achieves the equilbrium status when the mean arrival time rate of the customers to a banana store is λ and the mean arrival time rate to a cucumber store is b12λ . In this model, since the value of b12 happens to be one, the mean arrival time rate to a cucumber store is also λ .

Now we apply the procedure described in equations (8) through (13) to our simulator and give the necessary parameters we look for an optimal solution for the system. In this project, we will take a macro-to-micro approach, which requires us to find an optimal solution among four components first. Then we find out what kind of parameter's combination yields such an optimal solution. The analysis is done for each component of the systems. The illustration is given in Table 6. Let Ni be a stage of evaluation. Thus Ni = n1, n2, n3, n4 where n1 is context evaluation, n2 is input evaluation, n3 is process evaluation, and n4 is product evaluation. Thus we have four dimensions in the evaluation systems. From (8)

$$p\ (N) = \lim_{t \to \infty} (N,\ t) \text{ at } t = t \text{ , is given by}$$

$$p\ (N1,\ N2,\ N3,\ N4) = p1(N1*p2(n2)*p3(N3)*p4(N4) \quad (14)$$

where

$$pi(Ni) = (1 - pi)pi\lambda_i \quad (i = 1,\ 2,\ 3,\ 4) \quad (15)$$

with

$$pi = \frac{\lambda_i}{\mu_i} \quad (16)$$

(where λ is the average arrival rate in context evaluation. In other words, λ is the average time to input educational unmet needs into context evaluation. In a simulator, it is assumed that μ_i = 1.0. $\lambda 1$ is the average time needed to achieve the objectives of context evaluation--assessment of the unmet needs by our definition of querie, and u1 is the average utility time which is assumed to be 1.0. In the same way λ_2 for input, λ_3 for process, λ_4 for product evaluation.)

and

$$\begin{aligned} \lambda 1 &= \lambda\, n1 + b21\lambda 2 \\ \lambda 2 &= b12\lambda 1 \\ \lambda 3 &= \lambda\, n3 + b43\lambda 4 \\ \lambda 4 &= b34\lambda 3 \end{aligned} \quad (17)$$

In solving (17), the following are to be considered: (a) There is a time lag between the group of context and input evaluation and the group of process and product evaluation. That is, if the former group achieves their objectives, they are deactivated and the latter group is activated. Thus the former group is statistically independent of the latter group. (b) Input evaluation is statistically dependent on context evaluation because there is an interactive simulation flow between them. The same is true of process and product evaluation. The solutions are:

$$\begin{aligned} \lambda 1 &= \frac{\lambda n_1}{1 - b_{12} b_{21}} \\ \lambda 2 &= \frac{b_{12}\lambda n_1}{1 - b_{12} b_{21}} \\ \lambda 3 &= \frac{\lambda n3}{1 - b_{34} b_{43}} \end{aligned} \quad (18)$$

$$\lambda 4 = \frac{b_{34}\lambda n3}{1 - b_{34} b_{43}}$$

where
- b12 -- the probability of the specific objectives going to input evaluation.
- b21 -- the probability of the rejected objectives returning to context evaluation.
- b34 -- the probability that a specific strategy will be evaluated in process evaluation because of detected invalidity.
- b43 -- the probability that a specific strategy would be reconsidered and modified

In running a simulator, we will obtain the values for the above-mentioned variables and calculate $\lambda 1$, $\lambda 2$, $\lambda 3$ and $\lambda 4$ to estimate an optimal solution of the system. By inputting a set of variable parameters to a simulator, we will obtain a set of optimal solutions. Then the user is left to choose which optimal solution may be best-fit model to the situation faced. For example, a user may choose an optimal solution which will minimize the average interarrival time of context and of input evaluation, and maximize the average interarrival time of process and product evaluation.

RESULTS:

The simulator experiment was conducted on the computer programmed model. In order to simulate various situations under the same conditions, a set of random numbers was changed in each of ten runs. A set of probability density functions was also changed in each of three models to simulate different conditions under the same theoretical framework.

This section presents the results of these experiments. The statistical analyses are divided into an instrumental aspect of the model; the model's susceptibility to the change of probability weights and random initiators; and a theoretical aspect, the relationships among the first three stages of the CIPP model and the average program efficiency observed in three experimental models. The average program efficiency is determined by dividing the accumulated achievements of students by the number of days spend in process evaluation. Here it is assumed that the longer the project continues, the students' increased learning may occur in the accumulated manner. The last part of this section describes the calculation of an optimal solution for each of the experimental models. This optimal solution demonstrates what is the best-fit model of evaluation in a given educational setting.

THE MODEL'S SUSCEPTIBILITY: The model's susceptibility implies whether the model exhibits any distinguishable behavior different from the other models, given the unique set of probability weights and density function which realize viable educational settings. If one experimental model behaves significantly differently from another experimental model which has the same structure as the former due to changes occuring in the parameters, the model tends to possess the before mentioned susceptibility.

For the sake of detecting the model's susceptibility, the behaviors of process evaluation were extensively analyzed. Because the behaviors of process evaluation were most affected by the probability weights determined in context and input evaluation, they set the course for process evaluation. The behaviors of process evaluation were represented by the day-to-day accumulation of the program's effectiveness measured by the degree of student's learning, the estimated sources of internal invalidity which jeopardize a project, and the estimated true effectiveness of a program. Thus, statistical comparisons were done on these three variables.

In statistical comparisons the Kolmogorov-Simirnow two sample two-tailed test (5, p.127), which is concerned with the agreement between two cumulative distributions, was used in mode of the 'within' a model and 'among' models comparisons. The 'among' models comparisons are concerned with model's susceptibility to change occurring in the probability weight which describe a specific educational setting, whereas the 'within' a model comparisons are concerned with model's susceptibility to changes occurring in the random initiators which simulate a specific condition in which a specific educational setting is situated. In both comparisons, the test was conducted on a pair of the same programs ;each of which was from different experiments, in regards to the program's effectiveness, the estimated invalidity sources in the implemented strategy, and the estimated true effectiveness of the program. Table 7 presents a summary of the tests. It reports the proportion of the model's susceptibility to the changes in an educational setting (the probability weights) and of a condition (the random initiators). The percentage calculated by dividing the number of significant pairs (P l.t. .05) by the number of all possible comparison pairs, was used to indicate the model's susceptibility.

The among model's comparison yielded low percentage figures. For example, the susceptibility of the models is 1.2 percent in the program's effectiveness, 10.6 percent in the estimated invalid sources and 0.3 percent in the estimated true effectiveness of the program. These figures indicate that our simulator is insensitive to the given change of an educational setting.

The within a model comparison produced relatively high percentage figures compared with the among model comparisons. The test showed that the model's susceptibility to change in the educational conditions of the same educational setting is 13.6 percent in the program's effectiveness, 30.7 percent in the invalidity sources, and 13.6 percent in the estimated true effectiveness of the program. Model 1 demonstrated 28.1 percent in the program's effectiveness, 31.9 percent in the invalidity sources, and 28.1 percent in the estimated true effectiveness of the program. Experimental Model 2 is unique in obtaining 44.4 percent of the susceptibility in the invalidity sources. Thus, our simulator is relatively sensitive to the change of an educational condition.

THE RELATIONSHIPS AMONG CONTEXT, INPUT, PROCESS, AND PROGRAM AVERAGE EFFICIENCY. As discussed in the previous sections; context, input and process evaluation exists in time lag dimensions in evaluation systems. That is, context evaluation precedes input evaluation which precedes process evaluation. Context evaluation, therefore, determines the scope of input evaluation, which determines the direction of process evaluation. In such cases, the correlation coefficients among them may allow cause interpretation. To calculate the coefficients; the number of days spent in context evaluation, input evaluation, and process evaluation; and the average program efficiency were used. Thirty cases were tested from three models with ten runs each. Table 8 reveals a negative relationship (-0.92, p=.001) between input and process evaluation and a positive relationship (0.62, p=.001) between context evaluation and the average program efficiency. The rest of relationships are low and nonsignificant at p=.05.

The negative correlation between input and process evaluation may be interpreted to mean that if the project staff spends shorter days in input evaluation, then they tend to stay longer in process evaluation. The reverse is also true. This indicates that the project staff should utilize time to the minimum in input evaluation, so that the project may accomplish the student's learning to the maximum. The positive correlation between context and the average program efficiency is an indicator that the more time the project staff spends in context evaluation, the higher the average progr am efficiency will be. In other words, well planned evaluation strategies can promise better student learning.

OPTIMAL SOLUTION. To obtain an optimal solution, one should estimate the equilibrium state probability density distribution, in which the waiting time (queue) to be served, becomes constant. In our simulation model, the estimated constant queue implies the average days required to achieve on state of evaluation after infinitive simulation runs. Table 9 presents the obtained data and apriori probability weights used for the estimation of an optimal solution. Based on these data and the solution (18) from the previous section, the optimal solution was estimated. In Table 9, the expression of 'product evaluation' does not imply the last stage of the CIPP model, but the transient stage of process evaluation in which daily activities of the project are evaluated and modified if anything irregular or unplanned is detected. The number of days obtained from a simulation run and an optimal solution indicates the weight of efforts that the evaluation project staff should consider in their activities.

Comparisons between the estimated models and the experimental models reveal the following differences. The estimated best-fit models spent nearly equal amount of days for context and input evaluation, whereas the experimental models spend more days in input than context. For example, the ratio of 1 and 2 of experimental Model 3 is 19.2/1.0, and the ratio of op 1 and op 2 of estimated Model 3 is 0.89/1.27. The same kind of tendency prevails in the other two models. The ratio of op 3 and op 4 for estimated Model 1, 2, and 3 is about 1/5, whereas in experimental models, no model detected the situation that the accumulated students' learning was less than the accumulated invalidity. This optimal solution suggests that in the infinitively repeated runs of the experimental models, we tend to have one day in every five days to be spent for a consideration of whether the implemented strategies should be modified or not because of the accumulated invalidity which jeopardizes the project. The ratio of 1 and 2 for experimental Model 1 is 3/9, but for experimental Model 3 the ratio is 1/19.2 in spite of using the same probability weight of .7 for the administrative preferential attitudes. This result may be traced back to the change of random initiators rather than the difference in the functions used.

With the information of optimal solutions, the efforts of an evaluator are directed to create the condition in the actual settings which generate the optimal solution. For example, if an evaluator is given the actual educational settings as parameters of experimental Model 1 shown in Table 9, then one makes an attempt to realize the estimated Model 1 situations, in which the ratio of spending days in context and input evaluation is about 1/1, and every sixth day an evaluator checks and adjusts the implemented strategies for better results.

Several suggestions may be made for future research on a simulation model as a result of this study. The first suggestion relates to a priori probability weights. In order to approximate realistic situations as closely as possible in a simulator, a special effort to find the empirical probability weights must be made. The empirical weights may be calculated from a wide range opinion survey of the relevant substance to a simulator. The second suggestion relates to a further study of probability density functions. For example, a careful investigation must be done regarding the learning propensity of students under a specific program. The inaccurate representation of propensity by probability functions may mislead the simulation results. Finally, the results from the simulator should be compared with the results from an actual evaluation project in order to establish the validity of simulation.

REFERENCES

(1) Ferrari, D. Computer Systems Performance Evaluation 1978, New Jersey: Prentice-Hall, 123 p.

(2) Jackson, J. R. "Job Shop-like Queueing Systems," Management Science, 1963, V. 10, p. 131-142.

(3) Ohara, T. Based on the author's review of twenty evaluation reports submitted to the Illinois Office of Education in 1977 and in 1978.

(4) Provus, M. M. "The Discrepancy Evaluation Model." Reading Curriculum Evaluation. (Taylor & Cowley eds.) 1971, 231 p.

(5) Siegel, S. Nonparametric Statistics for the Behavioral Science, 1956, New York: McGraw Hill, 136 p.

APPENDIX

TABLES 1 - 9

TABLE 1
Simulation Flow of Context Evaluation

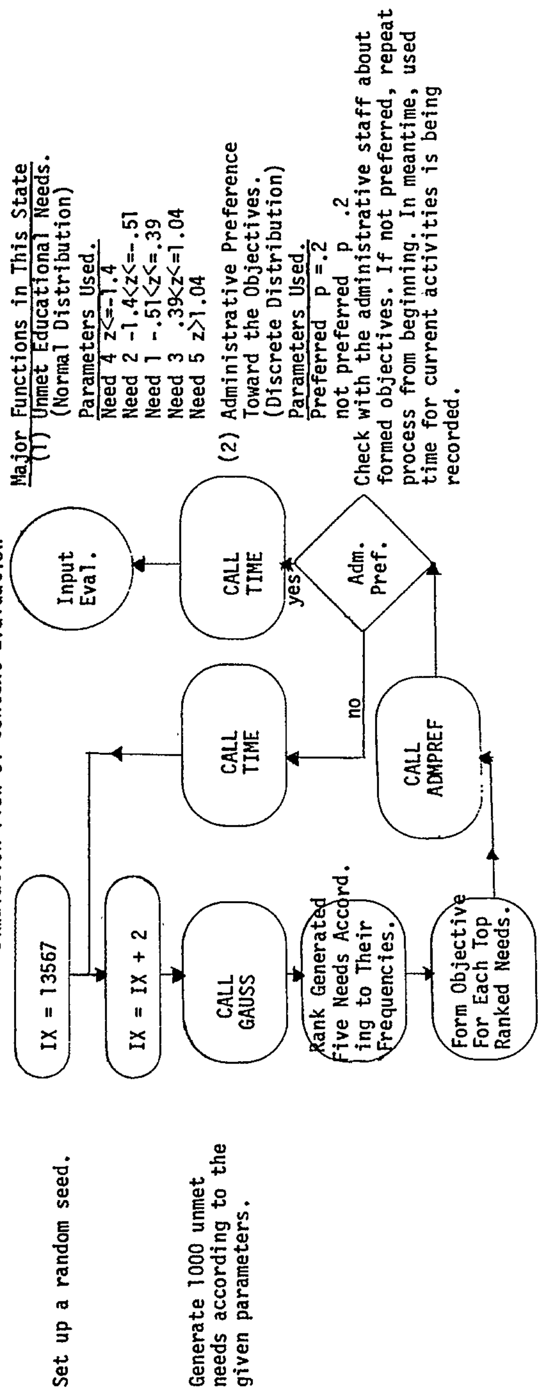


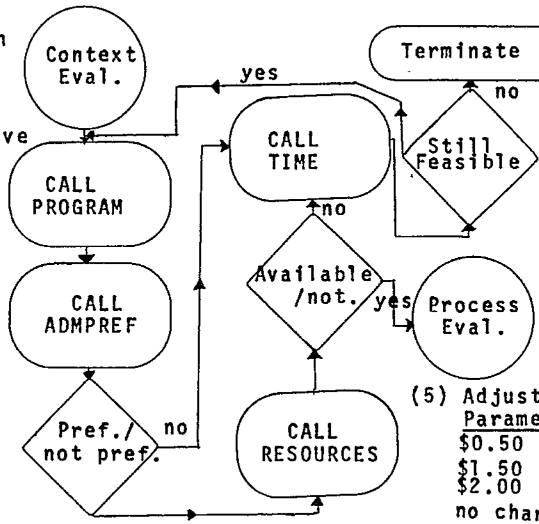
TABLE 2

Simulation Flow of Input Evaluation

Choose one program out of three for each three objectives with information of necessary cost and requirements.

Check with the administrative staff regarding the choices of these programs. If they are not satisfied with the choices, repeat the same process.

If the administrative staff is satisfied with the choices, then look for the funding agencies for the project.



(5) Adjustment
Parameters Used.
\$0.50 p<=.6
\$1.50 .6<p<=.85
\$2.00 .85<p<=.95
no change p> .95

Major Functions in This State.

(1) Programs.
(Discrete Distribution)
Parameters Used.
Prog. 1 p<.5
Prog. 2 .5<=p .8

(2) Available Resources.
(Discrete Distribution)
Parameters Used.
Available p>=.4
Not available p<.4

(3) Administrative Preference.
(Discrete Distribution)
Parameters Used.
Preferred p<=.7
Not preferred p>.7

(4) Time
(Discrete Distribution)
Parameters Used.
Time unit 1 p<=.5
Time unit 2 .5<p<=.8
Time unit 3 p>.8

TABLE 3

The Correlation Coefficients Among Invalidities (n=20)

	History	Maturation	Testing	Regression	Selection
History	1.00	-0.15	0.06	0.38 (*)	-0.08
Maturation		1.00	0.13	-0.03	0.31
Testing			1.00	0.40 (*)	0.00
Regression				1.00	0.23
Selection					1.00

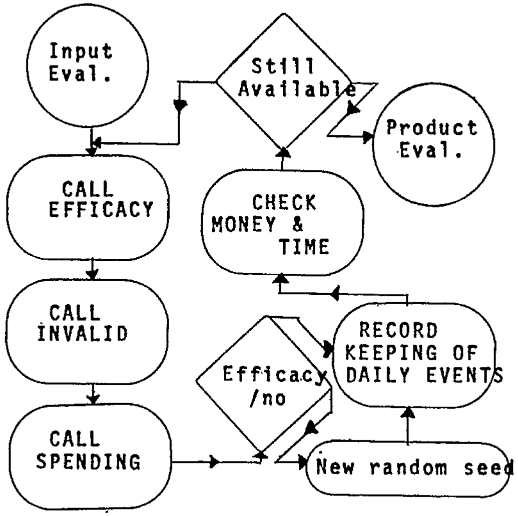
(*) indicates a significant pair at p<=.05

TABLE 4

Simulation Flow of Process Evaluation

Generate daily effectiveness, invalidness and cost of each project per person. If the accumulated effectiveness is less than invalidness of projects, then new strategy is implemented under the same programs.

Record keeping is also an important function at this state.



Major Functions Used.

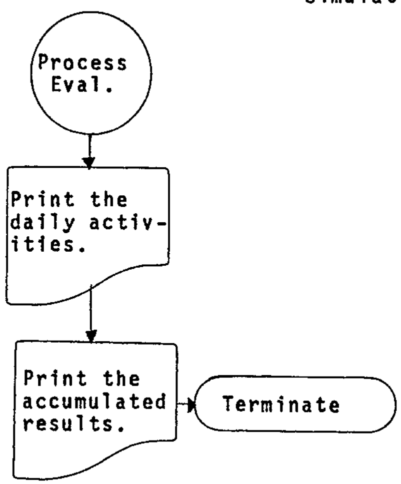
(1) Efficacy of Programs
(Normal Distribution)
Paramters Used.
effect unit 1 $z < -1$ 28
effect unit 2 $-1.28 \le z < -0.52$
effect unit 3 $-0.52 \le z < 0.52$
effect unit 4 $0.52 \le z < 1.28$
effect unit 5 $z > 1.28$

(2) Invalidness of Programs
(Discrete Distribution)
Parameters Used.
invalid unit 1 $p < .21$
invalid unit 2 $.21 \le p < .42$
invalid unit 3 $.42 \le p < .65$
invalid unit 4 $.65 \le p < .77$
invalid unit 5 $p \ge .77$

(3) Money Spending
(Discrete Distribution)
Parameters Used.
spending unit 1 $p < .3$
spending unit 2 $.3 \le p < .70$
spending unit 3 $p > .70$

TABLE 5

Simulation Flow of Product Evaluation



The daily activities of evaluation project are reported here. They are reported in the manner of per student, per project, and per day with regard to learning achievements of a student. The learning achievements are analyzed in terms of real achievement and 'face' achievement of a student.

In summing up the evaluation project, the accumulated achievements of all students involved are reported daily per program. Frequency distributions of face effectiveness, real effectiveness, and internal invalidities are also reported.

TABLE 6

Macro Optimal Solution of the Evaluation Systems

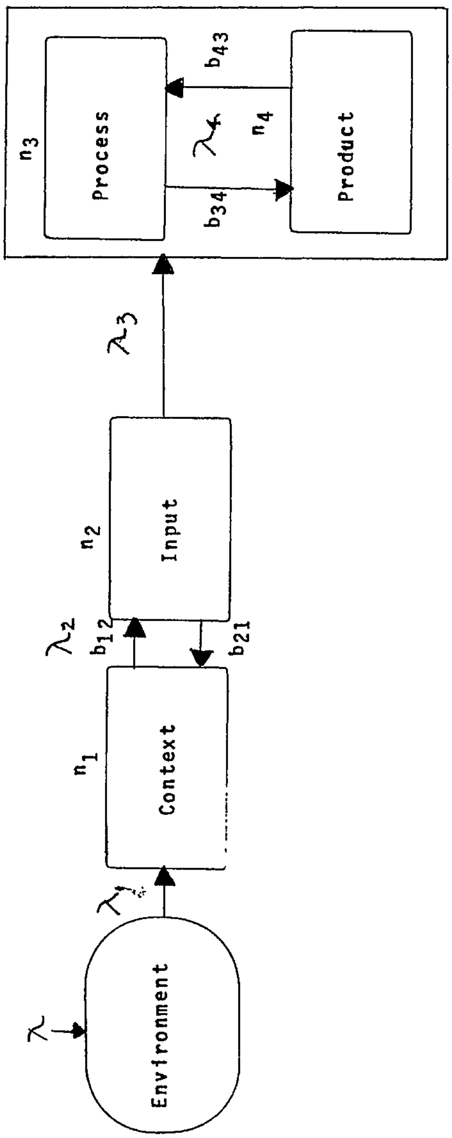


TABLE 7

The Percentage of Statistical in the Within and Among Comparisons

Among Models			
	Program's Effectiveness:	Model 1 and 2	0.0%
		Model 1 and 3	0.66%
		Model 2 and 3	3.0%
		average	1.2%
	Invalidness:	Model 1 and 2	1.0%
		Model 1 and 3	0.67%
		Model 2 and 3	30.0%
		average	10.6%
	Real Effectiveness:	Model 1 and 3	1.0%
		average	0.3%
Within a Model			
	Program's Effectiveness:	Model 1	28.1%
		Model 2	2.2%
		Model 3	10.4%
		average	13.6%
	Invalidness:	Model 1	31.9%
		Model 2	44.4%
		Model 3	15.5%
		average	30.7%
	Real Effectiveness:	Model 1	28.1%
		Model 2	2.2%
		Model 3	10.4%
		average	13.6%

TABLE 8

The Correlation Coefficients Among Context, Input, Process and the Average Program Efficiency

	Context	Input	Process	Program Efficiency
Context	1.0	-.213	.167	.619 *
Input		1.0	-.922 *	-.205
Process			1.0	.216
Program Efficiency				1.0

(*) indicates p less than .05.

TABLE 9

Optimal Solutions of Three Experimental Models and the Data Used for the Calculations

	Exp Model 1	Exp Model 2	Exp Model 3
1	3	1	1
12	.7	.9	.7
21	.3	.1	.3
2	9	8.2	19.2
3	46.7	46.4	40.4
34	.2	.2	.2
43	.8	.8	.8
4	0	0	0
	Optimal Solution		
	Exp Model 1	Exp Model 2	Exp Model 3
op 1	3.80	1.10	1.27
op 2	2.66	0.99	0.89
op 3	55.60	55.24	48.10
op 4	11.12	11.05	9.62

Note:
1 .. average arrival rate of context evaluation.
2 .. average arrival rate of input evaluation.
3 .. average arrival rate of process evaluation.
4 .. average arrival rate of product evaluation.
12 .. probability of objs. in input evaluation.
21 .. probability of rejected objs. to context
34 .. probability of specific program in product evaluation of process evaluation.
43 .. probability of specific program to modified and returned to process evaluation.
op 1 .. estimated arrival rate of context evaluation.
op 2 .. estimated arrival rate of process evaluation.
op 3 .. estimated arrival rate of process evaluation.
op 4 .. estimated arrival rate of product evaluation in process evaluation.

DR. TAKESHI OHARA is an associate professor who teaches business computer methods and management information systems. He received his degree from Southern Illinois University, Carbondale.

DR. KENNETH PICKARD is an associate professor who teaches business communications and business data processing. His degrees are from Indiana State University, Ball State University, and Northern Illinois University.

School of Business Administration
California State University
Long Beach, CA 90840
213 494-7691