DOCUMENT RESUME

ED 224 457 IR 010 438

AUTHOR	Wholeben, Brent Edward
TITLE	MICROPIK: A Multiple-Alternatives, Criterion-Referenced Decisioning Model for Evaluating CAI Software and Microcomputer Hardware Against Selected Curriculum Instructional Objectives. Paper and Report Series No. 73.
INSTITUTION	Northwest Regional Educational Lab., Portland, Oreg.; Washington Univ., Seattle.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	May 82
CONTRACT	400-80-0105
NOTE	129p.
PUB TYPE	Guides - Non-Classroom Use (055) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Computer Assisted Instruction; Computer Managed Instruction; *Computer Programs; Criteria; Decision Making; Educational Objectives; *Evaluation Methods; Instructional Materials; *Media Selection; *Microcomputers; *Models; Operations Research
IDENTIFIERS	Multiple Alternatives Program

ABSTRACT

This report describing the use of operations research techniques to determine which courseware packages or what microcomputer systems best address varied instructional objectives focuses on the MICROPIK model, a highly structured evaluation technique for making such complex instructional decisions. MICROPIK is a multiple alternatives model (MAA) whose overall goal is to formulate an evaluation and decision-making procedure and to model or simulate this evaluation framework as closely as possible, involving the school environment's established needs. An overview of the technical workings of the modeling framework and its performance of the evaluative comparison and final selection of alternative functions is followed by an explanation of the primary and secondary goals of the model. Alternatives evaluated by the modeling framework are discussed, as well as the criteria necessary to evaluate and compare these alternatives. Additional topics covered include constraints, the execution, results, and the general utility of such a model together with common advantages and potential pitfalls. (LMM)

Paper and Report Series

No. 73 MICROPIK: A Multiple-Alternatives, Criterion-Referenced Decisioning Model for Evaluating CAI Software and Microcomputer Hardware Against Selected Curriculum Instructional Objectives

Brent Edward Wholeben

University of Washington

May 1982

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue, Portland, Oregon 97204

NOTE

Published by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The work upon which this publication is based was performed pursuant to Contract No. 400-80-0105 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

PREFACE

The Research on Evaluation Program is a Northwest Regional Educational Laboratory project of research, development, testing, and training designed to create new evaluation methodologies for use in education. This document is one of a series of papers and reports produced by program staff, visiting scholars, adjunct scholars, and project collaborators--all members of a cooperative network of colleagues working on the development of new methodologies.

How can one reliably and efficiently match a desired instructional design to available computer hardware and software components? This report describes the use of operations research techniques in determining which of a vast array of instructional objectives are best addressed via which courseware packages or what microcomputer systems. The report presents a highly structured evaluation technique for making such complex instructional decisions.

Nick L. Smith, Editor
Paper and Report Series

CONTENTS

MICROPIK: A Multiple-Alernatives, Criterion-Referenced Decisioning Model for Evaluating CAI Software and Micro-Computer Hardware Against Selected Curriculum Instructional Objectives

THE CONTEXT

The evaluation of instructional objectives, available software for instructional implementation, and compatible hardware units rank among some of the more complex decisional problems facing educational professionals today. Like the issues of selecting school sites for closure and determining potential program units for roll-back based upon declining enrollments and dwindling fiscal revenues respectively, a highly structured and premeditated evaluation technique is required in the reliable assessment of valid criteria for determining which of the vast array of instructional objectives are best addressed via which packages on what microcomputer systems. We will explore this evaluation-decisional environment more closely before presenting a means for resolving the p oblems associated with matching the desired instructional design to computerized hardware and software components.

[Introduction]

The use of data processing technology within the educational domain has over the past several years restricted itself almost totally to such database management efforts as maintaining student and teacher personnel files, purchasing and inventory control, and other accounting/bookkeeping activities. More recently, and with the advent of affordable microcomputers for individual building-level use, electronic data processing activities have taken a firm hold of the instructional realm of the educational enterprise. This has been especially true in the such classroom-oriented activities as computer-assisted instruction (CAI) and computer-managed instruction (CMI).

Acting as the classroom teacher's adiutant, the microcomputer provides the hardware (electronic equipment) and software (actual instructional program materials) components necessary to promote a reasonably valid and reliable relay of information to the user student for the purposes of instruction, drill and evaluation (CAI). With additional sophisticated software, this electronic "right-hand person" is also able to track the performance and progress of each individual student, compare that progress with both local as well as prescribed norms, and schedule each student for either remedial, normal or enriched instructional activities based upon as assessment of the student's performance (CMI).

Since the overall cost of possessing an integral microsystem has become more reasonable over the past two-years, such standalone microcomputers as distributed by the Apple Corporation, Tandy Corporation (Radio Shack), Ohio Scientific, Texas Instruments, Pet-Commodore, and Health Kit, are becoming as commonplace as the standard ten-key adding machine was some few decades ago. And with the onslaught of hardware machines, has also come the proliferation of 'readv-made' software programs and packages available for use within the each particular system being marketed. Until recently, software designed for one system has been unusable on another system; and thus, selection of one particular microprocessor brought with it the forced acceptance of the philosophy, goals and related activities of the software supported by the particular operating system involved.

Today however, the days of system-restricted software are numbered, with software materials being coded for accessibility to many of the more popular hardware models on the market. And, as were once the textbook publishers concerned with usable workbook materials to complement their major instructional texts, their same research and development energies are now directed toward designing micro-software compatible with major hardware systems, and parallel to their more popular text-series. Available also recently, are diverse coded packages for use in the adminstrative arena of the school setting. Software packages designed to perform such school management applications as salary planning, student data recording, property management, accounting, payroll, personnel data recording, mark reporting and mailing label generation, are now available to the principal as readily as CAI and CMI packages are to the classroom teacher.

With the initial introduction of microprocessors on the educational scene, the more logical decisional sequence for selecting a machine remained in determining the utility of available software first, and then the parallel utilizability of the hardware compatible with the software chosen. Many schools nevertheless chose a reciprocal course of action -- that is, purchased a hardware unit for whatever reason, and then reviewed the availability of appropriate software for instructional and management activities -- unfortunately discovering that the more readily accessible machine was useless unless in-house software could be developed using one of the compiler languages; and also finding that few if any school personnel had the training or ability to program the required application(s).

The emerging wide-spread availability of software packages compatible to many of the more popular hardware systems on today's market, precludes many of the limitations involved in the 'chichen and egg' controversy illustrated above. However, the sophistication and regimen of today's hardware-software decisions are no less complex or complicated based upon the diversity and versatility associated with the software compatibility and hardware accessibility. In fact, the decisioning structure could be said to now be more complex, since such a wide diversity of potential choices -- mixes and matches -- are possible in the final design of a school-based data processing system utilizing the microcomputer hardware foundation.

[The MAA Situation]

The optimal choices associated with matching existing and/or desireable curricular objectives and instructional activities with available CAI/CMI software, and the array of microcomputer systems compatible with the useable software -- exists as one of the more complicated applications for which the utilization of the MULTIPLE ALTERNATIVES MODEL (Wholeben, 1980a) is specifically suited. Such a multiple alternatives analysis (MAA) situation is really a combination of six underlying sub-decisional systems which integrally represent the mix-match solution required. These sub-decisional systems can be defined as:

[1] the curriculum subsystem -- that is, assessing the differential strengths of various instructional activities in providing the foundation for valid satisfaction of curricular goals and objectives, and the ultimate accomplishment of the specific concept learning desired;

[2] the program software subsystem -- that is, assessing the differential utilities associated with each of the available instructional CAI or CMI packages in promoting the instructional activities underlying the purported design and development of each individual software unit, and its emphasis upon concept introduction, activity drill and practice, and assessment of learning which results;

[3] the hardware machine subsystem -- that is, assessing the differential utilities associated with each of the available microcomputer systems, and their concurrent support of such required peripherals as CRTs, printers, disk storage units, central memory capacity, graphic plotters, and interfacing potential larger, mainframe systems;

[4] the curriculum/software subsystem -- that is, assessing the degree or extent of capability in matching some subset of the instructional goals (activities, objectives) with defined characteristics of software packages, and the ultimate accomplishment of the specific concept learning required for 'normed' performance and progress;

[5] the software/hardware subsystem -- that is, assessing the degree or extent of capability in matching some subset of the availability software packages to the operating characteristics of the various hardware systems, and assuring that the program software units defined will be compatible to the hardware units selected; and

[6] the curriculum/software/hardware subsystem -- that is, assessing the total instructional system impact associated with particular 'mixes and matches' of the three major decisional systems incorporated within the multiple alternatives analysis setting.

Thus within each of the three major systems related to curriculum, software and hardware individually, there exists a sub-MAA model inherent to the overall multiple alternatives decision to be made. We will not dwell upon the obvious, but to illustrate the concept of multiple alternatives analysis, and its reliable means in modeling this CAI-related decisional situation.

Within the curriculum subsystem, the multiple alternatives are defined by the various 'alternative' activities which might be executed to satisfy stated instructional objectives; and in turn, the 'alternative' objectives which might be satisfied in order to bring about the desired conceptual learning. The mix/match of potential solutions to this delimma is illustrated by the various combinations some activities may form with other activities in satisfying the ultimate conceptual learning demanded of the instructional subject area or grade-level defined.

The software subsystem provides a different form of mixing and matching for final solution, since different software packages may or may not complement each other -- but do portray varying measures of effectiveness, efficiency, satisfaction and cost which are internal to the individual packages themselves. Thus while a particular package may in fact promote rapid and effective learning, the cost of this same package may be an ultimate factor in precluding the software unit from forming part of the solution.

A final example of the applicability of MAA and its utility in modeling these CAI requirements can be witnessed within the software/hardware subsystem. Here, the compatibility of each individual software unit for the particular hardware (operating) system included as a potential purchase, is controlled for. As was a common mistake some few years ago, the model for evaluating the multiple alternatives involved in choosing the best match between desired learning outcomes, available software packages, and compatible hardware systems must certain assure any selected software program will be functional on the hardware system purchased.

[Initial Assumptions]

As with all modeling situations, wherein some aspect of a decisional environment or milieu is to be simulated (i.e. tested for potential impact based upon expected occurrences), there exist some basal assumptions which the modeler must address, and be permitted to acknowledge in the final development of the decisional model. For the MICROPIK model, key assumptions will involve the availability of (and/or accessibility to) quantified criterion measures for comparing the various subsystem mixes of instructional activities, software and hardware, the degree to which the classroom teachers will submit to defining their courses and subject matter into specific, differentiated instructional units (observable activities), and the extent to which different instructional disciplines (mathematics, language arts, science, industrial arts, health education, etc.) can be co-terminously model (together).

The first assumption -- the availability of valid and reliable criterion measures suitable for evaluative comparison -- is integral to design of the MAA modeling framework; and therefore a sine qua non requirement for continuation with further model construction. However, these measures do not have to exist in the a priori sense to model design, but of course must be available for successful modeling execution and decision formulation. Such data gathering requirements will involve a quasi-experimental situation, in which measures of effectiveness, etc. are collected based upon observed (or perceived) performance. Since many of the criterion measures related to software and hardware will have to be assessed by the model builder, a related assumption exists that the number of software packages and hardware units for the intended modeling evaluation be limited to a set of likely candidates; and thus reduce the necessary complexity of the model to be constructed.

The second assumption, and often the most difficult to realize, is the delineation of instructional concepts and goals into a finite set of observable and performance-related instructional activities. Although the recent rebirth of demanded specificity and measured accountability for the classroom teacher via such implements as the student learning objective (SLO) in assuring the performance output associated with learning, many teachers seem reluctant to specifically identify which activites are definitively associated with which desired learning outcomes. Over the decades, the classroom teacher has evolved through such rhetoric as academic freedom and instructional autonomy into a 'not-to-be questioned' professional, with an internal code of ethics but without the presence of an external monitor. The collapse of the yearly teacher evaluation into a 20-minute observation of classroom tactics; and the absence of in-service instruction for improving the performance of the "experienced and tenured" staff person -- point to many of the failings of the educational domain as a managed and controlled environment.

To successfully model the evaluation of instructional activities against available software and hardware components however, requires that such a delineated framework of instructional objectives exist. Again, such delineation does not need to be in existence at the commencement of modeling construction -- and may proceed as the remaining parts of the model are developed.

The third and final key assumption on the part of the modeler as this CAI-related model is constructed, remains the extent to which the total instructional system (i.e. all disciplines) are modeled within the same formulation. For most purposes, it will be necessary (and acceptable) to model each discipline separately; and thus not constrain the decisional solution to be a resolution compatible to all aspects of the instructional milieu. This has many advantages as well as disadvantages; but remains a more workable format, and one which can be more easily descriptive of the particular disciplinary area.

[Projected Expectations]

The MICROPIK modeler is cautioned to remember, that the resulting criterion-referenced simulation of selecting the most appropriate software and hardware mix for optimal satisfaction of pre-stated instructional objectives and pre-defined instructional activities -- due to the complexity of its structure, and the naive face validity given its processes -- will often lead the general population (administrators and teachers) to believe its results (i.e. decisions of match) as the "gospel according to MAA". Although this author certainly does not discourage such discipleship, it is reasonable and prudent to understand the output of the CAI-MAM designed system as the best-likely decision based upon the criteria defined, and the modeling formulation constructed.

Oftentimes, certain specific requirements of a particular decision will not (or can not) be sufficiently modeled (that is, incorporated within the decision model design). If the modeler recognizes this fact, no compromise to the system is realized. However, the expectations of individuals effected by their understanding (albeit rudimentary) of the modeling framework will often be impacted by such a conscious (or unconscious) omission. Many criterion references may have to be applied to the formation of the final solution after surveying the results of the model's execution based upon the criteria input. Such a subjective addition to an otherwise 'objective' model is not compromising to the model, as long as the subjective criteria is agreed-upon as valid input to the final decision; and as long as such additive processes are consistent and visible for examination.

The classroom teacher in particular, must be brought to understand the decision model as a 'best match' of multiple alternatives. Teachers are often hesitant to adopt or accept a decision which is not 'perfect' -- and therefore have some difficulty in accepting the idea of optimality in problem resolution. Nothing has provided more of a barrier to the adoption of CAI and microcomputers within the instructional setting, than exactly this feeling of CAI being 'not as good as' the flesh-and-blood teacher -- and therefore additional expenditures should be directed towards greater teacher recruitment and concomitant reduction of teacher-pupil ratios, rather than the acquisition of microcomputers and packaged software.

[Expected Difficulties]

Several barriers and/or pitfalls can be expected during the initial design and formulation of the decisioning model, and during the examination of its output (modeled decisions). Some of these are model-related while others are user-related, and have been alluded to earlier in this paper.

The major, and probably most 'key' problem to be overcome by the modeler for acceptance of the MICROPIK framework, refers to the use of quantifiable measures (i.e. numbers) for measuring everything from effectiveness through perceived satisfaction, and required revenue expenditure. Mathematicians have long since given up on the critics who having claimed that 'not everything can be related to numbers', proceed to maintain that (therefore) 'nothing should be'. However, each modeling situation will not be devoid of such criticism, nor will any acceptable response or retort be useful. Obviously, all things can not be modeled in a quantitative sense -- but those that can, should not be ignored because of the conflict which may arise. Valid referencing and scaling of criteria, and their reliable measurement -- are the best (optimal) defense to the numbers-critic.

Other difficulties have been referenced in preceding sections, including the reluctance of teachers to definitively specify the relationships of activities to concepts learned (and objectives satisfied), the perception of compromise based upon optimization, and the acceptance of modeling-by-discipline rather than including the full needs and demands of the school setting -- although this last problem can often be a strong factor in the acceptance of results on the part of the individual disciplines or subject areas.

An additional difficulty to be faced by the modeler will concern itself with the concept of 'collective exhaustiveness' regarding the inclusion of criteria impacting the final solution or decision. It is a favorite technique of the modeling critic to announce, "... but what if this particular criterion had been included in the final design of the solution ... would a different decision have necessarily resulted?". The simulation design must be ready to incorporate additional criteria for re-execution of the original modeled framework; and thereby be able to detect any differential solution formulation based upon the existence of new criterion measures. And at times, the modeler must also be ready to state categorically, and be ready to defend the position, 'enough is enough'.

A final major difficulty to be faced by the modeler and the eventual acceptance of modeling results will concern: first, the validity of the criteria selected for impacted and constructing the solution, and the parallel validity of the references (sources) defined as producing these measures; and second, the reliability of the procedures utilized in gaining these required measures. Data will sometimes be available via records, other times via standardized instruments, and sometimes only through the administration of a subjectively-based opinion questionnaire. Advance planning and careful implementation of the data gathering portion of the model building sequence, will have great rewards in the end. In the same vein, nothing so completely nullifies and destroys an otherwise careful modeling effort, than the inclusion of invalid criteria or use of unreliable measurement techniques. Even though rectified, the subsequent results of the modeling solution will be viewed with distrust and non-acceptance.

THE MODEL

Before proceeding to demonstrate a sample construction of the MICROPIK model for evaluating various software and hardware packages across desireable curricular and instructional objectives, it is necessary to examine the rudiments of the 'multiple alternatives' modeling framework in greater technical detail. The colloquial 'garbage in, garbage out' remonstrative exists as especially pertinent to the development and implementation of the MAA modeling setting. Choice of alternatives and definition of their inter-relationships, the inclusion and specific referencing of certain criteria (and the exclusion of others), and finally the control fostered by what we will come to call the "RHS vector" (the 'right-hand-side') -- will force the model to execute in a manner either consistent with the situation being simulated, or in compliance with decisions already made by policy bodies, and now requiring pooled support and accompanying data.

Before building the specific CAI-MAM model, let us now in a very brief fashion begin to view the technical workings of the modeling framework; and how it performs the intended evaluative, comparison and final selection of alternatives function.

[Introduction]

The complex issue of multiple alternatives decision-making is no stranger to the educational analyst. The selection of some number of schools from a relatively large pool of potential candidates for closure is a MAM problem. Each school site represents varying measures of effectiveness, efficiency, satisfaction and expenditure for each of a number of criterion references (e.g. capacity of building, heating requirements, building age, projected enrollment change over future years, safety factors of neighborhood, and proximitv of other schools and their ability to absorb transferees in the event of the first school's closure). Some of these measures will be adjudged satisfactory (or nonsatisfactory) to varying degrees, and will be comparable with other schools across the district.

However, to include one site for closure as opposed to another site means, that "good" aspects of a 'to-be-closed' school must be sacrificed in order to keep the other school operational, even though the 'to-be-kept-open' school may have certain unsatisfactory measures on the same criterion variables which the now closed school exhibited as satisfactory. Such modeling of this decisioning situation is known as interactive effects modeling (Wholeben, 1980a), and represents the necessity of constructing solutions sets which will invariably include some form of 'controlled' preference/trade-off mechanics as the various alternatives are evaluated.

The issue of complexity is also represented in the statement of the problem: to select some number of schools for closure in order to promote certain defined goals of the district; and thus to determine how many schools will be closed and which ones. Obviously, such a model must in effect be simultaneously performing these two inter-related decisions: "how many?" and "which ones?".

The determination of which program unit budgets will be decisioned for continued funding (versus deallocation) is another example of the multiple alternatives framework, and its superior contribution to the realm of accountable and criterionreferenced evaluation and decision-making (Wholeben and Sullivan, 1981). In the fiscal deallocation model, criteria represent the projected expenditures within each object cost code for each of the units under evaluation; and in addition contain perceptual measures of administrative level of expendability. Once again of course, exists the dual responsibilities for determining how many program budgets will be discontinued, and which ones -- based upon the interactive modeling effects of the various criterion weights across unit alternatives.

[The Criterion Vectors]

The multiple alternatives model is simply a system of simultaneous linear inequalities and equalities which collectively represents the problem to be solved. Such an algebraic linear system is portrayed in <Figure 1>. Note how each linear combination represents a vector of values (viz., coefficients) which identifies the total, measureable impact to a system of the alternatives being modeled. Thus there exists a unique (normally) combination of coefficients for each of the criterion references used as input to the decisioning process. The alternatives themselves are further defined as binary variables (that is, taking on the value of either 0 or 1 (to be excluded in the final solution set, or to be included, respectively). Vector formulation for each criterion reference,

$$[\ a_{i1}x_1 \quad a_{i2}x_2 \quad a_{i3}x_3 \quad \ldots \quad a_{ij}x_j\ ]$$

portraying $\underline{i}$ criterion references across $\underline{j}$ alternatives, will then provide a basis for measuring total impact to the system as a whole attributable to the solution set constructed. Bounds (or limits) to what is allowable as a total impact to the system are expressed as vector entries within the conditional vector (or normally named, RHS, the right-hand-side). The RHS-values are the constants of the equations and inequalities modeling the system. <Figure 2> presents a listing of the four generic types of criteria to which each model should address content validity; and <Figure 3> depicts these criterion entries as members of the modeling framework previously illustrated within Figure 1.

[The Objective or Optimality Vector]

The remainder of the modeling process concerns the use of an additional vector to assist in determining from the potentially hundreds (or millions, in some exercises) of possible alternatives, that <u>one</u>, <u>best</u> mix for which the best, possible solution exists. This process is called the search for optimality, and the vector is known as the objective function (or sometimes, the cost vector). Geometrically, the objective function is a n-1 dimensional figure passing through the n-tuple space (convex) which is feasible (that is, includes all of the constraints postulated through the use of the linear equalities and inequalities) and which seeks a minimum point within the feasible region (if the goal is to minimize the impact of the objective function's values upon the system) or a maximum point within the feasible region (if the goal is to maximize the defined objective function's impact to the system as a whole).

[The Goal of MAA and MAM]

Simply stated, the multiple alternatives model is a technique which seeks to construct a solution set (a vector of 1's and 0's), such that this same solution vector represents the solution of the simultaneous system, constrained by a series of competing criterion measures (vectors), and based upon the optimality demands of the objective function.

THE MISSION

As with all complex applications of planning, design and development in the construction of systematic evaluation and decision-making models, the MICROPIK framework is built upon a delineative, deductive base. The overall goal or mission of the MICROPIK model is to formulate an evaluation and decisioning procedure, based upon the criterion-referenced assessment and comparison of various optional alternatives regarding curriculum goals, available software and compatible hardware; and to model this evaluation framework as closely as possible (i.e. simulate) with the established needs and demands of the school environment involved. In a more simple sense, "to do what needs to be done, and what the properly ordained decision-makers would do, if they only could". Sounds straight-forward enough, do you not agree?

[Mission of the MICROPIK]

It is the mission of the MICROPIK modeling framework to design and develop:

> a multiple-alternatives, criterion-referenced modeling structure -- evaluating and comparing potential microcomputer instructional software and related machine hardware -- resulting in an informed decision as to which software packages and hardware units are most optimally suited for enhancing the established instructional objectives for computer-assisted (CAI) and computer-managed (CMI) instruction within the educational enterprise.

A secondary statement of mission is also possible, dealing more generally with the CAI-MAM aspect of the modeling framework, yet more specifically with the notion of decision modeling; that is, to design and develop:

> a decisioning simulation structure -- capable of incorporating the desired, potential decisioning alternatives of the major policy bodies, and the relevant, valid criteria admissable to the needed comparison of alternatives -- and in full accord with established policy, consistent practice, and mandated legal principles and individual rights.

While the primary statement of mission (above) deals more directly with the framework and constructs of the MICROPIK application of multiple alternatives analysis (MAA), the secondary mission addresses specifically the foundational constructs of the underlying multiple alternatives model (MAM) itself.

[Major Secondary Goals]

As with the primary and secondary statements of mission defined in the preceding section, design and development of the MICROPIK modeling framework will encompass several delineative levels of goals, objectives, activities and tasks before the final MAM structure is ready for execution. The construction of such a systemic model is itself an exercise in implementing the usual constructs of a more generic "planning model". A developmental paradigm (roadmap or blueprint, if you wish) is essential for the controlled construction of a reliable decisioning technique; and that technique's inclusion of valid datum and algebraic relations.

Parallel to the normal (major) goals which would accompany such model construction (e.g. planning, historigraphic review, general design, field-testing, implementation, and assessment) certain secondary goals are of demonstrative important within the modeling episode; and bear illumination and clarification at this time.

The first, major secondary goal within the design and development of the MICROPIK framework, is to maintain vigilance upon the mutual-exclusiveness construct -- regarding both alternatives included for comparison, and criteria chosen for performing that comparison. Alternatives should be separate and independent (i.e. mutually-exclusive) of other alternatives within the model. This of course will not always be desireable; and at times, the modeler will seek to correlate the usefulness of one alternative based upon the parallel existence of another alternative. This would especially be true of an instructional objectives and activities model, where sequential and progressive learning and reinforcement must be available for optimal concept learning.

Parallel vigilance upon the mutual-exclusiveness of the criteria included within the modeling framework is a matter of model efficiency, rather than a source of unreliability. As in the past 'dark history' of evaluation and decision modeling, the model builder has not always maintained the highest professional standards; and has therefore constructed the model to best depict the specific decisions desired. This procedure of 'stacking the model' is not possible with the MAM framework, in terms of including a mass of 'stacked' criteria to weight intended decisions in a certain direction. However, this is a major concern when addressing the construct of criterion collective-exhaustiveness.

The next, major secondary goal within the design and development of the MICROPIK framework, is to insure the collective-exhaustiveness of both alternatives compared, and evaluative criteria utilized. Completeness or systemic totality of the modeled simulation is of primary importance; and exists as one of the most potentially compromising circumstances regarding the possible nullification of model results.

Without the collective-exhaustiveness of the multiple alternatives represented within the model, immediate criticism will be directed towards the model as not comparing 'all possible' decisional alternatives. And, even though some alternatives may be a priori determined to be a necessary part of the final solution (regardless of their attributes as measured by the criteria), these same alternatives must be included within the model in order to summarily include the impact to the system as a whole, based upon their 'forced' inclusion within the solution vector.

As mentioned above, the collective-exhaustiveness associated with the criterion-references must be a major concern of the model builder. Simply stated, if a particular criterion is not a part of the MAM framework, then neither its impact upon the various alternatives involved nor its effect upon the system as a whole can be represented and controlled. Unfortunately, the construct of collective-exhaustiveness applied to criteria is also one of the primary nemeses of the modeler. Without a doubt, demands will exist to include 'new' and 'different' criterion measures in order to survey their resulting impact to the model's decisioning process; the "... but, what if ...?" situation has been mentioned previously. Reconstruction of the model, and the related summary of new results can be very tedious, time consuming, and moreover nerve-racking for the modeler. Because of the time and expense (both fiscal as well as mental) involved, the actual independence or non-collinearity of additional criteria can often be addressed via such available techniques as parametric or non-parametric bi-variate correlation methods, and/or the use of a oneway analysis of variance procedure (to assess relative bias).

The third, major secondary goal associated with design and development, pertains to the referencing, scaling and measuring of these mutually-exclusive and collectively-exhaustive criteria. Oftentimes, a criterion will be defined (e.g. satisfaction) which defies direct, physical measure, and must therefore be referenced and measured via more synthetic techniques (e.g. opinionnaires or surveys) to obtain modeling input (Wholeben, 1980a; 1980b; Wholeben and Sullivan, 1981). In other cases, the method of scaling the sought criterion measure (that is, how quantified) will provide declarations of potential unreliability from model critics. For example, witness the ongoing controversy concerning the use of the agreement-continuum wherein proponents of the five-point:

STRONGLY DISAGREE		DISAGREE		NO OPINION		AGREE		STRONGLY AGREE
	.		.		.		.	

continuum scale "strongly disagree" with the six-point scale:

STRONG DISAGREE		DISAGREE		MODERATE DISAGREE		MODERATE AGREE		AGREE		STRONG AGREE
	.		.		.		.		.	

whose proponents state categorically, that "everyone has some degree of opinion, no matter how small or truly uninformed".

The controversy associated with referencing of course can be often only marginally defensible by the model builder. For example, if you want to know if parents are dissatisfied with the management and instruction of their neighborhood elementary school, as a measure of potential for the site to be closed in an era of declining enrollment -- you may not wish to ask the question via a survey, "Are you satisfied with your children's school?", in a climate of potential elimination of school sites. Other 'backdoor' methods will be necessary to obtain measures of satisfaction, without pre-biasing the respondent's input.

A final, major secondary goal to be addressed within design will concern the possible, desireable weighting of some criterion measures over others. Several techniques are possible for this within the MAM framework (weighting individual vector entries, modifying the RHS-vector, and weighting various solution vectors from the solution tracking matrix of cyclic optimization). Not only must be valid and reliable technique be utilized in the event that weighting is necessary; but so also must the procedure for obtaining the direction and extent of these weights from the policy bodies be beyond reproach.

[Selected Major Milestones]

As with all planning activities which include a systematic approach to design and development as well as a heavy time comittment for implementation and evaluation, several 'points of potential concern' sine qua non can be identified by the modeler. This points or decision junctures are important in that if any delay to the activities preceding the juncture is experienced, the whole process will be delayed; or in the parlance of the planning and networking theoriest, a 'bottle-neck' formed. For the reader additional understanding of the developmental aspects associated with model design and implementation, the following list of selected major milestones has been formulated.

M-01 : ACCEPTANCE OF THE MODELING ENVIRONMENT
M-02 : REVIEW OF ESTABLISHED POLICY/PROCEDURE
M-03 : DEFINITION OF CONTEXTUAL NEED/DEMAND
M-04 : STATEMENT OF MISSION/GOALS/OBJECTIVES
M-05 : FORMULATION OF ALTERNATIVES (w/ REVIEW)
M-06 : DEFINITION/REFERENCE OF CRITERIA (w/ REVIEW)
M-07 : DATA COLLECTION/SCALING (w/ REVIEW)

M-08 : EARLY FIELD-TEST OF MODEL (COMMUNICATED)
M-09 : FULL-SCALE EXECUTION OF COMPLETED MODEL
M-10 : ANALYSIS AND SUMMARY OF FINDINGS
M-11 : VALIDITY AND RELIABILITY TESTING
M-12 : PUBLIC HEARINGS (w/ FINAL REVIEW)
M-13 : SELL HOME AND LEAVE TOWN (w/o REVIEW).

This is hardly an exhaustive list; and with even a minor clarification and delineation of topic could results in several hundred milestones -- each as important as the more relevant 12 expressed in the above listing.

Finally, the 'non-planning theorist' reader must also understand, that the above milestones need not be addressed (and planned for) in an independent, separate fashion. Many facets of the modeling process take place in parallel order (as opposed to serial); and so several phases of the modeling process will be ongoing simultaneously. One of the best and most illustrative examples of such simultaneity occurs during the alternatives' development phase. As alternatives are defined and explored, the modeler will find it hard not to (in parallel) also explore the types of criteria which would be useful in evaluating the various alternatives, how these criteria might be defined, references, scaled and measured -- and even how they might be formulated within a criterion constraint vector for input into the MICROPIK decisioning model. Of course, some aspects are truly serial, and can not be performed simultaneously; for example, the serial order of the field-test versus the full-scale implementation.

We will now examine in specific detail, the illustration of the MAA and MICROPIK missions, and the implementation of their stated secondary goals.

THE ALTERNATIVES

The first major phase of MICROPIK design and development cocerns the identification, definition and development of the multiple alternatives to be evaluated by the MAA framework. The reader will recall, that the mission of this MAM-CAI modeling exercise seeks to evaluate stated curricular objectives and instructional activities (and their projected influence upon the desired degree of related concept learning), the appropriateness of available CAI/CMI program software for implementing these instructional learning exercises, and the correlated compability of existing microcomputer hardware (including peripherals) to execute the various program software packages. We will develop the alternatives-portion of the MICROPIK modeling framework within this current section; and reserve the next section for an exploration of the necessary criteria to evaluate and compare these alternatives.

The reader will also recall, that although such an evaluation of curriculum-software-hardware alternatives, and their interrelationships could very well be an end in itself, the author's over-riding concern is to posit a decisioning model by which schools and service districts will be able to make 'intelligent' decisions regarding the acquisition of computer software and hardware, and its utility in fulfilling stated computer-assisted and computer-managed instructional objectives.

[A Tri-Partite Hierarchy]

As was demonstrated in the initial development of the "curriculum activity packaging" (CAP) model (Wholeben, 1980b), a MAA modeling of curriculum objectives and instructional activities as related to concept learning could be demonstrated via a delineative or hierarchical framework. Consider the usual representation of the concept-objectives-activities environ:

CONCEPT 1.0

Objective 1.01 | Objective 1.02

Activity 1.01.01 | Activity 1.01.02 | Activity 1.02.03 | Activity 1.02.01 | Activity 1.02.02 | Activity 1.02.03

The multiple-alternatives formulated MICROPIK model seeks to satisfy to some optimal degree, all concept and objectives-related learning as specified by curriculum requirements. The existence of multiple-alternatives for MAA evaluation exists in the formulation of the various activities "which might" be implemented in order to meet instructional (learning) needs and demands. In the MICROPIK setting therefore, <u>all</u> concept and objectives learning must be satisfied -- it remains the activities which will evaluated for their relative utility or appropriateness in fulfilling this required satisfaction.

In a more advanced formulation of the MICROPIK model, where objectives are to be considered alternatives available for comparitive assessment as well as the underlying activities which demonstrate the execution of the objective's intent, it is still the evaluation of the activities which will not only demonstrate their utility for inclusion within the final curriculum package, but also inductively determines whether the particular objective which defines their presence will be itself associated with the final solution set. The reader should also see therefore, that alternative 'concepts' could also be modeled in this way.

This three-level or tri-partite hierachical formulation of the multiple-alternatives structure is extremely useful to the modeler, should such defined sophistication become necessary based upon the situation being simulated. As we will see, this delineative structure within the alternatives definition will become one of the major modeling constructs to emerge from the design of the MICROPIK framework.

[The Sectional Alternatives Vector]

Because the MICROPIK model seeks to evaluate the corresponding relationships between curriculum, software and hardware -- as well as comparisons within each of these three groups -- the structuring of modeling alternatives may be classified into the three major groups:

[1] curriculum/instructional alternatives;
[2] CAI and other program software alternatives; and
[3] hardware and peripheral(s) alternatives.

As with the tri-partite hierarchical development of the curriculum objectives and instructional activities, the design of both the software and hardware alternatives will assume a hierarchical setting.

Structuring the second section of the alternatives vector (recall that the first section refers to the curriculum entries) will be primarily concerned with different aspects of the same curricular or disciplinary framework being modeled in the first section. For example, language arts may require CAI packages which related to various types of instructions, such as: reading, spelling, vocabulary, sentence structure, and analogies. Several software packages may exist for each of the above five required areas which will summarily require evaluation both in terms of their variable values between each other (package), and in terms of their utility in addressing the stated instructional activities. The hierarchical design for this section of the alternatives vector may be constructed as:

READING		SPELLING		VOCABULARY			SENTENCE STRUCTURE			ANALOGIES	
Pkg 1	Pkg 2	Pkg 1	Pkg 2	Pkg 1	Pkg 2	Pkg 3	Pkg 1	Pkg 2	Pkg 3	Pkg 1	Pkg 2

The above seemingly bi-partite design could easily assume a more tri-partite status is difference grade-levels for CAI within the elementary school setting became a new, confounding variable for modeling within the language arts portion of the MICROPIK model.

The third and final part of the sectional alternatives vector will contain the various multiple alternatives related to the utility of various hardware machines (and their peripherals) in implementing the evaluated comparisons between the curriculum desired, and the software packages which best instruct the related instructional activities. This particular collection of hardware alternatives can be greatly simplified if the modeler in advance agrees upon 'hardware packages' for inclusion within the MICROPIK formulation. Thus, a certain model of APPLE (e.g.

APPLE II PLUS), a certain type of printer, and a certain number of disk drives might become the "APPLE" package, and therefore a single alternative for comparison against the "TRS-80" package, or the "OSI" package, etc.. In comparison with the tri-partite hierarchical structure of the instructional activities, and the bi-partite structure of the software alternatives, the hardware section of the alternatives vector would become a uni-partite or single-level collection of multiple alternatives:

APPLE PACKAGE	TRS-80 PACKAGE	OSI PACKAGE	ATARI PACKAGE	TI PACKAGE ... etc.

However, should different models of the same microprocessor be required for alternatives decision-making, and should varying types of peripherals be required for inclusion within the full MAM formulation -- a tri-partite (manufacturer-model-peripheral) relationship reappears. Because some manufacturers have refused to keep their software model independent (e.g. some TRS-80 II packages will not work on the TRS-80 III; and likewise for the latest problems between APPLE II PLUS and compatibility with the APPLE III), a higher-order decision may need to be made concerning not only the type of software and peripheral required, but also the compatibility of the 'level of model' needed to execute the compatible software. The discerning reader can easily see how a quad- or even quint-partite hierarchical structure may be necessitated by such a complex multiple-alternatives setting.

[Summary]

Thus the alternatives' vector for exposition of the MICROPIK model is divided into three main sections: the tri-partite curriculum section, the bi-partite software section, and finally the (hopefully) uni-partite hardware section. However the reader is cautioned regarding the true partitioning of the hardware section of the alternatives vector. It is very likely in consideration of the problems with the lack of upward-compatibility of a particular system's software, and indeed in the quality-differential between peripherals and the type of peripheral (e.g. graphics plotters), that the hardware section could easily take on quad-partite characteristics.

In summary then, the alternatives vector can be illustrated as follows:

CURRICULUM COMPONENTS					SOFTWARE COMPONENTS					HARDWARE		
ACT	ACT	ACT	ACT	ACT..	PKG	PKG	PKG	PKG	PKG..	PKG	PKG	PKG..
1.1	1.2	1.3	2.1	2.2	1.1	1.2	2.1	2.2	2.3	1	2	3

A more detailed representation will be presented at the conclusion of this paper.

THE CRITERIA

To fulfill the stated premises of the MICROPIK model in determining the appropriate microcomputer hardware and software in terms of stated instructional requirements, the various sections of multiple alternatives described in the previous topic must be evaluated across various competing criteria. As was mentioned in a previous section, the MICROPIK decisioning model requires a total of six 'types' of criterion formulations: three to address the intra-relationships existent within each of the three sectional areas of curriculum, software and hardware -- to allow cross-comparisons of the various alternatives within each of the main alternatives' sections. Two additional criterion sets are required to measure those inter-relationships which will need to be controlled between the sections of curriculum versus software, and software versus hardware. It is assumed, that the third possible bi-sectional criterion set which would relate curriculum versus hardware can be based upon the trichotomous inference resulting from the first two bi-sectional comparisons. Finally, a criterion set will be reserved for an overall, tri-sectional evaluation of 'curriculum v. software v. hardware' inter-relationships.

[Generic Criterion Indices]

As with all planning and development activities, the modeler will find the utilization of a 'philosophical' model most helpful in identifying and defining 'types' of criteria which may prove useful in discriminating between the multiple alternatives. This is of greater importance within the CAI-MAM framework due to the complicated relationships both between the three general alternative sections (curriculum, software and hardware) as well as <u>within</u> each of these general sections. Before a general listing and discussion of more specific criterion indices which will be of some benefit to the MICROPIK modeler, a more genus-oriented discussion of criterion-type will be presented.

Three categories of generic criterion indices seem to exist for all problems of evaluation and decision-making when dealing with multiple alternatives:

[1] index of contextual need based upon performance;
[2] index of relative worth or value; and
[3] index of general resource or expenditure.

The <u>index of contextual need based upon performance</u> is itself a relative comparison between the measured states of perceived need, current performance or use, and observed demand. Such contingencies as where demand is greater than need suggests either an unrealistic understanding of the enterprise, or an equally unrealistics understanding of the characteristics of whatever is declared 'in demand'. Of course, a contingency of need greater than demand might also point to a lack of understanding of the context within which the organization exists. Indicators such as might indicate waste (demand greater than performance and/or need) or intervention (need greater than performance) must also be addressed in some fashion as part of the contextual need set of criteria.

The <u>index of relative worth or value</u> is often more easily modeled into an evaluation framework due to its more 'esoteric' issues of: effectiveness, efficiency and satisfaction. To be effective, whatever is performed (or in our case, selected) must "do the job". To be efficient, the selected alternative solution must do the job as quickly as possible, and within the stated operational limits of the enterprise (or less). And to be considered satisfactory, the solution must portray 'good' feelings on the part of all parties involved; or at the least, be consistent in this regard.

The <u>index of general resource or expenditure</u> is a more direct relating of alternatives to those elements of capital, revenue and/or expenditure which might be required in the final implementation of the selected alternative. Such resources as: time, space, facility, personnel, cost, supplies and materials, and equipment -- will all be a potential part of this particular criterion set.

With these ideas in mind, we can now move to a more specific development of sample criterion references for the MICROPIK model.

[Identification and Definition]

To explore the various criterion indicators which will be of use in evaluating the curriculum, software and hardware multiple alternatives associated with the MICROPIK framework, an outline format will be presented for the reader's perusal. This outline will examine each of three major alternative sections first, and then examine potential criteria for performing the aforementioned bi-sectional and tri-sectional comparisons.

Set 1.0 <u>THE CURRICULUM SECTIONAL</u>
(examining relationships both within curriculum objectives, and between the various, multiple instructional activities)

1.01 measures associated with performance, need and demand

1.01.01 perception of school personnel (administrators, teachers, students)

1.01.02 observed time spent on classroom instruction for various topics, group instruction versus individualized or remedial requirements

1.01.03 relative importance of the curriculum unit based upon district level syllabus standards

1.01.04 relative importance of the curriculum unit compared to all other required curricular units

1.02 measures associated with worth or value

1.02.01 perception of effectiveness, efficiency and satisfaction on part of classroom teachers and students, for each curriculum unit

1.02.02 perception of related worth or value of current implementation structure for each curriculum unit

1.02.03 observed measures of effectiveness as relate to learning and retention

1.02.04 observed measures of efficiency as relate to time required for different instructional strategies

1.02.05 related utility of each unit for success in the adul' or occupational world-of-work

1.03 measures associated with general resource or expenditure

1.03.01 related requirements for equipment, supplies or other materials in implementation of each unit

1.03.02 necessary space and/or facility requirements

1.03.03 related personnel staffing needs

1.03.04 measure of relative impact upon other programs based upon resource allocation

1.03.05 related costs and/or expenditures for each unit based upon text books, work books, etc.

Set 2.0 <u>THE PROGRAM SOFTWARE SECTIONAL</u>

(examining relationships between the various software packages available to perform computer-assisted instructional/managerial efforts)

2.01 measures relating to the availability of various CAI/CMI and other administrative software packages, and their comparative utility

2.01.01 compiler languages

2.01.02 word processing

2.01.03 operating system languages

2.01.04 data analysis programs

2.01.05 database management programs

2.01.06 management planning programs

2.01.07 time/project/personnel scheduling programs

2.01.08 accounting software

2.01.09 specific CAI/CMI courseware packages

2.01.10 CAI/CMI author pilots

2.01.11 graphics packages

2.01.12 system editors

2.01.13 information retrieval service (communications multiplexors)

2.02 measures relating to the results of sample field-tests or use by other individuals, concerning effectiveness and efficiency in presentation and drill, and related satisfaction on part of using parties

Set 3.0 THE MACHINE HARDWARE SECTIONAL
(examining relationships between the various hardware packages available for CAI/CMI and other administrative utilization)

3.01 availability of, and relative performance in executing certain desireable functions

3.01.01 mainframes

3.01.02 peripherals

3.01.03 operating system specifications

3.01.04 interface compatibility

3.01.05 networking

3.01.06 expansion

3.02 measures of system specification

3.02.01 clock speed (MHz)

3.02.02 keyboard type

3.02.03 video display
(resolution, character width and line length, line height)

3.02.04 internal central memory

3.02.05 internal expansion

3.02.06 external expansion

3.02.07 internal baud rate

3.02.08 external interface baud rate

To perform the bi-sectional comparison which will relate the curriculum and software sections, and the software and hardware sections, the modeler is concerned with establishing tautological linkages between various parts of each section, based upon the final assessment of the criteria within those sections themselves.

These linkages are of the usual, 'logic-reasoning' specification, and will basically control for the existence of (for example) a particular software package in the final solution set, if and only if: (1) the curriculum sectional presents a favorable criterion picture of the instructional activities involved; (2) the software sectional also presents a criterion-related picture which suggests the package is useful; and (obviously) (3) that such a particular software package exists. Co-relating the software and hardware sections is identical in procedure to that just described for the curriculum and software sectionals.

An additional and somewhat more complex implementation of the constructs supporting bi-sectional comparisons, exists in the utilization of 'slack' variables. Although this treatment is beyond the scope of this particular paper, it will be illustrated for the more experienced reader.

Recall the algebraic relation (inequality or equality) within the criterion vectors as they describe their measures across all of the multiple alternatives. Given that there exists some criterion measure appropriate for evaluating <u>both</u> curriculum sectional units and software sectional units (that is, same reference and same scaling), the measures across first the instructional activities can be summated and stored within a defined slack variable; likewise for a sum across the various software packages. Such a representation would exist as:

Act 1.1	Act 1.2	Act 1.3	Act 2.1	Act 2.2	Act 2.3	Pkg 1.1	Pkg 1.2	Pkg 1.3	Slk 1	Slk 2	RHS
a_{11}	a_{12}	a_{13}	a_{21}	a_{22}	a_{23}				-1		= 0
						P11	P12	P13		-1	= 0
									+1		$\geq b_a$
										+1	$\geq b_p$
									+1	+1	$\geq b_{12}$

and in this example, assumes that the measures represent a score of positive benefit to be maximized (thus the reason for the $\geq$ requirement within the algebraic inequality).

Many other possibilities exist of course for the modeling of criterion references in the comparison of multiple alternatives, but are particular to specific situations; and therefore not easily generalized. Once the reader masters the concept and constructs involved, the adaptation of the method to other settings is (normally?) straight-forward.

[Illustration of Criterion Use]

Before moving on to a discussion of the various referencing, scaling and measurement techniques associated with data generation techniques for the CAI-MAM framework, it may be useful to provide a structured example of how a specific type of data might be collected and input to the MICROPIK model. The 'type' of data for this illustration is called "synthetic", because the source of its values is individual perception -- and not a physcially-rigid measurement of some kind (like for example, weight, height or age).

Synthetic measurement is nevertheless a most valid source of data for the evaluation of multiple alternatives; and therefore for input to the MAA modeling framework. These measures normallv come from one of two sources, and usually must address the issue of 'measurement reliability' as a more subjective, intuitive judgement. The usual source is the survey or opinionnaire, where a respondent's perceptual judgement or opinion is sought concerning certain issues. For example, the respondent might be presented with a declarative statement concerning the issue of priority for microcomputer acquisition for an organization who is currently within a state of fiscal depression. The declarative statement might be formulated as:

> THE SCHOOL SHOULD ASSIGN A HIGHER PRIORITY TO THE ACQUISITION OF MICROCOMPUTERS, THAN TO NEW EQUIPMENT FOR THE PHYSICAL EDUCATION CURRICULUM;

and might ask the survey recipient to respond by choosing a position on the 6-point agreement continuum (where 1= strongly disagree and 6= strongly agree). As an optional procedure, the surveyer could list (for example) ten competing activites which require funding, and ask the respondent to rank-order (1,2,...,10) the activities from most important to least important (of the ten listed). Here, a '1' might represent 'most important', and a '2' represent 'least important' (relative to the ten presented). The <u>important</u> thing for the reader to understand (you might have already guessed)is, that the first option positions a high-value as a more positive response (i.e. positive in benefit to the acquisition of micros), while the second option posits a low-value as the more positive response (1st is best, etc.). The stated importance lies of course in the structuring of the criterion vector containing either the 1-6 or 1-10 values; and additionally in the fact that the decision-maker will discriminate between the high and low values in opposite ways depending upon the option chosen.

The second source of the synthetic measure approximates the first so closely as to beg a differentiated description. This additional synthetic 'type' describes the results of a prior, often physically reliable assessment or measure; and which now requires the 'respondent's' opinion or judgement as to whether the initial physical measure is "good enough", and to what extent. This form of measurement is often the perceptual results of a product field-test in a controlled, environmentally-related setting -- where the product is put under the same conditions as will be expected to exist under normal user conditions upon sale. While physical measures such as time, amount of work done, type of performance, and versatility or flexibility may be the physical measures, the user's perception of utility and acceptability will also prove to be very important criteria for evaluative consideration.

The following criterion references were included in a recent evaluation of microcomputer courseware by the Northwest Regional Educational Laboratory of Portland, Oregon. (For more information, see the periodical "microSIFT News", Vol. 2, No. 1, October 1981). Responses were from a panel of evaluators who tested the software, and then offered their judgement via a 4-point agreement continuum. Although the reader may wish to discuss the varying degrees of non-specificity associated with the 21-items, they remain still illustrative of the means of data generation, and the source of quantitative input to the MICROPIK model.

The "criteria for evaluation" were separated into two categories, content and instructional quality; and were presented as follows:

CONTENT

[01] The content is accurate;
[02] The content has educational value; and

[03] The content is free of race, ethnic, sex, and other stereotypes.

INSTRUCTIONAL QUALITY

[04] The purpose of the package is well-defined;

[05] The package achieves its defined purpose;

[06] Presentation of content is clear and logical;

[07] The level of difficulty is appropriate for the target audience;

[08] Graphics/color/sound are used for appropriate instructional reasons;

[09] Use of the package is motivational;

[10] The package effectively stimulates student creativity;

[11] Feedback on student responses is effectively employed;

[12] The learner controls the rate and sequence of presentation and review;

[13] Instruction is integrated with previous student experience;

[14] Learning is generalizable to an appropriate range of situations;

[15] The user support materials are comprehensive;

[16] The user support materials are effective;

[17] Information displays are effective;

[18] Intended users can easily and independently operate the program;

[19] Teachers can easily employ the package;

[20] The program appropriately uses relevant computer capabilities; and

[21] The program is reliable in normal use.

The reader can easily witness, that the 1,2,3,4 options from the evaluator's assessment could be modeled for inclusion with the software sectional part of the MICROPIK. A criterion constraint would be constructed for each of the 21 items of judgement, and the 'mean-value' responses across all evaluators would be the the entries for each of the vector components; such that:

$$\sum_{k=1}^{K} (x_{ij}) \geq \text{MINIMUM}_i$$

for each of the $i=1,\ldots,21$ criteria; across each of K-possible packages; where x_{ij} is the mean response.

All criteria -- physical, synthetic or otherwise -- will be similarly modeled, and input into the MICROPIK framework.

[Reference and Source]

Having identified and defined the criterion measures which will be utilized within the MICROPIK modeling of the CAI software and hardware decisioning problem, the modeler must next turn attention to determining 'what' will be measured in order to provide a quantified value based upon the construct of each of the variables or criteria defined. In this context, the 'what' of criterion measurement is known as the <u>criterion reference</u> -- that is, what the modeler refers to in order to obtain a valid measure of the criterion point identified. Then of course, the modeler must determine 'where' such a measure will be available and/or from 'who' if other people must be involved. The 'where' and 'who' in this context of criterion measurement is known as the <u>criterion reference source</u> or <u>data-point source</u>. References will always involve a determination of validity of the particular measure, while sources will always necessitate an analysis of reliability. The reader must recognize, that potential non-reliability can related to the people involved, the place or time of the measurement, and the procedure utilized in the measurement process -- that is, the who, where, when and how. The remaining interrogative adverbs of what and why relate more closely with the determined validity of the measured criterion point.

References may be categorized (loosely, I admit) into the three areas of: physical, definitional and synthetic. A physical reference or measure is one in which a fully acceptable tool of measurement is utilized to determine the value or weight of the reference involved. In science, degrees of temperature, miles of distance, and knots of wind velocity are acceptable determinants of their associated references (temperature, distance and wind velocity).

Definitional references are simple or complex transformations of physical measures in order to obtain a new datum to address a defined criterion which can not be measured directly. For example in the determination of school closures, a total of nine definitional criteria were designed and tested for their usefulness in discriminating between elementary school buildings in order to determine their reasonableness for operational discontinuance (Wholeben, 1980a). Three were found to adequately perform this discrimination: thermal efficiency, energy waste, and thermal utility -- by algebraically combining a particular combination of such physical measures as follows:

<u>thermal</u> <u>efficiency</u>: BTU consumption (natural gas, #2 fuel oil, and electricity), capacity and current enrollment of the sites;

<u>energy</u> <u>waste</u>: BTU consumption (natural gas, #2 fuel oil, and electricity),

capacity and current enrollment of the sites, and the total dollar-expenditure for such utilities; and

<u>thermal</u> <u>utility</u>: BTU consumption (natural gas, #2 fuel oil, and electricity) and the total dollar-expenditure for such utilities.

For example, the definitional measure for energy waste resulted from the algebraic representation:

$$\left[\frac{\dfrac{\$\ \text{UTILITY}}{\text{BTU}}}{\text{CAPAC}}\right] - \left[\frac{\dfrac{\$\ \text{UTILITY}}{\text{BTU}}}{\text{ENROL}}\right]$$

We have already dealt with synthetic measures in some detail in the preceding section of this paper. Recall that synthetic measures are normally data points of perception or subjective judgement based upon personal opinion; and thus has all of the reliability problems associated with subjective bias. However, it must be reiterated, that synthetic criterion references are still very much an important 'source' of data for evaluation and decision-making. As is the case in all evaluation, the problem is seldom the intent; but too often the.content and process used in carrying out that intent.

Specific criterion references for quantifying usable MICROPIK data input will generally involve the use of several procedures or tools. Measures related to the curriculum sectional must be demonstrative of not only the content and process of the various instructional activities, but also the relative importance and degree of duplication existing between these activities in the promotion of individual concept learning. Such criteria as the degree of achievement, amount of time required to implement the particular activity, and amount of retention by student will be directly related to the references of performance testing via a number of valid items or problems, clock time, and some form of longitudinal testing utilizing similar problem item, respectively. Criteria related more directly to opinion or perceptual judgement on the part of students and teachers concerning the various instructional activities will be referenced by (for example) some number of statements which describe an opinion concerning some aspect of the activity, and via a survey format gain a measure of 'degree of agreement' by the respondent with respect to the particular individual items.

Gaining responses to the same item (via survey techniques) from two different though related populations is a direct example of how synthetic measures can be transformed in a definitional composite, much as the physical illustrations earlier in this section. Given responses from both students and teachers to an identical item on two different surveys:

> "Learning how to spell using a 'spelling bee'
> is better than using the class workbook."

Obviously, high agreement on the parts of both teachers and students is preferred. However to control not only for degree of agreement to the item, but also for the criterion identified as 'degree of consistency' between teacher and student responses, the following transformation may be utilized to provide a definitional measure of consistency:

$$\text{MINIMIZE} \left| \text{(teacher response mean) - (student response mean)} \right|$$

where this formulation controls for between-groups consistency of response. A similar method for controlling the measure of consistency 'within-groups' is to utilize the standard deviations computed for each of the populations; and formulated as:

$$\text{MINIMIZE} \left| \text{(teacher response standard deviation)} \right|$$

and

$$\text{MINIMIZE} \left| \text{(student response standard deviation)} \right| .$$

Measures of degree of achievement by students using different types of CAI software will be referenced similarly to those ideas expressed above for the instructional activities. Perceptual measures (synthetic) can also be referenced via the administration of valid questionnaires concerning feelings toward the experience of executing the various packages.

Criterion to permit the evaluation of the components of the hardware sectional will normally fall within either physical or synthetic references. Such physical references as clock speed of the CPU (central processing unit) in mega-hertz (MHz) equivalents (i.e. how many millions of cycles per second are performed), and of internal expansion capability in bytes of storage equivalents (a byte being a single character of input as defined by either an alphabetical character, a numeral (single-digit) or a special character (#,%,*)) -- provide readily understandable (?) illustrations. More subjective judgements are also possible concerning the 'esthetics' of the terminal face, or the quality of the printer. A survey format of the 'check-list' variety is a useful tool in gaining such information.

Through our addressing the issue of criterion references (that is, the 'what' of our needed criterion measure), we have paralleled the issue of reference source, or from where (whom) and how such information can be found (or be forthcoming). The data for the curriculum sectional will come from students, teachers and parents -- depending upon the type of criterion being measured. The process may involve the use of observation, a pencil-and-paper questionnaire, standardized achievement test, or a structured interview. Information for the criterion to permit comparable evaluations of the software packages will be measured in a similar fashion. Additional data for the software sectional however can also be gathered via the "dead data" technique of reviewing brochures and records, as well as the more "live data" techniques of observation and survey response.

Much of the information required to quantify the criteria of the hardware sectional will be found via the "dead data" search. Manufacturer's brochures and available technical product reports provide such reference sources. Journal articles may be also helpful; and so also the findings of such periodicals as the 'Consumer's Report'. Whatever reference and source the modeler utilizes for the generation of data points, the cautions concerning reference validity and source (procedureal) reliability must be ever present in the modeler's consciousness.

Except in more complicated MAA models related to the matching of instuctional activities to available CAI software and compatible hardware, the criterion reference for modeling both the curriculum-software and software-hardware sectional will that of 'availability' of the appropriate software package or hardware unit. The source of course will always be the manufacturer and distributor.

[Scaling and Measurement]

Scaling refers to the type of numeric which will represent the measure of the defined criterion reference; and may be one of four types: nominal, ordinal, interval or ratio. The reader is referred to any standard tests and measurement, or introductory statistics text for operational definitions of these scaling types. As a summary however, the types may be distinguished as follows (apologies in advance to those measurement specialists and/or statisticians among the readers):

[1] the nominal scale is pure categorical classification measure of group distinction only; the relationship between groups is one of difference without reference to either direction or extent; examples are sex (male v. female) and minority (minority v. non-minority);

[2] the ordinal scale is one-step-up from the nominal type in that direction or order is now distinguishable for different responses or measures; however, the extent between these directional differences is unknown, and provides a classic potential for interpretative error; examples are assigned ranks and achievement grading as defined by 'excellent v. good v. fair v. poor';

[3] the interval scale is an improvement upon the ordinal type in that both direction and extent (or degree) are now distinctive under interpretation; the intervals between each of the unique measurement points are equal throughout the scale; examples are age expressed in whole years, and off-spring expressed in whole units (normally); and

[4] the ratio scale exhibits all qualities of the interval type, and in addition allows infinite divisions between any two points on the scale's continuum; in fact, the

ratio scale is the only real continuum since it provides of the most finest of possible approximations available; for example speed expressed in cycles-per-second units.

- The measure (of course) is simply the numerical quantity which results from use of the scale in determining the value of the criterion from the selected criterion reference.

The reader should note, that different measures (and often different scalings) can take place with respect to the same criterion reference -- or different references with respect to the same criterion identified. Measures such as these are often the result of a survey of opinion which attempts to gain insightful data concerning various issues of interest or aspects of current endeavor.

The MICROPIK model will accomodate any of the scaling types dependent upon the intent of measure (indentified and defined) being sought. Availability of certain software and hardware units will often be identified as a '1' (availability = yes) or a '0' (availability = no); and therefore uses a nominal scaling type for final measurement. Presenting a group of respondents with a list of instructional activities concerning the satisfaction of a specific curricular objective, and asking them to rank-order their importance in promoting the learning defined by that objective, results in the ordinally-scale measure of ranks (1 = most important, 2 = next most important, ...). The interval type of scaling is assumed with such extended continuum frameworks as the 6-point agreement continuum. And finally, the ratio scale is most usable with the more physical measurements associated with system specifications, cost of various software and/or hardware units versus the salaries of additional classroom teachers, and achievement performance measures on the part of the students.

Valid criterion definitions and references, and reliable sources and measures, are of course not very useful if there exists no technique for entry in the multiple alternatives analysis model. Before discussing the formatting of measured criterion data points in such a way, that the MICROPIK model will be able to evaluate the various sectional options associated with choosing software packages and hardware units compatible with desired CAI/CMI objectives, it may be best to once again review the 'guts' of the MAM framework, and the model's criterion-referenced, decisioning-simulation needs.

Recall the design of the MAM framework as that of a matrix, where rows represent criterion measures across the various options or decisional alternatives, and columns represent the array of criterion measures for each of these decisional alternatives. We will be concerned by the 'row point-of-view', and address each row as the criterion vector of values or simply (?), the criterion constraint. Since each criterion vector (i.e. row) represents the values of a specific criterion across all alternatives, the reader can easily understand how these values will be capable of validly evaluating the various alternatives (against themselves). And, since each criterion constraint can be said to therefore constrain the solution process (i.e. arrive at a decision), each criterion vector can be thought of as an 'objective' or 'goal' of the modeling situation, in that certain limits will be placed upon the values which each criterion vector can assume (as a composite summation) before finally deciding upon a final, most optimal solution set.

Each criterion vector will be constructed to represent either a linear equality or inequality (although the inequality is often the more useful representation); and will therefore assume the general form of:

$$a_{i1} + a_{i2} + a_{i3} + \ldots a_{ij} \; (\leq , = , \geq) \; b_i$$

where i-criterion vectors have been constructed to evaluate the relative appropriateness of j-alternatives, and based upon a RHS-limit to the final composite (i.e. sum) of the particular i-th criterion vector of the value b_i. Note that b_i therefore will exist as an upper-bound in the '$\leq$' inequality, a lower-bound in the '$\geq$' inequality, and an "identity' via an '=' equality.

Thus, each a_{mn}, for m=1,2,...,i criterion vectors across each n=1,2,...,j decisional alternatives, will represent a particular, consistent scaling of value for each of the i-criterion vectors. And, since each b_k, for k=1,2,...,i RHS-values, delimits the total (summed) composite which each criterion vector can assume dependent upon the solution set formulated (x_n equaling either a '1' or a '0' depending upon the x_n's inclusion or exclusion for the final solution set), the particular scale utilized will determine the <u>type of objective</u> which the particular vector is attempting to satisfy.

For the time being, let us set our total confusion aside, and attempt to examine each scaling type via the criterion constraint framework explained (?) above. For the reminder of this particular discussion, we will adopt that convention that a value of '1' for the x_n decisional alternative will denote <u>inclusion</u> within the final solution set; and that a value of '0' will represent <u>exclusion</u> of that particular x_n option from the final solution.

The <u>nominally</u>-scaled criterion constraint vector can also be called the <u>frequency</u>-constraint or <u>counting</u>-constraint vector, due to its use in controlling for the various frequency of a particular type of category within the final solution. One particular type of nominal control is that of assuring representative-bias -- that is, assuring the inclusion of certain amounts of specific types of alternatives within the final solution set. To illustrate this, consider a MICROPIK problem which has defined five software package alternatives within the software sectional, and denotes the first two as basically 'grammerically oriented' and the remaining three as 'vocabular oriented' in terms of a proposed language arts CAI curriculum. And futher assume that constraints are required in order to model the following three, separate objectives:

[1] exactly one of the grammar packages must be a member of the final solution set;

[2] not more than two of the vocabulary packages are allowed inclusion within the solution set; and

[3] at least three software packages must construct the final solution set, overall.

The resulting sub-matrix of the full constraint matrix (collection of all criterion vectors) would appear as follows:

(Objectives)	x_1	x_2	x_3	x_4	x_5		(RHS-value)
[01]	1	1				(=)	1
[02]			1	1	1	($\leq$)	2
[03]	1	1	1	1	1	($\geq$)	3

(can you see that the final solution set must contain exactly <u>3</u> entries? and that only a total of <u>3</u> possible, feasible solutions exist? and why additional data would be needed in order to determine the final solution?). This example emphasizes the utility and necessity of the <u>objective function</u> in resolving which of the three potential solution sets will in fact be the most optimal set.

The modeling characteristics of ordinally-scaled criterion constraint is an extension of the nominally-scaled constraint vector. Since the terms "mean order" and "sum of order" are examples of the premise, "You can do any thing with numbers, meaningful or not.", ordinal constraints are modeled within MICROPIK as a type of indicator-variable as would be found in the modeling of dummy variables within multivariate regression procedures. For each of the desired 'ordering points' (e.g. ranks; or those points which would be associated with 'excellent-good-fair-poor' responses), a separate criterion constraint vector must be developed in order to control the inclusion of various 'ordered' alternatives within the final solution set. Consider the MICROPIK curriculum-sectional in which two sets of four instructional-activity alternatives are to be modeled. Each set of four alternatives has been ranked by a panel of experts as to their relative importance to the successful implementation of curricular goals, assigning '1' to the most important, and '4' to the least important of the four such that the following assignments result:

OBJECTIVE	ACT-1	ACT-2	ACT-3	ACT-4
1	3	1	4	2
2	2	1	3	4

and must be modeled consistent with the following stated objectives:

[1] each objective must be satisfied;

[2] at least two activities from each objective set must be members of the final solution set;

[3] at least two of the final solution activities must be of rank=1;

[4] only one activity of rank=3 is allowed within the final solution; and

[5] no activities of rank=4 are to be included as final solution components.

The final modeling framework for these five objectives will include a maximum of seven constraints, but could be identically constructed with five constraints (can you see the duplication?):

The modeling characteristics of ordinally-scaled criterion constraint is an extension of the nominally-scaled constraint vector. Since the terms "mean order" and "sum of order" are examples of the premise, "You can do any thing with numbers, meaningful or not.", ordinal constraints are modeled within MICROPIK as a type of indicator-variable as would be found in the modeling of dummy variables within multivariate regression procedures. For each of the desired 'ordering points' (e.g. ranks; or those points which would be associated with 'excellent-good-fair-poor' responses), a separate criterion constraint vector must be developed in order to control the inclusion of various 'ordered' alternatives within the final solution set. Consider the MICROPIK curriculum-sectional in which two sets of four instructional-activity alternatives are to be modeled. Each set of four alternatives has been ranked by a panel of experts as to their relative importance to the successful implementation of curricular goals, assigning '1' to the most important, and '4' to the least important of the four such that the following assignments result:

OBJECTIVE	ACT-1	ACT-2	ACT-3	ACT-4
1	3	1	4	2
2	2	1	3	4

and must be modeled consistent with the following stated objectives:

[1] each objective must be satisfied;

[2] at least two activities from each objective set must be members of the final solution set;

[3] at least two of the final solution activities must be of rank=1;

[4] only one activity of rank=3 is allowed within the final solution; and

[5] no activities of rank=4 are to be included as final solution components.

The final modeling framework for these five objectives will include a maximum of seven constraints, but could be identically constructed with five constraints (can you see the duplication?):

(Objectives)	x_{11}	x_{12}	x_{13}	x_{14}	x_{21}	x_{22}	x_{23}	x_{24}		(RHS)
[01.1]	1	1	1	1					$(\geq)$	1
[01.2]					1	1	1	1	$(\geq)$	1
[02.1]	1	1	1	1					$(\geq)$	2
[02.2]					1	1	1	1	$(\geq)$	2
[03.0]		1				1			$(\geq)$	2
[04.0]	1						1		$(\leq)$	1
[05.0]			1					1	$(=)$	0

While intervally-scaled constraints can be modeled similarly to the ordinal type, careful preparation of the interval-based response continuum will often yield measures closely related to those of the ratio-variety, and thus permit ratio-type construction. For this reason, the following presentation will relate to both occurrences of interval and ratio measurement scaling of the criterion constraint vectors.

Unlike the previous discussion, ratio-scaled constraint entries are the actual criterion measure resulting from the data point on the criterion referenced identified. For example, in the case of a physical measure related to clock time (measured in MHz of cycles per second), a hardware sectional of five package alternatives would contain a constraint whose a_{ij} entries for the particular constraint vector would be the actual, recorded MHz quantity from system specifications. As an illustration, assume these five hardware package alternatives have been evaluated on two separate criteria, the first on clock time, and the second on the mean response obtained from field-test users who responded to the item:

"Response time for the unit was satisfactory."

utilizing a 6-point agreement continuum scale which itself assumes ratio-qualities. The tabular results of these measures were as followed:

CRITERION	Unit-1	Unit-2	Unit-3	Unit-4	Unit-5
"clock"	1.2	0.4	1.7	0.9	0.1
"response"	3.5	2.4	1.6	3.7	4.7

and will require modeling as follows:

[1] no more than two units must be selected as solutions;

[2] the total sum of 'clock time' within the final choice of units for solution must not exceed the value b_c; and

[3] the total sum of 'response satisfaction' within the final choice of units must be at least the value b_r.

This sub-matrix related to the hardware sectional will be modeled as follows:

(Objectives)	x_1	x_2	x_3	x_4	x_5		(RHS)
[01]	1	1	1	1	1	($\leq$)	2
[02]	1.2	.4	1.7	.9	.1	($\leq$)	b_c
[03]	3.5	2.4	1.6	3.7	4.7	($\geq$)	b_r

and once again illustrates the utility of the objective function which will be explored in a future section.

We will now deal more specifically with the development of the RHS-values especially needed for the successful computation of the exampled b_c and b_r used above; and of the various methods for controlling desired system impact.

THE CONSTRAINTS

It was necessary within the previous section concerning criterion definition, referencing, scaling and formatting to illustrate the utility and credibility of the criterion-input to the MICROPIK model by structuring 'criterion constraint' examples. For the more experienced reader, it may now seem redundant and after-the-fact to commence a formal presentation on the ideas, structure and utility surrounding the utilization of such a vector within a mathematical modeling framework.

As we have already witnessed, the constraint vector is one of two algebraic types: either inequality or equality. This algebraic format serves to input specific criterion values of a defined criterion reference across the available alternatives into the model; and further utilizes the algebraic relational (i.e. $\leq$, =, or $\geq$) as the control over the final alternatives' selection (solution) procedure. In this section, we will examine in greater detail how this control works; and how the modeler can vary such control in order to structure a most versatile and flexible alternatives evaluation setting.

[Direction and Valence]

The reader will recall, that the numerical values associated with each particular criterion reference are input to the MAM framework as coefficients of a linear inequality or equality, in the vector form:

$$[\ a_{i1}\ \ a_{i2}\ \ a_{i3}\ \ \ldots\ \ a_{ij}]$$

where the i-th criterion (model objective or decision constraint) has distributed specific values across j-alternatives. In full algebraic linear form, the vector of coefficients represent a series of operands of either positive or negative values due to the actual numerical coefficient (e.g. a $+a_{ik}$ versus a $-a_{ik}$ for some k-alternative) whose linking operator is always the arithmetic operation of addition. For example:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j$$

where each of the x_j independent variables represent the various multiple alternatives being evaluated for selection or inclusion within the final solution set. In using arithmetic addition to form a composite of the a_{ij} values whose related x_j's take on the value of '1' (i.e. inclusion within the solution set), we assume the coefficients to be additive, and thus representative of some summed effect of the particular criterion reference being modeled.

We have seen, that the coefficients will assume different modeling roles dependent not only upon the reference being modeled, but also upon the type of scaling which was utilized for quantifying the criterion-referenced measurement itself (i.e. nominal, ordinal, interval or ratio). In addition, the modeler must also determine exactly what effect the sum of each of the criterion vectors will represent for the problem being constructed. That is, will a larger sum of coefficients (viz., of higher value) be seen as more positive (benefit) or negative (undesireable). For example, if a survey item which seeks high agreement from respondents on the effect of each of several CAI/CMI packages upon student learning is to be input to the MICROPIK model, and the 6-point agreement continuum (where 6= strongly agree) was the response format used for data collection -- then the various software packages which will finally form the solution set should be such that they display "higher" agreement value than their evaluated companions. In this example, the sum of those criterion vector coefficients which modify the solution software alternatives will take on a larger value, since the coefficients themselves should be of higher 'agreement' weight. As we will soon see, such a criterion vector constraint will be called a 'maximizing' constraint, since the maximum sum of coefficient-values possible is desired.

To examine a different type of vector constraint, consider that criterion constraint whose coefficients represent the purported unit cost for each of the hardware packages being evaluated. Our goal of course, is to maximize all positive aspects of the packages possible while minimizing the expenditure required to obtain these same packages. In this case, the final sum of the cost coefficients would be preferably a small as possible without compromising quality and utility of the various alternatives included within the final solution; and so, the 'smaller' the sum, the better. Such a criterion vector constraint will be call a 'minimizing' constraint. And as we will soon see, a third type of constraint, the 'identity' constraint, will also be useful when exact-value sums are required from the modeling of the particular vector constraint.

[The 'Maximizing' Vector Constraint]

The vector which seeks a higher-valued sum of the available evaluative coefficients modifying the potential solution alternatives is known as a maximizing-vector or maximization constraint. It is assumed, that the coefficients within the vector represent a desireable, positive influence upon the decisioning process; and that (therefore) the higher the coefficient value of any particular alternative being evaluated, the more likely that same alternative will be selected as a member of the final solution set.

To assure this desireable event, the algebraic inequality relational 'greater than or equal to' ($\geq$) is utilized to construct the criterion constraint, such that:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j \geq b_i$$

is the resulting inequality member of the MAM modeling framework, where the value b_i is considered a lower-bound of the modeling constraint summation. That is, b_i is that quantity which must be matched or surpassed by the summation of coefficients, in order for the particular x_k alternative solutions to be members of the final solution set. Until some combination of x_k's from the available x_j-alternatives can be found which will produce a sum greater than or equal to the listed b_i value, the modeling framework is considered not solved; and if the combination can not be found, the problem setting is considered infeasible -- no solution is possible within the constrained decisioning setting as designed.

[The 'Minimizing' Vector Constraint]

The vector which seeks a lower-valued sum of the available evaluative coefficients modifying the potential soltuion alternatives is known as a minimizing-vector or minimization constraint. It is assumed, that the coefficients within the vector represent an undesireable, negative influence upon the decisioning process; and that the lower the coefficient value of any particular alternative being evaluated, the more likely that same alternative will be selected as a member of the final solution set (assuming of course, that a low value correspondingly means low negative impact).

To minimize as much as possible the undesireable aspects of this particular criterion upon the final solution, the algebraic inequality relational 'less than or equal to' ($\leq$) is utilized to construct the criterion constraint, such that:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots a_{ij}x_j \leq b_i$$

is the resulting inequality member of the modeling framework, where the value b_i now represents an upper-bound of the modeling constraint summation. That is, b_i is that the highest value which the vector sum is allowed to assume -- and therefore allowing the sum to take on as low a value as possible in its formation of the final solution set. As with the maximizing vector, if such a minimum standard can not be satisfied by the summation across this particular vector, the problem is declared infeasible.

[The 'Identity' Vector Constraint]

The third and final type of constraint which may be utilized within any MAA modeling setting is the identity-constraint. This vector is constructed as an algebraic equality, in the form:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j = b_i$$

where now the specified b_i quantity is neither (or both if you want to be cantankerous) an upper or lower bound on the possible sum of the coefficients, but rather the exact quantity which that same sume must achieve for admittance of the modified alternatives into the final solution set. As we witness in a previous section, the identity constraint is very useful in controlling for the modeling of nominally-scaled criterion variables, and/or for the criterion vectors which rep esent the dummy (indicator) vectors of a previous ordinally-scale criterion reference. In addition, the identity-constraint is best suited for controlling for those stringent standards which impact upon the decisioning process, as might be required by affirmative action regulations, or the imposition of stratified-group comparisons.

[System Impact Control Via RHS-Bounds]

Control for the construction of the final solution set is based upon the criterion coefficients which modify the multiple alternatives being evaluated for inclusion within that solution; and the value of the specific bound placed upon the linear inequalities or equalities being modeled within the MAM framework. As the criterion coefficients which modify the solution alternatives are summed for the combination of alternatives comprising the solution set (where $x_k = 1$), this arithmetic sum is compared to the b_i value (RHS-bound) to assure compatibility with the desired impact sought (i.e. $\leq$, $=$, or $\geq$). When a particular set of alternatives can be found, such that:

[1] those 'maximizing' criterion vector coefficients modifying the members of the solution set produce sums which for each such criterion constraint, are greater-than-or-equal-to the established RHS-value(s);

[2] those 'minimizing' criterion vector coefficients modifying the members of that same solution set produce sums which for each such criterion con-

straint, are less-than-or-equal-to the established RHS-value(s); and

[3] those 'identity' criterion vector coefficients modifying those same solution alternative members produce sums which exactly display the values of their associated RHS-bounds;

then a solution exists which satisfies the established constraints placed upon the decisioning process as identified via the various criterion inequalities and equalities. Such a solution is known as 'feasible', and may or may not be the optimal (i.e. best) solution possible based upon the constraint matrix and RHS-vector. The determination of optimality is a function of an additional vector of values, known as the objective function -- which will be discussed later in this section. First however, we shall examine in more detail this issue of controlled impact and the RHS-vector.

The values of the RHS-vector are of course those bounds which when placed upon the sum of the coefficients of the various criterion constraint vectors control the selection of potential solution alternatives via upper or lower bounds, or identities. Simply stated, an upper-bound represents the highest value which is acceptable based upon the sum of 'solution' coefficients; and therefore most often represents a control for undesireable or negative effect as defined by the particular criterion vector. Similarly, a lower-bound represents the lowest value which is acceptable based upon the sum of these same 'solution' alternatives' coefficients; therefore most often represents a control for desireable of positive impact as defined by the particular criterion vector.

Such control based upon criterion vector coefficients sums is a form of <u>generalized system impact</u> control, in that (with the exception of the identity) the only requirement is to meet the upper and lower bound restrictions placed upon the inequalities. Because the restrictions are based upon the composite values of a summation, it is likely that the interactive-effects relationship between criterion values and solution alternatives will produce a solution set where some members may display 'less than acceptable' criterion weights on one or more criterion references. Such a circumstance should come as no surprise to the reader, as a particular alternative's strength on several other criterion vectors may outweigh its associated weakness on a single measure. Since the vector sum will not distinguish its individual members (coefficients), this particular method of control is known as identifying impact to the system as a whole.

The reader may also need to be reminded at this point, that seldom do decisioning situations present such simplistic settings as will be remediated by solutions which are clearly full-positive in scope -- that is, have no negative by-products or effects associated with them. Complex situations will always require the conscious knowledge of both the positive and negative impacts associated with the solution(s). The decision-maker must be ready to establish the required preferences in order to perform the necessary 'secondary choices' which will be required when alternative decisions present both positive as well as negative aspects to the system; and then be prepared to acknowledge those trade-offs which are associated with the solution's related negative effects.

Specific system impact (as opposed to general) is capable of being modeled within the MAM setting, via such techniques as: selective sub-vector summations (controlling for marginal values of particular multiple alternatives), and individual single-independent-variable inequality (constraint) construction where j-inequalities would be required for modeling each of the j-alternatives for a particular criterion reference. In most cases, the modeler will be able to a priori detect if a particular alternative has a criterion measure which makes it undesireable as a solution (regardless of its other measures), and therefore can be excluded from the MAM procedure completely.

Generalized system impact (which is the preferred procedure) can itself be modified or varied in order to study the changes in the selection of potential solution alternatives. Such a technique is known as the restriction or relaxation of the RHS-values in their constraint of the decisioning process.

The restriction associated with the control of the RHS-vector over the selection of solutional alternatives is basically a procedure of placing more difficult demands upon the constraint vectors in their formulation of a final solution set. For the maximization vector, this will normally mean an increase of the lower-bound which the final coefficient sum must meet or surpass. For the minimization vector, a more restrictive environment will mean a decrease in the upper-bound which the coefficient sum must satisfy. Restriction of the RHS-values is usually executed in order to detect at what level of individual constraint control will the same solution set be constructed regardless of the reference of the objective function.

On the other hand, the relaxation of the individual RHS-values places less demand upon the constraint summations as they measure the generalized impact of particular solution sets to the sytem as a whole. For the maximization vector, a relaxed state is usually associated with a decrease in the value of the particular lower bound -- thus making the attainance of a sum more easily accomplished (and therefore more accepting of less positive impact by some alternatives). Similarly, the RHS-value related to a minimization vector will be increased in a state of relative relaxation -- allowing more negative impact to be acceptable to the final solution set. Relaxation of the RHS-values is usually executed in order to generate a diverse array of solution alternative vectors dependent upon the respective influence of different objective functions.

[The Objective Function]

In additional to the criterion row vectors we have already examined as they relate to the modeling of multiple alternatives for the multiple alternatives analysis setting, another vector is necessary to force the formation of a solution set which is 'optimal' as defined by some a priori standard. Unlike the vectors of the constraint matrix, this new vector does not have an algebraic equivalent in the sense of an inequality or equality. Called the objective function, this vector provides the basis for constructing a solution set vector which not only is deemed acceptable to the criterion vector constraints of the constraint matrix, but which optimizes (maximizes or minimizes) the value of an additional vector of values or standards.

Thus, while the various constraint inequalities and equalities evaluate the multiple alternatives for the existence of a feasible solution (i.e. whether any solution is possible), the objective function vector chooses which of those alternative solution sets best (most optimally) addresses a particular issue. For example, the objective function may strive to prepare a solution within the constraints of the problem, such that: the satisfaction of the students involved as measured by their attitude is maximized; or, the additional expenditures which would be required to purchase additional equipment is minimized.

Choice of the objective function is itself a function of the overall objective(s) of the system comprising the problem area. Some modeling strategies will incorporate only a single objective function in the execution of the decision model; and others may use several in order to examine the impact upon the construction of the solution set. As we will see in the next section, the preferred technique is to utilize each of the constraint vectors serially as the objective function; and to record the differential impact to the formulation of the solution set associated with each vector's ultimate guide of the decisional process.

[Construction of the RHS-Bounds]

The composition of the RHS-value will depend simultaneously upon the intent of the criterion constraint it modifies, and the type of scaling utilized in designing the criterion constraint's coefficients. We will examine each of the types of RHS-bounds by its association with scaling types. This discussion will apply to both maximization and minimization vectors (as well as in most cases, the identity vector).

Both nominally-scaled and ordinally-scaled constraints will normally be represented as 'frequency' or 'counting' coefficients, and will therefore require a RHS-value which controls for the total frequency associated with a particular criterion within the final solution set. As was discussed previously within the criterion section, potential solution alternatives can often be criterion-addressed via measurement scales which indicate distinct type or membership, rather than a arithmetically computable value of both direction and degree.

For example, consider the situation wherein the construction of the MICROPIK model requires crossreferencing of various software packages with compatible hardware units, for utilization within the implementation of CAI/CMI strategies. Five software packages are being evaluated which present instructional activities related to the mildly-handicapped, in the area of reading comprehension. Two of the packages can be implemented on one of the hardware systems available, while the remaining three software packages are compatible only to another hardware unit (which must be purchased if chosen), The problem has been designed to include the already on-line system with the evaluation of the not-yet-purchased system, in order to fairly compare the attributes of each system in relation to the potential software purchases.

For illustration, the software-hardware cross-references will exist as follows:

HARDWARE UNITS	SOFTWARE PACKAGES
A	1,2
B	3,4,5

and where an additional constraint of 'only a single hardware unit' must result as the preferred solution in terms of the hardware sectional itself.

Utilizing tautological constraint vectors as developed for the modeling of internal constraint logic subcategories for contingency allowance under specific inclusion (Wholeben, 1980a), that is:

"If A 'OR' B, Then C (possible)."

the resulting MICROPIK framework would exist as follows:

S-1	S-2	S-3	S-4	S-5	H-1	H-2	S_a	S_b	.	RHS
1	1						-1		=	0
					-2		-1		≤	0
		1	1	1				-1	=	0
						-3		-1	≤	0
					1	1			=	1

To model the situation above, it is necessary to utilize slack variables as temporary storage locations to denote whether any of the evaluated software packages were chosen by the model as acceptable to the curriculum instructional activities within the curriculum sectional (not shown). These slack or temporary storage variables are denoted above as S_a and S_b; and will denote the selection of any of either the software 1,2 or software 3,4,5 packages, respectively. It is acknowledged, that discussion of the use of slacks (and indeed,tautologicals) is beyond the scope of this present paper. However, the reader should be somewhat aware of the potential for such manipulations of nominally-scaled criterion entries; and be able to at least rudimentally understand their utility. The third constraint subset, which relates the constraint of 'one, and only one' hardware unit is to be a member of the solution to the hardware sectional, is a more direct and easily verifiable use of the nominal-scale.

In consideration of both interval and ratio measurement scales as providing the basis for the arithmetic operations of multiplication and division -- not acceptable to the nominal or ordinal measure -- the construction of RHS values assumes a completely different perspective and rationale. Cognizant of the desire to control for 'general system impact' as opposed to specific alternatives values (allowing the model to generate internal preference and trade-off decisions), the development of RHS-values will now follow the generalized goal:

> ‹to design, formulate and quantify specific b_i component entries of the RHS-vector for each modeled i-th criterion reference; such that the individual b_i values establish bounds which the algebraic inequality or equality relational of the criterion vector must seek to satisfy; and where these individual b_i values denote 'general system impact' as that measure which is defined as the sum of the individual criterion measures across the potential solution alternatives being evaluated›.

If we equate 'general impact' with the more arithmetic term of 'mean impact', then the goal becomes controlling the evaluation and final decision-making (selection of alternatives for membership within the solution set) via the structuring of some 'mean value' for controlling the summation of criterion values across the various potential alternatives. In general, one might think of this goal as follows:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j \quad (\leq, =, \geq) \quad k(MN_A)$$

where:

MN_A represents the mean of all a_{ij} summed across all potential x_j solutional alternatives; and

k is some constant factor (multiplier) of the mean(a_{ij}) RHS-vector entry.

If in fact, our goal is to model the selection of alternative solution via the control of their 'mean impact' to the system as a whole, then the chosen mean term must equate the role of 'mean impact' to a specific numerical quantity. The value of MN_A does not satisfy this need alone, since the sum across component entries will often result in a quantity greater than their computed mean value. However, if the modeler could identify some constant number which would approximate the perceived number of alternatives which would in turn reside in the solution set, then the use of the term $k(MN_A)$ would itself approximate the average impact to the system of a select k-number of solutional alternatives.

Accepting the above as a useful methodology for developing numerical quantities for describing mean system impact, a new problem arises. If in fact the MN_A value will denote average impact, than a 'high outlier' of the modeled criterion distribution (a_{ij}) could be as large as two or three times the size of that same distribution's mean (MN_A). Thus the use of the term MN_A alone would also bias the quantity of alternatives chosen for the solution set, since one alternative with a 2-times the mean value weight for its specific criterion value entry would add a double-factor to the final criterion constraint sum of that particular criterion vector across the selection solution alternatives. In addition, the MAM framework seeks to model average impact, which assumes preferences and trade-offs existing. The computer value MN_A is an absolute quantity, with no such flexibility inherent within the structure of the arithmetic summation.

To resolve this dilemma, the use of the computed standard deviation for the specific criterion distribution is warranted. Identified as SD_A, the addition or subtraction of the standard deviation to (or from) the mean of the distribution -- that is, $MN_A + SD_A$ or $MN_A - SD_A$ -- provides a readibly usable technique for numerically modeling the concept of mean system impact as references each particular criterion. It remains now to address the two situations which warrants the use of addition or alternatively, the use of subtraction in developing the RHS-value.

Addition of the SD_A term to the criterion vector computed MN_A term is required for the existence of the interval or ratio scaled minimization ($\leq$) constraint, where the RHS-component represents an upper-bound; that is:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j \leq k(MN_A + SD_A).$$

Alternatively therefore, subtraction of the SD_A term from the criterion vector computed MN_A term is necessary when using the interval or ratio scaled maximization ($\geq$) constraint, where the RHS-component represents a lower-bound; that is:

$$a_{i1}x_i + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j \geq k(MN_A - SD_A).$$

Recalling that the constant k represents the expected number of decisional alternatives which will be finally selected as members of the solution set, the multiplication of either the ($MN_A + SD_A$) or ($MN_A - SD_A$) terms by k represents the 'mean impact' to be entered into the RHS-vector for controlling the objective of that particular criterion vector constraint.

The reader may now ask how such an approximation technique could ever be useful for modeling the algebraic relational (=) of the identity constraint, since the potential of relating some specific sum to a computed flexible mean is remote. To actually model the identity relational, the decision-maker uses a matched pair of maximization and minimization constraints; and thereby attempts to double-bound the specific criterion vector's sum.

Construction of the RHS-values for modeling identity constraints will obviously depend upon the specific criterion being referenced, but will nonetheless approximate the following paradigm:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j \leq k(MN_A + \tfrac{1}{2}SD_A)$$

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{ij}x_j \geq k(MN_A - \tfrac{1}{2}SD_A)$$

where the use of the '½' factor relating to the standard deviation term is arbitrary. In general, as the particular criterion reference being modeler requires more or less convergence upon the identity of the RHS-value, the model builder will choose to use lesser or greater fractional parts of the SD_A term, respectively.

[Weighting Via Co-Related Vectors]

We have discussed previously how individual criterion constraints could be more or less influential upon the selection processes of the MAM framework through the restriction or relaxation of the constraint's RHS-value. Having discussed the computation of the RHS-values in the preceding topic, the reader should now be able to visualize the RHS-values; and therefore how the increase of a particular RHS component would restrict the maximization constraint while relaxing the minimization vector's process. Similarly, the decrease in a particular RHS component would thun relax the maximization constraint while restricting the process of the minimization vector. The author cautions the reader however, to employ such varied and most useful techniques only after attaining initial integer feasibility (i.e. assuring that at least one solution exists as the problem is currently constructed).

Another technique for weighting differential effects upon the final solution set membership's contribution to the measured general system impact, exists in the use of co-related vectors. This procedure requires a form of stratification of the available decisional alternatives into groups of relative impact, based upon the values for the individual criterion constraint being referenced. The general idea is to select separate alternatives as being more (or less) desireable for inclusion within the final solution set, based upon their individual criterion values. Of course, an alternative may be differentially 'desireable' due to relatively positive values on one or more criterion references, while containly correspondingly negative values on other vectors. Since this is almost always true, the construction of the co-related vector(s) for modeling weight will often require different co-related vector(s) across different constraints for the same alternative.

As an illustration, consider the problem where ten alternatives are being evaluated for determining the final solution to which alternative instructional activities will be implemented to satisfy curricular objective 'O'. A panel of expert teachers have reviewed the activities, and certified each to be of value sufficient to warrant their inclusion within the multiple alternatives modeling framework. This panel has also stated, that depending upon the criterion reference involved some alternatives are not only of more positive value but also should somehow be weighted for greater potential entry into the final solution set. To understand their position, the panel has identified three separate groups of preference (high, moderate and low) for the ten alternatives; and has for two specific criterion references segregated these ten alternatives into one of the three classes of preference as follows:

CRITERION	0-1	0-2	0-3	0-4	0-5	0-6	0-7	0-8	0-9	0-10
-A-	1	2	2	3	3	2	1	1	1	3
-B-	2	3	3	3	1	1	2	2	3	1

where a matrix-entry of '1' represents high preference, while an entry of '3' relates to correspondingly low preference. The panel also assumes that at least 6 instructional activities will be required, and prefer at least 4 of these activities be of preference factor 1 or 2 on at least one of the criterion vectors, and at least 2 of these 4 be factor 1 or 2 on both.

To illustrate the constraint matrix design, we will again call upon the use of slack variables as we did in the previous section, utilizing them as temporary storage locations for within-matrix summations. This particular example will require two of these slacks -- one for the preference indicators associated with criterion A, and the other associated with criterion B. The modeling design would then exist as follows:

CO-RELATE	01	02	03	04	05	06	07	08	09	010	S_1	S_2		RHS
-A-	1	1	1			1	1	1	1		-1		=	0
-B-	1				1	1	1	1		1		-1	=	0
"at least 4"											1	1	$\geq$	4
"at least 2"	1					1	1	1					$\geq$	2
"at least 6"	1	1	1	1	1	1	1	1	1	1			$\geq$	6

(What would have happened had an additional objective been required which stated "at least 4 of these activities be of preference factor 1 or 2 on both of the criterion constraints"?).

THE EXECUTION

Now that the areas of theoretical intent, conceptual design, and technical development have been discussed at some length and with illustrations, it is time to explore the execution or implementation phase of MICROPIK modeling -- how to gain the required results of the model formulated, and what to do with those data elements once collected. It is beyond the scope of this paper to discuss in detail the mathematical software programs which facilitate the evaluation of the MAM framework. The more discerning reader is directed to be vigilant for an upcoming manuscript publication of the author entitled, "Multiple Alternatives Analysis for Educational Evaluation and Decision-Making," -- scheduled for release in late 1982 or early 1983.

This section will deal with the major four facets of the execution phase: cyclic optimization, the development of the solution tracking matrix, the creation of the various types of solution vectors, and criterion reference weighting techniques based upon the various iterations of the cyclical objective function. The individual post-hoc analyses (statistical or otherwise) which are recommended for the results of the MAM execution, will be examined in a succeeding section entitled 'Results'.

[Cyclic Optimization]

Although the multiple alternatives modeling framework requires only a single objective function for implementation of a related multiple alternatives analysis, the suggest, preferred execution technique is to employ a cyclical optimization procedure, wherein each of the criterion vectors utilized within the constraint matrix portion of the MAM is cycled through the model sequencially as the objective function. In other words, given a problem of one hundred multiple alternatives modeled across twenty criterion constraints, the constructed model would be executed a total of twenty-times, once for each of the criterion constraints, where the objective function would be composed of those a_{ij} values also existent within the particular i-th constraint.

The utility of cyclic optimization can best be witnessed in the statement of its goal:

> to generate a separate set of solution members based upon each individual criterion reference modeled within the full model, such that the selection of these members is based upon the same set of criterion constraints as modeled via criterion vectors and RHS-values, but where the objective function is varied according to the reference of the individual criterion vector entries.

For the above example therefore, a total of twenty solution sets would results, where the variability of membership would depend totally upon the utilized maximization or minimization of the particular criterion vector acting as the objective function for that execution.

Each criterion vector would of course be either maximized or minimized as relative to it respective positive or negative emphasis regarding the criterion values of its vector components. That is, the objective will always (or at least should) be to generate a solution which maximizes the positive or minimizes the negative characteristics of the associated criterion vector. There will moreover be occasions when selected criterion vectors will be both maximized and minimized (on separate runs) during the stage of cyclic optimization (see Wholeben and Sullivan, 1981).

The implication of cyclic optimality techniques within the setting of the MICROPIK model for selecting appropriate software and hardware packages in accordance to desired CAI/CMI-related instructional applications, illustrates a special case for the application of a cycling-executable procedure.

Recall the structure of the constraint matrix for the full MICROPIK model, composed of criterion references for each of the required five sectionals: curriculum, software, hardware, curriculum-software and software-hardware. Since each sectional is concerned with a sub-matrix portion of the full constraint matrix, a series of zero-submatrices or empty submatrices result. That is, when concerned with the curriculum sectional alone, the related row portions of the software and hardware alternatives' columns will be devoid of any data entry; and thus, 'empty'. Likewise for consideration of the software-hardware sectional, the associated rows of the curriculum alternatives' columns will be empty -- and therefore by convention, contain zeroes for each of the matrix cells within that particular submatrix portion. Imagine this potential problem setting as follows:

```
CURRIC  ALTER    SFTWAR  ALTER     HRDWAR  ALTER

(-sectional-)    /////////////     /////////////
/////////////    (-sectional-)     /////////////
/////////////    /////////////     (-sectional-)
(----------sectional---------)     /////////////
/////////////    (----------sectional---------)
```

where the various sectionals (or subsystems) relate data evaluation points either within or between decisional alternatives.

Since some of the criterion vectors (row-wise) will contain major segments of zero-entries (e.g. the curriculum sectional, where the software and hardware portions of the curriculum vectors will contain only 0's), use of that vector as an objective function poses the problem of how the MAM execution will interpret the large number of 0's. For example, if the criterion reference is such that the objective of optimization should be minimization, the zeros will have greater influence than the actual non-zero entries of the curriculum sectional portion of the criterion row vector. On the other hand, the objective of maximization will be somewhat more reliable in that the zero entries will not have as great an influence as the non-zero components; however, such non-influence is only conjectural, and really depends upon the inner-workings of the various vectors.

The author has developed another technique which seems to provide not only the reliability required of cyclic optimization techniques, but also assures the related validity of the non-zero criterion entries which might be used as the objective function entries. To explore use of this technique, consider the following circumstances related to the use of cyclic objective functions when the modeling framework (viz. the constraint matrix) contains numerous zero-submatrices or empty subsystems. The objective of the objective function is to provide an array of values which the MAM system can either maximize or minimize depending upon their measured criterion (positive or negative impact, respectively). Seldom will the values of the criterion vectors be numerically larger than three or four digits, since large numbers can be expressed in decimal units (234,556 dollars = 2.35 thousands of dollars) and smaller numbers (e.g. 1,2,...,6 of the agreement continuum) can be easily accomodated. Seldom also will negative numbers be required. Therefore, the discrimination between these smaller positive numbers and the value of '0' has great potential for being confounded, when the sum of vector entries is controlled by the RHS-vector entries.

However, if the value of a relatively large number (e.g. the value of 999999999) was substituted for the zero-entries associated with empty submatrices, and the remaining non-zero, valid entries left the same -- the ability to discriminate between valid non-zero entries and the simulated zero-entry of '999999999' is certainly enhanced. The true test is of course whether such conjecture will be viable under both maximized and minimized optimality. Minimization holds the the least potential for confounding effects, as the sum of entries within the objective function is attempting to attain a optimal minimum value relative to the x_{ij} alternatives selected for inclusion within the solution set. If in fact, the sum of all valid, non-zero entries was still less than the simulated '999999999' (zero) entry, the chance of a '999999999' entry within the final solution set would be extremely small (and maybe impossible!).

For the case of requiring the maximization of the composed objective function (vector displaying positive impact values), the use of '999999999' will obviously be as disasterous as the use of '0' with minimization. However by multiplying the entire vector by '-1' -- that is, changing its valence structure -- the new value of '-999999999' becomes as foreign to maximization as it positive counterpart was to minimization. For the remainder of this paper, the use of a simulated '999999999' or '-999999999' vector entry to control for empty submatrices will be referred to as '*' and '-*' subvectors, respectively.

[Solution Tracking Matrices]

Each full execution of the cyclic optimization technique will of course provide a solution to the problem being modeled; and therefore will denote which decisional alternatives were included as members of the solution set, and which were not (i.e. excluded). Depending upon many factors (e.g. the degree of RHS restriction and/or relaxation; and the criterion influence of the particular criterion reference utilized as objective function), it is not uncommon to construct a variety of solution sets as a result of the various criterion vectors utilized in cyclic optimization. In some cases (in fact), a separate and distinct (unique) solution vector may result for each of the separate and unique criterion vectors, especially under a condition of relaxed RHS-values (Wholeben and Sullivan, 1981).

The attainance of unique solution vectors based upon the implementation of cyclic optimality is more than just an interesting result. Indeed, the existence of different solutions based upon different objective functions is exactly "what the doctor ordered", when demand exists to study the effect of bias upon the formulation of a particular solution. The reader should now be able to understand how three approaches to the never ending "... but, what if ..." problem can now be examined.

The first as we have explored within the criterion section deals with the introduction of new criterion references within the modeling framework; and then carefully examining the results of the varied solution formation. The second as examined in the preceding section on constraints, discussed the varied restriction and or relaxation of RHS-values as another method for analyzing the impact of criterion bias and decisioning intervention. The third technique of understanding the effect of new criterion references upon the solution set formation process is now available in the form of 'tracking' the varying solution set vectors as resulting from a cyclic optimization procedure. As we will see moreover, the use of 'solution vector tracking' goes beyond the identification and recognition of criterion impact and bias; and provides the main foundation for promoting such techniques as: the <u>integral solution composite vector</u>, the <u>progressive criterion frequency vector</u>, and the <u>stepwise reformulation strategy</u> for an iterative, sequential decision-making format.

In order to study the impact of cyclic objective functions, and their effect upon the formulation of a solution set vector, the construction of a <u>solution tracking matrix</u> is necessary. Structured as a rectangular dataset, where rows represent the array of multiple alternatives being evaluated and columns depict the individual criterion references for each of the cyclic objective functions -- cell entries are simply either 1's or 0's reflecting which alternatives were included (=1) within the final solution vector based upon the maximization or minimization of the particular criterion reference. As an illustration, consider the problem where eight alternatives have been evaluated across five criterion-referenced objective functions (i.e. the results of five separate executions of the MAM framework); the simulated results might have existed as follows:

	C-1	C-2	C-3	C-4	C-5	ISCV
A-1	1	0	1	0	0	2
A-2	0	1	1	0	1	3
A-3	1	0	0	1	0	2
A-4	1	1	1	1	1	5
A-5	0	0	0	1	0	1
A-6	0	0	0	0	0	0
A-7	1	0	0	0	0	1
A-8	0	0	1	0	0	1
PCFV:	4	2	4	3	2	

where:

ISCV =: the <u>integral</u> <u>solution</u> <u>composite</u> <u>vector</u>, and represents the frequency with which each individual alternative was included within a solution set across all cyclic optimizations (i.e. the sum of the row vector); and

PCFV =: the <u>progressive</u> <u>criterion</u> <u>frequency</u> <u>vector</u>, and represents the total number of solution entries which comprise the solution vector based upon the particular criterion reference of the objective function (i.e. the sum of column vector).

Use of the solution tracking matrix not only summarizes the individual results of the cyclic optimizations, but also provides two additional and necessary ingredients for a more valid and reliable decisioning process. Summing the frequency of solution inclusion (ISCV) constructs a 'weighting' network for the various potential solution alternatives. In our example, alternative '4' with a weight of '5' inclusions has a distinct advantage over alternative '2' with a weight of '3' inclusions,..or certainly alternatives '1' and '3' with weights of '2' inclusions each. The ultimate goal of course is to select the final solution set as that set of decisional factors which best models (or is modeled by) the criterion input for the evalution process. The integral solution composite vector provides the necessary data for just that evaluative need.

Summing the number of solution entries based upon the type of criterion objective function, constructs a analogical time-series mapping (or tracking) of the potential for further solution inclusion based upon a reiterative, stepwise solution reformulation technique. This summation of the column vectors (PCFV) has b· n found to a reliable predictor of the modeling framework's potential for generating additional decisions (solution sets) based upon the identified RHS-valued constraints (Wholeben, 1980b).

To translate the aforegoing paragraph into english, a practical illustration might be helpful. Consider the situation in which some number of schools need to be identified for potential closure according to a set of 24 agreed-upon criterion references. The use of cyclic optimization (cycling each of the 24 criterion vectors through the MAA model as objective function) is utilized, and the technique of reiterative, stepwise solution generation executed. Simply stated, this stepwise procedure will choose one and only one school for closure based upon the initial construction of the ISCV; then update those criterion vectors which will change value due to the closure of the school selected (e.g. enrollment, average walking distance, energy waste); and then re-execute in order to construct a second ISCV to determine the second school site for potential closure. Of course, the question is how many sites will require closure in order to meet the modeled district needs (constraints), and how will the modeler know when that limit has been achieved?

The 1980 (Wholeben) study on school closures found, that on successive iterations of the stepwise process, the values of the PCFV (the progressive criterion frequency vector) declined in a consistent fashion. That is, the individual sums of the column vectors decreased as each additional school was closed, and the original database progressively updated to reflect each of those closures. Obviously, the approach of such sums to the value of zero represents the inability of the school closure MAA model to select additional sites for closures; and thus interprets the goals of the district for site closure as having been satisfied.

Application of the solution tracking matrix, and its related components of integral solution composite (selection tally) and progressive criterion frequency vectors -- to the MICROPIK model and its need to select appropriate software packages and microcomputer hardware units compatible with desireable CAI/CMI instructional objectives -- presents a special case (with special problems) to the modeler in terms of data interpretation.

In the previous illustration, the value(s) of the selection tally vector were shown to be a result of summing across each row of a solution tracking matrix, where rows represented each solution alternative; and columns, each of the criterion vectors used as a cyclic objective function. With the MICROPIK model, the alternatives are split into three sectionals: curriculum, software and hardware -- representing different though obviously related decisions regarding the selection of appropriate CAI/CMI software packages and compatible hardware devices to match with a parallel selection of instructional activities whose needs can be met with these same software and hardware decisions. Having constructed the solution tracking matrix for the MICROPIK problem, the modeler in summing the decisional 1's across each of the inherent sectional rows must keep in mind, that three subsets of decision-making have been analyzed by the multiple alternatives analysis model:

[1] those curricular objectives and instructional activities which will be satisfied in the CAI/CMI mode of instruction;

[2] those curriculum software packages (i.e. courseware) which will accomodate these above selected instructional activities and curricular objectives; and

[3] the particular computer hardware devices (and peripherals) which will operationalize these above curricular courseware packages as they satisfy the desired CAI/CMI instructional objectives.

Accordingly, the display of the MICROPIK solution tracking matrix will be better demonstrated as follows:

$$\| \; C_C \cdots \mid C_S \cdots \mid C_H \cdots \mid C_{CS} \cdots \mid C_{SH} \cdots \mid C_{CH} \cdots \mid C_{CSH} \cdots \; \|$$

$$\mid (1,\ldots)(1,\ldots)(1,\ldots)(1,\ldots)(1,\ldots)(1,\ldots)(1,\ldots) \mid$$

C_1-1	
C_1-2	
C_1-3	The curriculum subsection of the MICROPIK
.	solution tracking matrix; where each of
.	various C_i curricular objectives and the
C_2-1	related C_i-r instructional activities are
C_2-2	tested for their inclusion within each of
.	the solution vectors as formed by the
.	cyclic C_C ... C_{CSH} objective functions.
C_i-p	
S_1-1	
S_1-2	
.	The software subsection of the MICROPIK
.	solution tracking matrix; where each of
S_2-1	various S_j curriculum and the related
.	S_j-s courseware packages are tested for
.	solution vector inclusion.
S_j-q	
H_1	
H_2	
.	The hardware subsection of the MICROPIK;
.	testing H_k hardware inclusion decisions.
H_k	

where the appropriate criterion sectionals are represented as:

$[C_C]$ =: CURRICULUM SECTIONAL

$[C_S]$ =: SOFTWARE SECTIONAL

$[C_H]$ =: HARDWARE SECTIONAL

$[C_{CS}]$ =: CURRICULUM - SOFTWARE SECTIONAL

$[C_{SH}]$ =: SOFTWARE - HARDWARE SECTIONAL

$[C_{CH}]$ =: CURRICULUM - HARDWARE SECTIONAL

$[C_{CSH}]$ =: CURRICULUM - SOFTWARE - HARDWARE SECTIONAL.

The appropriate row summations across the applied cyclic objective functions will now present selection tally vectors (ISCV) for each of the three "C", "S" and "H" (curriculum, software and hardware) subsections; and thus denote the array of instructional activities which can be satisfied via the parallel inclusions of courseware and hardware devices. It is important to note also, that with the structure of the selection tally vector denoting a 'range of inclusion', the <u>extent of satisfaction</u> is available for modeler evaluation and decision-making.

[Solution Vectors and Stepwise Reformulations]

The construction of the 'final' solution vector, as a <u>binary</u> representation of the "integral solution composite vector" (ISCV, or selection tally vector), is a rather straight-forward procedure in most cases. The problem usually encountered will involve the arbitrary decision to determine what degree of inclusion for any particular alternative will signal that alternative's selection as a decision (=1) or a non-decision (=0).

Consider the ISCV which has resulted from a cyclic optimization of a ten-alternative, twenty-criteria MAM execution; and may be simulated as follows:

$$[\ 02\ 07\ 04\ 09\ 11\ 18\ 00\ 15\ 06\ 12\ ]$$

where the first alternative was chosen as a solution a total of two-times, the second alternative a total of seven times, and so forth. The sixth alternative (inclusion = 18) was found to have the highest selection factor, the eighth alternative with the second highest (inclusion = 15); and the seventh alternative never entering any of the cyclic optimizations as a probable solution to the modeled problem. Based upon the range of the inclusion frequencies as shown, the final solution vector would be constructed by serially including (one at a time) each of the solution alternatives, starting with the one with the highest inclusion frequency first. Thus, the final solution vector would display serial development as follows:

$$[\,0\;0\;0\;0\;0\;1\;0\;0\;0\;0\,]$$
$$[\,0\;0\;0\;0\;0\;1\;0\;1\;0\;0\,]$$
$$[\,0\;0\;0\;0\;0\;1\;0\;1\;0\;1\,]$$
$$\vdots$$

until 'some' ultimate criterion was satisfied (e.g. minimum new equipment expenditures, maximum school sites left opened, or minimum additional bus-stop requirements).

The necessity of such a 'test-retest' procedure for final solution vector formulation stems of course from the lack of control upon solution formation by the various criteria <u>after</u> the maximization (or minimization) of the last criterion vector during cyclic optimization. The reader may also detect problems with the notion of 'testing-retesting' using only an 'ultimate' criterion reference -- instead of utilizing 'all' of the criteria within the original model. That is, the potential exists for the fourth serial configuration of the final binary solution vector,

[0 0 0 0 1 1 0 1 0 1]

to satisfy the ultimate criterion (or criteria), but coterminously violate one or more of the original criterion references which were utilized within the execution of the MAM-comstructed problem. Other obvious problems might arise, where the k-th serial configuration of the solution vector violates a single criterion vector, halting serial solution construction -- and thus preventing the further development of a more optimal solution (i.e. better value on the ultimate criterion).

The final major problem with serial solution formulations exists with the existence of 'inter-dependence' between the various potential solution alternatives. For example, the closing of a particular school could logically cause a most positive effect upon a neighboring school whose own enrollment has been decreasing. The student transferees from the closed school who live within walking distance to the other school left open will obviously serve to alleviate some of the vacancy problems associated with the second site. Without taking this into consideration however, a serial construction of the final solution as described above might erroneously include that site as a site for immediate closure -- most embarassing to say the least.

To control for such invalid decisions (and unreliable decision-making), the use of a stepwise solution generation system is suggested, in lieu of the serial system discussed above. The stepwise solution technique incorporates many of the valid parts of the serial approach, but utilizes the serial system in a more sophisticated way.

Using our previous example of a simulated selection tally vector,

[02 07 04 09 11 18 00 15 06 12]

the 'stepwise approach' to constructing the final solution set would exist as follows. Since the sixth alternative (inclusion = 18) clearly outdistances its competitors, it would be chosed as the initial 'solution'; or,

[0 0 0 0 0 1 0 0 0 0]

would become the first iteration result of the final solution vector. The original MAM-framework would then be redesigned to signify the loss of alternative-six as a potential decisional alternative to be evaluated across the criterion references. In addition, those criterion references which are related by vestige of this decision (e.g. the enrollment of neighboring schools which would have to absorb the student transfers) would be recalculated to denote value changes (e.g. relationship of new enrollment to total capacity of the site, or the amount of vacancies). Having completed these recomputations, the reduced n-1 alternatives' model would then be re-executed, and a totally new solution tracking matrix constructed and selection tally vector designed.

The result of this n-1 (or nine alternatives) cyclic optimization might be simulated as follows:

[01 05 04 08 09 -- 00 14 05 10]

where alternative-six has been deleted from further consideration. As was found in the aforegoing 'serial construction' procedure, the eighth alternative (now = 14) becomes candidate for inclusion within the final binary solution vector, or:

[0 0 0 0 0 1 0 1 0 0].

Once again, the criterion vector constraint matrix is redesigned to denote the effect of choosing the eighth alternative as a solution; and the now n-2 MAM-model re-executed once again. However this third re-execution now results in the selection tally matrix,

[00 03 04 05 09 -- 00 -- 03 07]

where alternative-five is the third stepwise candidate for the final solution set (instead of alternative-ten as found with the serial procedure). Thus the final solution vector becomes,

[0 0 0 0 1 1 0 1 0 0],

and not

[0 0 0 0 0 1 0 1 0 1].

The reader is cautioned to the dangers of not subscribing to the idea of a stepwise solution strategy; and for the reasons which are hopefully very apparent above.

The use of the idea of stepwise solution generation may not play a primary role in the tracking of the various cyclic solutions for the MICROPIK formulation. It is reasonable for example, to expect, that only a single type or brand or computer hardware will be purchased by an individual school depending upon the results of the MICROPIK execution(s). Thus, the hardware sectional of the model would not be subject to a stepwise strategy. It is also reasonable to assume, that the software sectional results (solution vector entries) will be of such a nature, as to require only 'sight-verification' for final decision-making and selection. And if the decision-maker is satisfied with the particular degree to which each of the curricular objectives is met, the results will exist as chosen by some arbitrary selection from the tally vector's initial formulation.

However, the stepwise technique can play a most important role in the MICROPIK setting <u>if</u> the initial selection tally vector displays the model's determination that more than a 'single type' of hardware manufacturer is required for optimal CAI/CMI implementation (e.g. both APPLE <u>and</u> TRS-80). It is suggested in such instances, that the model be re-executed a total of two additional times -- once where the system is constrained to choose APPLE, and only APPLE as the hardware unit; and then where the system must choose TRS-80 as the single device compatible with other decisions from the courseware and activities portions. It can be expected, that the selection tally vectors with respect to the first 'two-device' solution will change based upon first the exclusion of TRS-80 as a candidate, and then secondly the exclusion of the APPLE. For the sake of review, the constraints effecting each of these suggested 'restraints' to model inclusion would exist as follows:

<for the existence of APPLE, and only APPLE>

APPLE	TRS-80	ATARI	OSI	TI	...		RHS
1	0	0	0	0	...	(=)	1

<for the existence of TRS-80, and only TRS-80>

APPLE	TRS-80	ATARI	OSI	TI	...		RHS
0	1	0	0	0	...	(=)	1

where (as the reader will hopefully recall) the forced inclusion of the hardware unit will be reflected to the curriculum and software sectionals via the tautological constraints within the curriculum-hardware and software-hardware sectionals.

With each of these two re-executions, the modeler must then evaluate not only the differential extent(s) to which the new two selection tally vectors solve the originally modelled issue, but also the degree to which a forced, single-device restraint upon the implementation of a CAI/CMI strategy retards the actual satisfaction initially desired.

[Weighting Solutions Via Tracking Vectors]

Although the technique of weighting particular decisions has been discussed in the previous section (see CONSTRAINTS), it seems appropriate to briefly demonstrate the potential benefits of declaring particular solutions as more important than others.

As demonstrated via the use of solution tracking matrices, a sub-matrix vector exists for each of the results of a cyclic optimization (maximization or minimization) of the individual criterion references. This sub-matrix vector, or tracking vector, demonstrates which alternatives were determined both integer feasible and optimal based upon the values of the full constraint matrix and the cyclic objective function, respectively. That is, each column vector of the solution tracking matrix shows the particular weight of that cyclic objective function's criterion reference upon the final solution (binary) constructed. As the value of the objective function changed (i.e. different criterion reference used), so often (in most cases) does the configuration of the resulting solution vector. We have found this circumstance to be especially true, where the RHS-values have been constructed in what we have previously named the 'relaxed' state -- that is, giving the solution process more 'lee-way' in selecting the best solution combination for final inclusion.

These various tracking vector results can be utilized to produce desired (or undesired) weights for the final selection tally (integral solution composite) vector. By determining the factor-weights to be used in the weighting process (e.g. identifying the base-criterion objective function, and then assigning factors of related importance to the other criterion references in the form of 2-times as important, 1.5-times as important, etc.), a weighted selection tally vector can be formulated.

Consider the following problem, where the solution tracking matrix has been formulated for the results of a five-alternative, five-criteria model:

WEIGHT:	(2.0)	(1.5)	(1.0)	(2.0)	(3.0)	(n/wgt'd)	(wgt'd)
	C-1	C-2	C-3	C-4	C-5	NW/ISCV	W/ISCV
A-1:	1	1	0	0	1	3	6.5
A-2:	1	0	0	1	0	2	4.0
A-3:	0	0	1	0	0	1	1.0
A-4:	0	0	0	0	0	-	-
A-5:	1	1	1	1	1	5	9.5

where the use of weights does not change the direction of the selection tally vector, but does in fact change the ultimate degree to which each member of the tally vector is deemed of comparable value.

Now consider a different problem, where the various tracking vectors are not so demonstrative in their selection of potential solution alternatives:

WEIGHT:	(2.0)	(1.5)	(1.0)	(2.0)	(3.0)	(n/wgt'd)	(wgt'd)
	C-2	C-2	C-3	C-4	C-5	NW/ISCV	W/ISCV
A-1:	1	0	1	0	0	2	3.0
A-2:	0	1	1	0	0	2	2.5
A-3:	0	0	0	1	1	2	5.0
A-4:	0	0	0	0	1	1	3.0
A-5:	0	1	0	1	1	3	6.5

where weighting has provided a discriminant pattern for further evaluation by the modeler of the differences between the first three alternative decisions.

The reader will also note, that weighting can also take place with the MAM framework, prior to the initial execution of the model; and therefore be representative implicitly within the selection tally vector. The reader is also cautioned of the 'opportunity' for <u>double-weighting</u>, where weights are factored into the various criterion constraints before execution, and then utilized again as weights for each of the solution tracking vectors as described above.

Application of weighting techniques to the MICROPIK setting has obvious benefit for the evaluator and decision-making. By not weighting prior to the construction of the initial selection tally vector, the modeler has the opportunity to witness the differential effect (if any) weighting has upon membership within the solution vector(s). This is especially so, when the analyzed compatibility between activities, courseware and hardware has been determined initially without weighting; and then various weights are applied to instructional activities and/or software to note the effect upon the composite vector's structure.

However, the modeler must be extremely cautious of weighting strategies, and their impact upon tautological requirements, via the curriculum-software, software-hardware and curriculum-hardware sectionals. It is apparent, that indiscriminant weighting could not apply its effect universally across the entire MICROPIK model; and therefore not provide reliable results within the various differential selection tally vectors (weighted). For this reason, the use of 'weighted' solution tracking vectors is <u>discouraged</u>, except in cases where clear control over tautological cross-impacts is possible.

THE RESULTS

In the previous section on execution of the multiple alternatives model in general, and the MICROPIK modeling formulation in particular, we have been concerned with the generation of a variety of solution vectors which would provide the evaluator and/or decision-maker useful information for discriminating between multiple decisional alternatives as solutions to some pre-defined problem. Specifically, this problem is the conscious acquisition of CAI/CMI instructional software (i.e. courseware) and compatible hardware (i.e. micro-computers) for satisfying an array of identified curricular objectives, and their delineated instructional activities.

We have discussed in some detail (or for some of the readers, too much technical detail) the application of criterion references in the form of inequalities and equalities to the final selection of a set of decisional alternatives -- not only feasible in terms of solving (i.e. modeling the desired characteristics of) the problem, but also optimal in terms of providing the 'best solution' as defined by some one (or more) objective functions. We have then witnessed how the individual results of each solution set have been incorporated into a solution tracking matrix, the final composite solution identified, and differential weighting applied as desirable.

In this section, the concept of criterion strength will be be explored as it impacts upon the MICROPIK problem resolution. We will examine the related concepts of decision validity and decisioning reliability, and demonstrate how it can be applied to the MICROPIK setting. Finally, the use of various statistical procedures will be evaluated for their utility in providing the basis for some ultimate 'professional statement' concerning the validity of the results, and the reliability of the procedure utilized in determining these results.

[Criterion Strength and Decisioning Reliability]

Evaluation and all decision-making resulting therewith, demand a high degree of accountability, visibility and responsibility. Today's complex issues require equally complex methodologies to assess both content and process of such issues, and to provide an understandable environment within which to simulate potential decisions and measure resulting effect or impact. As important moreover, is the secondary demand for providing a means for post-hoc evaluating not only the results of the simulated decisions, but also the influence (singularly as well as collectively) which the criterion references lend in making the original decisions. The clear need for the criterion-referenced decision-maker therefore is to satisfy the following five objectives:

[1] to <u>validate</u> the sophisticated decisioning methodologies which are so necessary for addressing complex problems -- yet so often ignored, discounted or feared;

[2] to <u>study</u> criterion effect upon the decisions made, and the impact which the system receives via those decisions; and thereby understand differential criterion weighting and influence -- "what" made a difference in constructing the decisions, and the varying impact resulting;

[3] to <u>provide</u> a high degree of visibility, and therefore accountability, to the public interests

served and affected via those decisions -- generating a milieu of trust within which the decisions, no matter how unexpected, can be trusted and accepted;

[4] to <u>simulate</u> the variable impact upon the decisions made by introducing additional criterion influences into the model, and thereby perform a path analysis from solution to solution as different criteria are utilized to construct each decision or solution -- satisfying the innate need of some individuals who must always ask, "... but, what if ...?"; and

[5] to <u>permit</u> easy and quick decisioning replication within an ever changing environment -- knowing the relationships between past successful decisions and the criteria used to construct those solutions, in order to understand the potential of future decisions based upon the new values of more current criterion measures.

Generally, the notion of <u>criterion strength</u> refers to the identification of those measures which in effect constructed the final decision or solution to the modeled problem; and furthermore provide a 'factor' measure of ordinal value or weight within that same group of 'solution-formation' variable measures. Specifically, criterion strength will address <u>three</u> fundamental questions existent within all decisioning evaluation:

[1] which criterion references most clearly defend the decisions made?

[2] to what extent are the criteria individually representative of the decisions made?

[3] how do the most discriminating criteria within this decision setting relate to each other in terms of importance and influence?.

A later part of this section will illustrate the utility of discriminant function(s) formulation for answering these questions of criterion strength, respectively, by evaluating the following rudiments of discriminant analysis:

[1] criteria included within the formation of discriminant functions -- that is, which references were 'entered' into the composition of the prepared functions;

[2] order-of-entry of each of the variables which discriminate the final solution vector; and

[3] weight (or factor strength) relationship between the standardized canonical discriminant coefficients.

Generally, the notion of decisioning reliability refers to the degree of trust which is implicit to the decision model (in this case, the "multiple alternatives model" -- MAM); implicit in the sense, that the decision-maker can accept the results of such a criterion-referenced technology, both in terms of content (viz., effect of the criterion references within the model) as well as process (viz., effect of the model upon the criterion references). Specifically, decisioning reliability will address two fundamental questions existent within all decisioning evaluation:

[1] to what extent are the criteria collectively representative of the decisions made?

[2] to what extent can the defined matrix of criterion references re-predict the original binary (include v. exclude) solution?.

An additional part of this section will illustrate the utility of discriminant function(s) formulation for answering these questions of decisioning reliability, respectively, by evaluating the following charactistics of discriminant analysis:

[1] canonical correlation coefficients which offer a measure of relationship between the 'set' of discriminating criterion references and the 'set' of dummy variables which are used to represent the solution vector; and

[2] the frequency of mis-inclusions and/or mis-exclusions (or over-estimations and/or under-estimations) discovered when the classification coefficients constructed to predict a solution with the known relationships among the discriminating criterion variables, are utilized to re-predict the original dependent variable (original solution).

[Tools for Validity and Reliability Testing]

To construct discriminant functions from the relationships between the model just discussed above and the resulting solutions formulated, require the use of linear vectors and combinations of vectors (matrix). Only those vector and matrix formulations most germane to this paper will be discussed below. The reader is invited to be patient until the scheduled publication of the manuscript, "Multiple Alternatives Analysis for Educational Evaluation and Decision-Making" in late summer of 1982, for a detailed illustration of all vectors and matrice pertinent to MAM.

Solution Set Vector. In order to distinguish between alternatives included or excluded as members of the final solution to the system modeled, a vector of binary-decision representations is required, in the form:

$$[\ 1\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ \ldots\ 1\ ]$$

where '1' means that the criterion values associated with that particular x(j) will be computed to measure resulting system impact; and '0' means that the underlying criterion values will have no impact upon the system.

Selection Tally Vector. To observe the effect of each criterion reference upon construction of the system solution, a method called cyclic optimization (Wholeben, 1980a; Wholeben and Sullivan, 1981) is used. Under this regimen, the model is executed once for each unique criterion being used to constrain the model, where each unique criterion is cycled through the model as the objective function. For example, during one execution in the case of the school closure model, the intent may be be prepare a solution set whereby existing capacity of the remaining schools will be maximized; in another cycle, the model will be executed such that the schools remaining open within the district will minimize the amount of energy expended for facility heating requirements. The selection tally vector is basically a frequency summation vector, compiling the number of times each alternative was chosen as part of the solution vector, across all cyclic optimizations. Such a vector will be represented as:

showing that the first alternative was selected as solution a total of 3 times, the second alternative a total of 7 times, and so forth. This vector is extremely important when the MAM procedure requires a step-wise decisioning process such as the school closure model -- evaluating a revised database after closing a single school such that the effects of closing each individual site is summarily incorporated into the next decision for determining additional site closures.

Discriminant Criterion Inclusion Vector. This vector simply represents another binary entry vector of 1's and 0's, signifying which particular criterion references were utilized via discriminant functions to develop the canonical classification coefficients, and the standardized canonical discriminant function coefficients.

Discriminant Criterion Entry Vector. This vector contains 1,2,...,k entries, where k criteria were utilized in the development of the discriminant functions, and the 1,2,...,k entries represent their order of entry into the discriminant formulation. Criterion variables not entered into the function(s) receive a value of '0', by convention.

Discriminant Weighting Summary Vector. Applying discriminant procedures to the binary solution vectors will result in the computation of standardized canonical discriminant function coefficients. These coefficients will reflect the utility of entered criterion vectors if those vectors contain standardized measures in lieu of the normal raw scores. By dividing each of the standardized canonical coefficients by the smallest of the standardized canonicals, the quotient will provide a factor of importance for each of the criteria as relative to the other criterion entered in the discriminant formulation. The discriminant weighting summary vector is a linear representation of these factors (quotients), where the minimum entry value is always '1.00' (smallest standardized coefficient divided by itself). Non-entered criterion locations receive a value of '0.00' by convention.

Other 'tools' have been referenced in the proceeding section of this paper: criterion constraint matrix, condition limits vector (RHS), objective function vector, and the cyclic optimization tracking matrix. Other formulations are currently under study by the author (e.g. the optimality weighting matrix) to investigate new relationships which may allow greater accountability and useful reliability of the multiple alternatives modeling framework.

[Criterion Strength Via The Optimality Weighting Matrix]

The explicit check on procedureal (model) reliability by way of the discriminant functions (and their re-predictability of set membership), and the more implicit check on criterion validity by noting the type (which ones?) and strength (how much?) of the various criterion variables entering the discriminant analysis -- are not the sole measures of post hoc evaluation available to the MAM decision-maker. A further check on validity and reliability is afforded the modeler via the construction of the optimality weighting matrix.

The optimality weighting matrix is simply a summary of the the preponderance of each criterion-referenced variable utilized within the MAM procedure, as measured within each subset of the solution v. non-solution multiple alternatives. The measures of preponderance (direction, strength and weighting` result from the application of analysis of variance (ANOVA) procedures to each of the modeling criteria, based upon an alternative's membership in the final binary solution set. Successive ANOVA procedures can also be applied to the criteria based upon each of the results of the cyclic optimizations.

Denoting an alternative as either a member of the solution set (that is, =1) or not a member (therefore, =0), two separate data distributions can be constructed and summarily evaluated for both the statistical and magnitudinal significance(s) of their computed mean-value differences. Since this (0,1) analysis of variance procedure can be applied to each criterion reference, and for each of the cyclic optimality solution set results, a matrix format can be utilized to display, and furthermore evaluatively summarize the ANOVA results. This matrix is called the optimality weighting matrix, where each row represents the individual criterion reference modeled within the MAM framework, and where each column denotes the particular criterion-modeled cyclic maximization or minimization based upon a single criterion focus. For example, a 32-alternative and 24-criteria model would enable the composition of a 24 x 24 dimensional matrix with a total of 1056 cells (impressed?; or beleagured?). All such matrices will always be square matrices.

Each of these m^2 cells will be composed of the results of that particular oneway analysis of variance which utilized solution set membership (0,1) as an independent variable, and the individual criterion reference (constraint vector values) as a dependent variable. The specific statistics resulting from such a procedure which are of importance to our matrix are as follows:

[1] means of both the solution and non-solution distribution; and their individual standard deviations;

[2] the statistical significance of the set membership mean-differences; and

[3] a non-parametric check (usually the use of chi-squared) of those criterion mean-differences which result from non-ratio-scaled criterion references.

With this summary information, the evaluator or decision-maker is able to view the frequency of statistically-significant differences between the solution and non-solution sets, the direction of these differences and their conformance to initial constraint demands, the relative strength or magnitude of these differences with respect to degree of difference between the distribution means, and finally the extent to which the integral solution composite vector reflects the intent of the modeling framework -- and thus the intended solution to the original problem.

You might be thinking, that the above procedure will operate correctly for a ratio-scaled criterion variable, and also provide a check on well-constructed interval-scalings -- but not be at all useful for summarizing both the nominal and ordinal criterion vectors. And you would be most correct. Unfortunately, nominal data must be analyzed via contingency analysis procedures (or what most people call cross-tabulation or chi-squared techniques). Obviously, mean differences and standard deviations are not a function of this analysis (or even meaningful). The evaluator will substitute the statistical significance of the chi-squared statistic, and some summary of the differences between observed and expected frequencies, for the usual cell entries of the optimality weighting matrix.

The use of a non-parametric, numerically-ranked, oneway analysis of variance procedure (e.g. Kruskal-Wallis) works well for ordinally as well as interval scalings. Some readers might think the above ruminations an adroit hassle; but other than the validity and reliability test benefits of such statistical techniques, it does support the use of interval and ratio scalings as often and as completely as possible, without compromising the modeling framework.

As you might also have already guessed (or feared), the application of the optimality weighting matrix design to the MICROPIK setting is (once again) a special case.

Because of the use of sectionals (curriculum, software, etc.) in the MAM construction, subsets of criterion references exist which apply only to specific subsets of the multiple alternatives being evaluated. Therefore, some criterion vectors will apply only to the evaluation within the software sectional, while other criteria apply only to the evaluation between the software and hardware sectionals; or curriculum and software sectionals. Application of the ANOVA procedures to the various cyclic optimizations and the resulting relationships with the full criterion set within the constraint matrix, should therefore (it is suggested) be directed towards each of the sectionals, rather than a system total approach.

THE INTERPRETATION

As we approach the end of our sojourn through the world of mathematical modeling and multiple alternatives analysis, and their role in the evaluation of potential decisions concerning the selection of microcomputer software and hardware for CAI/CMI applications, there remains the need to discuss the less-technical aspects of modeling -- albeit no less important. It is easy to become enamored with the process of the MAM framework, and its role in the MICROPIK setting, and unconsciously ignore the potential difficulties of the model -- both content and process --and their impact upon the resulting alternatives evaluated and decisions selected.

We have taken a great deal of time in exploring first the conceptualization of the multiple analysis framework, and second its application within the MICROPIK structure. This was necessary in order for the reader to fully understand the vast utility of the model as well as lend credence to the postulates presented. As one colleague stated some several weeks, "How can you possibly explain an application of your model to the CAI setting, if the general reader does not first understand the model itself?". This morning, I received his evaluation as to the utility of this paper, and its satisfaction in resolving just that issue he asked of some weeks ago. His response was, "Oh.". But was it declarative, interrogative or exclamatory?

This last section will deal with the underlying premises of the MICROPIK modeling structure, and their related positive and negative influences upon the decision-making required. We will initially examine the general utility of such a model, and the advantages to be enjoyed. In addition, some of the more common disadvantages and potential pitfalls of this model will also be discussed; and their role in arriving at erroneous conclusions. Finally, a totally unsolicited and thoroughly unbiased statement of the implications for this technique in future decision-making will be made.

[Utility of the MAA Modeling Procedure]

It nearly suffices to state, that the multiple alternatives analysis framework adds to the evaluation and decision-making setting those components which often seem non-existent in the realm of educational decisioning: visibility, responsibility, accountability and credibility. In reviewing the aforegoing 114 pages of this technical paper, what specific references have been made which would allow the reader to adopt a trusting attitude towards the MAM modeling procedure in general, and the MICROPIK application specifically?

Multiple Alternatives. The responsibility of the evaluator and decision-maker is to examine feasible alternatives in resolving a dilemma, and then determine the most optimal approach to follow. The problems associated with not identifying and defining all available alternatives are well documented in situations where a solution to a particular problem was declared unobtainable. Other problems, concerned more with controlling for decision-maker bias and the likelihood of pre-arranged decisions, have also proved the utility for adopting a multiple alternatives' orientation.

Criterion References. Accountability in evaluation and decision-making is inextricably linked to the data utilized in formulating, analyzing and selecting the decisional alternatives in remediating a particular problem situation. The process invoved in identifying and defining the criteria for a required decision, the choice of datum points for correlating a criterion reference, and the measurement of these points for quantifying the necessary comparative values of these defined criteria -- lends a visibility to the evaluatoin and decision-making process which is fully open to public (and private) scrutiny and critique. Constituents may not agree with the decisions made, but they must understand the bases for these decisions, and the validity of these underlying criterion foundations.

Solution Membership. The singular, most indefensible aspect of decision-making in a multiple-alternatives environment is the determination of size and identity of the final solution set. The questions of "how many" and "which ones" must be answered in a structured, scientific sense; and as discussed above, reflect both the intent and demand of the criterion references imposed upon the decisioning framework.

Interactive Effects. Seldom does an alternative action possess such qualities as to be an obvious choice for membership within the final solution set. More often, alternatives will display positive characteristics on many criterion references, only to denote one or two negative by-products which may be undesireable to the system being modeled. The application of a main-effects modeling design allows positive attributes to cancel the displayed negative features; and thus nullifies any control over such negative impact to the system as a whole should those alternatives be selected as solutions. Interactive effects modeling on the other hand controls not only the impact of particular subsets of alternatives upon the system, but also individualizes the effect of each alternative across all of its criterion measures.

Focused Optimality. The questions associated with 'what is possible' and 'what is best' address all aspects of decision-making.

The consideration of feasible alternatives, and the selection of some optimal alternative of 'all available alternatives' requires the parallel choice of an overall discriminant criterion reference. That is, once all of our demands have been met (constraints), the one, single best choice (optimal) must be found based upon some predefined point of reference (objective function). Several such points of reference (cyclic optimization) allow the decisioner to examine the impact of potential alternatives upon the environment being modeled.

Trade-Offs and Preferences. Since alternative solutions will often display both positive and negative attributes regarding their probable impact(s) to the system, decision-making must be able to reliably monitor both the direction and extent of effects to the system be remediated. While many side-effects may be undesireable, the quality of each alternative's positive characteristics must be allowed to model the desireable benefits of that alternative. Simultaneously, positive and negative characteristics must be allowed to co-exist and therefore be measureable, in order to truly model the real-world situation.

Stepwise Solution Formation. Since some aspects of any decision impacts upon other decisions which may be forthcoming, preparations must be made to control for the effect of such preceding decisions upon potential succeeding decisions which may be necessary to completely satisfy stated constraint requirements. With a criterion-referenced dataset as the basis for comparative evaluation among alternatives, the selection of a single decision will obviously effect the criterion values in some way (assuming of course, that the decision does in fact provide some degree of remediation to the system be modeled). In order to evaluate the 'remainder' of the system problem, this dataset must be updated to reflect the degree of solution already imposed by the choice of the previous decision (alternative selected). Subsequent analyses will then be able to provide a valid and reliable 'next' solution to impact upon the extent of problem 'remaining'. The final entry to the solution set is reached, when a subsequent analysis fails to detect a new member; and the last dataset update reflects system, criterion-referenced components as desired.

Simulation (Before) and Interrogation (After). The final measure of utility for the MAM formulation lies in its ability to provide both inductive and deductive reasoning mechanics to the system evaluator and decision-maker. Based upon a carefully derived set of criterion-references which are deemed representative of both the system being modeled, and alternatives which possess a varying degree of potential to resolve an identified problem within this system -- the multiple alternatives model is able to simulate the problem setting, and thus derive (viz., induce) the necessary solutions which reflect the demands and needs of the system. Moreover in the case where decisions have already been presumed based upon some set of criterion measures, the MAM framework is able to interrogate the problem setting, and thus derive (viz., deduce) the demands and needs of the system which reflect the a priori solutions made. Even in the event of a set of decisions without the benefit of an identified criterion-referenced database, reasonable criteria can be postulated and subsequently measured against the proposed solution set.

[Advantages and Disadvantages of MICROPIK]

The need to test the content of decisions for validity, and the process utilized in arrived at this content for reliability, suggests the rather superfluous assertion that any technique for making decisions has its problems in addition to its laudable benefits. The MICROPIK modeling formulation is (regretably) no exception to this existential assertion.

Recall the main goals of the MICROPIK structure: to provide a criterion-referenced, multiple-alternatives decisioning model for evaluating CAI/CMI software and micro-computer hardware for its compatibility with desired curricular objectives and instructional activities. That the modeling framework as exposed within the preceding paper actually accomplishes this task, posits the main advantage of the model over any other decisioning tool known to this author. As stated within the preceding section concerning the utility of the MAM procedure in general, specific advantages are assignable to the MICROPIK framework in terms of its ability to:

[1] provide an evaluation framework for the development of a set of decisions (solution set) from a larger set of potential, multiple alternatives;

[2] utilize a criterion-referenced dataset as the basis for comparing the direction and degree of positive and negative attributes associated with each of the potential, multiple alternatives;

[3] control for the interactive effects between the measured criterion values, the various groupings (combinations and permutations) of the multiple alternatives, and their resulting impact upon the system as a whole -- and thus determine the members of the solution set in terms of "how many", and "which ones";

[4] investigate the effect of varying the optimality design for each separate execution (solution set formation) -- and thus examine sequentially the biasing factors associated with each criterion vector;

[5] prepare a database revision strategy for implementing a stepwise-procedure in developing the final solution set; and

[6] simulate the impact upon the system as a whole of potential decisions for remediation of the defined problem, as well as interrogate the relationship of past and/or current decisions to the agreed-upon criterion references supporting those decisions.

However, the implementation of the MICROPIK model also has a number of disadvantages associated with its utilization -- as therefore does the MAA model in general. These disadvantages can be encapsulated within three general headings: model-related, user-related, and equipment-related.

Model-related disadvantages are probably obvious to the reader at this point. The development of all possible or feasible solution alternatives, the definition of all sufficient and necessary criteria, the scaling and measurement for each coefficient entry to the criterion contraint vectors, and the conceptualization and computation of the appropriate RHS-values -- are enough to divert evaluator interest to other less-sophisticated evaluation techniques; and have been known to drive even the most adroit educational administrator to fits of manic depression.

User-related disadvantages are foreshadowed by the initial use of the terms 'mathematical modeling' and 'simultaneous linear inequalities', and the tendency on the part of the administrator to request immediate psychotherapy. These exist sufficient historical references to past evaluators who have utilized quantifiable evaluation techniques to maske the real missions of their endeavors, or to provide post hoc support to a priori decisions devoid of a valid criterion-referenced framework. The use of any new terminology is greeted with the criticism of "jargon"; the use of mathematical techniques with the criticism of "... not everything can be quantified ... and therefore, nothing should be ..."; and the use of sophisticated decisioning strategies and (as we will soon see) electronic computers with the criticism of "... too hard to understand ... too technical for consumption by the general public ... and therefore, not useful ...".

Equipment-related disadvantages are in reality the true insurmountable barriers to the acceptance and subsequent use of any MAM design. Although the author has on occasion (but infrequently) solved MAM problems "by hand" -- this is not the preferred technique. Therefore, the use of computers is the modeler's salvation. But to utilize these computers, specific software packages must themselves be available (or written) to correspond with the required mathematical programming algorithms needed for MAM solution. While such packages are available (e.g. IPMIXD, MPOS, EZLP, LINDO), they are not a usual software component on most computerized hardware mainframes. And unfortunately, most evaluators have not been instructed in their use, let alone their existence and utility.

Finally, the MICROPIK formularization of the general MAM model is (unfortunately though not apologetically) a complex variation of the multiple-alternatives, integer programming system. The use of alternative sectionals, and separate criteria to relate various sectionals for cross-evaluation, adds to the potential confusion and conflict on the part of the user and the public whose needs the modeler is attempting to satisfy.

Currently, the evaluation of curriculum, courseware and hardware for CAI/CMI implementation proceeds in an often undefined manner -- hardware is purchased; the existing compatible software is examined; and curricular objectives or instructional activities redesigned to fit the available courseware perogatives. In effect, the classroom teacher is 'locked-in' to a hardware device, which in turns narrows the choice of software, and thus ultimately defines the satisfaction of particular objectives. More and more, school districts are examining software first, then the hardware devices for compatibility, and so forth. While the newer trends in evaluating CAI/CMI are producing more satisfying results, the ability to control for all multiple alternative instructional activities while satisfying (to some degree) all curricular objectives, and relate these to the available courseware and hardware -- has not been possible (until MICROPIK, obviously).

[Major Pitfalls and Erroneous Conclusions]

Within the consideration of advantages versus disadvantages, we a priori assumed a successful design, construction and execution of the MICROPIK model. Now however, some time must be expended in discussing the potential problems associated with the inappropriate design, invalid construction and/or unreliable execution of the modeling framework.

As has been reiterated throughout this paper, <u>inappropriate design</u> is usually associated with the exclusion of some alternatives (for whatever reason) from the modeling framework. For example, the absence of various instructional activities and their relationships to potential courseware availability will automatically preclude the model's potential in satisfying their needs. Likewise, the absence of a particular criterion from consideration will preclude the model's ability to control for that criterion's impact upon the system -- which may be positive or negative, and maybe even disasterous.

The appearance of invalid construction as a major pitfall often takes the form of problems associated with the scaling and measurement of the criterion coefficients (vector components); and is thus a secondary problem stemming from inappropriate criterion referencing. Problems will also arise based upon the criterion's measure and its utility in describing system impact based upon a row-vector summation.

Unreliable execution is a frequent problem associated with the construction of the RHS-vector, and the complex restriction versus relaxation effect these values have upon the summations of the individual criterion constraint vectors. The use of the cyclic optimization strategy also provides difficulty for the maintenance of reliability; indiscriminant maximization (or minimization) can introduce conflicting demands to the system, and produce solution sets in direct opposition to one another. In addition, compilation of the various cyclic solution vectors into a final selection tally vector (though valid) can also provide a new source of unreliability to the final determination of the actual binary solution vector.

In general however, once all of the procedureal, technique-oriented, and sequentially-defined prerequisites have been met, the major problems associated with the MICROPIK modeling situation remain: first, its interpretation for decision-making; and second, its incorporation into practice.

The interpretation of MICROPIK results must include a firm understanding of the MAM process, and its evaluation structure. This is the reason for expending the time and energy in the current development of this research paper. Individuals who accept the premises upon which the MAM technique is built, and the postulates of multiple alternatives evaluation and criterion-referenced control -- must also accept the notion of trade-off and preference structure, and optimality. A common problem has frequently been the erroneous conclusion, that the model's decision concerning solution membership is devoid of any negative impact. Other misinterpretations surround the idea of 'what was the problem as defined (?)', and therefore does the solution truly solve the problem, or merely cope with the actual problem's negative impact. For example, closing schools does not solve the problem of declining enrollment, but does permit a rational and accountable means of coping with its effects. Successful execution of the MICROPIK model will provide the best fit of courseware and hardware with desired activities -- but may not be able to meet all of the desired needs. A limitation of a single hardware device, and a particular preponderance of courseware on a particular hardware unit, may require the sacrifice of a single discipline's CAI requirements due to non-compatible software on the preponderant device chosen.

The incorporation of MICROPIK results into practice must never be the result of solely following the binary indicators of the final solution set vector. The modeler must recall, that the membership of the solution vector resulted from a mathematical analysis of a number of criterion-oriented inequalities, which themselves were products of definition, referencing, scaling and measurement -- and therefore all of the problems associated therein. The decision-maker must look upon the MICROPIK results as structured, controlled "suggestions"; and in many cases, just further "input" to the decisioning process which always rests in final form with a flesh and blood person. Contrary to public, wide-spread predictions of doom, technology will never replace the human decision-maker -- although the potential is there to make that decision-maker more valid, reliable and honest.

[Implications for Future Application]

In closing this most laborious but very satisfying project, the forthcoming criticism from individuals who believe nothing (or at least, choose not to) unless it is accompanied with reams of data print-outs, must be addressed, and their concerns fully acknowledged.

A full piloting or field-test of the MICROPIK model, and its resulting effectiveness and efficiency in selecting micr-computer hardware, compatible instructional software (i.e. courseware), and related CAI/CMI curricular objectives and instructional activities -- has as of the date of this paper not been accomplished. In fact, the author is currently developing a greater diversification of criterion needs and references for input to the model. Field-testing of the model is currently scheduled for the autumn of 1982; and is expected to involve a large number of school districts in order to obtain sufficient frequencies of observation to afford the necessary cross-comparisons between model types, and supported software packages. It is also the intent of this author, to involve each of the major hardware and software distributors (as much as possible) in the design, development, construction and final implementation of the MICROPIK model. Obviously, such coordination requires a great deal of lead-time; and much to my chagrin, can not be modeled in a multiple-alternatives setting (or can it?).

Another obviously major portion of the intended piloting of the MICROPIK formulation will depend upon the ability of school districts to define their desired CAI/CMI needs; and then relate these needs to specifically definable and measureable instructional activities. States such as WASHINGTON which have begun concerted efforts to direct each school district to develop "student learning objectives" (SLOs) for each disciplinary or curriculum area, will provide greater facilitation in the final derivation of CAI and CMI curricular objectives and instructional activities. And of course, only those school districts which have the necessary microcomputer hardware and courseware will be included in the project if they so desire. Large purchases for data processing technology is not a priority item for districts who are currently forced to RIF classroom teachers due to budgeting problems.

The interested reader is invited to contact the author, and begin communications which might provide a basis for cooperative ventures in satisfying the uncoming requirements for a full-scale field research. Others are invited to stay tuned to further developments in the MICROPIK process, and its impact upon the general evaluation and decision-making structure currently found in most school districts ... same BYTE time ... same BYTE channel.

Dr. Brent Edward Wholeben
14110 81st Place, N.E.
Bothell, Washington 98011
(206) 821-4468 / 543-4940

University of Washington
College of Education
January 19, 1982