AD-785 668

DEVELOPMENT OF A TEST BED FOR EVALUATION OF SMALL UNIT DOCTRINAL ALTERNATIVES IN THE COMBAT ARMS

Robert T. Root, et al

Army Research Institute for the Behavioral and Social Sciences
Arlington, Virginia

1970

DEVELOPMENT OF A TEST BED FOR EVALUATION OF SMALL UNIT DOCTRINAL ALTERNATIVES IN THE COMBAT ARMS

ROBERT T. ROOT, Ph.D.
LARRY E. WORD, MAJ, USA
U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
ARLINGTON, VIRGINIA

To gain an understanding of human performance in Army jobs it is necessary to study men on the job in a systematic manner in a realistic setting (1). In some instances a job can be abstracted and be brought into the laboratory. However, in the case of the combat arms -- and particularly the maneuver arms (infantry and armor) -- the research laboratory must go to the field. To take the laboratory to the field, there must be a vehicle for realistically simulating what occurs on the battlefield in order to study the contribution to unit success made by the human component. There must also be a methodology for objectively measuring behavior if this contribution is to be determined.

In 1972 the Army Research Institute initiated a program of research directed at developing improved unit training techniques. Initial concentration was placed on techniques for squad and platoon training in the maneuver arms. A complementary research effort was directed at investigating improved methods for unit performance assessment. In both cases research direction was provided by the fundamental principle that training and evaluation must be directly related to the individual's job and the unit's mission. It was a conscious attempt to continue the current movement toward performance-oriented training and evaluation in a realistic job setting.

Two products of this research program provide the basis for a facility for field research, a test bed for simulating the tactical environment. The first is the REALTRAIN Method, a technique for realistic small unit tactical training permitting two-sided, free-play engagements. The second is the Unit Performance Assessment Model, which not only looks at a unit's achievements (in terms of selected objective criteria) but also at the "costs" incurred in carrying out an assigned mission.

THE REALTRAIN METHOD

For men in units to learn to work as one and for tactical training to be meaningful, units in training should engage in two-sided, free-play simulations of the battle situation. The REALTRAIN Method for infantry squads and armor platoons is built around field exercir , permitting a high degree of tactical realism. In infantry and armor situations where it would be natural for two opposing units to be in direct contact, engagement simulation techniques for achieving battlefield casualties were developed. These engagement simulation techniques permit the simulation of the effects of individual and crew-served weapons -- e.g., the M-16A1 rifle, the M60 machine gun, the M72 Light Anti-tank Weapon, and indirect fire weapons. The REALTRAIN Method has introduced a realistic and objective methodology for casualty assessment from which a meaningful determination of unit success or failure can be made.

Objective casualty assessment with the M-16A1 rifle is obtained through the use of six-power telescopes mounted on an individual's weapon and numbers attached to each participant's helmet (see Figure 1). Hits are achieved when an individual engages a member of the opposing force through the telescope and correctly identifies the number on his helmet. This information is passed to the opposing force via a radio net operated by controllers with each group.

Close-in ability to engage armored vehicles and bunkered positions is achieved by the rifle squad using the M-72 LAW (light anti-tank weapon), a shoulder-fired, recoilless weapons system discarded after firing. Assessment of hits with the LAW is obtained by use of an auxiliary sighting system of clear plexiglass discs attached to the ends of an expended casing (see Figure 2). This auxiliary sight, which is boresighted with the gunner's sight, can be used by a controller looking through the tube of the weapon. Thus, by first identifying the range to the target, the controller can objectively assess whether the gunner has a correct sight picture when he fires. Hits are transmitted to the opposing force by a controller-operated radio.

Engagement simulation techniques provide each soldier immediate feedback on his behavior in a simulated tactical environment. Development of appropriate stimulus/response relationships is the key to the optimization of learning which occurs in these exercises. Good tactical behavior is strongly reinforced; poor decisions are immediately evident to all participants involved.

While techniques for simulating casualties such as these have probably sparked the most interest in the method, the model underlying the method, though simple, includes several other important features. As may be seen in Figure 3, realistic situational exercises may be used for both training and evaluation. A particularly important aspect of the method for training is the After Action Review, involving men from both sides of a two-sided exercise, which provides exercise participants further feedback on their performance in the field, much of it from their peers. Exercises may be repeated a number of times. In two-sided exercises the situation will never be exactly the same, providing soldiers the opportunity to practice tactical skills until they are thoroughly proficient. When simple exercises have been mastered, the complexity of an exercise may be increased to include training on additional tactical skills.

Typically leaders and soldiers find that REALTRAIN exercises require all of the skill they possess just to "survive" and ultimately achieve unit success. They are eager, even impatient, to learn more advanced skills. The training stimulates all those involved to originate and implement better tactical procedures to gain an advantage over their opponents.

The motivation found is the result of the application of a number of principles of behavioral science. Participating soldiers perceive that they are learning tactical skills related to their very survival and not just going through some meaningless drill. There is the built-in element of competition which fits the cultural pattern of the average American GI who doesn't like to lose. Then there is the element of credibility; participants know that winning or losing is not based on sheer luck or the subjective judgment of some umpire but on their own behavior.

When participants are thus motivated to learn, the effect of some of the learning principles of performance-oriented instruction is maximized. During an exercise they quickly learn appropriate <u>stimulus-response connections</u>, learning when and where to apply previously-learned skills. Participants pay attention to the <u>feedback</u> received on the quality of their performance, especially as much of it comes from their peers and is most often couched in the vernacular of the present younger generation. <u>Hierarchical learning</u> of tactical skills can be achieved rapidly as the complexity of an exercise is increased. <u>Stamping-in of correct responses</u> is achieved by repetition of an exercise.

The use of field exercises for unit evaluation or for experimental purposes, however, additionally requires a method for objectively measuring behavior observed on the simulated battlefield. Without a measurement methodology, REALTRAIN exercises have limited utility.

THE UNIT PERFORMANCE ASSESSMENT MODEL

The development of the Unit Performance Assessment Model was based on three considerations. The first was that evaluation criteria should be objective and quantifiable, and should be tied as directly as possible to performance on the battlefield. If it were possible to conduct evaluations of unit performance in the field, what would one want to observe and measure?

Present ATT and ORTT exercises tend to stress the planning of a mission rather than the unit's conduct of the mission. This, in part, has been due to the feeling of many that a well-planned mission would logically lead to a well-executed mission. It should be readily apparent, however, that such a notion does not reflect consideration of an intelligent enemy on the other side. In any two-sided battle, planning is important but a unit must execute the plan and be willing, and able, to modify it in response to the actions of the opposition. The emphasis on planning rather than execution has also been due, in part, to the fact that the Army has not had a method for conducting realistic, two-sided field exercises.

In carrying out a mission a maneuver arms unit achieves certain objectives; however, there must necessarily be some expenditure of unit resources in attaining these objectives. It is felt that there are two conceptually independent dimensions to be considered in the evaluation of unit performance. One describes the quality, "goodness", or merit of the unit's achievement; the other describes the "costs" of that achievement. Both achievements <u>and</u> costs should be considered in unit evaluation. An attacking platoon may overrun a defended position. But if in doing so, it suffers extensive casualties, expends most of its ammunition, loses key weapons such as its machine guns, and no longer has available its tactical radios for communication, it would probably not be able to regroup for a subsequent engagement. The costs are too high!

When achievements have been identified, the attempt should be made to define these achievements in terms of observable and/or measurable variables which are the necessary and sufficient defining characteristics for each of the achievements. Cost criteria should similarly be defined in terms of objectively assessable factors. (Less directly assessable factors of cost such as fatigue, fear, and motivation have not been considered for inclusion at this point in model development.)

The second consideration was that the evaluation method should be "product" rather than "process" oriented, working backward from final products (mission objectives) to intermediate products. The analysis of "outcomes" -- achievement products and cost products -- rather than processes is largely a reflection of the fact that in a dynamic battlefield situation with an "intelligent" enemy there may be a number of equally effective ways of achieving a mission objective. The idea of a single "school solution" for a given scenario does not appear to coincide with the realities of battle.

We are concerned with whether a unit was effective, in terms of its achievement, not necessarily how it arrived at these achievements. However, we are also concerned with the unit's "efficiency" relative to effectiveness; i.e., what were the costs incurred by the unit relative to its achievements.

The third consideration was that the model should permit the inclusion of prescribed standards of performance providing a pass-fail ("go-no go") decision capability. Recent advances in instructional technology emphasize the systematic analysis of tasks or knowledges for which training is to be provided. A major step in the "systems engineering" of the tasks to be performed by individuals or by groups of individuals requires the specification of the standards to which and the conditions under which the task is to be performed.

Complementing this trend in instruction is the developing theory of criterion-referenced testing. Persons are evaluated against prescribed standards of performance, not against other individuals as in norm-referenced testing. If the model presented below did not include a provision for entering standards of performance (for both achievement and cost), the resulting index of performance would be in absolute performance terms. It would not be possible to determine how a unit performed relative to desired levels of performance.

If a pass-fail decision is to be made, the performance index should allow the commander in the field to specify a minimum acceptable standard of achievement and a maximum acceptable standard of cost for a given performance situation. Standards do not necessarily have to refer to final on-the-battlefield behavior. They may be prescribed by a field commander for the level of performance he expects a unit to achieve at different stages in the unit training cycle. Differential standards may also be applied for varying conditions on the simulated battlefield; e.g., standards may be changed for a given tactical exercise run at night as opposed to during the day.

The Unit Performance Assessment Model developed on the basis of these considerations provides a performance index, P, that is an objective measure of how well a unit performs in an assigned mission task. If the quality of performance is assumed to be a monotonically increasing function of achievement, for constant cost, and a monotonically decreasing function of cost, for constant achievement, then

$$P = \left(\frac{A - A'_{min}}{A_{max} - A'_{min}}\right) - \frac{C}{C'_{max}} \qquad (1)$$

where A is the obtained achievement score, A_{max} the maximum possible achievement score, A'_{min} the minimum acceptable achievement score, C the obtained cost score, and C'_{max} the maximum acceptable cost score.

Both achievement and cost are measured on a 0 to 1 scale. A'_{min} and C'_{max} are standards of achievement and cost, respectively, which are set prior to an exercise. A_{max}, the maximum possible achievement score, is set at 1. When demonstrated achievement, A, is less than the minimum acceptable achievement, A'_{min}, P is less than zero, and unit performance is not acceptable. Similarly, P is always less than zero when C is greater than C'_{max}. Expressed graphically, the line joining A'_{min} and C'_{max} represents the P = 0 line (Figure 4). Any obtained P to the upper left of the line is positive and indicates passing ("go") performance; an obtained P to the lower right of the line will be negative and indicates failing ("no go") performance. For a unit with unacceptable performance, the unit commander can determine if this performance is due to: substandard achievement (below A'_{min}), too high a level of cost (above C'_{max}), a combination of both, or because of an improper balance between achievement and cost ($A > A'min$, $C < C'max$, but $P < 0$).

The question may be asked whether unit personnel can or should set minimum acceptable achievement (A'_{min}) levels without at the same time considering some acceptable non-zero cost (C'_{min}). And, conversely, can or should they set maximum acceptable cost (C'_{max}) levels without at the same time considering some expected level of achievement (A'_{max}) less than A_{max} for those costs. If A'_{min} is determined relative to C'_{min} and C'_{max} is expressed relative to A'_{max}, the equation for calculating P becomes:

$$P = \left(\frac{A - A'_{min}}{A'_{max} - A'_{min}}\right) - \left(\frac{C - C'_{min}}{C'_{max} - C'_{min}}\right) \qquad (2)$$

The graphic expression for the revised model is shown in Figure 5. The line joining points a and b represents the P = 0 line. Acceptable (pass) performance would be any value for P lying on or above the P = 0 line when A is above A'_{min} and C is below C'_{max}. This modification to Equation 1 has the effect of increasing the area of passing performance, as may be seen by the position of the P = 0 line for Equation 2 in relationship to that for Equation 1 (shown as a dashed line in Figure 5). Although standards for A'_{min} and C'_{max} have not changed, new standards have been introduced which relax the constraints on satisfactory performance. For A'_{min}, some non-zero cost is permitted; for C'_{max}, some level of achievement less than A_{max} is also permitted.

Unit personnel may find that establishing two sets of standards is a difficult task. Therefore, the procedure to be used in the field may not be selected from a consideration of what would be desireable psychometrically, but on the basis of an empirical determination of what officers in the field are willing or able to do.

The model is not intended to be a sophisticated solution.

As the user of the model is to be the field commander with the responsibility for training and evaluating his unit, any assessment technique must be simple and, with procedural aids, easily used. The performance index, P, in the form shown in Equation 1 above should have a certain intuitive appeal to a commander in the field as it considers costs subtracted from achievements and the setting of performance standards for both achievements and costs.

Before the commander in the field can use this method, it is necessary for experienced military personnel at some higher level to select appropriate evaluation criteria for a given type of exercise, following the procedure outlined below. Computations must then be carried out to derive weights for each criterion, reflecting its relative contribution to mission success.

Candidate achievement and cost criteria for a given scenario (for example, a rifle platoon in the attack) are selected on the basis of objectivity, quantifiability, and orientation to "real world" situations. Experienced military personnel rate each of the candidate achievement and cost criteria as "Very Important" (in terms of successful mission performance), "Important", or "Relatively Unimportant". Candidate criteria ranked as most important are selected for further consideration. This step provides an indication of the perceived importance of each criterion measure, but does not give a quantitative measure of their relative contribution to successful mission accomplishment.

To obtain these data military experts rank order cases of possible achievement (or cost) outcomes for hypothetical values of criteria selected in the previous step. They are asked to rank order the cases in terms of the desireability of the outcome. Table 1 shows an example for sixteen achievement "cases".

TABLE 1. Example of Situational Cases for Attacker's Achievements to be Rank Ordered by Military Experts in Terms of Desireability of Outcome

Achievements	Possible Outcomes					
	1	2 ...	8	9 ...	15	16
Objective Taken (1=Yes, 0=No)	1	1	1	0	0	0
% Def. Casualties	100	80	0	60	0	0
% Def. Key Weapons Lost	100	0	100	100	50	0
% Def. Ammo. Expended	100	90	15	90	0	0
% of Obj. covered by fire	100	100	100	85	95	0

For scaling purposes an unlikely desireable outcome (Case 1) as well as an unlikely undesireable outcome (Case 16) is included in the set of cases. When a set of outcomes has been ranked by expert military judges, the rankings are scaled using the normalized-rank method (2) to develop ratio scales of achievement (or cost). A multiple regression analysis is used to derive linear equations expressing achievement (or cost) scale values as a function of the selected criteria. Linear equations for obtaining achievement and cost scores for use in Equation 1 will appear as shown in the following examples:

$$A = .50x_1 + .30x_2 + .10x_3 + .05x_4 + .05x_5 \qquad (3)$$
$$C = .40y_1 + .20y_2 + .20y_3 + .10y_4 + .10y_5 \qquad (4)$$

The variables represent the values for each criterion measure obtained during an exercise (or those established *a priori* by a unit commander in setting A'_{min} and C'_{max}). The numerical weights in Equations 3 and 4 reflect the relative importance of each criterion measure. If the variables are measured on a scale between 0 and 1, A (or C) will be 1 when all of the achievement (or cost) criteria take their maximum value of 1.

A brief example will help to clarify the procedure for calculating P, as it might be done in the field. If the objective were to evaluate, through a situational exercise, the performance of a rifle platoon attacking a position defended by a rifle squad, the following results shown in Table 2 might be obtained.

TABLE 2. Achievements and Costs for a Rifle Platoon Attacking a Position Defended by a Rifle Squad

Attacker's Achievements

x_1: Objective taken (1)
x_2: Defenders's casualties (50%)
x_3: Defender's key weapons lost (100%)
x_4: Defender's ammunition expended (80%)
x_5: Percent of objective covered by attacker's fire (100%)

Attacker's Costs

y_1: Attacker's casualties (40%)
y_2: Attacker's ammunition expended (70%)
y_3: Attacker's key weapons lost (50%)
y_4: Attacker's other weapons lost (60%)
y_5: Duration of engagement (2 hours = .25)

Prior to the exercise the unit commander establishes a minimum acceptable performance standard for each criterion. In this example when these are entered into the equation for A (Equation 3), a minimum acceptable achievement, A'_{min}, value of 0.17 might be obtained. Similarly, he establishes a maximum acceptable cost for each cost criterion and a maximum acceptable cost, C'_{max}, of -0.70 might be obtained from Equation 4. (Since costs are losses, they are given negative signs.)

Entering exercise results in the equations for A and C (substituting decimal equivalents for percentages) one obtains A = 0.81 and C = -0.51. Then entering the above values in the equation for P, the following result is obtained:

$$P = \left(\frac{.81 - .17}{1.00 - .17}\right) - \frac{-.51}{-.70} = +0.04$$

Thus, for the standards set before the exercise, the hypothetical unit performed at an acceptable level.

The final step in the full development of the Unit Performance Assessment Model is to validate it in the field. Ironically, the criterion against which the model will be validated (in a series of training exercises, for example) will be judgments of proficiency obtained from military experts. However, care will be taken to ensure their judgments are as objective and quantitative as possible and to minimize subjective bias.

THE TEST BED

Tactical doctrine for small maneuver arms units has changed slowly over the years. Many of the changes that do occur result from experience gained in the most recent conflict in which our armed forces have engaged. Combat has been the "test bed" for shaping unit techniques and tactics. In this sense, then, a great deal of present maneuver arms doctrine represents a remnant of the past. Small unit doctrinal alternatives appropriate for possible future conflicts have generally not been open for test.

It has been known for some time that task force organizations composed of infantry, armor, and anti-armor elements offer the best chance of success on a conventional, mid-intensity battlefield. Despite the fact that the cross-attachment of weapons systems and men to support them is commonplace in the Army, tactical doctrine has been developed for each type of unit in relative isolation, with little attempt to achieve an integrated approach to tactical employment and deployment. With little opportunity to try out various alternatives for optimizing reaction time and weapons effectiveness for a combined arms force in a situation representative of the battlefield of the future, doctrine in this instance has developed through committee action.

Placing a representative combination of units in a realistic, simulated performance environment will permit those charged with developing new doctrine to study the effects of different techniques of unit employment and deployment. Interactions among personnel, equipment, and tactical conditions can be observed. Limited try-outs in the field will lead to the development of alternative hypotheses for courses of action. For instance, hypotheses related to alternative techniques for employment of infantry and anti-tank personnel in advance of armor elements in the attack could evolve during the developmental stage. Hypotheses could then be experimentally tested under controlled conditions. Control of certain factors, such as terrain or time of day, while varying technique of employment and force ratio, for instance, would allow the specification of the variance in the situation attributable to each independent variable. Attention to the development of hypotheses and the use of efficient experimental designs for testing them would result in a recommendation of practical significance based on sound empirical data. An additional consideration in any doctrinal investigation must be an attempt to determine the influence of non-cognitive variables on unit performance.

While this discussion of the test bed has emphasized its utility in the evaluation of doctrinal alternatives for purposes of example, there are a number of other areas of potential utility for the test bed. Very briefly, they are:

- empirical determination of unit performance standards
- field assessment of the training value of training devices and other simulation techniques
- validation of paper and pencil aptitude tests for maneuver arms samples
- validation of simulated (paper and pencil) performance tests for the maneuver arms
- experimental determination of the influence of non-cognitive variables on battlefield performance
- empirical study of the role and influence of leaders on unit performance
- development and evaluation of new engagement simulation techniques (e.g., improved laser engagement simulation).

This last research area, it may be seen, in turn will serve to increase the sophistication of the test bed, and hence, its utility, in addition to providing new, innovative training, evaluation, and doctrinal techniques to units in the field.

In summary, an experimental laboratory in the field is possible if men and equipment in the simulated environment are required to perform actions demanded on the real battlefield; if the sights and sounds of battle are realistically simulated; and if two-sided engagements may be played when the situation calls for direct confrontation between opposing forces. Methods of measurement of unit performance must similarly take into account the critical factors, or criteria, that determine success or failure in battle. REALTRAIN exercises provide the tactical environment; the Unit Performance Assessment Model offers promise as an objective and quantitative means for measuring unit performance.

BIBLIOGRAPHY

1. Uhlaner, J.E., Human performance, jobs, and systems psychology. BESRL Technical Report S-2. October, 1970. (AD 716 346)

2. Guilford, J.P., Psychometric methods. 2d ed., New York: McGraw-Hill, 1954.

Fig. 1: M-16A1 Rifle Simulation Technique

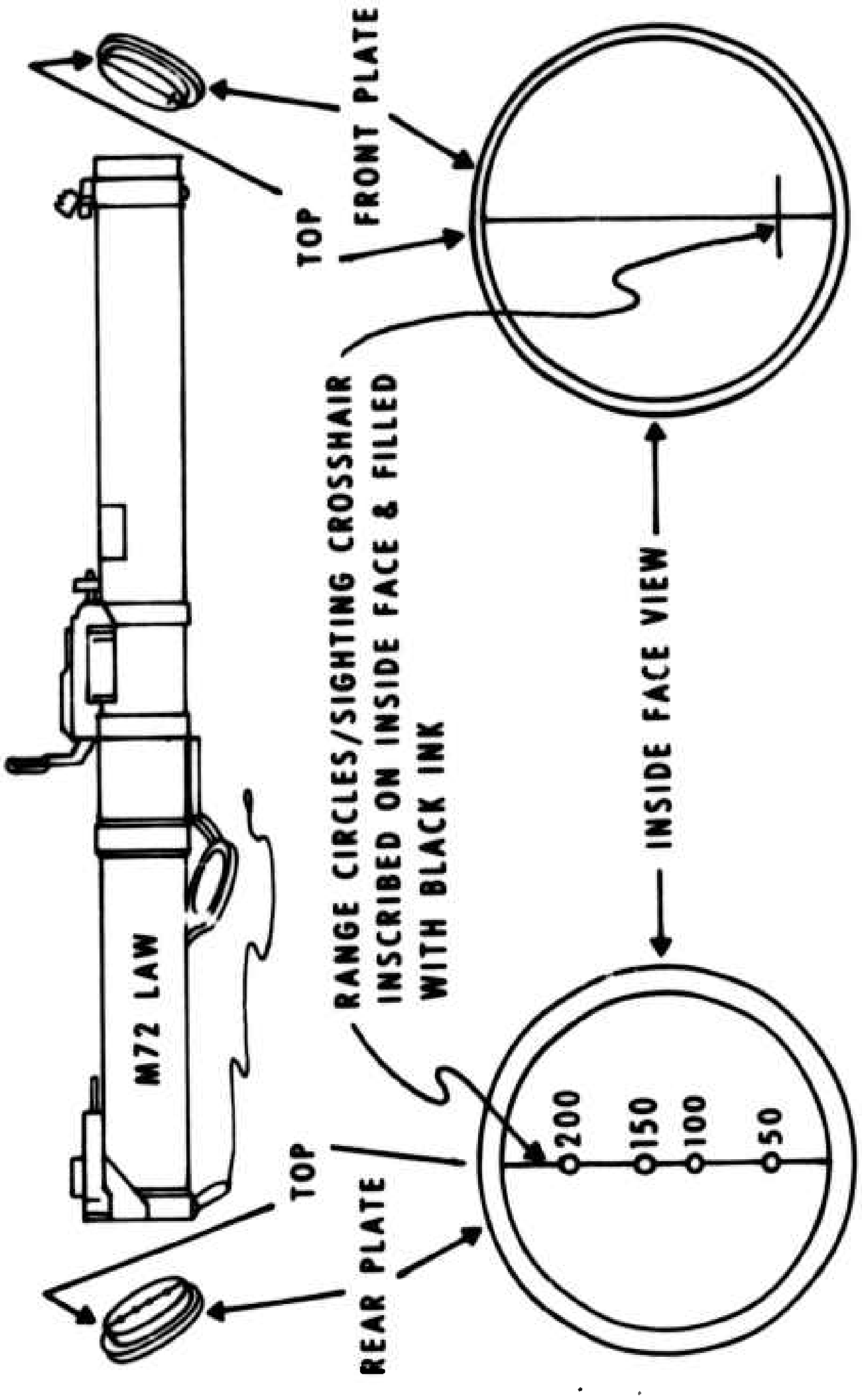

Fig. 2: Light Anti-Tank Weapon Simulation Technique

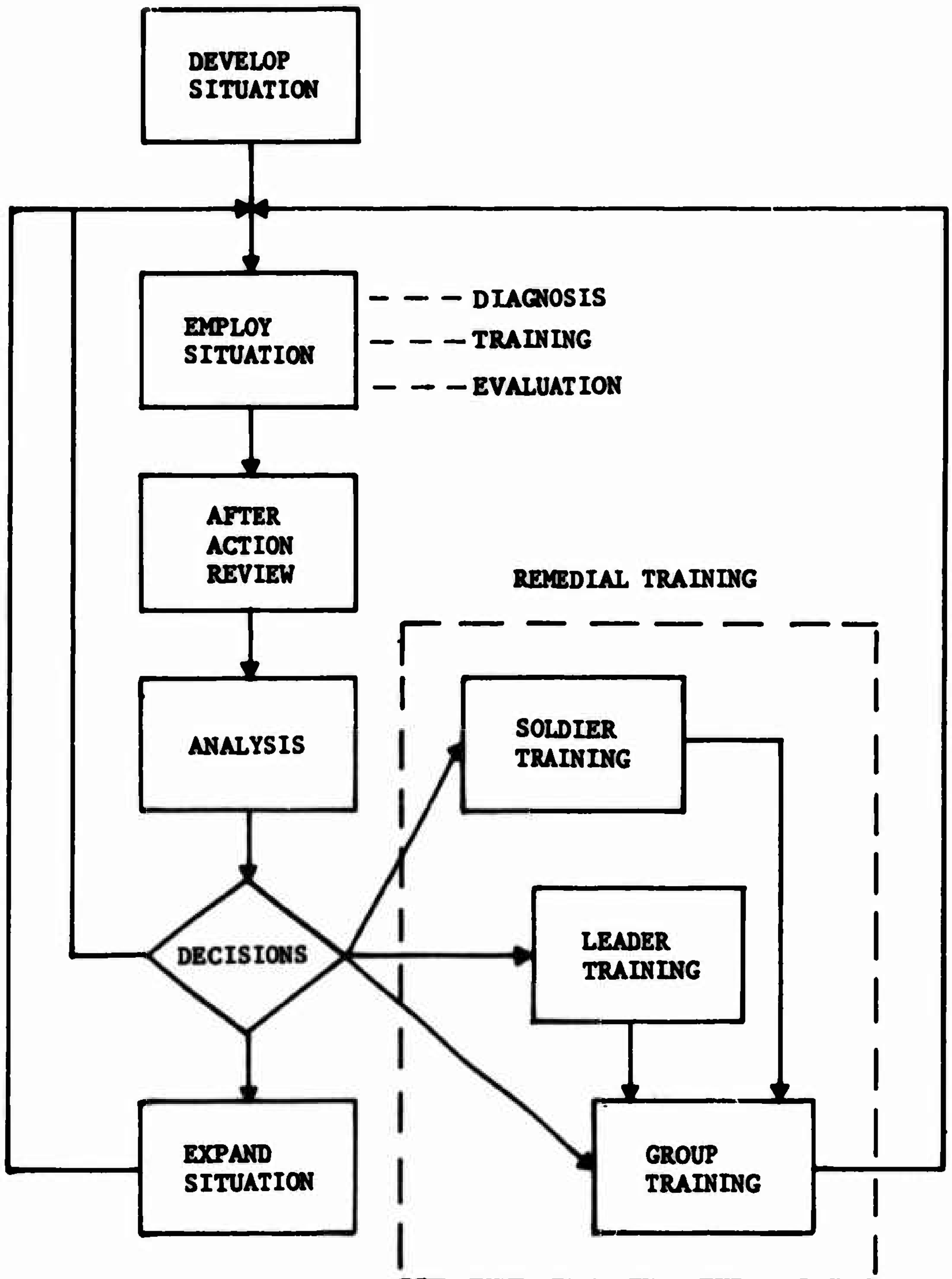

Fig. 3: The REALTRAIN Model

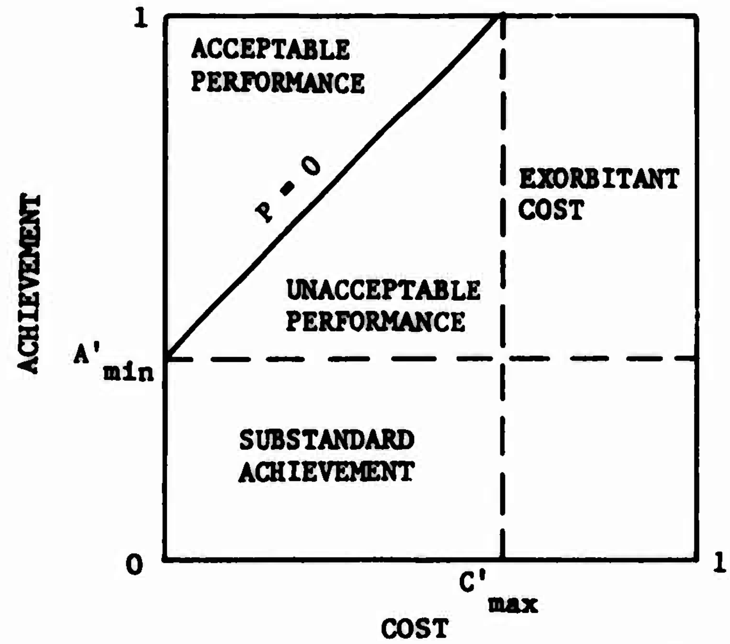

Fig. 4: Graphic Representation of the Unit Performance Assessment Model

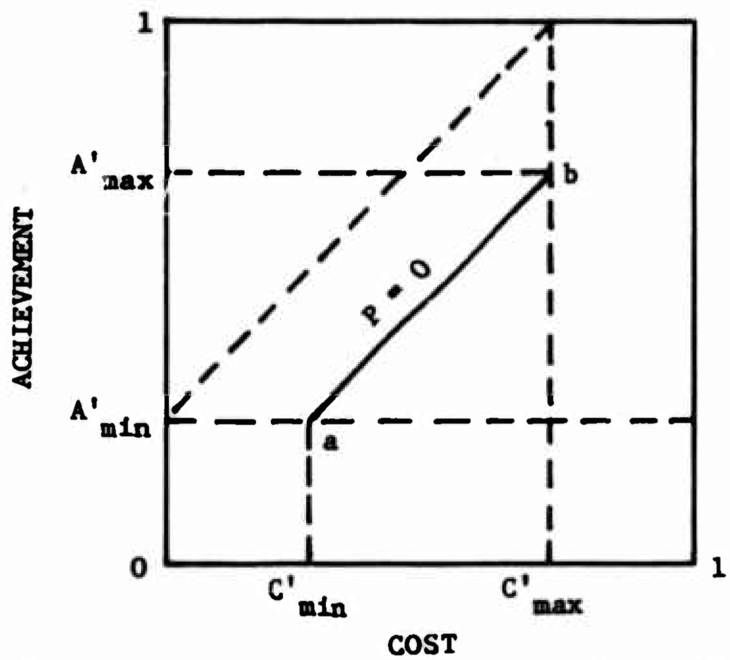

Fig. 5: Graphic Representation of the Revised Unit Performance Assessment Model