

ARI Research Note 99-18

Training-based Requirements for Semi-Automated Forces

Jim Kornell
Syukhtun Research, Inc.

Research and Advanced Concepts Office
Michael Drillings, Chief

March 1999



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

Approved for public release; distribution is unlimited.

DTIC QUALITY INSPECTED 1

19990310 034

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

A Directorate of the U.S. Total Army Personnel Command

**EDGAR M. JOHNSON
Director**

Research accomplished under contract
for the Department of the Army

Syukhtun Research, Inc.

Technical review by

Tony Papa

NOTICES

DISTRIBUTION: This Research Note has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This Research Note may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this Research Note are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE March 1999		3. REPORT TYPE AND DATES COVERED FINAL 2 Aug 93 - 30 Jan 94	
4. TITLE AND SUBTITLE Training Based Requirements for Semi-Automated Forces				5. FUNDING NUMBERS PE 0605502A 2O665502M770 TA 6901 WU C05 Contract No. MDA 903-93-C-0144	
6. AUTHOR(S) Jim Kornell (Syukhtun Research Inc.)					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Syukhtun Research 2740 Williams Way Santa Barbara CA 93105				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute ATTN: PERI-BR 5001 Eisenhower Avenue Alexandria, VA 22333-5600				10. SPONSORING/MONITORING AGENCY REPORT NUMBER Research Note 99-18	
11. SUPPLEMENTARY NOTES COR: Frank Moses					
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for Public Release; Distribution is Unlimited.				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words): Effective use of current and planned semi-automated forces (SAF) capabilities is important to maintain Army readiness. To guide commanders in specifying training requirements using SAF, a model of how instruction and training can use SAF is needed. The objective of this research reported here has been to build a foundation for such a model. The long-term goal is to construct a knowledge-based system to aid commanders in translating training objectives to training requirements.					
14. SUBJECT TERMS semi-automated forces instruction training management					15. NUMBER OF PAGES
					16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited		

93014410212 0/1

Technical Report xxx

Training-based Requirements for Semi-Automated Forces

Jim Kornell
Syukhtun Research, Inc.

August 1994

For Contracting Officer's Representative
Frank Moses, Ph.D.
US Army Research Institute for the Behavioral & Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

The views, opinions, and findings contained in this report are those of the author and should not be construed as an official Agency position, policy, or decision, unless so designated by other official documentation.

TRAINING-BASED REQUIREMENTS FOR SEMI-AUTOMATED FORCES

CONTENTS

Title Page

Foreward	i
1 The Problem	1-1
1.1 Motivation	1-1
1.2 Opportunity	1-1
1.3 General problem	1-2
1.4 Phase I research objective	1-4
2 Phase I Scope	2-1
2.1 The nature of the problem	2-1
2.2 Users	2-2
2.3 What Phase I isn't	2-2
3 Context	3-1
3.1 Instructional systems design.	3-1
3.2 Epistemology of knowledge outcomes	3.2
4 Structure & theory of the problem space	4-1
4.1 Bounding the problem: Content and abstraction	4-1
4.2 The theory: Assumptions and axioms	4-5
4.3 The theory: Constructs and features	4-7
4.4 Critique of the constructs	4-18
5 Taxonomy	5-1
5.1 Structures of constructs	5-1
5.2 Content => Instruction heuristics	5-7
5.3 Representative construct clusters & heuristics	5-17
5.4 Critique of the taxonomy	5-24
6 Method	6-1
6.1 Ontology	6-1
6.2 Overview	6-4
6.3 Stage 1: Characterize objective	6-5
6.4 Stage 2: Develop training approach	6-6
7 Reading topographic maps: The method applied	7-1
7.1 Stage 1: Characterize the task	7-1
7.2 Stage 2: Develop training approach	7-3
7.3 Analysis	7-4
8 Conclusions & recommendations	8-1
8.1 Critique of the theory	8-1
8.2 Recommendations	8-1
8.3 Finis	8-3
References	R-1
Appendix A: Brigade/Battalion task decomposition	A-1

List of figures

Figure 3-1	Gagné's knowledge outcomes	3-3
Figure 5-1	How we learn	5-1
Figure 5-2	How we train	5-3
Figure 5-3	Content	5-4
Figure 5-4	...And more content	5-5
Figure 6-1	Events, agents, and objects, as a problem space .	6-2
Figure 6-2	The Chi/Slotta ontology	6-3

Semi-automated forces (SAF) are human-controlled computer-based simulations built to work in real time with manned exercises. Use of SAF in Army training is expanding, especially for collective tasks (e.g., STOW-94). SAF offers reductions in cost, improved traceability, and far more scenario flexibility than corresponding live fully-equipped group exercises.

Effective use of current and planned SAF capabilities is important to maintain Army readiness. To guide commanders in specifying training requirements using SAF, a model of how instruction and training can use SAF is needed. The objective of the research reported here has been to build a foundation for such a model. The long-term goal is to construct a knowledge-based system to aid commanders in translating training objectives to training requirements.

Trainers looking here for specific results, such as a new metric for the interpretation of National Training Center data, will be disappointed. We believe that before such metrics should be codified and promulgated in a knowledge based decision aid, an explicit model of training with SAF should be developed. This does not mean that discovery of such metrics is not a critical task. But a knowledge-based system is a *knowledge-level model* of human reasoning (Newell 1982). Constructing this model helps such metrics be productively integrated and applied (Clancey 1992).

This kind of knowledge-level model building is difficult, and may seem very distant from the day-to-day concerns of the training community. The principal author recalls a talk at an expert systems conference in 1984. A speaker glibly claimed, to great laughter and applause from the mostly-industrial audience, that the ivory-tower types who wanted to talk about epistemology should be ignored. The ensuing decade has shown the speaker to be exactly wrong. The knowledge systems that have worked have been just those that have taken epistemology seriously (McDermott et al 1986; Hayes-Roth et al 1992). This report addresses the intellectual foundations needed for a successful knowledge based system. Some trainers and training designers might consider the issues discussed here to be peripheral or irrelevant. There is strong sentiment in the knowledge systems community that they are not.

Another spur to careful development of intellectual foundations is the dual-use mandate of the SBIR program. There is no commercial market for SAF-based training. There is a commercial market, though, for translating complex training objectives to specific requirements. A general foundation increases the potential for dual Army and commercial use.

Three assumptions are made:

- There are fundamental regularities in human learning
- There are corresponding regularities in how effective training and instruction are carried out
- These regularities can be articulated and applied productively to novel instructional and training circumstances and needs.

A fourth assumption is also made, namely that there is a continuum linking training and instruction. Training is generally taken as in-the-field, hands-

on practice. Instruction by contrast is thought of as a classroom activity. Some Army commanders might emphasize this distinction as the difference between talk and action. However, there are many intermediate cases--the training of command decision-making, the increasingly technological nature of weapons, and the provision of sophisticated automated support to lower and lower levels of the command hierarchy, even to the individual soldier. To argue about whether a tank commander in a SIMNET exercise is being instructed or trained seems pointless. We assume the continuum view.

This work was performed under Contract MDA903-93-C-0144 for the Army Research Institute for the Behavioral and Social Sciences (ARI). It was very helpfully supervised by Dr. Frank Moses. Dr. Angelo Mirabella, also of ARI, also generously devoted time, effort, and insightful commentary. James Kornell, of Syukhtun Research, Inc., is the principal author. Contributors include Dr. James Geiwitz, of Advanced Scientific Concepts, Inc.; Dr. Susan Fischer, of Anacapa Sciences, Inc.; and Lt. Col. Ken Bell, ret., of THETA Technologies, Inc.

All conclusions and procedures in this report are solely those of the authors and are not in any way an official position of the US Army or of ARI or its employees.

Selection, implementation, and validation of training activities are a commander's most important peacetime activities. Constantly improving the results of those activities is a high priority.

In the Phase I investigation detailed in this report, the feasibility of developing a method to generate training requirements from training objectives has been explored. This section describes the motivation, opportunity, and general outlines of the problem. It closes with a definition of "training requirement" in the context of semi-automated forces and a succinct statement of the Phase I objective.

1.1 Motivation

Training can separate victory and defeat. The odds of winning an offensive "increased 30 to 1 by belonging to a high as opposed to a low training-ratio population" (Hart & Sulzen 1988). The Army officially states that "the training of soldiers, leaders, and units to win in combat will remain the Army's single most important task" (US Army FY93 Posture Statement). Estimates vary from \$5 to \$20 billion a year in Department of Defense spending on training.

Demands for intensity and amount of training are increasing because of changing mission requirements. For example, there may be much more "peacekeeping", e.g., trying to mediate civil wars or warlord conflicts. There may be significant border control. Migration management could be needed to avoid serious regional destabilization. Restoration of civil order may be required. We may be called on to perform "paramilitary" actions, e.g., drug production control. Future missions may force us to confront shifting and chameleon-like enemies, guerilla warfare, city fighting, blurred lines between combatants and noncombatants; even blurred lines between military and civil actions, e.g., as in the Rwanda relief effort. Some influential analysts, such as Dr. Andrew Marshall of the DoD Office of Net Assessment, predict even more far-reaching and radical changes in mission (Wall Street Journal, 15 July 1994).

But budgets are shrinking and personnel being reduced. To reconcile this discrepancy, we need to link training objectives more precisely and efficiently to available training resources. The goal of this research has been to develop a method for designing efficient training that satisfies training objectives.

1.2 Opportunity

While more will be demanded of training, it is also true that techniques and tools are emerging that greatly expand the potential power available to designers of training. There is a growing understanding of subtler levels of human performance, and there are emerging technologies that make possible previously unaffordable or impossible training approaches.

Improved methods of task analysis allow improved training design. Cognitive task analyses can lead to better understanding of the roots of performance, performance breakdown and stress points, and better design of appropriate training materials (Lesgold et al 1986). Metacognitive analyses and explicit

training of domain-specific metacognitive techniques may lead to significant improvements in task performance [Geiwitz 1994]. Over the next decade, group cognitive task analyses will emerge to provide tools and techniques for better understanding of, and therefore training of, group behaviors and performance. There is increasing interest in and study of the factors that determine success in field exercises [Mirabella 1993]. We have barely begun to scratch the surface in applying this emerging knowledge to improve battlefield competence.

Complementing the advances in cognitive science and the study of decision-making are advances in computing technology. Developments in both hardware (e.g., easily available processing and visualization capabilities, available for \$2- to \$3-thousand dollars, far beyond those of a mid-80's VAX 780) and software (e.g., object oriented programming, cross-platform object interoperability, powerful software tools) make creation of interactive training-focused simulations and their use in the field both more affordable and more powerful.

SAF is a visible and important product of this convergence. The Defense Modeling and Simulation Office (DMSO) credits the term "semi-automated force" to ARPA's Advanced Simulation Technology Program's effort to provide large software-driven forces to interact with manned simulators on the SIMNET battlefield [Brooks et al, 1989]. SAF is also called SAFOR, as in the title of this report. (Because SAF seems to be replacing SAFOR in conventional usage, SAF rather than SAFOR has been used in the body of the report.) CGF (computer-generated forces), IFOR (intelligent forces), and AFOR (automated forces) are other synonyms. DMSO characterizes SAF as having four basic attributes (DMSO 1993):

- Real-time interface
- Entity-level representation of combat elements
- Distributed Interactive Simulation (DIS) compliance (DIS is a set of technical protocols for real and simulation entity interactions)
- Credible surrogate for manned forces

Examples of SAF include the BBN SAF 4.3.3, ModSAF, BDS-D CGF, IST SAF, IFOR/WISSARD, CCT SAFOR, SWEG/SUPPRESSOR, and Janus A. There are many SAF-related programs in development: IBM Blackboard Research Study, UCCATS/JCM/AMOEB, BBS/DIS, Eagle, RESA, and MARS. CASES, JSIMS, and WARSIM 2000 are also related.

Further, projects currently in infancy may become important SAF elements. In particular, the Joint Warfare Simulation Object Library (JWSOL), a DMSO- and ARPA-sponsored effort to develop general-purpose simulation objects for use in a wide variety of contexts, has the potential to dramatically improve the flexibility and speed with which training simulations can be constructed and fielded. The current Principal Investigator is also knowledge engineer for JWSOL (Kornell et al 1994).

1.3 General problem

The context is increasingly complex missions and shrinking budgets on the one hand, and new analytic techniques and the capability to implement their results on the other. The core problems of training remain: how to determine what to train, how to efficiently and effectively train it, and how to validate the training.

The Army emphasizes battle-focused training. Obviously, training should result in maximal positive change in behavior with minimum negative transfer. To achieve this, the Army provides a comprehensive set of training programs and guidelines for training strategy.

Three positive training results are desired: improvement in competence, with minimal negative transfer; efficiency and generality of transfer; and extended duration of normal retention. At the same time, three resource issues should be optimized: minimizing training time; reducing needed training facilities or substituting more easily available for less easily available facilities; and minimizing the need for training personnel or the time required of training personnel (FM 25-100).

For all the reasons discussed in the previous two sections, training approaches need continual revision. Determining the best way to train is difficult. While performance data is central to designing training, the data themselves are complicated, ambiguous, incomplete, and may unintentionally incorporate observer bias. Data interpretation is also problematic. Mental and methodological tools for testing and validating hypotheses about causal relationships are unreliable. Validating the effectiveness of training is also difficult.

Collection of performance data requires integration of partial evidence from a large number of sources, with varying degrees of similarity to the actual conditions for which performance is to be trained. Warfare is not static nor mechanistic, so fundamentals like "concentration of force" may not only mean different things in different circumstances, but may have widely differing military values in different circumstances. (Contrast operations in the Gulf War with those in Somalia.) Since a wide variety of missions may be required, the significance of performance shortfalls in one area or another may not be at all clear. For example, a propensity toward independent action may be extremely good in a guerilla situation and extremely bad in a massed attack situation. Since most military actions require a wide variety of skills, knowledge, and competences to be brought together, accurately discriminating the source of breakdowns is not easy. Dr. James Banks of ARI/Monterey summarizes some of the issues:

The battlefield is a tightly coupled, interactive, nonlinear world.. Usually there is not a simple single reason why a task is performed well or badly. There may be a constellation of contributory factors. Even if a task is performed badly, frequently it is possible to later compensate for it. Or the enemy action can cause you difficulties, regardless of how well or badly you did your previous tasks...

I think the ability to diagnose causes for performance also varies by the size and complexity of the team. For a complex team with lots of resources, there are lots of ways to perform a task. For example, if I postulate "attrit the enemy at long range" as a defense mission, a Battalion task force can do this with its maneuver elements, tactical air support, attack helicopters, artillery, and so on. And, if I don't succeed with one method, if I am aware enough and flexible enough I may be able to succeed with another method. The causes of success or failure may, therefore, not at all be obvious and quite distal from a given event... There is a large chance element on the battlefield that enormously complicates interpretation of performance and results.

Even if the data were better behaved, the problem of interpretation would

remain. In any engagement scenario, real or simulated, multiple interacting events produce multiple outcomes. Even assuming accurate and complete measurement of events (which of course has just been ruled out), connecting events and outcomes is very difficult using traditional approaches. Dr. Angelo Mirabella of ARI/Alexandria points out that "... The whole is not necessarily the sum of its parts... Even if it was, we would still face the problem of how to recombine the pieces for analysis, validation, improvement ... of training strategies" [Mirabella op cit.].

Training validation is problematic for three reasons. First, and most important, the only true test of combat performance is actual combat. The second is self-reference, and is a result of the constraints imposed by the first. Self-referential training can occur because the training validation measures are based on the content of the training rather than the real-world performance to be trained, or because the training is shaped to fit preexisting measures of effectiveness. For example, if the goal of training is allowed to become too specific to good performance at the NTC, the the Commander responsible for that training is training the test, not the subject. When training is self-referential, good training results become detached from good combat results. Third, for complex skills and for those involving group behavior, performance is very difficult to measure. Observation and data-gathering are difficult because of the number of variables involved and the conditions in which observation must take place. Inference about underlying causes of performance are similarly difficult, because multiple competing hypotheses may all to a greater or lesser extent be true, and because observers' biases may predispose the same evidence to be seen, remembered, and interpreted very differently by different evaluators (Loftus 1975).

1.4 Phase I research objective

In this section, we have briefly put forward the unexceptional arguments that:

- Missions are increasing in complexity, with the promise of at least near-term and possibly long-term acceleration in this increase
- Resources are limited
- Characterizing competence and identifying its sources is very difficult in a broad mission context, especially for group tasks
- Determining how to train effectively is therefore fundamentally difficult.

In response to these circumstances, this research is aimed at *development of a method to translate training objectives into training requirements, in the specific context of SAF*. "Training requirements" are conceived as the intermediate between training objectives and training implementation requirements. A training objective is, for example, to develop command, communication, coordination, and execution skills for Brigade-level emergency humanitarian relief delivery (as in Rwanda and Zaire). Corresponding training requirements would include a detailed scenario of events, players, and roles, and an inventory of available resources. Training requirement implementation would then identify and schedule specific Brigade elements, computer-based simulations, and support elements to perform the exercise.

Since SAF is clearly at the implementation level, why should it be explicitly made part of a method aimed at requirements? The reason is simple. SAF enables scenarios that would otherwise be unrealistic. Explicit awareness

of current and projected SAF capabilities changes what can be included as potential requirements. Therefore, it is intrinsic to the research.

The Phase I objective has been to lay a stable intellectual foundation for the objectives-to-requirements translation method. In Phase I, we have attempted to establish epistemological and ontological foundations for the method, and to taxonomically decompose training requirements to make them methodologically tractable. In the next section, this objective is described in detail.

The goal of this research has been to develop a method to support the translation of training objectives into training requirements. The Phase I approach has been to model the foundations of the problem.

2.1 The nature of the problem

The problem is to generate training requirements from training objectives. There are two subproblems:

1. Characterize the structures of the subject task germane to the training objective.
2. Determine how best to train the objective.

The context is as specified in the previous section.

2.1.1 Characterize the structure of the task

"Task structure" comprises the various characteristics of the task, the degree to which different parts of the task hold these characteristics, and the relations between the characteristics along with how those relations are resolved in expert task performance. Determining task structure is a diagnostic problem. It is also the place for framing the domain problems to which the method (described in Section 6) will be applied, i.e., selection of some characteristics and values as salient, rejection of others. Framing a problem is often the single most important element in its solution.

For example, in interpreting topographic maps, the two principle components are (1) correctly reading the physical features from the topographic representation, and (2) interpreting these features with respect to military advantage and vulnerability. There are a number of visual/spatial patterns germane to (1); good performance in (2) is contingent on accuracy in (1). True experts appear to "see through" the visual features of the maps and the spatial features they represent, directly to military significance. Framing the topographic map reading problem as "find the minimum fuel expenditure route from A to B" leads to very different training than framing it as "find the minimum fuel route from A to B, taking cover and concealment and potential enemy positions and intentions into account."

2.2.2 Determine how best to train the objective.

This is a prescriptive problem on two levels, the requirements and the implementation. Training *requirements* and training *implementation* requirements are different. A training objective for which improvement follows the Power Law—learning to calculate force ratios, for example—may have a training *requirement* for basic instruction in technique followed by repetitive practice. That requirement may be *implemented* via classroom instruction and written materials, or by manipulation of physical tokens, or by a computer-aided instruction course, or by a computer game in which progress depends on correct answers, or in some other way entirely. The factors that influence choice of *implementation* are situational, but the factors affecting the *requirement* are intrinsic to the nature of the task and the human competences for which it calls.

2.2 Users

Target users for the Phase II knowledge based system product are training designers and commanders charged with training their subordinates and units. Program managers considering allocation of training simulation development/modification funds might also use it. These groups have different needs. Commanders are indicated because that is where the need is most acute; Army doctrine assigns direction of training to commanders. Each commander must assess training needs, select training objectives, develop training strategy and plans, execute training, and perform post-training review. However, commanders are not trained in training: "We don't train them to train" (Lt. Col. Ken Bell, former CGSC instructor). Program managers have a strong stake in achieving maximal return from their simulation investment dollars. There are currently no generally accepted criteria by which program managers can judge whether a potential simulation is likely to result in actual improvement in battlefield performance. Nor are there criteria to judge which elements of a proposed simulation are essential and which are incidental (or even potential vectors of negative transfer).

2.3 What Phase I isn't

In the following subsection, we go into some detail regarding what is not included in this research. For most research activities, such detail would be superfluous. In this case, the bounds demarking what should and should not be included are somewhat fuzzy. How much needs to be known, in how many subject areas, to generate a high-quality set of requirements? Because this report is intended to serve as the foundation for Phase II knowledge-based system development, explicit justification for choices about inclusion and exclusion is appropriate.

The diagnostic step of selecting what task(s) to train is not included. This is a very difficult diagnosis, as indicated in the previous section: causes of breakdown or impasse can be proximate or distal, of intention or execution, even completely outside the competence of the unit which experiences the breakdown. The diagnostic process is rich with potential for investigation. How does a commander discriminate among and select the best training objectives? By what criteria are they judged? How are the various elements of collective performance assessed to make the diagnosis? There are (can be) no statistics, but it is plausible that, unit by unit, a very high proportion of the training performed in the Army is *not* directed at the skills most in need of work. Considering that for what is in many respects a simpler task, highly trained medical doctors using expensive diagnostic equipment are wrong roughly 50% of the time in diagnosing what prove to be fatal diseases in hospital settings (Eddy 1982), it would be unexceptional if faulty diagnosis by commanders of training needs is the norm rather than the exception. This is not a condemnation of commanders' performance so much as a statement that there is inadequate support for an extremely difficult job. This is potentially a rich area for follow-up, but it is outside the scope of the present work.

We wish to emphasize this decision. Saying we will develop requirements but not support training need objective diagnosis and generation is analogous to suggesting construction of a medical expert system to support prescription but not diagnosis. That would be an unusual system. In the present case, we

believe exclusion of diagnosis is justified by its complexity and scope.

It is possible that the method, described in Section 6, might be used as an aid to validation of training objectives. That is, mismatches between observed behavior and the internal structures of the task(s) selected for training could highlight errors in task identification. For example, if there are observed behaviors for which there are no associated structural or procedural elements, then the identified training task does not account for the observed behavior, and is unlikely to improve performance. Alternately, if it is seen that a breakdown in only a small part of the fundamental task structure can account for observed behavior, then a great reduction in the total training needed might be possible.

Characterization of the learners is also omitted. Understanding the target learners is critical to designing training. However, the scope of the current effort is already adequately broad. We assume that any enabling objectives that result from student assessment can simply be considered as subobjectives to which the method can be applied.

There are levels of ambition associated with or applied to training objectives. For some training, the goal is to replicate expert-level performance. Command and medical education are examples. For some training, it is to achieve a functional but non-expert level of performance; military driving is an example. For some training, it is simply to install a floor under which competence will not fall; e.g., "park the Jeep right away if the oil light comes on". We assume the target level is known to the user of the method, and that the training objectives (both enabling and terminal) are appropriate to the target level.

The method is not a replacement or substitute for task analysis and cognitive task analysis. In reading through the method, it may seem at times like task analysis/cognitive task analysis approach. It is not. Task analysis should articulate the *operational* content of a task. We have tried to develop a *descriptive* characterization of the task—one level up in abstraction—that will allow task classification and selection of appropriate training techniques.

Nor is this a replacement or substitute for instructional systems design methods. Rather, it is intended for use with them. ISD methods are reviewed in the next section.

The method is is not intended as a process re-engineering tool. "Process re-engineering" occurs when information systems analysis of group tasks reveals correctable inefficiencies, and shows new ways of organizing work processes to achieve desired ends. Process re-engineering is an issue emerging with increasing frequency as computers are embedded ever more deeply into military and civilian tasks. The factors and processes described in the next sections could support process re-engineering efforts, but the work is not intended as a process re-engineering tool *per se*. In other words, the intent is not to change the task.

Understanding the current state of the art in instructional systems design (ISD) is necessary context for the current work, and will help in establishing the requirements for the method. In this section, ISD--the systematic analysis of instructional needs and development and testing of instructional methods--is described and critiqued, and the relation of the current work to the ISD process is specified.

3.1 Instructional systems design

Gagné , Briggs, and Wager [1992] propose a nine-step instructional design sequence.¹ Each step is reviewed. Then, the relationship of the current investigation to the ISD framework is described.

1. Define the instructional goals.

This is normally achieved by considering and specifying the gaps between the desirable state of affairs and the current or observed state of affairs. Presence of new ideas or introduction of new technology can also trigger creation of instructional goals.

2. Conduct instructional analysis

Describe the task(s) to be trained. Instructional analyses can include task analysis, cognitive task analysis, and learning analysis (e.g., discrimination between terminal and enabling objectives, sequencing of enabling objectives). Additional analyses might include knowledge-level and metacognitive analysis.

3. Identify entry behaviors and learner characteristics

Determine which required enabling skills the learners already have, and note any other characteristics of the learners germane to developing instruction.

4. Identify performance objectives

Translate needs and goals into specific, detailed objectives. Detail is necessary both to development of training materials and to creation of measurement tools to assess student progress and determine the degree of success in achieving the objective(s). A properly prepared learning objective has:

¹ Meister [1991] suggests thirteen steps: analyze job, select tasks for training, construct performance measures, select instructional settings, develop objectives, develop tests, develop entry behavior, determine sequence and structure, specify learning events and activities, specify instruction methods, develop instruction, validate instruction, and implement instruction. The arrangement of steps is different, but the overall structure and content is similar.

- A statement of behavior (action) the trainee must exhibit, i.e., the training outcome. This may be either positive or negative, e.g., for situation Y, the goal may be to perform X, or it may be to avoid or refrain from performing X.
- Specification of the conditions under which the behavior must take place.
- A statement of the standards for satisfactory performance.

Some kinds of training goals are more conducive to characterization via the behavior/ conditions/standards approach than others. For example, "teamwork" is harder to specify and measure than "operate a radio". Getting the statement of a training objective right is critical. Also, conditions and standards are frequently inadequately specified in training system development.

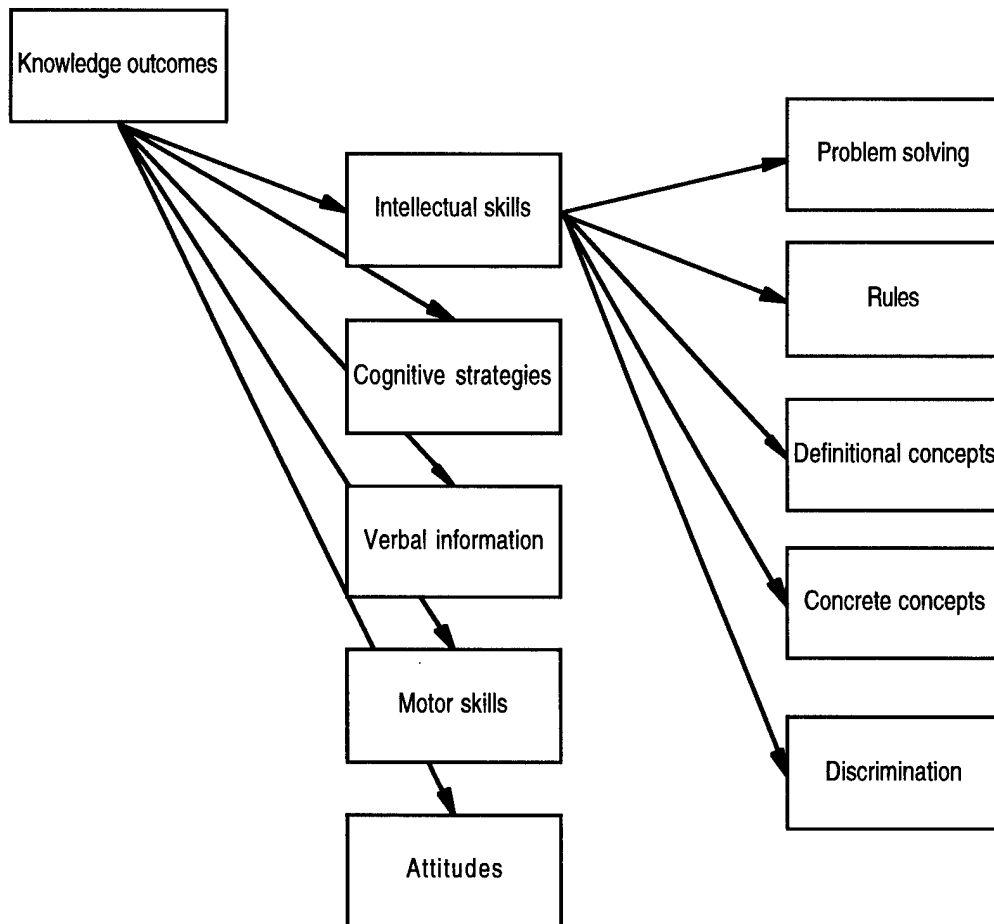
Gagné et al identify five classes of outcome.

- *Intellectual skills*. Learning how to do something of an intellectual sort. The ability to interact with the environment (broadly construed) in terms of symbols and concepts, and/or the learning of procedural knowledge.
- *Cognitive strategies*. Skills for governing thinking. This includes both operational and predictive knowledge. For example, having appropriate expectations, and knowing how much certainty to invest in them, is a critical skill of competent commanders.
- *Verbal information*. Declarative knowledge, and the ability to integrate it with other knowledge and to recall it appropriately.
- *Motor skills*. Physical skills, and (to the extent not covered by other categories) knowledge of how and when to apply them.
- *Attitudes*. A persisting state within the affective domain that consistently influences or modifies an individual's choice of actions. Attitudes are immensely important in the military.

Within intellectual skills, there are five subcategories: problem-solving, rules, defined concepts, concrete concepts, and discrimination. In reverse order:

- Discrimination involves distinguishing features of objects or situations as similar or different, e.g., steeper or shallower slopes on a topographic map.
- Concrete concepts involve identifying a class of object characteristics, objects, or events, e.g., a hilltop, a ridge, or a saddle.
- Definitional concepts involve classification of an object or event in accordance with a definition, e.g., a potential enemy AoA, a kill sack.
- Rules require knowing rule content, conditions, and in most cases exceptions, along with the competence to apply the rule correctly, e.g., "T-72s have a shorter firing range than M1A1s, so covering this ridge will allow us to prevent enemy forward progress."
- Problem-solving involves generating a solution to a novel problem, e.g., selecting the most rapid route for tank platoon advancement with adequate cover and concealment.

Figure 3 - 1. Gagné's knowledge outcomes.



5. Develop criterion-referenced test items

Criterion-referenced test items determine whether the students have acquired the desired skill, rather than merely showing they remember the instruction. These are necessary for training evaluation. They can also be used to place students within a training program, by determining which of the enabling and terminal objectives they have already mastered.

6. Develop instructional strategy plan

Instructional strategy is a plan for assisting learners with their study efforts for each performance objective. It is a mapping of instructional activities to accomplishment of objectives, as well as diagnosis and prescription of the most apt instructional approach for the objective.

7. Develop instructional materials

The selection and development of instructional materials.

8. Conduct formative evaluation

This is the instructional version of alpha and beta testing of software--formative means that the course is being revised and tuned through this evaluation. A three-step approach is advocated by Dick and Carey (1990). First, try the materials one-on-one. Then, try the materials on a small group of students. Then, perform a field trial in which the instruction is revised based on the whole class. (This is of course subject to the kind of objectives and material involved.)

9. Conduct summative evaluation

Evaluate the course or training program after it has been implemented and put in place.

There is more than these nine steps to instructional design. In actually developing a training course, resources, other constraints, teacher preparation, installation, and diffusion have to be considered.

3.2 Epistemology of knowledge outcomes

Knowledge is the core element of instruction and training. Why not simply categorize task material with respect to the knowledge outcomes listed above and base decisions on training approach on the results?

Gagné's knowledge outcomes are problematic. It is not consistently clear what they refer to: the task, the cognitive process, or the result of the cognitive process. The name suggests the result, but the specific category names sometimes suggest the operation of the cognitive process (discrimination), sometimes what the process operates on (concrete concepts), sometimes the nature of the thing to be learned or applied (definitional concepts), sometimes the formal representational structure of the knowledge (rules), sometimes the characteristic things people seem to do in managing certain classes of task (cognitive strategies). Both Anderson, *op cit.*, and Newell, *op cit.*, along with many others, would claim that *all* of Gagné's categories (excluding the actual execution of motor skills) could be implemented as declarative memory plus procedural rules. Gagné's knowledge outcomes are epistemologically.

But suppose we take them as fuzzy concepts of "things people do" when they perform tasks, e.g., apply discrimination, execute rules, understand definitional concepts, etc., are they then adequate as guidelines for selecting approaches to training?

No. They suffer from what Anderson (1990) calls the *identification problem*. Statements regarding the nature or structure of knowledge require induction of the mechanisms of thought from observable behavior in specifiable circumstances. The circumstances are the input, the observable behavior the output. Formal systems theory shows that there are an infinite number of functions (in this case, cognitive mechanisms) that map a particular input to

a particular output. Even if only two functions produce the same behavior from the same input, there is no way to identify which one is psychologically real and which one is not. This fundamental uncertainty makes strong statements about mechanism impossible.

Anderson makes this point with respect to the entire enterprise of cognitive science, but it is true in the more restricted case of Gagné 's knowledge categories. Membership in a given category can arise from multiple sources and have multiple implications, which category membership in itself does not specify. This is the critical failure.

These alternate sources and alternate implications can be significant for training. Many people were fooled by Joseph Weizenbaum's ELIZA program (1969), which used a simple keyword pattern matcher to mimic a Rogerian psychotherapist. ELIZA used rules and appeared to perform problem solving. But clearly, the training Weizenbaum gave ELIZA is radically different than that one would provide a real therapist trainee. Taking another example, "attitudes" are so defined that anything persistent involving affect is an attitude, no matter what its origin, i.e., biological, learned, chemically-induced, or socially or circumstantially determined. Yet understanding the source of an attitude is potentially critical to developing training. Not to beat this into the ground, but logic is another example. The research of Johnson-Laird and Wason on syllogisms is persuasive that they are normally solved with a mental model, which can be represented as being at the rule level of knowledge outcome. The mental model approach to solving syllogisms isn't very good (our cognitive capabilities didn't evolve to solve syllogisms), so people who want to become experts learn the Venn diagram approach, which works very well, and which can also be expressed as rules. But teaching the mental model approach and teaching the Venn diagram approach are different: the first implies repeated exposure, simple to complex, with feedback, and substantial memorization. The second suggests teaching operationalization of a concept.

In general, the identifiability problem means that substantial experimental evidence is needed to support an ontology or epistemology of teaching learning. Even for the best studied areas, such results are controversial (e.g., diSessa 1993) and often ambivalent (e.g., Patel et al 1993).

Gagné 's categories seriously underspecify the requirements that an effective training method must satisfy. In the next section, such requirements will be proposed.

The problem is to characterize a training objective or cluster of related training objectives and from that, to formulate requirements for training. It is not to select or validate a training objective, or to perform other ISD functions.

By requirements, we mean specifying sequences of events or circumstances for an individual or a group so that a desired performance capability arises and persists. This definition focuses on *learning as goal-driven adaption to the environment*. The training method is the design of an optimal environment for learning. The definition constrains the concerns of this theory to behaviorally-evidenced capacities and facilities for learning.

The persistence of learning contrasts with transient changes in performance due to drugs (e.g., coffee), fatigue, and other short-term influences. These factors are implicitly included in task-specific definitions of performance capability, as task conditions that must be managed to achieve the desired level of performance.

In this section, a theory of the problem space is developed. The minimal requirement for a theory is descriptive adequacy. That means that the phenomena of interest are encompassed by the theory, and that the regularities in their structures, relations, and operations are accounted for.

In addition to descriptive adequacy, predictive validity is needed. The capability to predict the most productive training requirements for a training objective is the goal. The precise predictive accuracy of physics or chemistry is *not* the goal; hard sciences like these have definitional concepts direct observation of the material of interest (Gould 1981). ("So far as the laws of mathematics refer to reality, they are not certain. So far as they are certain, they do not refer to reality"—A. Einstein.) Natural sciences don't "linearize" as much, have fuzzy concepts (in the fuzzy logic sense), and the material of interest must be induced from the observable evidence, not defined. To state that a student understands a topic is to make an inductive statement. John Haugeland (1978) calls hard science theories nomological; Micki Chi (1992) calls natural science theories mental model based. A good natural science theory provides persuasive explanatory power and guidance in predicting productive avenues of research. The psychology of learning and the practice of teaching/training are natural science, not hard science. Our goal is a natural science theory.

4.1 Bounding the problem: Content & abstraction

The first question for a theory is, what is it about? The question of what should be explicitly in the theory, what implicitly in, and what to leave out is determined by (1) the goals of the theory and (2) its level of abstraction. The first hurdle for prospective content, relation to the goals, is pragmatic:

if a construct ("construct" is meant include associated operations) proximally affects method, content, or presentation of instruction or training, it passes. Within that very large category, formal descriptive qualities of constructs are explicit, while specific content is implicit. For example, that a task requires a repertoire of plans and subplans is a formally descriptive aspect. The specific content of the plans is implicit (with respect to the theory, not to an particular training plan.) The explicit content may be at the level of decision making and social judgment, or at the more basic cognitive function level, like discrimination, recall, and physical response.

The problem of what "proximally" means in the phrase "proximally affects training" is answered by the theory's level of abstraction. What choices of level are there? What do the choices mean? What are the selection criteria, and how do they relate to the overall goal of the theory? We start with Newell's levels (1990): biology, functional architecture, algorithm, and knowledge.

Biology is used in the sense of what it is and how it works as biology (as contrasted with how it works in the world, e.g., as biomechanics, for example). It can be eliminated from consideration. It might be necessary, as a source of constraints, were this a theory of selection of training goals. The theory assumes that the training objectives are achievable by the subject learners, and leaves concern with biological constraints to the selection process.

Pylyshyn (1986) defines the next level, functional architecture, as the set of functions provided by the biology that are used to get things done: "storing and retrieving symbols, comparing them, treating them differently as a function of how they are stored, (hence, as a function of whether they represent beliefs or goals), and so on, as well as such basic resources and constraints of the system, as a limited memory." The question immediately arises, since these are clearly basic-level cognitive functions, how far do they extend upward in complexity, e.g., into mental algorithms and procedures? Pylyshyn's answer is, not far. He proposes the principle of *cognitive impenetrability*: the operations at the functional architecture level are not affected by the organism's goals and beliefs. If this is accepted, clearly this level can be eliminated from consideration for our theory. It can't be trained, and it is subject to the same argument as biology, i.e., it's implicitly present in the training objectives.

Pylyshyn's principle, as stated, is problematic. Recently, the principal author talked with a mountain-bike racer. He described some of the amazing feats of top-flight riders, who tear down steep, rocky, technically challenging trails at speeds exceeding fifty miles per hour. Consider the vision problem such riders solve. They must identify the best prospective micro- and macro-routes in real time while their eyes are mounted in exceptionally bouncy heads, working through dust-covered, slightly-distorting

sunglass lenses (which themselves bounce slightly). The basics of vision would seem to be subject to cognitive impenetrability. Yet, the training that such riders put themselves through almost certainly affects their visual processing. Therefore, though it is neither direct nor immediate, there is an effect on functional architecture due to goals and beliefs. In the same way that physical capacity can within limits be altered by goals and beliefs (when translated into action), mental capacity can also be so altered.

If Pylyshyn's principle is refined to specify *direct* alteration, this problem almost disappears. But not quite.

What about automaticity, which is frequently a goal of learning and practice, often even considered a sign of expertise? When automaticity is achieved, which is often unconsciously, does that mean that the cognitive skills and knowledge it encodes migrate from the higher to the functional architecture level? They are no longer directly affected by goals and beliefs, e.g., the ability to compose grammatically correct sentences in English is automatic for native speakers, and while the content of those sentences, and even the grammar, can change to clothe a particular idea, the fundamental ability doesn't.

The notion of cognitive impenetrability is subject to the classic problem of trying to make a continuous phenomena discrete: the boundaries are problematic. Probably the best answer to this dilemma is to recall the earlier acceptance of fuzziness. There is a center to the notion of functional architecture, and that center is outside the space of our theory.

This does not mean that goal and behavior-influenced changes to functional architecture, and optimal approaches to such changes, are not important or are ignored. It does mean that functional architecture is not the level of discourse at which the theory is set. Training motor skills for example is encompassed, but at the level of articulate goals and the intentional structuring of environment and behavior, not at the level of specification of changes to functional architecture.

Both Pylyshyn and Anderson call the next level the algorithm level. (Newell calls it program symbol.) This is the basic level at which modeling of the world is performed. It is also the level at which steps of cognition² can be correlated with observable behaviors. While such correlation will always be conjectural, it can be developed to a degree that has reasonably strong explanatory and predictive power. This level is clearly in the theory.

The highest level is the knowledge level. (Pylyshyn: semantic level [distinct from Tulving's semantic memory, though]; Anderson: rational level.) Newell's knowledge level hypothesis (1982) posits a level above the algorithm

² Note that steps in cognition are not isomorphic with task steps as they might be elicited during a task analysis.

level for which knowledge is the medium and the "principle of rationality" is the law of behavior. The principle of rationality is, "if an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select that action." Goals, selection, and implication (i.e., that knowledge content can be of a necessary or likely future state) are defined conventionally. Knowledge is defined as "Whatever can be ascribed to an agent, such that its behavior can be computed according to the principle of rationality."

This definition avoids epistemological swamps, but it doesn't restrict travel very much. Suppose I bump into a hot stove and jump away. It wasn't knowledge, but reaction, that caused me to jump; that is, sub-symbolic, not symbolic, processing. However, a perfectly good explanation in terms of knowledge and goals resulting in behavior can be constructed. Further, the definition obscures distinctions of knowledge that may be significant for training. For example, one claim to know the melody to Gershwin's *Rhapsody in Blue*—by which is meant that, when heard, the melody is correctly identified. It is possible to know the melody in this sense, but not be reliably able to recall it unaided. In designing training, these two different kinds of knowledge, comprehension and production, are likely to be treated differently.

The principle of rationality is also slightly unfortunate, in that by assuming implication it ignores what Fodor (1986) calls epistemic boundedness. In simple terms, that we know *A* and that we know that *A implies B* does not necessarily mean that we know *B*. It takes time and cognitive resources to work out the implications of our knowledge. Recognizing too late the opportunity you gave your opponent in chess is a familiar example. A second misfortune is that the principle of rationality seems to assume a uniform level of valence for all goals. It does accommodate goal hierarchies, subgoals, and sequencing. But different goals are pursued with differing degrees of intensity, and that valence can alter subgoals or even terminal goals. The relative importance of the goal and the motivation with which it is sought are important aspects in developing training, but this is obscured in the principle of rationality.

Despite these problems, from a fuzzy point of view, the center of the knowledge level aligns directly with the level at which training is normally specified. Therefore, it is within the scope of the theory.

Having considered these four levels (biology, functional architecture, algorithm, and knowledge), it is an interesting aside to note that while the conventional interpretation is that the lower levels constrain the higher—if only we were smarter, faster, our senses more acute, weren't losing our damn memories, etc.—in fact it is the higher levels that constrain the lower. We are subject to evolutionary selection. It is the high-level goals and capabilities that the lower-level parts evolved to serve. For example, stereo color vision conveys survival advantage for predators like ourselves. Understanding what color stereo vision must accomplish highly constrains the

investigation of how human vision works. David Marr suggested the analogy of trying to understand bird's flight by studying only feathers. Without understanding what feathers are supposed to do, what they are makes little sense. Similarly, without understanding what knowledge and algorithms do, understanding functional architecture and biology is impossible.

4.2 The theory: Assumptions and axioms

To recap: the goal of the theory is to characterize training objectives and training requirements in such a way that the requirements can be strongly associated with the objectives. I am concerned with the algorithm and especially with the knowledge levels. We are also, of course, concerned with the meta-goals of the theory: utility, generality, parsimony.

The stance taken throughout is that learning is an active constructive process, in which the learner is the principal actor. The learner actively builds, tests, repairs, and refines mental models and physical skills for problem solving competence. The notion of learner as a passive receptacle of delivered knowledge—the "factory" model, in which knowledge is some sort of complicated "assembly" for teachers to "install" in students' heads—is wrong. This should be an unremarkable position.

We make two further assumptions. First, we assume that normative models underlie any tasks or objectives to which the method might be applied. That is, for a given task there is as a minimum a reasonable prior consensus on what the task is and is not, and what constitutes a good or satisfactory solution. In most cases, it will also mean that there is an accepted and practiced standard approach to generating solutions. (Note that "approach" is not the same as "procedure". An approach may specify only what kinds of information should be gathered or present, or what kinds of possibilities should be considered, without providing the more detailed structure a procedure encodes.)

Second, we assume a meaningful connection between task characteristics and the effectiveness of various approaches to training. The evolutionarily determined fit between our minds and the world has discernible and describable consequences for the way we best learn particular concepts and competencies. Human learning is not infinitely plastic. What the connection is for a particular task may be controversial; that there is a connection should not be.

A theory has five essential components: axioms, constructs, attributes of constructs, relationships among constructs, and operations.³

³ For example, in mission planning, the *axioms* are Own Unit and Friendly versus Enemy, and pursuit of mission success. The basic *constructs* include the mission, the operational order, mobility corridors, avenues of approach, courses of action, battle positions, and (usually) the forward edge of battle area. "Cover and concealment" is an *attribute* of avenue of approach. Courses

The theory has two axioms:

1. People learn.
2. Different ways of learning are more or less effective with respect to different things to learn.

"People" and "learn" are both fuzzy concepts. That is, there is a center and a periphery, and things are a lot more clear at the center than at the boundary. Given that, we can agree that people learn.

Regarding (2), physical mimicry of physicists will not cause you to learn classical physics. Dropping apples while looking at the moon won't work, either. (Unless you're as brilliant as Newton.) Academic study of physics followed by supervised practice probably will work (in most cases).

There are two important notes about (2). First, it doesn't matter whether the connection is "causal" or "correlational". We're concerned with productive inferences about training approaches.

Second, it—the relation between how we learn and what there is to learn—is not the only factor. There may be myriad reasons for learning to go well or poorly, unconnected to specific ways we learn and the nature of the material. Illness, life circumstances, workload, self-concept, even lighting conditions or time of day; many more could be named. It is possible, for almost any example, to construct a rationalization for why it should be considered part of "the way we learn". ("Illness saps energy and attention, which are enabling conditions of the way we learn.") However, this leads to "all and everything-ness": anything that might *influence* learning gets included in the *definition* of learning.

A note about both axioms taken together: without the second of the assumptions above, namely that there is a correlation between how a subject is taught and how well it is learned, the axioms would have little implicative force.

The axioms are satisfactory, in the sense that if either one is invalidated, the theory collapses. We infer from them that with the right characterization of how we learn and what there is to learn, how we should teach or train should be able to be "read off" (using assumption 1) from their relationship. It looks like:

How learning works => *Requirements for instruction* <= *What there is to learn*

From this view, the requirements decompose and organize what there is to learn into the most facile form and sequence for learners.

of action relate battle position, mobility corridors, and avenues of approach to missions and operational orders. Maneuver to change battle position is a typical operation.

4.3 The theory: Constructs and features

What are the constructs for training? We propose seven constructs, organized roughly as how we learn, what there is to learn, and how to train. We present them in this section, and provide a detailed justification—a taxonomy specifying the links we claim—in the next.

Before enumerating them, it is appropriate to ask, isn't there a simple version of this theory that works, from which the more complex version could be built? The general law of complex systems is, *working complex systems arise from working simple systems*. But unfortunately, no simple system is adequate.

The constructs are:

How we learn:

1. Ways of learning

What there is to learn:

2. Kinds of tasks, domains, methods, and solutions

3. Classes of content and meta-content

4. Sources of error

5. Notable qualities of expert performance

6. Conditions of performance

How to train:

7. Approach to instruction

How we learn, as it is germane to instruction and training

What do we mean by "how we learn"? We've said before that biology and functional architecture are not within our scope. (By implication, the subject is adult learning, not developmental.) Still, at the algorithm and knowledge levels, we need to specify, at least operationally, possibly fuzzily, what we mean. Style? High-level mechanisms? Methods? Processes?

By "how we learn", we mean *how we construct content and operators to achieve our goals within the performance environment*, within the level of abstraction bounds set for the theory. So, for example, indexing and retrieval are critical cognitive operations which must be learned in conjunction with new content and operators, but they are generally below the algorithm level of abstraction (except of course where indexing and retrieval are explicit goals of instruction, as in the teaching of mnemonics.) What is indexed and retrieved is (presumably) on the algorithm or knowledge levels.

At the knowledge level, we distinguish similarity-based learning and explanation-based learning, two concepts which had their genesis in machine learning research (Mitchell 1982). At the algorithm level, we name rehearsal, organization, and elaboration.

How much is encompassed by these mechanisms? Gardner (1983) claims that there are six "intelligences": linguistic, logico-mathematic, spatial, bodily-kinesthetic, self-knowledge, and knowledge of others. We assert that all are learned using one or more of the above mechanisms. Learning is parsimonious

with respect to mechanism. However, we condition this claim by emphasizing that there are internal structure and learning variations for all of them; this is as true of spatial and bodily-kinesthetic as it is of the more obvious linguistic and logico-mathematic kinds of knowledge. (Not to mention self- and social knowledge.) Further, overly facile identifications are probably not justified, e.g., visual with spatial (Nissen 1985). It should also be noted that there may be different levels or aspects of the same task, which may require different teaching approaches. The difference between qualitative and quantitative understanding in physics is an example (Smith *et al* 1992).

Before examining how we learn, we should consider the role of psychological conditions for learning. Mayer (1987) suggests three fundamental conditions for learning: reception (attention, basic communication), availability (common domain vocabulary, prerequisite knowledge), and activation (information/knowledge assimilation, integration). These certainly seem plausible; it's hard to imagine successful learning without all three. On the other hand, these are assumed about training at our level of discourse. Our theory is primarily interested in instructional strategy; it is not a complete ISD method. All training methods require all three, so this isn't a distinction that makes a difference for us.

Motivation is the key learning condition variable from the point of view of training. The level of motivation present can have a major effect not only on how training goes but even what is trained. It's the "hill climbing" issue: there are multiple routes to the target hilltop, and motivational "fitness" can play a big part in choosing which one to take. For example, developing proficiency at some tasks requires arduous practice. High-motivation students can be set fairly quickly to work on such tasks, and thereby achieve proficiency in a relatively shorter time than lower-motivation students, who may need more intellectual "rest-stops" (e.g., attractive subgoals off the direct route) on the way. If the motivation is great enough, it may even be that a higher hilltop is chosen as the goal.

However motivation is decomposed (e.g., into goals, values, self-efficacy, control processes), it more properly belongs in the characterization of the learner than in method selection *per se*. Specifying the "hilltop" is part of developing a training objective, as is selecting the "viewpoints" (enabling objectives) along the way. It is symptomatic of the dense interconnection of training influences that to some extent this makes the theory more fragile. While the best method may be determined by the structure of the material independent of learner characteristics, the best presentation may be highly sensitive to motivation level. Since we claim that presentation and method are really poles on a continuum, this is a legitimate issue.

Pylyshyn's cognitive impenetrability principle provides relief. While motivation can certainly change, and can be taught by creating an appropriate environment, it is not at the knowledge-level or the algorithm-level of

abstraction. One cannot directly change one's motivation without constructing a training program much more analogous to a physical training program than to a case of intellectual choice. Therefore, while motivation is important both in selecting training objectives and characterizing the learner, it is outside the scope of the present theory.

1. Ways of learning

First the knowledge level is discussed, then the algorithm level.

Similarity-based learning (SBL) induces theory or an abstract general rule from particular examples. This can be with respect to categories, operators, or both. Explanation-based learning (EBL) maps a theory or abstract rule to an example, then remembers the mapping⁴; again, this applies to categories, operators, or both.

Examples of SBL include induction of decision trees from multiple examples, development of general categories from multiple instances, classification of new information with respect to existing categories, and learning of new patterns of physical movement.

Examples of EBL include induction of (elements of) decision trees from single examples, classification of novel examples, and induction of use (of an artifact, concept, or situation) from knowledge of function.

These examples are not meant to be comprehensive. For instance in some problem-solving situations, it may be profitable to remember both abstract categorizations and operators, and specific individual cases. At the knowledge level, the latter would be SBL, even though the fit is imprecise, and the reason for indexing an individual case may be explanation-based.

SBL and EBL are not identical with their reasoning cousins. Explanation-based reasoning is a common way of resolving impasses during problem solving, e.g., by constructing an explanation of how the current situation could have arisen or could be sustained, such that the most appropriate action is embedded in the explanation. In effect, explanation-based problem-solving applies weak knowledge when there is no strong knowledge available. (One doesn't need to construct an explanation based on first principles if the engine is on fire.)

Explanation-based reasoning should also be distinguished from explicative tasks. There are critical Army tasks that are explicative; preparing mission statements and orders is a prime example. Describing and explaining things is an important skill, and may depend on explanation-based reasoning, but it may also depend on similarity-based reasoning. And, it is not uncommon for a practitioner in a particular discipline to be very competent at explanation-

⁴ Specifically, it remembers a declarative justification of the newly-learned material in terms of its existing knowledge.

based reasoning, for diagnosis, for example, but not especially competent at communicating his or her reasoning.

What are the instructional implications of the distinction? EBL requires significant prior knowledge about the domain to guide search; therefore, a strong domain theory is preferable. In addition, because the core of the method is mapping, domains where surface features and underlying structures are distal are likely candidates, as are event-poor domains. SBL requires inductive bias for learning, i.e., some bias toward relevant features and away from spurious or irrelevant features (Mitchell *op cit.*). However, this is often not a problem, since for most problems there are abundant cues to distinguish useful features. SBL is appropriate for many classification tasks, case-based reasoning, for event-rich domains, and for motor skill and pattern recognition tasks.

There are many hybrid SBL/EBL learning tasks. Construction of an explanation can rely upon having a similarity-based goal concept (e.g., answer) toward which to map. And, SBL may be embedded in a larger EBL task.

Rehearsal, organization, and elaboration are the three mechanisms at the algorithm level. Rehearsal is mental practice; in traditional psychology, it is called "strengthening". Organization and elaboration correspond roughly to SBL and EBL.

We say "roughly" since, like nearly everything concerned with learning, even these basic concepts are controversial. For example, traditional views on teaching physics hold that naive views must be replaced by formal models and methods. From this view, rehearsal applies to learning the basic (definitional) concepts, organization to creating the model, and elaboration to applying it to actual problems. But diSessa, *op cit.*, argues strongly against the replacement theory, claiming that theoretical physics, to the extent that it is grounded in physical intuition, elaborates on rather than replaces naive physics. In this view rehearsal and elaboration are the key learning mechanisms, with organization being much less important.

What there is to learn

This cluster of constructs is organized roughly as, what there is to learn (kinds of tasks, content and meta-content), what goes wrong (sources of error), what goes right (characteristics of expertise), and where it's done (conditions of performance).

2. Kinds of tasks, domains, methods, and solutions

One common approach to task classification is to use a fundamental aspect of the problem space, namely whether it is convergent, divergent, or transformational. Search through a convergent problem space leads to a single best or correct answer. Mathematics is an example; diagnosis is another. A

divergent problem space does not have one best answer, either because the constraints on the solution are weaker or because of the goals of the process don't lead to single solutions (as for example having interpretive goals or multiple inconsistent goals requiring trade-offs). Planning and design are typical examples. Transformational problems are those where the complexity of the input and the complexity of the output are roughly equal and problem-solving is finding a suitable mapping between the two. Computer-based modeling is a representative example.

In the knowledge acquisition and expert systems communities, it is more common to refer to convergence as classification or structured selection, divergence as construction, and to view transformational as a hybrid involving both.

Manipulation, monitoring, identification, interpretation, diagnosis, and debugging are typically viewed as convergent. Prescription, design (arranging objects under constraints), planning (arranging actions under constraints), and rationalizing (explaining, describing, motivating, etc.) are seen as divergent. Generally, other common tasks, e.g., control, repair, configuration, scheduling, prediction, instruction, and evaluation, can be composed from these primitives.

The power of task characterization is in constraining generic content. For example, all diagnostic tasks will include plans, hypotheses, some device/system/situation model, error recognition and avoidance strategies, specific methods and skills, and in most cases systematicity, e.g., how to pursue maximal payoff paths, how to return to the last valid branch point from a dead end, etc. In addition, diagnostic tasks will have a pre-enumerable (at least in principle) list of goal states. Once such generic content is identified, methods for instruction can be more appropriately selected.

There is another aspect of the problem space of significance, relative wealth and poverty of information and possibilities. There are three manifestations of this: data-rich and data-poor problem spaces (e.g., predicting Soviet Armor tactics contrasted with predicting Somali insurgent tactics), event-rich and event-poor (firing a rifle, firing an ICBM), and hypothesis-rich and hypothesis-poor (situation assessment involving guerrilla war, situation assessment involving frontal assault).

Malleability of the problem space is another attribute. By "malleability", we distinguish domains where experimentation is possible, or even normal, from domains where actions, once taken, are irrevocable. Different content and different attitudes are appropriate when experimentation is possible.

"Standard deviation"—the degree to which task episodes vary from a central core—can also be consequential. If there is a strong central prototype the training investment should be in the teaching and practice of basic skills. If there is a great deal of variation (and there isn't a strong theory to account for it), elaboration of basics and techniques for probing the periphery need

to be taught.

How good must solutions be? Optimal or satisficing? Fallesen (1993) found that a common error in tactical decision making was to search for optimal solutions, rather than to accept a satisficing solution and devote the saved time to higher-payoff planning and execution.

Is the desired performance comprehension or production? Evidence from developmental and rehabilitation studies both show marked differences in comprehension and productive performance. Recognizing a good plan is a different matter from developing one. Facility with comprehension may be a necessary condition for the desired performance, but it is far from sufficient. This may seem so obvious as to not require mention, but failure to make this distinction is commonplace.

Some tasks are fairly homogeneous with respect to the knowledge and skills required. Electronic circuit design is an example. Others are more heterogeneous; battlefield command, for example. If expertise in a domain requires knowledge from a variety of fields, or at multiple levels of abstraction, or both, it is heterogeneous. When this is the case, the difficulty of learning the task increases and often explicit training in integration of material from multiple relevant areas is required (Patel *et al*, *op cit.*).

The last task/domain issue is the role of distributed representation (Zhang & Norman 1994). Since this is a fairly new subject in psychology (or at least a new formulation), we will first define what we mean, then discuss its consequences for training.

The bases for studying distributed representation are the insight that complex behavior can be accounted for by positing relatively simple mechanisms in a complex environment (Simon 1981; Brooks 1992), and the empirical fact that different representations of the same problem can have dramatic effects on problem-solving performance (Hayes & Simon 1977).⁵

"Distributed representation" refers to problem solving in which part of the way the problem is represented is internal to the problem-solver, and part is external in the environment. From this view, an external representation is not just a peripheral aid to cognition, but a necessary component of a distributed cognitive task. Zhang and Norman found five properties of external representation:

1. External representations provide memory aids
2. External representations can provide information that can be directly perceived and used without being interpreted and formulated explicitly.

This is consistent with Gibsonian affordances (Gibson 1979).

3. External representations can anchor and structure cognitive behavior,

⁵ A frequent example contrasts arithmetic with Arabic and Roman numerals.

especially to the extent that the external representation is instantiated in physical structures and artifacts.

4. External representations can change the nature of the task as seen from the perspective of the problem solver, i.e., externalization of representation often makes tasks easier (Zhang & Norman, *op cit.*)
5. External representations are an indispensable part of the representational system of any distributed cognitive task.

The last property is definitional, but brings out the critical insight that problem solving involves representational systems, the components of which can be spatially, temporally, and qualitatively distributed.

While the long-term impact of distributed representation analysis may be greater for task analysis and process re-engineering than for training, it certainly is a dimension that must be considered. In some cases, learning to properly use the external representation and to integrate internal and external representational elements are in themselves training tasks. This is true for the topographic map reading example given in Section 7.

3. Classes of generic content and meta-content

We subdivide this category into what there is, what it means, how it works, how it's managed (meta-knowledge), and how it's integrated (meta-cognition).

What there is: Categories and category structures (concrete, prototypical, and definitional), attributes and attribute values, operators (includes both actions and preference-functions), action repertoires, patterns and meta-patterns, pattern-category associations, attitudes, proprioceptive knowledge, and motor skills.

Concrete categories map abstract ideas to concrete implementations, as for example the concept of "4" to four tanks, four battalions, etc. Prototypical categories are defined by best or prototypical members, relative contrasts between categories, and polymorphy (neither necessary nor sufficient membership rules) (Lakoff 1987). Ducks and blackberry pie are examples. Definitional categories are those, like force, mass, acceleration, logic, and algebra, whose existence is by definition rather than observation.

Pattern recognition is defined as classification of stimuli that are "perception or somewhere near perception". This definition is obviously imprecise, but it is meant to convey the notion that pattern recognition is convergent problem solving, but that not all convergent (classificatory) problem solving is pattern recognition. For example, one would not consider either formal logic or elimination-by-aspects as pattern recognition, although both are used for classification.

What it means: Goals and goal structures, data structures, implicative associations, and uses (of objects).

How it works: Indexing and retrieval, procedures and meta-procedures, plans and subplans, hypotheses and hypothesis structures, goal satisfaction markers,

data gathering methods, data identification/validation methods, and errors, error structures, and error detection markers.

How it's managed (metaknowledge): Metaknowledge skills include:

- Problem discernment and identification
- Problem representation
- Validation of problem representation
- Planning strategies for problem solving
- Validating the plan
- Applying the selected problem solving method
- Implementing the solution
- Monitoring implementation and the developing solution, and
- Evaluating the solution.

How it's integrated (metacognition): Along with content, there is metacognitive content. Metaknowledge, i.e., domain-specific knowledge of how to best organize and apply specific problem-solving knowledge, is included above. Here, we add:

- Managing coherence, self-consistency, consistency with related knowledge, completeness, embeddedness
- Assigning, monitoring, and adjusting the salience of task events, goals, and consequences, including goal conflict management and urgency management
- Strategies for managing confidence
- Uncertainty management, applied to data, knowledge, knowledge application, and projected results
- Truth maintenance⁶
- Systematicity
- Allocation of cognitive resources, including managing focus of attention, degree of attention, and duration of attention.

To the extent that any of the above map to internal representations, content can be either real or virtual. Virtual content can be used to build more constructs. Some examples:

- Prediction = Category + virtual attribute value + virtual operator
- Standards = Virtual attribute value(s) + virtual operator(s)
- Decisions = Goals + (sometimes) plans
- Constraints = Goals + attributes and values + virtual operators
- Consequences = Goals + virtual action(s) + prediction
- Error amplitude = Error detection + error structures + consequences

Even something as apparently ephemeral (but actually crucial) as "a sense of elegance or beauty", of a design, for example, can be represented using

⁶ Truth maintenance is managing the status of states and propositions during nonmonotonic problem solving. Suppose I observe *A* and infer *B*; from *B*, I infer *C*. Observation of *D* also supports the truth of *C*. Now I learn that *A* is false. Properly revising the status of *C* is a truth maintenance problem.

these categories.

4. Sources of error

By "error", I mean avoidable mistakes. In some domains, mistakes are inevitable, either because data are necessarily incomplete and unreliable or because there is no reliable (or pragmatically applicable) predictive knowledge. Sources of error include:

Conceptual: Distance between surface features and underlying structure, presence of irrelevant features or operators, local optima, lack of relative contrast between categories or operators, confusion of correlation for cause, confusion of explanatory for predictive knowledge, unjustified induction.

Cognition: Cognitive biases, e.g., framing, regression, base rate, anchoring, availability, representativeness, false attribution, and/or numeric bias (Kahneman et al 1982); cognitive loading; plus a bias toward a particular category or operator ("to a kid with a hammer, everything looks like a nail.") While this latter can often be accounted for via combinations of prior cognitive biases, it is particularly important in a military context.

Knowledge: Lack of knowledge, presence of irrelevant knowledge, erroneous beliefs¹², volume of knowledge necessary to find a solution, incorrect solution identification.

Pragmatics: Lack of data, false data, ineffective or confusing external representation, inadequate physical resources, inadequate temporal resources, ineffective use of physical or temporal resources, inappropriate choice of goals, active (i.e., enemy) interference. (Enemy interference may not be avoidable, but one way military experts are distinguished from novices is by the degree to which they anticipate and plan for such interference.)

The point of this catalog is not to list all possible classes of error, but rather to show the range of error classes that can be identified and that have direct consequences for training. (Specific error-class/training consequence connections are shown in the next section.)

5. Characteristics of expertise

If the training objective is not expert performance, this construct may be optional; it may not be necessary to look closely at expert/novice differences. On the other hand, "expertise" may be provisionally defined as the target competence, and studied in that light.

Glaser (1989) characterizes the structure of competence as:

- Knowledge organization and structure
- Depth of problem representation
- Theory and schema change
- Proceduralized and goal-oriented knowledge
- Automaticity to reduce attentional demands
- Metacognitive self-regulatory skills

Mapping this to observable traits of expert performance, attributes can include quickness to detect potential breakdowns, exceptional flexibility, automatization of subtask elements, focus on underlying phenomena with filtering of irrelevant data, adaptability, immunity to "noise", etc.

The more detailed the available knowledge-level analysis of expertise, the more power is available to characterize and teach expert behavior. For example, it is strongly claimed that expert physicians use hypothetico-deductive reasoning (from hypothetical causes forward to observable effects) during diagnosis, while medical students tend to use abductive reasoning (from observables backward to causes). There is evidence that hypothetico-deductive reasoning can be explicitly and successfully taught (Patel *et al*, *op cit.*).

6. Conditions of performance

These include: time stress, consequence stress, number of concurrent tasks, number of concurrent performance demands, number of concurrent obligations, circumstantial stress (noise, light conditions, temperature conditions, fatigue, etc.), and interruptions.

Approaches to instruction and training

7. Approaches to instruction and training

Approaches include both strategies and tactics. These are points on a continuum, which are easy to distinguish in the extreme cases, e.g., structured state-space differentials as a strategy, writing on the blackboard as a presentation tactic. (Actually, even this example is unclear—"writing on a blackboard" can be a strategy.)

Teaching strategies can be roughly divided along two orthogonal dimensions: telling/doing, and modeless/modal.

"Telling" strategies involve discourse on the part of the instructor and reception on the part of the learner. Telling can include verbal and textual material, diagrams, pictures, and audio. (One supposes even odors could be included.) Halliday (1975) categorizes seven language functions. Four he groups as pragmatic: instrumental, regulatory, interactional, and personal. Three he groups as mathetic: heuristic, imaginative, and informative. Pragmatic functions orient and act. Mathetic functions create knowledge and move beyond the immediately referential. Bruner (1986) points out that another function must be added, the metacognitive capacity to reflect on language and telling as a subject in itself. Telling can apply any or all of these functions.

"Doing" strategies involve some form of demonstration or direction on the part of the instructor and active imitative or goal-seeking behavior on the part of the learner. Doing can involve physical mimicry, performing experiments, practicing motor skills, using simulations, or constructing artifacts. Of course, it can also involve telling; for example, teachback, in

which the learner plays the role of instructor, is a form of doing.

Telling and doing are meant to be reasonably broadly construed. It should not be thought that telling is somehow passive, in contrast to doing, which is dynamic. Both are active. Socratic dialog is a form of telling. The distinction is that in telling, internal constructive action is required of the learner, while in doing, both internal and external constructive action is required.

Modeless strategies present material in the "timeless present", i.e., as facts, assertions, or operations to be internalized and reliably reproduced. Operationalizing $F = MA$ with respect to various physics problems is an example. Modal teaching applies "stance marking" (Feldman & Wertsch 1976), that is, use of linguistic or other tokens of uncertainty and probability to invite negotiation of, and responsibility for, meaning. Bruner *op cit.* gives the example of his fifth grade teacher, Miss Orcutt, making the statement "It is a very puzzling thing not that water turns to ice at 32° Fahrenheit, but that it should change from a *liquid* into a *solid*." Instead of taking a modeless approach and presenting a fact, and backing it by authority, she invited her students to share her fascination with the (rather bizarre, if we weren't so habituated) phenomena of threshold-based state transition. Modeless strategies are concerned with truth, while modal strategies are concerned with meaning. Looked at from a different perspective, modeless strategies rely on authority, while modal strategies require at least some degree of learner initiative.

(As an aside, the traditional form for scientific reports—including this one—uses the modeless style. The implication is that we are concerned with *truth*, rather than the more modest and realistic goal of simply trying to accurately report the view of a particular phenomena from a particular human group's perspective.)

Organizational strategies are slightly simpler. We distinguish structural, discovery, and narrative. Meyer (1975) characterizes five common structural markers:

- Covariance, a causal relationship between antecedents and consequences
- Comparison, a reference to known structure(s), either metaphorical or analogical
- Collection, clustering related concepts or operators, in time, space, or function
- Description, provision or general statement followed by supporting evidence
- Response, pairing problem and solution, question and answer, or remark and reply.

Discovery techniques can be "pure", guided, or expository. In pure discovery, the learner is set in the learning environment to find on his or her own those regularities that comprise the training objective. In guided discovery, the learner is given cues to focus search. Both pure and guided

discovery are inductive approaches. Conventionally, in inductive approaches, the target rule, concept, or procedure is provided after the learner has discovered (induced) the underlying structure. In expository discovery, a deductive approach, the rule is provided at the start, and the discovery process lies in deductively finding its correct application.

The narrative approach provides a skeletal plan or schema which can be elaborated and refined. Narrative seems the most flexible of the organizational approaches.

Presentation tactics for instruction include classroom (verbal, written, diagrammatic, pictorial/image, auditory), structured environment (guided and/or discovery), demonstration, modeling and simulation (real-world, computer-based), field exercise (guided and/or discovery), on-the-job training (OJT), and hybrid approaches such as field class, SAF, or classroom role-playing (e.g., teachback).

4.4 Critique of the constructs

These are the constructs of our model, along with their attributes. Before considering how they fit together, two things that were left out, either partially or wholly, need consideration: knowledge representation and situated cognition. Then, three problems with the constructs need discussion.

Mechanisms of knowledge representation. The representation of knowledge is different from mechanisms of knowledge representation. By the former, we mean the general statement that people solve problems by creating representations, that knowledge is used both to create and to manipulate these representations, and that actions are based on the results of such representational activities. We can imagine what will happen if, say, we upend our coffee cup; we represent a possible event, rather than trying it out in fact. Representation of knowledge is assumed by the theory.

Mechanisms of knowledge representation, on the other hand, are things like rules, frames, schemas, etc. These mechanisms are suggestive—it seems like they must influence, even determine learning—but it's a major research program to demonstrate that (a) a satisfactorily precise statement can be made about how people actually represent knowledge about a particular subject, and (b) that useful predictive statements about learning and teaching can be made based on these statements. Is pure behavior the right level of abstraction (Brooks 1992)? Sub-symbolic processing (Rumelhart & McClelland 1986)? Symbolic (Anderson 1983, Newell 1990)? Fuzzy (Kosko 1994)? Emergent (Minsky 1985)? Model-based (Johnson-Laird 1983)? Narrative (Bruner 1984)? None, some, or all of the above? Real-world experiments that strongly support one of these in contradistinction to the others are extremely hard to construct. Empirical data on learning and teaching approaches based on one knowledge representation (to the interpretive exclusion of the others) are also hard to find (although Anderson has a better record than the others.) Our belief is that knowledge

representation must be important, but the field is not mature enough to guide training. (For a counter-view, see Shank [1986].)

Situated cognition. Also called situated action, this growing body of research asserts that understanding a task without understanding the situation in which it is performed is impossible (Suchman 1987; Glasser 1991). This stems from two factors, distribution of representation (and knowledge) in the world, and the nature of task performance as a group-embedded and group-responsible entity. The arguments of the situated actionists are powerful, and are to some degree represented in the above, i.e., in the notions of distributed representation and conditions of performance. However, as with apparently everything in the field, they are controversial and far from universally accepted.

There are three more problems with our constructs.

1. There are not enough empirical data associating task characteristics with adult training approaches; there are too many free variables; existing empirical data are ambivalent and subject to multiple competing interpretations.
2. There is so much plasticity in human learning that everything maps to everything else.
3. Bad separation, fuzzy lines, and messy level-of-abstraction problems exist between constructs; it's hard to tell what's an independent variable and what's dependent.

Response to 1. Many are working on further development of empirical data. Our question is whether the theory can account for, or more realistically, just encompass, the data that exist.

Response to 2. The plasticity of learning is why there isn't any generally accepted theory or taxonomy. If we were worse at learning, we'd have to be better at teaching. That is, if one and only one training method worked for a particular task, we would by now know it. However, for the great majority of tasks, any number of training methods seem to work to a greater or lesser extent. People will learn almost no matter what.

Many attempts at taxonomy construction have been made—Gagné's, described above is one. Within the Army, the NDTC's early efforts to link tasks to training device design; training device effectiveness models, such as TRAINVICE, and DEFT; and rule-based approaches, such as OSBAT. The Harliss Associates Method and Klein's Critical Decision Method can also be considered attempts to codify instructional design based on task classification.⁷ The

⁷ We thank to Dr. Angelo Mirabella for portions of this list.

Navy also sponsered a major recent taxonomic study [Hogan et al 1987]. The fact remains that no taxonomy is yet generally accepted.

Response to 3. We claim there is a discernible and testable heuristic structure uniting each of the above constructs. In the next section, we attempt to work it out.

How do the various constructs developed in the preceding section fit together? The position implicit in the axioms—that instruction is the bridge between how we learn and what we need to learn—can now be made more specific. In overview, this looks like:

<i>How we learn</i>	=>	<i>Instruction</i>	<=	<i>What we learn</i>
<i>Knowledge level</i>		<i>Method</i>		<i>Task</i>
<i>Algorithm level</i>		<i>Organization</i>		<i>Content</i>
		<i>Presentation</i>		<i>Errors</i>
				<i>Expertise</i>
				<i>Conditions</i>

The taxonomy that follows attempts to make the consequences of the constructs explicit. This section:

- develops heuristics associating the content constructs with the learning and training constructs;
- presents some typical clusters of features and associated heuristics; and
- critiques the taxonomy in its current form.

5.1 Structures of constructs

Figures 5-1 through 5-4 show the structure of the constructs developed in the last chapter, with some elaboration. Note that the leaf nodes shown are not mutually exclusive. Different knowledge levels can be present in a task; different training approaches may be employed; different elements will combine in actual tasks. In general, the leaf nodes of all of the charts are mutually inclusive. Selection of one does not eliminate others.

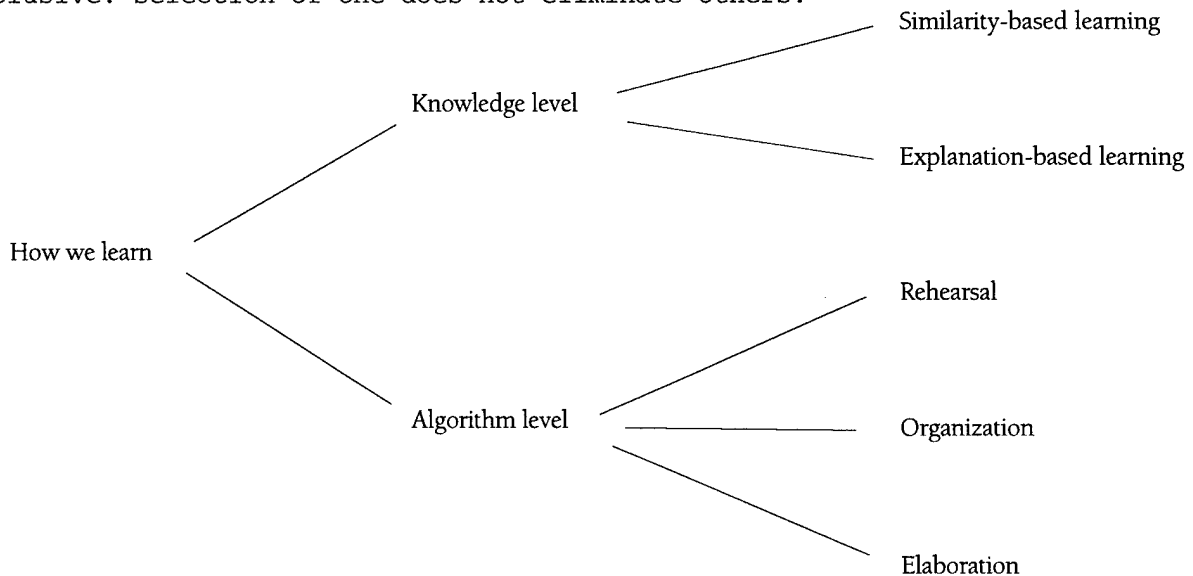


Figure 5 - 1. How we learn.

A few of the leaf nodes in Figure 5-2 may need explanation.

Model tracing assumes that there is a right way to perform a task, and that the learner is to be led through the model, provided opportunities to practice, and gradually given autonomy. Learning is seen as reducing the "distance" between the expert model (or the standard competence model) and the learner's model.

Issues recognition extends model tracing by identifying problematic decision points and the typical errors for each. Explicit instruction on error avoidance, recognition, and repair is part of training. In effect, a stereotypical learner model is built into the training.

Critical decision training focuses on those transition nodes in the process tree where significant branching occurs. Almost all tasks combine a relatively smaller amount of decision making with a relatively larger amount of task execution. The critical decision approach concentrates on understanding and developing facility in the decision space. A distinguishing assumption of the critical decision method, as contrasted with, say, focus gambling, is that a robust (or at least, sufficient) representation of all of the relevant path possibilities and their consequences can be learned and subsequently used in the field.

Abstract problem spaces include a prediction of student performance and compares that with a competence model. The "abstracted problem space" is a model of the domain; the student competence model is a goal/subgoal model. Wanderings or disjunctions are repaired based on the distance between the two. This, like false model training, is trying to get at "choosing the

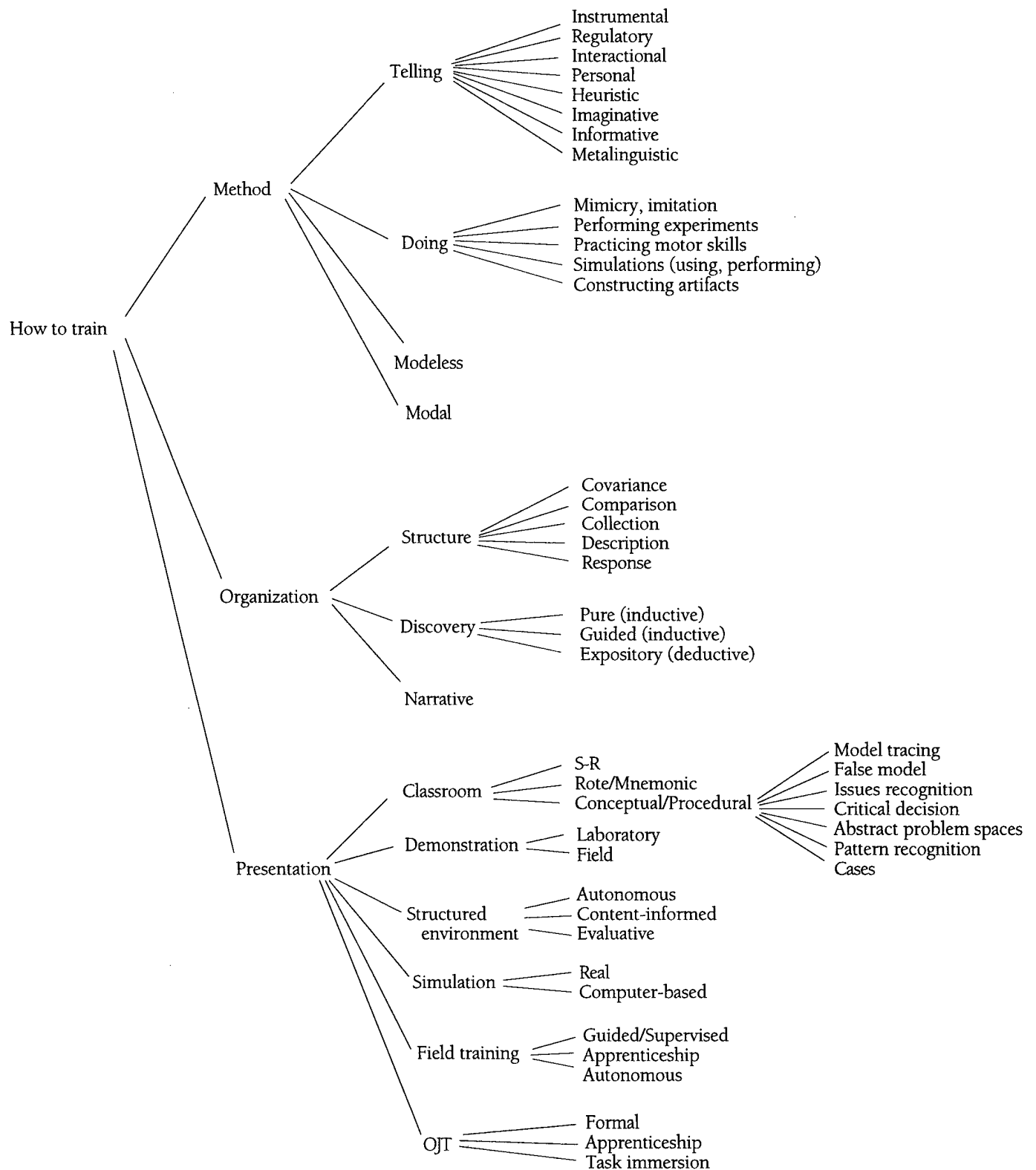
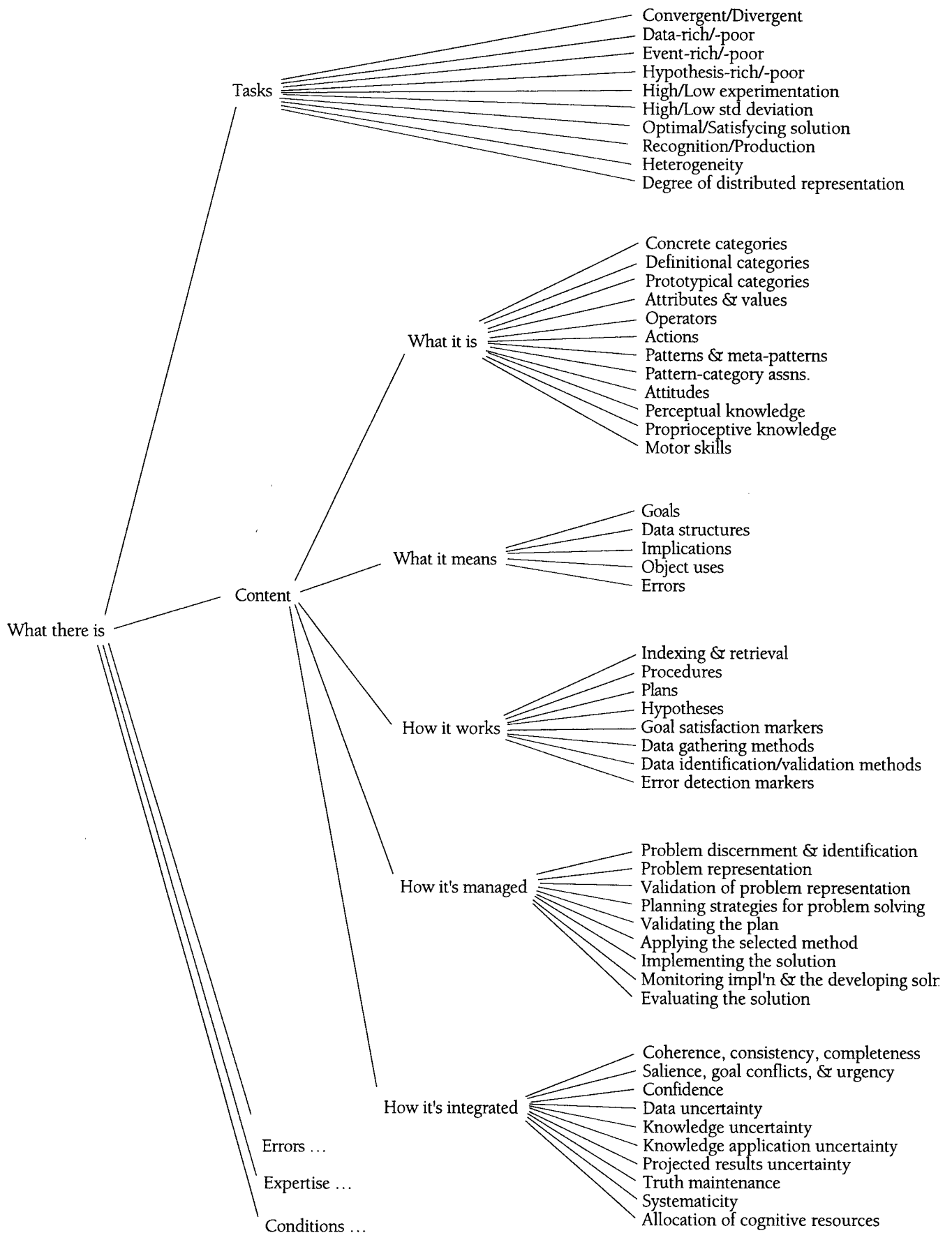


Figure 5 - 2 How we train.



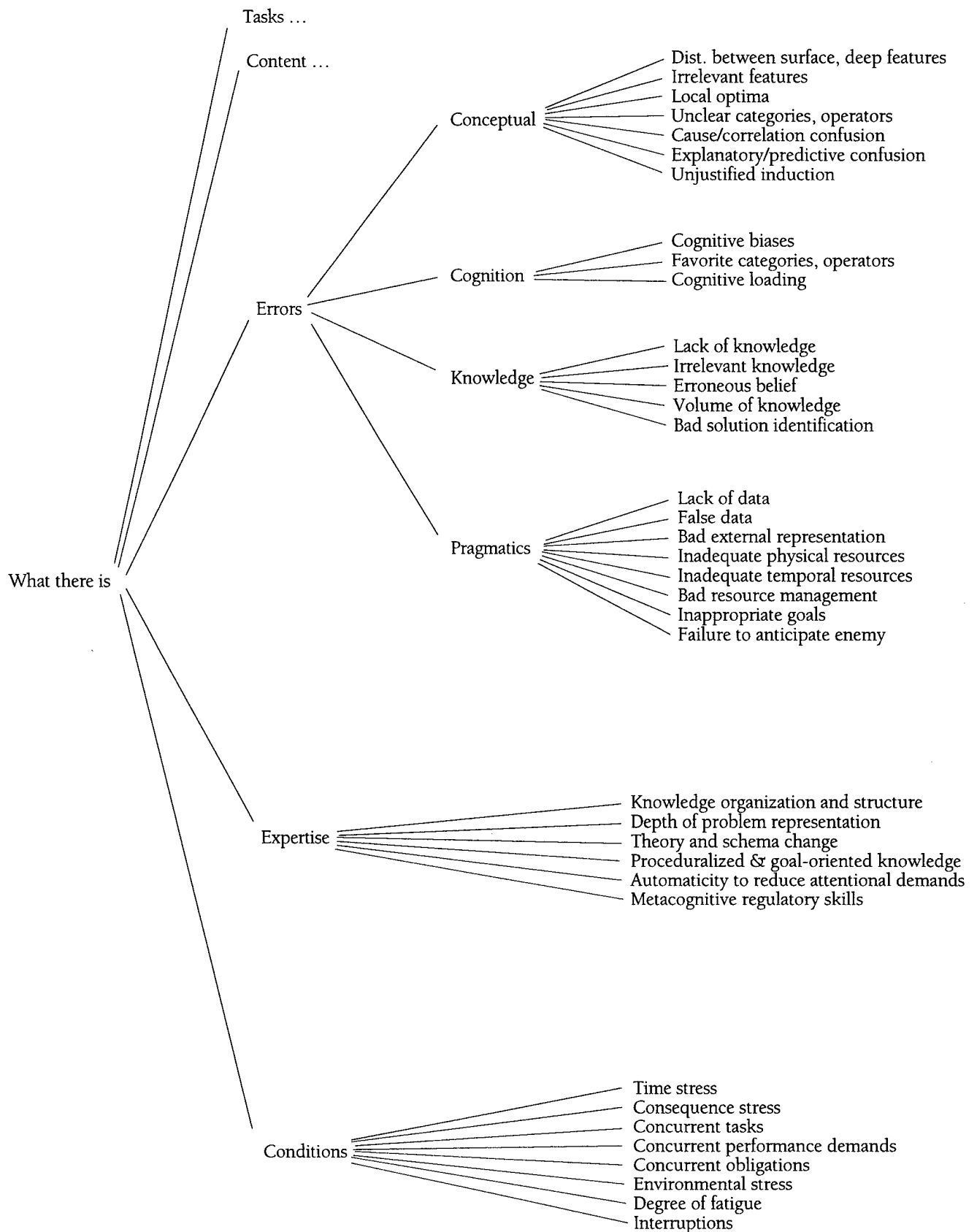


Figure 5 - 4 Content

wrong thing to do", while the previous two are more oriented toward "doing what you choose to do correctly". (This is a little strong-model tracing could certainly be applied to a planning [i.e., choosing what to do] task.) This higher level of abstraction makes abstract problem spaces more general than conventional model tracing. On the other hand, in a strongly data-driven task (e.g., classifying features on a radar screen), using a goal-distance strategy is probably not very efficient.

Pattern recognition training is new as a method (Fischer 1994), although of course people have been taught to perceive patterns since time immemorial. From careful analysis of expert performance, key schemas and triggering patterns are identified. Then, simple-to-complex organization is employed to present prototypical and boundary examples of the key patterns, each immediately associated with the relevant schema.

Cases are an adjunct of the other methods (Kolodner 1993). Their use depends on the nature of the task, especially the kinds of episodes normally encountered and the degree to which episodes are temporally and/or conceptually distinct from one another. Case use also depends on common errors and the availability of data or simulation. A number of kinds of cases, each with its own strengths and weaknesses, are cited in the literature (Geiwitz & Kornell 1992): typical, tough, interesting, prospective, retrospective, boundary, partial, simulated, actual, minimal, and aided.

Omitted from Figure 5-2 and our discussion is the pacing of instruction. Deciding whether learning a particular task or topic should self-paced or instructor-paced depends on the learners and on the preferences of, and possibly the constraints faced by, the instructor. Other than some rehearsal learning that may not be susceptible to time compression beyond a certain point, the structure of the content does not in principle constrain the pace of its instruction.

One other issue not in the chart is what to say first. By "what to say first", we mean whether to organize material as simple-to-complex or complex-to-simple, concrete-to-abstract or abstract-to-concrete, or component-to-composite or composite-to-component. For example:

- In a domain where the whole determines the structure of the parts, instruction may organize material from complex-to-simple, e.g., the student is taught the overall mission before concentrating on his or her individual role. A different instructor, though, might take a different perspective, seeing the complex (e.g., mission success) as an emergent property of the parts (e.g., specified and implied tasks), and may wish to teach the simple first.
- A hierarchical domain can be taught from the top down (abstract-to-concrete) or the bottom up (concrete-to-abstract)
- Some kinds of knowledge are easier to absorb in the whole than in the parts, e.g., the concept of a "mobility corridor" is easier to learn than

all the possible complexes of features from which it can arise; others are easier in discrete "chunks", e.g., route planning in terms of mission, AoA, cover and concealment, estimated ToA, etc.

The point of these examples is to illustrate the position that these ordering choices, like the choice of pacing, depend on the learners and the instructor(s) more than the intrinsic structure of the task.

5.2 Content => Instruction/Training Heuristics

The content constructs--task, generic content and meta-content, error sources, characteristics of expertise, and conditions of performance--are now considered, and a taxonomy developed. In real applications, the content features don't appear in isolation. It's as though one were to analyze the use of the color red in a painter's oeuvre. There might be interesting and insightful comments to be made, but one could hardly understand the paintings by viewing them through a red-only filter. Similarly, saying that a domain is event-poor, taken alone, doesn't yield much. Is it a supernova? Or a forest fire? The former has a well-understood theory, is predictable once begun, and is neither influenced by nor consequential for humans. Forest fires are the opposite on each count.

In that spirit, this section looks at constructs and blocks of features, and the next section takes three representative tasks and considers feature-clusters and their consequences. It is organized on the lines of Figures 5-3 and 5-4.

5.2.1 Tasks, domain characteristics, methods, and solutions.

Convergent/Divergent Convergent problems have single answers, and the truth of the answer is in the world. Examples include the accuracy of a diagnosis, the numerical solution to a problem, or the precision of a Shuttle rendezvous. Divergent problems have solutions of greater or lesser elegance, rather than answers; and as long as the minimum problem constraints are satisfied, the quality of solution is in the head as much as the world. Elimination (structured pruning of the search space) and analogical case-matching (comparing current circumstances to known situation/solution pairs) are contrasting convergent methods. Least commitment (exploring multiple paths concurrently as data are gathered) and generate-and-test (exploring individual paths fully) are contrasting divergent methods.

Because these are such broadly inclusive categories, and because so many tasks partake of both, only weak training heuristics are associated.

For convergent tasks, precision and efficiency of execution are training goals. Classification methods always have declarative categories and structures; they must also have rules of evidence acceptability and evidence salience. There may also be specific conflict resolution rules. One would expect these to fit closely with the goal/subgoal trees, perhaps especially with the consequences associated with particular goals.

For divergent tasks, rules for combining elements, and for finding "minimum energy" paths or constructs within the problem space are required. A sense of elegance (since this is what guides selection of otherwise equal branch-points during solution construction), styles of approach, some form of apprenticeship, and in Army command training, careful graduation of responsibility are appropriate.

Data-rich/Data-poor. In a data-rich problem space, data selection and filtering are going to be significant elements. If the domain is one in which the relative distinctions between concepts are not clear, there are likely to be data management cognitive bias errors like anchoring and numeric bias. Then, counteraction will need to be taught. In some tasks, a goal-driven approach is used as a strategy to deal with too much data; the goals focus the search and provide criteria for data selection. Alternately, in a data-poor environment, there are likely to be long and elaborate chains of reasoning, and errors of knowledge selection and application, along with reasoning-oriented cognitive biases like base rate errors, will be problematic and therefore require special attention.

Event-rich/event-poor. In event-rich domains, availability bias is unlikely, and similarity-based induction is more likely to be reliable, since candidate generalizations can be frequently tested. In event-poor domains, availability bias is much more likely and generalizations are more likely to be explanation-based. There is a subdivision of problem spaces, though, with respect to number of events, based on the nature of the theory of the problem space. In problem spaces with strong predictive theories, availability bias is much less likely; volcanic eruptions are rare, but the arbitrary or spurious feature associations have long since been eliminated from analysis.

Hypothesis-rich/hypothesis-poor. In hypothesis-rich domains, data gathering, calibration, and integration techniques will be central, since discovering a path through hypothesis-space will be an important part of problem-solving. In hypothesis-poor domains, greater cognitive resources will be devoted to selection among a few alternatives; prediction is likely to be more significant.

High/Low experimentation. When experimentation is possible, e.g., as in diagnosing an electronic circuit card, a key element of instruction will be strategies for selecting the best experimental path, operational skills, and data analysis methods. On the other hand, when experimentation isn't possible, accuracy of situation assessment and predictive modeling are more significant.

High/Low standard deviation. If there is a strong central prototype and relatively little deviation, the training investment should be in the teaching and practice of basic skills. If there is a great deal of variation, and there isn't a strong theory to account for it, elaboration of basics and techniques for probing the periphery need to be taught. In domains where there is a great deal of variation in surface features but little in underlying causes, a productive approach might couple presentation of cases moving from the center to the periphery with explanatory exercises in which learners anticipate errors and justify mappings.

Optimal/Satisfying solution. When solutions are satisfying, teaching this explicitly and providing "solution calibration metrics" are training priorities. On the other hand, if an optimal solution is required, emphasis should be on method, plus error detection and repair. The metacognitive issue raised above, regarding resource allocation between finding the "perfect" tactic and spending time on execution is an issue in Command training.

Heterogeneity. The degree of heterogeneity and the depth to which different knowledge sources must be compiled both shape the influence of heterogeneity on training. Empirical results are equivocal. In some, e.g., studies of electronic device troubleshooting, results suggest that teaching theory apart from practice is ineffective; experienced practitioners learn theory only and exactly in those places where it has direct practical consequences. On the other hand, experiments in medical instruction where basic science teaching has been joined to clinical practice, and both taught jointly, have not been as successful as hoped; students taught science and clinical skills together do not appear to apply science as effectively in diagnosis as students taught science as a self-contained subject prior to clinical teaching. A reasonable, if tentative, inference might be, if multiple knowledge sources are required for effective problem solving, teach them separately, but make knowledge integration a distinct and substantial part of training.

There may be an integration issue for domains with mixed formal and heuristic methods. "Numeric bias" is a tendency to prefer and overvalue numeric data as compared with non-numeric data (Josephs 1994) (possibly in the belief that numbers are more objective?) Numeric bias might cause the results of the formal methods to be given excess weight simply because of their surface precision, however well or badly it might be justified by its underlying assumptions. "This (product of a formal method) is *true*, while this (product of an heuristic method) is only a *belief*." (Recall the discussion of modeless and modal teaching methods, above.) If this is applicable, then integration and proper valuing should be an explicit training objective.

Comprehension/Production. The issue here is task specificity. If the task is to recognize, evaluate, or critique, comprehension skills are appropriate. If the task is to plan, design, or act, production skills must be taught. This may seem trivially obvious. But if, say, Army helicopter pilots are taught route planning with a focus on route evaluation skills, then comprehension skills are being mistaken for production skills. Deciding whether a route is good or bad is different than generating the route in the first place.

What about using production training to enhance long-term recall of comprehension knowledge? For example, would two hours of drawing topographic maps result in longer term retention of concepts than two hours of looking at topographic maps? This is consistent with the "multiple access paths" theory of recall, but it is very sensitive to the content/technique balance. If the production technique is a separate and substantial body of knowledge, then time spent on it is likely to be "off-task". A non-programmer spending two hours attempting to reprogram a battlefield command decision aid is not going to increase his or her understanding of how to use the aid. On the other hand, if the technical skills are already adequate, as with drawing rough approximations of topo maps, then recruitment of productive intelligence could be effective.

Degree of distributed representation If distributed representation is part of problem solving, it may need to be trained. In the case of specialized representations like topographic maps or complex representations like computer-based simulations, fluency with the representation may be a training objective in itself. When dealing with distributed representation that will be used in the field, at least some training with high fidelity-to-field conditions is necessary.

Content

Concrete categories. Concrete categories are similarity-based learned. Exposure and rehearsal are appropriate. If the category applies over a broad range of surface features, a suitably broad range of examples is called for.

Definitional categories. Definitional categories are normally motivated by or descriptive of an operational end. Strongly associating definitions with their operationalization is recommended: definition and effective procedure, taught as a whole, rather than the more common "memorize this" approach, which often leaves students able to reproduce the definition but not able to apply it to anything.

Prototypical categories. In training prototypical categories, training should begin with central examples and only after these are mastered progress to boundary patterns.

Operators. Kotovsky and Simon (1990) found that the difficulty of finding and sequencing operators to be a better predictor of overall problem difficulty than the size of the search space. Zhang and Norman, *op cit.*, had similar results. If discovering and sequencing operators (i.e., figuring out what to do next) is problematic, teaching problem decomposition, developing automaticity with operator subsequences, working from simple to complex, and calling on learners to explicitly articulate, debug, and integrate operators will be productive. If operator complexity co-occurs with a difficult distributed representation and/or large distance between surface features and underlying structure, then explicit teaching of components and practice at applying theory to resolve impasses (explanation-based learning) is needed.

Attitudes. Gagné *et al*, *op cit.*, in discussing the teaching of attitudes, contrast direct methods like contingent reinforcement with indirect methods like providing opportunities to emulate a credible, powerful human model. They also provide guidelines:

1. Provide learners with information about possible alternative choices.
2. Provide the learner with the pros and cons associated with the desired behavior.
3. Provide relevant models for the desired behavior.
4. Ensure that the environment supports the desired choice behavior.
5. Fit the desired behavior into a larger set of values, if possible.
6. Identify and teach the skills that make the desired choice behavior possible.
7. Recognize and reward choice behavior when it is exhibited.
8. Don't inadvertently punish the desired behavior.
9. Allow learners to set their own goals with regard to the desired behavior.
10. Use alternate instruction strategies such as simulations, role playing ... or other involving experiences in which the benefits of the desired behavior become obvious.
11. Don't inadvertently pair the behavior you want to change with one that is not related to that behavior.

Patterns and meta-patterns, pattern-category associations. Teach perceptual recognition skills with pattern recognition and strong association with affective effects. Perceptual training should have an explicit, articulated cognitive goal, e.g., "this pattern implies this consequential real-world state". The affect association comes from the consideration that all of our sensory activity is *motivated*. Most of what the senses and brain do is filter and inhibit. Successful sensory events need a limbic system "sponsor"—pleasure, caution, pain, whatever. The difference between "memorize this topo map pattern" and "this one is called a kill sack, and if you blunder into one

everyone under your command including you will be killed". The second one is vivid and memorable. The perceptual/affective link, besides being obviously true between people, makes sense from the evolutionary point of view that the arousal/importance of a sensory experience should be associated with its potential to alter one's state, and that potential for state alteration causes/is associated with an emotional response. Creating such an association in training would strengthen its effect. And the inverse seems almost more true, that separating sensory training from affect would greatly weaken the training. At a more abstract level, this is saying that we learn and remember what is meaningful, and that for perception affect is what determines meaning. This isn't saying that there are no cognitive consequences of sensory experience, or that affect is not important for cognition; just that there is a strong evolutionary and biological connection between affective associations and the memorability of sensory experience. When the meaning-words associated with a pattern are also value words ("better", "safer", "stronger"), affect is tightly bound to pattern. Training should combine pattern, association, and consequence or affect.

Most sensory pattern recognition will be concerned with prototypical categories, and should therefore progress from central to boundary examples.

Proprioceptive knowledge, motor skills. Teach proprioception and motor skills with demonstration, mimicry, visualization, repetition, and proximal reinforcement. Practicing boundary conditions may be productive. For motor skills, boundaries can include higher-than-task requirements for speed or accuracy of performance, or extreme versions (exaggerations, caricatures) of performance. Performance speed for motor skills can sometimes be taught "slowed down" with gradual increase to normal speed; not all motor skills are amenable to this, though. Part-skills-to-whole-skills training is common for motor skills.

Goals and goal conflict resolution. In data-rich domains where goals are a principle means of organizing search, goal trees and goal satisfaction criteria should be taught explicitly. This is also the case for domains where there is a great deal of uncertainty and goals are used to organize and calibrate data. Goal conflict resolution can in some cases be taught by training valuing some goal hierarchies over others. However, this won't always be true, and field conditions will inevitably result in conflicting goals. For example, defend in sector until ordered to retreat versus being killed. For this class of goal conflict, enculturation to the Army's better virtues--honor, courage, loyalty, steadfastness--is the only adequately deep approach.

Implications. A great deal of military command decision making is based on predicting the future. Some prediction is statistical, and calls for explicit

training in use and interpretation of statistics. (This is a currently undertrained area that will grow in importance in the coming decades.) For non-statistical prediction, use of computer-based simulations accompanied by explicit practice at making, articulating, justifying, and verifying predictions should be provided. This is not currently a typical use of simulations. If the position is accepted that prediction is a critical command decision making skill, explicit practice seems sensible.

Coherence, consistency, completeness. Knowledge coherence influences the extent to which inferences can connect different knowledge elements. This helps insure knowledge consistency, aids knowledge debugging, and helps regenerate partially forgotten knowledge. It can be taught by having students articulate their rules and hypotheses, examine their own work and the work of others to identify the sources of error, and to work out or describe the boundary conditions for the concept, operation, or process in question.

Urgency. Learning the urgency of different demands and the attributes and values that mark important situations is basic to military training. Misplaced urgency is damaging; failure to realize that urgency is needed can be even more damaging. While urgency is a response to a situation rather than an abiding mental state, much of heuristic guidance for attitudes applies. Apprenticeship, task immersion, and learner-directed post-event analysis are appropriate.

Confidence. Appropriate confidence can at one level can damp cognitive noise, and at another can guide knowledge application. Overconfidence can lead to incorrect knowledge application and inadequate attention to plan validation; lack of confidence can result in failure to apply available knowledge. The last stages of training should as closely as possible follow real-life performance difficulty, conditions, and consequences so that learners can develop appropriate levels of confidence. In young military commanders, overconfidence appears to be more problematic than its lack (Fallesen *op cit.*); this may need to be addressed.

Uncertainty. Uncertainty can be associated with data, knowledge, knowledge application, and projected results. Empirical results (e.g., Tversky *op cit.*) show that people are, in many surprisingly common cases, terrible. On the other hand, Anderson, *op cit.*, criticizes some of this research (i.e., base rates, representativeness) as being contrived, artificial, and inapplicable to real life. Training to avoid anchoring, to employ multiple perspectives (to break up framing biases), and explicit practice at assessing the coherence, completeness, and consistency of data are appropriate. As an aside,

framing/reframing practice can aid performance when indexing and recall are issues.

Truth maintenance. AI research suggests that both uncertainty management and truth maintenance are error-prone and resource-intensive. (This is consistent with the K.I.S.S. theory of planning and operations.) Distributed representations can aid truth maintenance, but the teaching of conditional probabilities for real field use is problematic.

Metaknowledge and metacognition. The issue is, is there an issue? If there is, then a (possibly implicit) task/objective has been identified, namely, train the appropriate competence(s). Repeating the list provided in the previous section, the metaknowledge steps are:

- Problem discernment and identification
- Problem representation
- Validation of problem representation
- Planning strategies for problem solving
- Validating the plan
- Applying the selected method
- Implementing the solution
- Monitoring implementation and the developing solution
- Evaluating the solution

Error sources.

Distance between surface features and deep structure. When this is the case, a number of heuristics apply (Collins & Stevens 1982):

- Systematic variation of cases, to explore/traverse the problem space
- Counterexamples and hypothetical cases, to reveal inconsistencies, incompleteness
- Entrapment cases, to illuminate errors
- Hypothesis identification strategies, to formulate/articulate predictions and rules
- Hypothesis evaluation strategies, to aid knowledge integration and concept refinement

Cause/correlation confusion. If there are causal relations, then the goal is to learn to recognize causal states, correctly fire appropriate rules, then check to see if you're done. With correlations, calibration/qualification/combination of evidence comes into play. With causal relations, uncertainty focuses on correct identification (of the presence of causal factors) and correct knowledge selection (e.g., firing the right rule). With correlations, the issues of combination of evidence and knowledge uncertainty are added. Confusion of correlation for cause can result in

unrealistic certainty factors, misplaced confidence, and misapplication of cognitive resources.

Cognitive biases. There are many cognitive biases; Silverman (1986) lists twenty-six. Three particularly important ones in a military setting are anchoring, availability, and framing. "Anchoring" is selecting a conclusion early in a decision process, and then filtering incoming data to fit the conclusion. This may include exaggerating supporting evidence, and suppressing contradictory evidence. Entrapment cases and learner solution critiques are useful training approaches. "Availability" is use of the most easily indexed or retrieved knowledge or experience to reason about a current situation. This can arise from time-proximal cues or, more pernicious in the military, from having only a small number of exceptionally vivid experiences to draw on. While pragmatically it is problematic, field training is the appropriate training approach. "Framing" is how a problem or situation is stated; in effect, the position from which the problem is viewed. Frames carry implicit, often unconscious, constraints and preferences. For example, people prefer a 95% chance of survival to a 5% chance of not surviving, although of course they are the same thing. A recent \$6 million Army effort framed a group planning task as "produce a plan", and as a result developed a very nice computer-based expert planning aid. However, there is some evidence from the previous way of doing the job--planners gathered around a table, with large maps, overlays, and grease pencils--that the real task for the commanders is to jointly develop, synchronize, and motivate a common *mental model* of a plan. Framed from this perspective, creating a single-seat, small-scale view of the problem is exactly wrong. Fortunately, framing and reframing can be practiced, and awareness of frames can be trained by the simple expedient of drawing attention to them and having learners practice generating and evaluating multiple frames.

Knowledge errors. The obvious cure for lack of knowledge, irrelevant knowledge, and erroneous belief is of course to "train better", so that these aren't problems. An issue in Army training, though is what might be called "instructor loyalty to text". For example, increasing interest in naturalistic decision making and evidence that real commanders use this approach effectively in the field (e.g., Klein 1992) has apparently been met by many instructors with a "well, they're doing it *wrong*" attitude. While it's not surprising that in a discipline with so much uncertainty attachment to doctrine and process might arise, special training attention may need to be paid to maintaining the distinction between the actual task and its description.

If solutions are ambiguous, difficult to discern, or transitory, solution identification needs to be taught explicitly. This may be particularly true in

situation assessment and monitoring tasks, particularly with regard to correctly assessing changes of state. (In monitoring tasks with few state changes, anything that does happen is *ipso facto* important.)

Bad resource management, inappropriate goals. Training both of these can apply both to problem solving and to problem solving management (metaknowledge and metacognition). For example, a problem solving procedure may be carried out correctly, but if too much time is consumed, there has been an error in process management. Or, to reverse this, realistic process management may compel alterations ("errors") in problem solving. Fallesen *op cit.* reports a seriocomic example. In a survey of Gulf War commanders, 96% reported abbreviating or abridging the Command Estimate process due to time constraints. The 4% who didn't explained that they omitted the Command Estimate process entirely.

Failure to anticipate enemy. Andriole et al (1987) found that more experienced commanders were more risk-averse than less senior commanders, and spent more time looking for ways that things could go wrong. The metaknowledge steps of representation validation, problem solving plan validation, and solution development monitoring were apparently better developed in the senior as compared to the junior commanders. The heuristic, then, is to explicitly teach the validation steps.

Task conditions

For the Army, performance and training conditions can range from real-actual combat-through, in reverse order of realism, NTC, field training, CPX (command post exercise), high-fidelity simulation, low-fidelity simulation, and of course, classroom.

High time- or consequence-stress. If expertise is routinely exercised in high time stress circumstances, it's almost certain that expert practitioners can't accurately describe their problem-solving behavior in other than operational terms (e.g., "when X happens, do Y") (Ericsson & Simon 1984). Any reports experts make of their thinking or decision making will be necessarily retrospective, and is likely to be a fabrication based on the individual's personal theory of how he or she decides. Conscious deliberation is likely to be reserved for (explanation-based) resolution of impasses, with maximum automaticity of frequently occurring tasks. Training should mimic this, practicing performance with feedback on routine sub-tasks, and reserving conscious decision making exercises for impasse-resolution.

Concurrent tasks, concurrent performance demands, concurrent obligations. In reviewing these, as well as explicit concurrent tasks or obligations, it

should be remembered that in all conditions, from actual warfare to classroom, there is always at least one concurrent task; namely, senior officers are expected to train junior officers.

Environmental stress, degree of fatigue, interruptions. These are all performance degradation factors: non-enemy interference with task performance. Particular attention should be paid in training to "catastrophic error" points, so that the points of highest risk of failure or of failure with the greatest consequences can be trained intensely. (In this connection, if workload analysis has not been performed, and cognitive loading is an issue, it ought to be.)

5.3 Representative construct clusters & heuristics

In this section, three common classes of problems are described in terms of constructs, implications, and heuristics. Applying definitional concepts to organize data, situation assessment via classification, and prospective reasoning were chosen to cover a broad range of Army tasks. Each will be characterized with respect to the constructs, then the heuristics will be used to formulate guidelines for a training approach.

5.3.1 Applying definitional concepts to organize data

Applying definitional concepts is a central activity of field engineering, artillery targeting, calculating force ratios, scanning a battlefield for nuclear, chemical, or biological contamination, planning logistics functions, and many other common tasks involving calculation. Therefore, the pedagogical goal is to unite definitions with their "operationalization", including both how to apply the concepts and validity constraints specifying when they can be used legitimately. Following is one way to frame the problem.

Applying definitional concepts is:

- Convergent
- Event-rich
- High experimentation
- Low standard deviation
- Both recognition and production
- High degree of distributed representation
- Data-rich
- Hypothesis-rich
- Optimal solution
- Homogeneous content

The only attribute assignment that might be controversial here is "low standard deviation". There may be significant variation in superficial features, and presumably to the extent that surface features are taken as a valid part of a distributed representation, these can strongly influence outcomes (Wason & Johnson-Laird 1972). On the other hand, normally definitional concepts are only robust if the underlying features are stable. So, low standard deviation in reference to the deep structure is a fair assessment, despite the high variation in surface features.

Generic content includes concrete categories, definitional categories, attributes and values, and operators. These are organized both by goals and by data structures. Key mental operations include:

- Indexing and retrieval • Procedures
- Plans • Hypotheses
- Goal satisfaction points
- Data identification/classification methods
- Error detection markers

All of the "management" issues apply:

- Problem discernment and identification
- Problem representation • Validation of problem representation
- Planning for problem solving
- Validating the plan • Applying the method
- Implementing the solution
- Monitoring implementation and the developing solution
- Evaluating the solution

Metaknowledge and metacognitive issues include:

- Coherence • Consistency
- Completeness • Systematicity
- Knowledge application uncertainty
- Truth maintenance

Error sources include:

- Distance between surface features and underlying structure
- Problem misrepresentation
- Irrelevant features • Lack of knowledge
- Irrelevant knowledge • Inappropriate goals

The four most striking features of expertise are knowledge organization, depth of problem representation, automaticity, and metacognitive self-regulatory skills, e.g., as in monitoring solution development, detecting false trails, and performing truth maintenance.

Conditions are typically ideal for practice, i.e., there is adequate, uninterrupted time, although of course in the college final exams of such warm memory, there are time and consequence stresses.

Stepping back from this overall characterization, what are the emergent patterns? In general, the hardest problems for students are (White 1983):

- Seeing the true nature of the problem beneath its surface features
- Representing the problem correctly
- Anticipating and/or detecting errors both of representation and implementation
- Deriving maximal benefit from problem solving success.

The first two, taken together, are problems of indexing and retrieving the correct knowledge and avoiding premature knowledge selection. The third involves operator selection and application, and solution plan monitoring and

tracking. The last involves integration of specific case knowledge into the larger problem-solving repertoire.

What heuristics can we read from the taxonomy presented in the previous section? Management of a data-rich, hypothesis-rich problem solving space requires well-understood goals and goal structures, goal/data relevance, and with clear goal satisfaction criteria. From the distance between surface and underlying features, systematic variation of cases, showing many sets of surface features with the same deep features, is indicated. From the likely degree of distributed representation, explicit practice to develop fluency with the representational tools is appropriate. Development of automaticity calls for isolation and practice of component ("between decision point") skills. Error management calls for explicit practice at articulating the causes, consequences, and markers of errors. Finally, since consistency and freedom from contradiction are critical features of the definitional concepts, explicit "consistency and coherence" practice in integrating both specific case success and general evolving knowledge is called for.

To organize these, four powerful heuristics for teaching complex definitional concepts from Reif & Allen (1992) are appropriate:

1. Formulate a clear definition of the concept and accompany it by an explicit procedure specifying how this concept can be interpreted in any particular instance.
2. Let students practice applying this definitional procedure to interpret the concept in a variety of special cases.
3. Let students use their explicit knowledge about the concept to confront situations that are error-prone, either because of intrinsic difficulty or because of confusion with prior knowledge. In each case, ask them to detect mistakes of concept interpretation, diagnose the likely reasons for them, and correct them.
4. Guide students to summarize and organize their accumulating case-specific knowledge so as to acquire a useful repertoire thereof--and then encourage students to use their case-specific knowledge in familiar situations.

5.3.2 Situation assessment via classification-elimination.

Situation assessment is one of the most important command responsibilities. There are a number of distinct intellectual elements in situation assessment: data gathering, filtering, combination, and interpretation, along with ongoing assessment revision. Classification-elimination is a data-driven classification technique in which potential end-states (answers, diagnoses, decisions) are found by successively eliminating candidates until only one remains. The strategy is to progressively prune the search space by propagating the consequences of constraints or intermediate conclusions; the problem becomes increasingly well-specified as problem solving proceeds. Elimination is used at a number of stages in situation assessment. Sometimes

its use is unconscious, as in selecting what initial categories and possibilities are of interest. Sometimes it is used intentionally, as in assigning missions to recon resources, in attempting to validate or rule out enemy plans or positions.

As a classificatory technique, it presumes prior existence of case-independent categories, attributes, and values, and reasonably strong power of evidential reasoning. These qualities are easy to see when elimination is contrasted with another common classificatory technique, analogical case-matching. There, attribute salience, evidence weight, and notions of similarity are more prominent.

As above, we first characterize elimination, then read off the general heuristics, then suggest a generic training plan.

Elimination is convergent, data-rich, and is oriented toward comprehension. In hypothesis-rich domains, there may be important distributed representation issues. It is neutral with respect to other task and domain characteristics.

As a method, elimination is concerned with:

- Concrete categories
- Operators
- Attributes and values
- Actions
- Pattern-category associations
- Definitional categories
- Prototypical categories
- Perceptual knowledge
- Patterns and meta-patterns

Data structures and their implications are the central source of meaning.

Key mental/procedural functions include

- Plans
- Hypotheses
- Data identification/validation methods
- Error detection markers
- Procedures
- Data gathering methods

Problem solving management places emphasis on:

- Planning strategies for problem solving
- Validating the plan
- Implementation
- Monitoring implementation and the developing solution
- Evaluating the solution
- Applying the selected method

Metaknowledge and metacognitive issues include data uncertainty, results uncertainty, truth maintenance, and systematicity.

Conceptual error sources include presence of irrelevant features, getting stuck at local optima (e.g., pursuit of unavailable data—"if only we knew X"), cause/correlation confusion, and unjustified induction. "Anchoring", a cognitive bias in which an early attachment to a particular hypothesis leads to misinterpretation of evidence, is a typical issue in elimination. Irrelevant knowledge—which can lead to a variation on the "availability" bias—and erroneous beliefs, especially regarding the implications of data, are more problematic with elimination than with other methods. Pragmatic error sources

(lack of data, false data, etc.) are normal for Army field operations, but their seriousness is circumstance-sensitive. It is easy to see where each of these breakdowns can and do occur in battlefield situation assessment.

Expertise is characterized by knowledge organization, theory and schema change, and procedural and goal-oriented knowledge. Conditions, like pragmatics, are domain sensitive.

As a convergent, data-rich process, data management and evidence calibration are going to be central issues. In some domains, data management can be achieved through familiarity with the structure of the problem space. When this is true, systematic variation of cases to explore and traverse the problem space is appropriate. If evidence calibration is an issue, as it is in battlefield operations, then two kinds of counterexamples and hypothetical cases are needed: one to reveal inconsistencies and/or incompleteness in the learner's knowledge, the other to illustrate data dependencies and connections.

On the theme of navigating through a data-rich, hypothesis-rich problem space, hypothesis identification strategies should be taught to aid the learner in formulating and articulating predictions and rules. Hypotheses should be integrated with data-gathering plans and procedures, and students should be taught hypothesis and plan evaluation strategies to help them integrate local plans and procedures into a domain-specific system. If there are restrictions on plan generalization, these should be brought out.

To the extent that reasoning depends on statistical distributions, explicit training in the appropriate formalism is indicated. Realistically, if base rates are known, their strong emphasis in training is justified. However, this is a double-edged sword. The Soviet "scientific" approach to battle resulted in explicit mathematical formulations, and since the Soviet command structure reserved authority almost pathologically to the topmost command elements, Soviet maneuvers were highly predictable. Their base rates were known in detail. This knowledge is apparently so seductive that training continues to focus on defeating the Soviet threat, even though there is no more Soviet army. (They don't exist, but if they did, we'd be ready.) The Army has been a little slow to change here.

Preventing local optima errors suggests metacognitive self-monitoring skill and training the parallel pursuit of multiple paths (so that fixation on a single subplan does not impede problem solving progress). Preventing anchoring errors suggests use of entrapment cases, that is, cases where the early evidence suggests one solution but further work shows it to be false. Some evidence suggests that current Army training is not effective in this regard [Thordsen et al 1991].

Developing expertise in procedural knowledge application and accuracy of data gathering and/or interpretive technique will be important in domains characterized by complex mappings between surface features and deep structure.

The higher the level of command, the more true this is. When minimization of data gathering and interpretation error is within the control of the practitioner, it must be fully developed to prevent false solutions resulting from correct inference based on incorrect data. If sensory or spatial pattern recognition is important to data interpretation, simple-to-complex, prototypical-to-peripheral pattern recognition training should be given.

In domains with high data uncertainty—namely, almost all situation assessment—the effective practitioner will need to know when to hold to normal constraints and when to relax them. Thus, there should be explicit training in recognizing impasses and identifying which constraints can be relaxed with the minimum damage to the solution process. In other words, for anticipated battlefield situations, the "degrees of safety" of different assumptions should be taught. (Of course, given the number of competing theories of how battlefield decisions are actually made, broad agreement on which assumptions are or are not safe is not very likely.)

With respect to generic content, as a classification problem elimination will have categories and category structures, plans, hypotheses and hypothesis structures, device, system, or process models, data gathering and calibration methods, and context-specific models of systematicity. Elimination, which specializes classification, will also have a constraint space model.

As a final complicating element, order effects are important in training. Depending on the domain, even apparently distinct elements may be learned differently depending on when they are introduced relative to the rest of the corpus. In the example documented by Patel *et al*, *op cit.*, teaching medical students biology in conjunction with clinical application results in different learning than teaching biology and clinical practice serially. By analogy, does teaching theory and historical analysis of warfare in conjunction with field exercises generate a different result than teaching the theory prior to the application would? Patel *et al* would answer, "yes." This question is, as far as we could determine, not yet researched in the command decision domain.

5.3.3 Prospective reasoning

Cohen, Greenberg, and Delisio (1987) characterize prospective reasoning by four attributes:

- Knowledge about the state of the world is incomplete
- Outcomes of actions are uncertain
- Decisions as to which action to take affects multiple, conflicting goals
- Actions can produce new information that can change the state of knowledge about the world and the utility of taking other actions.

Prospective reasoning is of obvious interest to the Army, since most battlefield decision making will be prospective.

Task characterization is not constrained by this definition, other than the inference that satisficing rather than optimal solutions are likely.

Similarly, the kinds of categories and operators, goals and data, and procedures and hypotheses are unconstrained.

Because of its "open" nature, how prospective reasoning is managed is key to success. From that position, all of the following may apply:

- Problem discernment and identification
- Problem representation • Validation of problem representation
- Choosing strategies for problem solving
- Validating the plan • Applying the selected method
- Implementing the solution • Data uncertainty
- Monitoring implementation and the developing solution
- Evaluating the solution

Metaknowledge and metacognitive issues include:

- Salience, goal conflicts, urgency • Confidence
- Allocation of cognitive resources • Knowledge application uncertainty

Error sources are not strongly constrained by the definition. Depending on the domain, local optima, false data, lack of data, cognitive loading, and cause/correlation confusion might all be significant. Inadequate attention to anticipating the enemy has been shown to be a novice mistake in military settings (Andriole et al, *op cit.*) Three common cognitive biases in prospective reasoning are anchoring, availability, and choice on more important attribute. All are unconscious techniques for simplifying decision making. Anchoring and availability have already been described. Choice on more or most important attribute manages complexity by blanking most of the factors out, then basing the decision on the one or few factors remaining. Preventing correlation-for-cause errors requires instruction to make common errors of this type explicit; depending on the domain, counterexamples and entrapment cases may be used.

Expertise can be roughly characterized as distinguished by extensive procedural and goal-oriented knowledge, automaticity to reduce attentional demands, and metacognitive self-regulatory skills. Conditions of performance are not constrained by the definition, although in an Army setting, time, consequence, and environmental stresses are all likely, as are fatigue and concurrent tasks and performance demands.

Because prospective reasoning is much more about effects and circumstances of reasoning than reasoning itself, it is difficult to select heuristics. If problem discernment is an issue, for an example, it needs to be taught. (Infiltration and sabotage is harder to detect than full frontal assault. Or, an alternate example, mortality and morbidity of troops is easier to discern than gradual decline in morale.) The same is true for problem representation, solution planning, and so on.

A key issue in the military domain may be training an explicit distinction between common-experience and military concepts. To take an obvious example, to a Brigade commander a hill is an objective, or an observation point, or a

potential enemy location, etc. The point is to make explicit the differences and similarities between common terms and concepts and their domain-specific uses.

Classroom training in goal conflict management and dynamic plan revision should be followed by either simulated or field practice (preferably, by both.) Depending on the cost of creating simulated or field training episodes, guided discovery may be effective. Apprenticeship is likely to be appropriate; the Army has institutionalized a degree of apprenticeship in response to this need.

Training to prevent choice on more important attribute can be achieved with entrapment cases. It is also important that the culture of learning and practice not promote a "magic bullet" attitude: the notion that the "Brilliant Strategic Insight" is the hallmark of the highest level of Command expertise. Battles--like all other complex human endeavors--are won by hard work, careful planning, and sound execution far more frequently than by unique Command insight.

Development of automaticity calls for repetitive practice of the target processes and procedures, with appropriate variation in input and output. As with elimination, above, sequencing of training goals depends on the domain.

5.4 Critique of the taxonomy

There is no generally accepted taxonomy such as is attempted above because:

- Learning theory is neither strong enough nor precise enough in its predictions to generate a taxonomy. Even John Anderson, who in ACT* (pronounced "act-star"; Anderson 1983) created the first comprehensive unified theory of learning and cognition, in his most recent book turned strongly to situation-based rather than learning theory-based explanations for cognition.
- Common sense is an unreliable guide, both in the small and in the large. In the small, apparently similar things, e.g., visual and spatial discrimination, turn out to be significantly different. In the large, every learning reform seems to be based on "obvious common sense", even when the proposed reform directly contradicts its immediate predecessor.
- Empirically-driven approaches suffer from lack of precise common vocabulary, contradictory data, and extreme difficulty in identifying and systematically varying independent factors. (Also, in the open literature, about half of the "real topic" data is based on fifth- and sixth-graders learning physics, and the other half seems to be on college freshman learning physics.) The "ideal" empirical approach, induction of decision trees from actual case data, requires clearly-defined categories, well-understood attributes, and significant values. Army battle data are far from this.

The method is a procedure to characterize a training objective and to generate strategies for training. It is not a method for predicting or evaluating performance, although evaluative factors are elicited, since a training objective is inadequately specified if its standards of performance are hazy.

The reliability of attribute value assignment is an issue for any classificatory method. Would independent users come up with the same decompositions and judgements when applying this method to the same task? Standard terms and meanings have been used to support consistency. Still, this is an empirical question that cannot be answered in Phase I.

6.1 Ontology

Before describing the method, its ontology—the way it views and construes the world—must be described. There are three reasons:

- The method characterizes training objectives and requirements. Characterization requires a point of view about what *is*, and about how it should be viewed. That's the essence of an ontology.
- Ontology is the source of the inevitable and highly useful inductive bias necessary to perform characterization. Differentiating the important constructs, attributes, and values from the unimportant is based on what the domain is perceived to be. Ontology is the foundation from which goals can be pursued.
- Explicit ontology allows clearer distinction between intensional and extensional elements. Intensional aspects are concerned with viewpoints, functions, and uses. Extensional aspects are concerned with the representation of observable "thing-in-itself" qualities. In most domains, maintaining clarity between these two is fairly easy. As soon as knowledge and learning are introduced, though, the boundary becomes very fuzzy.

Ontology at least attempts to describe the bounds.

Articulating and scrutinizing the ontology of a method is better than leaving it unstated. An unexamined ontology is an invitation to error.

Why was the ontology not developed as part of the theory? Because it is the vehicle for applying the theory to actual problems, but it is not intrinsic to the theory itself. Objects can be used in convergent and divergent tasks; events can be data-rich or data-poor; agents can perform physical or mental actions. In many domains, ontology is inseparable from theory. For this problem space and this approach, the ontology is applicative, not theoretical.

The basic model underlying the method is as shown in Figure 6-1. The impetus for this model came from a number of sources; most strongly from Ortony, Clore, and Collins (1988). There are four fundamental categories: events, agents, and objects, and the problem spaces with respect to which they are meaningful. By this definition, problem spaces are at a higher level of abstraction than events, agents, or objects. Put another way, events, agents, and objects instantiate problem spaces, but there are characteristics of the higher-level abstraction that are not particular to its components. It will be referred to below as the SEAO (Space/Event/Agent/Object) ontology.

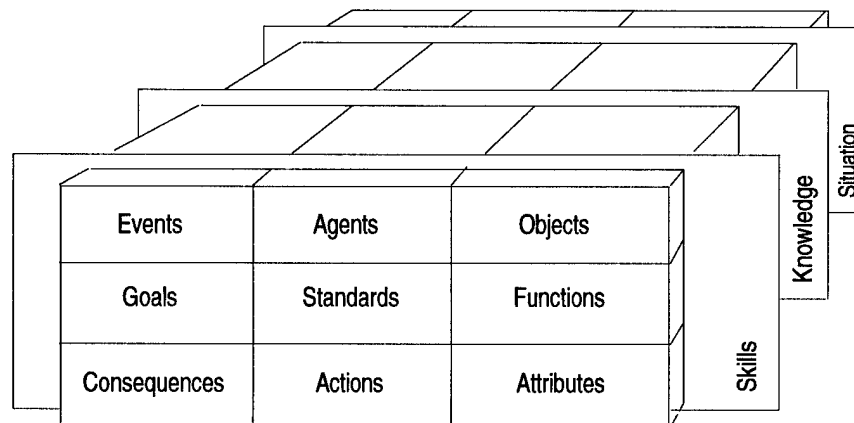


Figure 6 - 1. Events, agents, and objects as a problem space.

An event is a transformation of the internal and/or external task state or conditions, which may or may not include the generation of a product. States may be either objective (e.g., meet at assembly area), subjective (e.g., understand a Concept of Operations at a briefing), or both. Events are meaningful because of their consequences; these consequences are evaluated with respect to the central agent's goals. That is, when an event occurs, the agent judges it according to how it affects the achievement of his or her goals. Agents initiate events in order that their consequences can bring goal achievement closer.

Agents are generally individuals. According to context, groups or organizations can be treated as agents, i.e., as having unified intentions and actions. Agents take actions; actions can be evaluated by standards.

Objects are the physical things used in tasks. In some circumstances, agents or organizations may be treated as objects during task performance, e.g., a unit might maneuver around another unit. Objects have functions; agents interact with objects via object attributes.

Problem spaces are events, agents, and objects taken as a whole and including emergent qualities and features not present in the components considered apart.

Chi and Slotta (1993) propose an alternate ontology; so do Alexander et al (1986).

The Chi/Slotta ontology is shown in Figure 6-2. Its basic division of the world is into what's physically present, what it does, and what we think about it. While there is no doubt this is a psychologically valid decomposition, it is insufficiently differentiated to support the goals of the method. For instance, a person's body would be classed as matter. In a military context, though, people's bodies aren't treated as matter (except by logistics). Matter, even living animal matter, has the wrong features and associations. Another example: part of the Command Estimate process requires reasoning about enemy intentions. In mapping such a concern to the Chi/Slotta ontology, it is not clear whether it is about a potential intentional event or about an emotional or intentional mental state; it seems to be both. The SEAO ontology

is much more parsimonious: intentions belong to agents, and the distinction between intentions and actual events is kept clear.

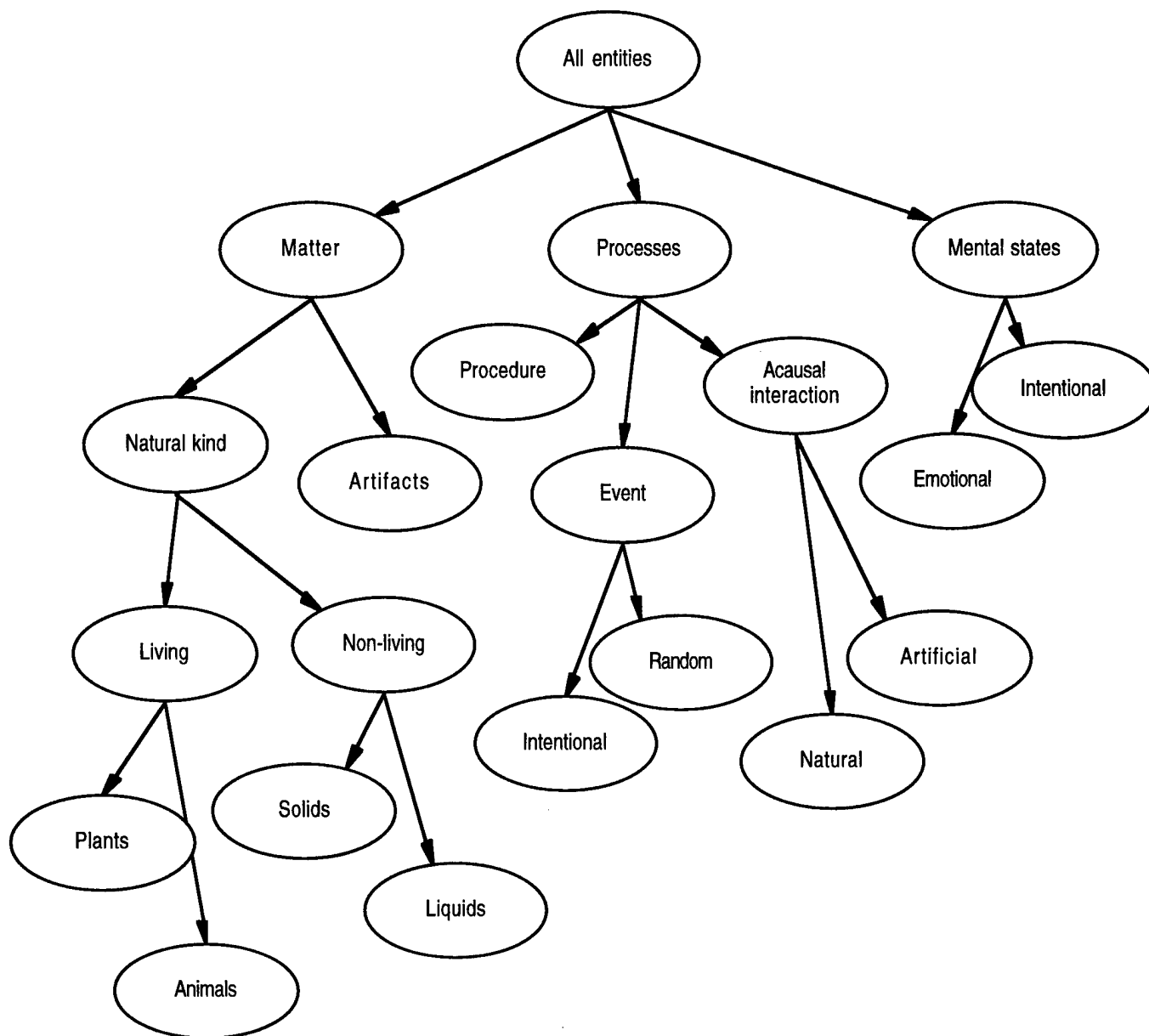


Figure 6 - 6 The Chi/Slotta ontology.

The Alexander/Freiling ontology provides a formal language with statements (domain equations and domain function declarations) and operators (discriminated union, cross product, domain mapping, set collection, ordered set collection). It distinguishes static, dynamic, and epistemic levels of ontology. The static includes the physical or primitive objects in the problem space; the dynamic the state space and transformations in the problem space; and the epistemic defines the constraints and methods that control knowledge use in the other two levels. If we translate from problem space to vernacular

language, it has essentially the same structure as the Chi/Slotta ontology: what's out there (static), what happens (dynamic), and what we think about it (epistemic). So, for the proposed method, it has the same problems as Chi/Slotta.

Clarity and useful differentiation are enough to recommend the SEAO ontology. There are two more reasons to prefer it. First, it is grounded in affect. Second, it appears to be immediately productive in an Army context.

Grounding in affect is crucial. While we aspire to objectivity, truth, and so on, it is the affective aspects of these goals that make them important. We pursue desirable goals and avoid undesirable consequences. High morale defeats low morale. Emotions organize sensation, cognition organizes emotion, sensation drives emotion, emotion drives cognition. Separating one from another is a necessary and useful artifice, but it's just that—an artifice. It doesn't justify a pretense that they are independent. Humans didn't evolve to pursue meaningless goals, and the root of meaning is emotion. The SEAO ontology evolved from the study of emotion, and like the implicit lines in a perspective drawing, all of its elements point to the experiencing and responding human at the center.

Military applicability can be seen by considering a recent ARI report (Fallesen *op cit.*) Analysis of this report suggests that the Army teaches objects reasonably well, events less well, and agents least well of all. Specific results supporting these conclusions:

- Of nine areas called out as sources of breakdown or diminished performance, none referred to use of objects
- Critical events, i.e., carrying out estimate procedures, managing the planning process, routinely fell short of training standards.
- Agent-oriented actions were in some cases extremely poor, e.g., information exchange rates as low as 17% for sharing of significant information.

6.2 Overview

The basic method is to examine in turn the events, agents, objects, and SEAO interactions involved in the training objective, first working them out according to the perspectives shown in the Figure 6-1 (events in terms of goals and consequences, etc.), then looking at each according to the constructs developed in Chapter 4. The second stage is to map the results to instructional and training requirements. How this proceeds is shown in detail in the next subsection, and in a fully worked-out example in Section 7.

Repeating an earlier assertion, this is not task analysis nor a substitute for task analysis, particularly with respect to functional decomposition. Task description should generate a statement of task steps. Task analysis should extend that by specifying the knowledge and skills necessary at each task step, along with the bounds of the task and the input/output requirements. A good task analysis should provide a thorough statement of the task. The method presented here is at a different level of abstraction. It is an attempt to develop statements about the task. The implicit ordering—task analysis, then this method—is structural, not incidental.

As suggested by the diagram, the method constructs a matrix, although because of the volume the information involved, a virtual matrix might make more sense than an actual one.

Unless explicitly called out, a human-centered perspective is assumed. That

should be clear from the descriptors used to characterize events and objects, but it is worth repeating.

Stage 1, Characterize objective, has three steps.

1. Build inventories. List the events, agents, and objects involved in the task, and for each, make the appropriate elaboration, e.g., goals and consequences for events, etc. This is in effect the descriptive raw material of the method.
2. Connect the lists internally. Flows and dependencies should be articulated. Most of this work will already have been done in the task analysis, especially if the flows are temporally ordered. However, since task analysis doesn't usually take the Event/Agent/Object perspective, and since not all tasks are amenable to temporal decomposition, there may be new information generated here.
3. Analysis. The results of steps 1 and 2, along with aspects of the problem space, are analyzed with respect to the constructs from Chapter 4.

Stage 2, Develop training requirements, has five steps.

1. Approach to events.
2. Approach to agents.
3. Approach to objects.
4. Approach to the problem space.
5. Integration. This step is more than sequencing and consistency checking. Coherence, that is, the degree to which both task objectives and training approaches are mutually supportive, is critical. Therefore, it must be explicitly analyzed and developed.

6.3 Stage 1: Characterize objective

Step 1. Build lists

- a. List significant events, agents, and objects involved in the task.
- b. For events, list goals; for each goal, list subgoals, if appropriate; then list consequences of reaching or failing to reach the goal(s)/subgoal(s). Consequences should include both direct (e.g., mission failure, death) and indirect (e.g., alteration in the space of possible next or future significant actions)
- c. For agents, list standards for normal performance—these may track the goal/subgoal structure, or they may not. For each standard, list the actions necessary to achieve that standard. Include constraints, if appropriate (e.g., a limit on the time allocated or allowed for a task or subtask).
- d. For objects, list salient functions. For each function, list salient attributes. These are task-sensitive. For Artillery, range, payload, and mobility are key functions; for logistics, object functions may be "consumes X volume of Y per unit time" or "takes up Z space".

Step 2. Connect the lists with respect to the task

Consider event flows, agent flows, and object flows, then map the flows to one another. Event flows are linear or parallel sequences of task-relevant events. (It should generally be possible to read event flows almost verbatim from task analysis documents.) Agent flows follow each agent over the course of task performance. Object flows track the use of each task-relevant object during task performance.

Experience in knowledge acquisition suggests that it may be appropriate to start with normal flows, then develop an appropriate number of breakdown or pathological flows. What constitutes "an appropriate number" will vary with the task and the objective.

The result should be a set of statements or diagrams encoding information of the form "this agent uses this object at this point in the process to initiate/continue/complete this (desirable) event." The intent is to examine the connections between events, agents, and objects as they combine within the task.

Step 3. Analyze list and connection content

This is an iterative step. First, each substep should be performed for the events; then the same substeps should be performed for the agents; then for the objects. As appropriate, they should be performed for the event/agent/object interfaces. Then, they should be performed for the problem space. Because in most cases redundancy will be substantial, successive iterations should be quicker and easier.

6.4 Stage 2: Develop training requirements

In going through this, it is important to continue the three-part (event, agent, object) approach. For example, a solution may be dynamic with respect to events but static with respect to objects. The event may be temporal (fire the weapon) but the object manipulation procedure may be atemporal (calculate range, azimuth). The object produced may be linguistic (a report) but the reasoning spatial. And so on.

At this point, there should be four inventories, one each for events, agents, and objects, plus one focusing on the problem space as a whole. The first four steps of this stage (approaches to events, to agents, to objects, and to the problem space) should check the appropriate inventory against the taxonomy, and the training heuristics and guidelines should be recorded.

Step 5. Integration

The last step proceeds in two parts. First, internal coherence and consistency must be checked. Second, the results—which are really the ultimate results of applying the method—need to be integrated with the task analysis results and the ISD process.

In checking internal coherence, judgments will need to be made about the degree of autonomy subtasks can or should have from a training requirements point of view. Making subtasks independent simplifies the training task (with respect to the subtasks), but adds an integration training task, and risks removing critical motivation or context information from task learning. And, as will be seen in the next Chapter, some tasks, even though they may have clearly distinct subtasks, resist decomposition.

Integrating results with the larger ISD process is straightforward on one level, problematic on another. Since the results of the method and the analysis described fits quite well into the standard ISD approach, integration is simple and direct. On the other hand, the idealistic recommendations and guidelines that may come out of this analysis will meet the cold realities of limited budgets, limited facilities, and possibly, trainers who are accustomed to doing things they way they have always done them. (Another way of saying

this is to acknowledge that ISD methods and the method presented here are idealizations. Good tools are important, but no substitute for good hands wielding them.) Not surprisingly, integration is likely to be difficult and to require hard work to implement. On the other hand, it might make a positive, possibly even significant, impact on training.

To test the method, it was applied to the Army problem of reading and interpreting topographic maps. The larger problem is planning force movement and/or laydown to achieve mission success within a Brigade/Battalion sector. Diagrammatic illustration of the the Brigade/Battalion commander's responsibilities and the place of topographic map reading among them is provided in Appendix A.

This is an individual, as contrasted with a group, task. An individual task was selected for simplicity, despite the long-term focus of this research on collective tasks. It was also recommended by the unusual circumstance of having access to a subject matter expert, Dr. Susan Fischer, who is also a psychologist with strong experience in training design and evaluation (e.g., Fischer *op cit.*) Applying the method with her improved the method-testing that accompanied the method trial.

7.1 Stage 1: Characterize the task

Problem-solving for this task is defined as complete and correct understanding of the meaning of topographic map features as they represent real-world features, for all such features as are recognized as having military significance. Further, problem solving requires recognition of those features among all that may be present on a particular map that are significant with respect to a specified mission.

A great deal of enabling knowledge is assumed in this definition. To constrain this, we assume that the learner is the commander of a mechanized or armored battalion, that the mission is to defend in sector, and that the enemy is also a mechanized or armored unit. There is still a huge amount of enabling knowledge necessary to use the map for situation assessment and force planning. However, it is possible to limit the objective to recognition of potentially significant features, leaving the actual force planning problem to a separate training process.

Step 1: Build inventories

The event is interpreting the map for its military significance, in the context of performing situation assessment and planning. The agent is the map-reader: the Brigade/Battalion Commander and/or his staff, for our purposes usually the S2 or S3. These are Colonels and Lt. Colonels, possibly Majors; not below. The object is the topographic map. There are two goals:

1. Understanding the terrain and surface features, as "things-in-themselves"
2. Understanding the military meaning of the terrain.

The latter is far more important than the former. The analogy with recognizing word shapes and understanding the meaning of prose is apt. There are several subgoals in (2). The commander wants to understand:

- The degree of enemy mobility
- The affordances of the terrain vis-à-vis the mission
- Cover and concealment
- Where the enemy attack might or could come
- Whether the enemy could move into or out of his sector from a neighboring friendly sector

- The possibilities of the terrain in front of his sector ("in front" meaning toward the enemy)
- The possibilities of the terrain behind his sector.

The consequences of map-interpretation error include mission failure, casualties, loss of resources, and cross-sector leakage. (With respect to this last, Dr. Fischer wondered if the Army has a real reward structure for taking this seriously. It has seemed to her that far greater emphasis is given to individual Command responsibility for a sector than command team co-responsibility. Confirming evidence comes from a Gulf War post-analysis in which commanders ranked sector coordination as their highest candidate for increased training attention.)

The function of the map—the only object—is to represent the terrain and surface features within and around the sector. The map's attributes include contour lines, numeric and iconic symbols, colors, and a legend guiding interpretation of the other content.

The agents' action—reading and interpreting the map—is not performed at a stereotypical location or time. The functional context for map interpretation is situation assessment or development of Commander's Guidance. As part of task characterization, the redundancy provided by other situation assessment activities helps identify and correct errors. Standards are qualitative; the Army currently does not have numeric standards for adequate or expert performance, e.g., accurate identification of some percentage of the significant features within a set time, with a false positive rate no greater than some specified percentage.

Step 2. Flow

This is as simple as it seems possible to get: the agent uses the object to perform the action.

Step 3. Analysis

Analysis of the problem space emphasized the dual-task nature of the problem. Reading the terrain features is convergent (a hill is a hill), but Dr. Fischer described interpreting the meaning of the features as "divergent within constraints". That is, both she and the SMEs with whom she has worked would say there are many solutions to a particular terrain/force laydown configuration (divergent), but that the terrain places strong constraints on what these are (convergent). Dr. Fischer emphasized strongly that the perspective of the expert is that the interpretation drives the reading, not the other way around. It is a strongly goal-driven, rather than data-driven, activity. Said another way, the task is not framed as, "what is there" (reading), and "what can we do with it?" (interpretation). Rather, it is framed as, "what do we want to do? And, then, how can we achieve it?"

Both reading and interpretation are data-rich, event-rich, and hypothesis-rich. Reading terrain from maps is experiment-rich; one can read many maps and view many kinds of terrain. Interpreting the maps is experiment-poor, in terms of verifiable experiments. Solutions are satisfying. Comprehension skills are applied both to reading the terrain and to interpreting it. The task is fairly homogeneous, but discussion of this construct brought out the lack of a clear boundary to the interpretive task. It is hard to draw clear lines between

"simple" reading (identification of features), initial reading/interpretation of military potential, and interpretive exploration of potential mission-specific plans.

Discussion of distributed representation brought out the lack of knowledge of how experts achieve their performance. We discussed two possibilities. One, experts manipulate the external representation to understand issues like mobility, analogous to the way physicists manipulate differential equations. The second possibility was that experts translate the external map representation into an internal mental model (e.g., Johnson-Laird 1983). Dr. Fischer didn't know. Her speculation was that experts probably do some of both, using the representation in its simple external form where possible, and constructing an internal representation only for places with notable risk, potential, or uncertainty. (Recall the discussion of the use of EBL to resolve impasses in application of SBL.)

Metaknowledge steps are not significant, e.g., as long as the task boundary excludes planning, there are no issues of problem discernment (the commanders know they're reading a map), problem representation, representation validation, etc. Coherence with respect to interpretation is the most important metacognitive concern. Other metacognitive issues, e.g., of uncertainty management or truth maintenance, are not significant. (There is significant uncertainty associated with enemy intentions and plans, but that is beyond the scope of the task as defined.)

Definitional (icons and symbols) and prototypical (content and interpretation) concepts are involved. In parallel, two classes of operators seems to be present, perceptual and inductive. Dr. Fischer originally identified two levels of patterns, direct (e.g., hills, swamps) and composite (e.g., kill sacks, canalizing terrain, mobility corridors). She has recently come to the conclusion that there is yet a third pattern recognition level achieved by experts; for example, "affordances for massing of forces".

Common novice errors include:

- Failure to properly use map legends to guide terrain content reading. Contour lines the same distance together on different maps can mean very different things in the real world, according to how the maps are drawn and scaled.
- Interference of irrelevant features. Experts filter efficiently; novices can get stuck on low-salience features.
- Confusion of correlation for cause. Novices are more likely than experts to become convinced that because the enemy *could* use a particularly appealing AOA, the enemy *must* use that AOA.
- Lack of larger-domain interpretive knowledge.

This leads to the characterization of expertise, which is notable for knowledge organization and automaticity. That is, experts see the significance of maps quickly, retrieve and apply relevant information, and devote conscious attention only to those places of greatest danger or advantage.

7.2 Stage 2: Develop training approach

What can be read from the taxonomy based on the characterization above? For analysis, we temporarily separate the reading from the interpretive task. Interpretation is the event; reading is use of the object.

Taking that perspective, the object—the map—requires both definitional and

prototypical concepts. Consulting the taxonomy on training the definitional concepts results in an immediate undercutting of the distinction just attempted. Definitional concepts should be taught with their operationalization; in the present case, that means teaching the definitions with their military interpretations. The prototypical patterns should be taught center-to-periphery. Paralleling the definition/operationalization link, the patterns should be associated with their meanings (affect) as well. This link should be cast in appropriate terms, e.g., "strength and vulnerability" or "advantages and disadvantages". We earlier claimed that when meaning is conventionally expressed in value terms, affect is strongly involved. Therefore, coupling prototypical patterns with their value-laden meanings is associating affect.

Failure to use the map legend can be identified and repaired in the classroom. To train efficient filtering of "noise", graded (simple-to-complex) whole topos should be trained, with increasing levels of irrelevant features. As with the definitional and prototypical concepts above, this calls for joined reading/interpretation training, since the nature of the mission (in this case, armor brigade defense) constrains what is signal and what is noise. To illustrate and reduce confusion of correlation with cause, entrapment cases can be used. The problem of lack of larger domain knowledge is crucial, but it is also at the bounds of this problem and the entrance to the larger military planning and command problem.

Since the mechanisms of expertise are not well understood, specifically the degree to which external representation is translated to internal model, it is not clear whether training aids to internalization are appropriate. For example, computer graphics to illustrate the translation from contour lines to 3D features or interactive simulation to simulate movement through terrain could aid visualization and spatial imagination; but if, relative to expertise, these are the wrong skills, this could have negative transfer.

7.3 Analysis

All the recommendations generated by the taxonomy are included in a training course currently under evaluation (Fischer *op cit.*), except the whole-topo signal-to-noise training. The course is producing good results. Students using the course are both more accurate and faster at recognizing significant terrain features and patterns on topographic maps than controls.

It is not clear whether their ability to interpret the military meaning of terrain features has improved. However, as stated above, that is much more dependent on larger domain knowledge, and its change will be correlated to change in that corpus. In other words, the objective studied by Dr. Fischer was the quality of topo map reading. A different approach might have been taken if the whole of situation assessment had been taken as the goal. But the task decomposition that generated the objective isolated topo reading as an independently-trainable component. Judging the research from the perspective of situation assessment as a whole is to misunderstand the objectives the training is meant to address.

The theory and method as they now stand are critiqued, and further action is recommended.

8.1 Critique of the theory

We naively hoped that the structures of learning, training, and generic task content would mesh into a crisp taxonomy with clearly defined constructs and readily testable links. Instead there is a much more fuzzy structure. The content of the taxonomy is plausible, but hardly definitive. For almost any entry, counterexamples could be found. Plus, the human influence of a great teacher can overwhelm technical considerations of method and approach. What Napoleon said of military leadership—"I'd rather have an army of sheep led by a lion than an army of lions led by a sheep"—holds by analogy for instruction: a great teacher with a mediocre method is better than a poor teacher with the perfect method. But of course a great teacher with the right method will outshine a great teacher with the wrong method.

There are two issues with the current theory that are unsatisfactory.

1. *SBL, EBL, and training*

The two-level theory of learning—similarity-based and explanation-based learning at the knowledge level, rehearsal, organization, and elaboration at the algorithm level—doesn't adequately focus or constrain training approaches. In the extreme cases, they are clearly applicable: teach motor skills and low-level pattern recognition with rehearsal, mathematical theorem proving with EBL. But in the great middle ground, they don't prescribe or delimit with enough precision.

We believe that with more work and more careful analysis, the theory of learning outlined in Chapter 4 could be used to classify and illuminate training approaches, and could be tied to task characteristics as well. Therefore, it could be a productive bridge between tasks and training. However, its current form is not sufficient.

2. *Lack of methodological specificity in the taxonomy*

There is focused and successful work in intelligent computer-aided instruction and elsewhere that links very specific kinds of tasks with equally specific approaches to training. For example, Lesgold's work with context-sensitive state-space differentials for the training of electronic troubleshooting has been shown empirically to be much more effective than the traditional approach. This is not in the taxonomy, but should be. The taxonomy should be able to say, "for a task with characteristics A, B, and C, use model tracing; but if it has characteristics E, F, and G, don't use model tracing." The taxonomy should be more specific, and should include proscriptive as well as prescriptive content.

8.2 Recommendations

We make six recommendations.

Knowledge acquisition with training experts

A program of knowledge acquisition with training design experts should be designed and executed. Repertory grids (Kelly 1955) could be used to characterize how experts understand training design. A second set of knowledge acquisition activities, again using repertory grids, should be devoted to how commanders generate training objectives from their observations, and what they do to validate them.

Further theoretical development

Those data should then be integrated into the theory and taxonomy. The taxonomy should be explicitly expanded to include group training. Despite the difficulty of finding adequately precise terminology, factor analysis and automated induction of decision trees on the data still seem worth pursuit. Quinlan's C4.5 machine learning program (1993) would be appropriate for this. It would be reasonable to pursue the relationship between theories of learning and specific training approaches.

Extension to include training objective diagnosis and articulation

As suggested by the second knowledge acquisition task proposed above, and by the discussion of the difficulty of the diagnosis and likely frequency of misdiagnosis in Section 1, an intelligent diagnostic aid to training objective development seems appealing. It is a natural extension of the current work.

Full automation of the taxonomy and method

Once the method and taxonomy have been extended and deepened in the ways just described, a knowledge based system incorporating them should be built. There are powerful knowledge system development tools available, e.g., Intellicorp's Kappa, and there are true multi-platform delivery environments, e.g., the Zinc Application Framework from Zinc, zApp from Inmark, and XVT from XVT. (Multi-platform delivery means that a single system is designed and built, then "cross-compiled". All three of the above toolkits allow a single set of source code to be compiled to run on Apple Macintosh, IBM-PC-class machines running any of DOS, Windows 3.1, Windows NT, or OS/2, and Unix/Motif, including Sun and H-P.) The complexity of the method suggests that a knowledge based system would make it more manageable. The quality and simplicity of the human-computer interaction component would be of particular importance for the system to be a success. The facility with which draft interfaces can be constructed with Zinc or zApp means that testing and validating the interface is practical. The generative component should draw on expertise in training design and requirements specification, especially automated scenario generation.

Empirical testing

The taxonomy and method should be tested by applying the automated method to a wide range of current Army training programs, selecting one or more current training programs with the greatest contrast to the recommendations of the taxonomy, constructing a taxonomy-based alternative, and testing it to compare the results to the current baseline.

Commercialization

Training in the commercial world is big business (estimated at \$5 billion per

year in the United States). There are obviously signal differences between military and nonmilitary training, but similarities are great as well. When the final Phase II proposal is submitted, a Phase III commercialization plan will be included.

8.3 Finis

A theory of the relationship between learning, generic task content, and training has been developed. A taxonomy has been constructed, based on the theory, linking generic content with training approaches. An ontology and a method for eliciting task descriptions and generating taxonomic recommendations has been presented. Three small examples and a formal knowledge acquisition exercise with a subject matter expert all supported the plausibility of the theory, taxonomy, and method.

Six recommendations for further work have been made; in sum, they are for more research, then development, testing, and commercialization.

This report documents Phase I research, which has been conducted intending to lay the foundation of Phase II construction of a knowledge-based decision aid incorporating the method here described.

REFERENCES

- Alexander, J., Freiling, M., Shulman, S., Rehfuss, S., Messick, S. (1986).
Ontological analysis. In J. Boose & B. Gaines, eds., *Knowledge Acquisition Tools for Expert Systems*. Academic Press, San Diego, CA
- Anderson, J. R. (1983). *The Architecture of Cognition*. Harvard, Cambridge, MA
- Anderson, J. R. (1990). *The Adaptive Character of Thought*. Lawrence Erlbaum, Hillsdale, NJ
- Andriole, S., Black, H., Hopple, G., & Thompson, J. (1987). Intelligent aids for tactical planning. In J. Boyes, S. Andriole, eds., *Principles of Command and Control*. AFCEA Press, Washington, DC
- Brooks, R. (1992). Intelligence Without Representation. *Artificial Intelligence* (57)
- Bruner, J. (1986). *Actual Minds, Possible Worlds*. Harvard, Cambridge, MA
- Chi, M. (1992). Conceptual change within and across ontological categories. In R. Giere, ed., *Cognitive Models of Science*. University of Minnesota Press, Minneapolis
- Chi, M., Slotta, J. (1993). The ontological coherence of intuitive physics. *Cognition and Instruction* (10:2-3), pp. 249-260
- Cohen, P., Greenberg, M., Delisio, J. (1987). "MU: A development environment for prospective reasoning systems". Proc. 6th National Conference on Artificial Intelligence, Morgan Kaufmann, Los Altos, CA
- Collins, A., Stevens, A. (1982). "Goals and strategies of inquiry teachers". In R. Glaser (ed.), *Advances in instructional psychology* (Vol. 2). Lawrence Erlbaum, Hillsdale, NJ
- Cummins, R. (1983). *The Nature of Psychological Explanation*. MIT Press, Cambridge, MA
- diSessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction* (10:2-3), pp. 105-226
- Eddy, D. (1982). Variations in physician practice: The role of uncertainty. *Health Affairs* (3), pp. 74-89
- Ericsson, K., & Simon, H. (1984). *Protocol Analysis*. MIT Press, Cambridge, MA
- Fallesen, J. (1993). Overview of Army Tactical Planning Performance Research. US Army Research Institute for the Behavioral and Social Sciences Technical report 984. AD# ADA273273
- Feldman, C., Wertsch, J. (1976). Context dependent properties of teacher's speech. *Youth and Society* (8), pp. 227-258
- Fischer, S. (PI), (1994). Student Text for Tactical Terrain Pattern Recognition Course. Anacapa Sciences, Santa Barbara (prepared for ARI Leavenworth)
- Fleishman, E., & Quaintance, M. (1984). *Taxonomies of Human Performance*. Academic Press, Orlando, FL
- Gagné, R., Briggs, L., & Wager, W. (1992). *Principles of Instructional Design*. Harcourt Brace Jovanovich College Publishers, Fort Worth TX
- Gardner, H. (1983). *Frames of Mind*. Basic Books, New York
- Geiwitz, J., Kornell, J. (1992). The Knowledge Acquisition Guidebook. Anacapa Sciences Technical Report 785-4, Santa Barbara, CA
- Gibson, J. (1979). *The ecological approach to visual perception*. Houghton Mifflin, Boston, MA
- Glaser, R. (1989). Expertise and learning. In D. Klahr & K. Kotovsky, eds., *Complex information processing*. Lawrence Erlbaum, Hillsdale, NJ
- Gould, S. J. (1981). *The Mismeasure of Man*. W. W. Norton, New York
- Halliday, M. (1975). *Learning How to Mean*. Edward Arnold, London

- Haugeland, J. (1978). The nature and plausibility of cognitivism. *Behavioral & Brain Sciences* (2), pp. 215-260
- Hayes, J., Simon, H. (1977). Psychological differences among problem isomorphs. In N. Castellan, D. Pisoni, & G. Potts, eds., *Cognitive theory*. Lawrence Erlbaum, Hillsdale, NJ
- Hogan, J., Broach, D., & Salas, E. (1987). A Task Information Taxonomy for Instructional Systems Design. Naval Training Systems Center, Orlando, FL
- Johnson-Laird, P. (1983). *Mental Models*. Harvard, Cambridge, MA
- Josephs, R., et al. (1994). "When quantity means quality", reported in *Science News* (145:12), March 19, 1994, p. 188
- Kahneman, D., Slovic, P., & Tversky, A., eds. (1982). *Judgement Under Uncertainty*. Cambridge University Press, Cambridge, England
- Kotovskiy, K., & Simon, H. (1990). "What makes some problems really hard: Explorations in the problem space of difficulty". *Cognitive Psychology* (22), pp. 143-183.
- Lesgold, A. (and fifteen others) (1985). *Guide to Cognitive Task Analysis* DRAFT. Unpublished, even now. I don't even recall how I got hold of the copy I have.
- Mayer, R. (1987). *Educational Psychology*. Little, Brown & Co., Boston MA
- Meyer, B. (1975). *The Organization of Prose and Its Effects on Memory*. North-Holland, Amsterdam
- Minsky, M. (1985). *The Society of Mind*. Simon & Schuster, New York
- Mirabella, A. (1993). The challenge of measuring group performance. Panel Presentation to the Human Factors and Ergonomics Society Convention
- Mitchell, T. (1982). Generalization as search. *Artificial Intelligence* (18), pp. 203-226
- Newell, A. (1982). The knowledge level. *Artificial Intelligence* (18), pp. 87-127
- Newell, A. (1990). *Unified Theories of Cognition*. Harvard, Cambridge, MA
- Nissen, M. (1985). Accessing features and objects: Is location special? In M. Posner & O. Martin, eds., *Attention and Performance XI*. Lawrence Erlbaum, Hillsdale, NJ
- Ortony, A., Clore, G., Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge, England
- Patel, V., Groen, G., Norman, G. (1993). Reasoning and instruction in clinical medicine. *Cognition & Instruction* 10(4), 281-333
- Pylyshyn, Z. (1986). *Computation and Cognition*. MIT Press, Cambridge, MA
- Quinlan, R. (1993). *C4.5: Programs for machine learning*. Morgan Kaufman, San Mateo, CA
- Reif, F., & Allen, S. (1992). Cognition for interpreting scientific concepts: A study of acceleration. *Cognition & Instruction* 9(1), pp. 1-44
- Rumelhart, D., McClelland, J. (1986). *Parallel Distributed Processing*. MIT Press, Cambridge, MA
- Simon, H. (1981). *The Sciences of the Artificial (2nd Ed.)*, MIT, Cambridge, MA
- Smith, C., Snir, J., Grasslight, L. (1992). Using conceptual models to facilitate change. *Cognition & Instruction* 9(3), 221-283
- Thordsen, M., Klein, G., Michel, R., and Sullivan, E. (1991). Methods for providing direct feedback about decision processes for command and control classroom exercises. ARI Research Note 91-20. Army Research Institute, Alexandria, VA AD# ADA231760
- White, B. (1983). Sources of difficulty in understanding Newtonian mechanics. *Cognitive Science* (7), pp. 41-65
- Zhang, J., Norman, D. (1994). Representation in distributed cognitive tasks. *Cognitive Science* (18:1), pp. 87-122

End User Manual for the Standard Army Training System, Volume 2: Training the Force. TRADOC TCA, Ft. Eustis, VA 23604. 1992

FM 100-5. Operations. Headquarters, Department of the Army. 1986

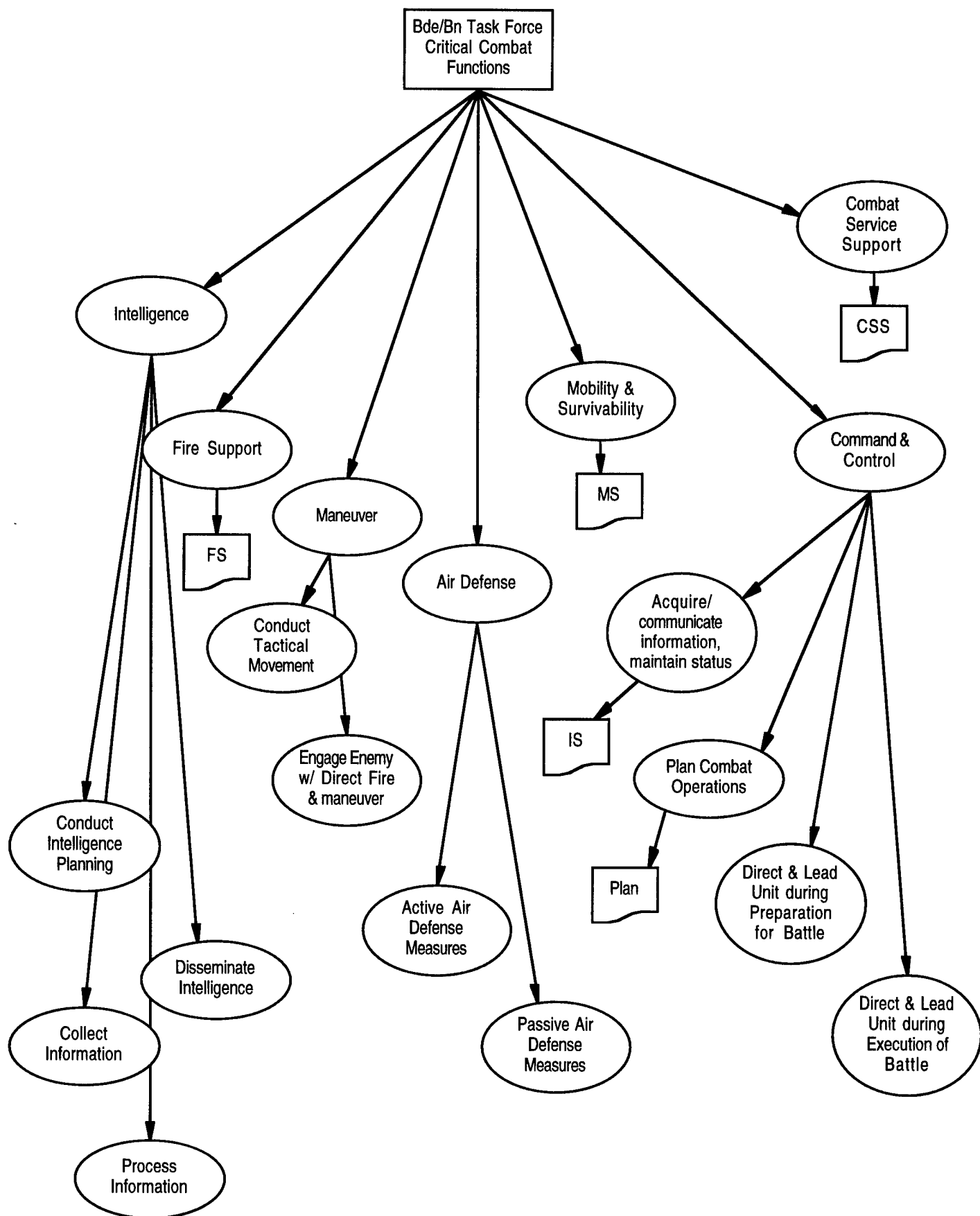
FM 25-100. Training the Force. Headquarters, Department of the Army. 1988

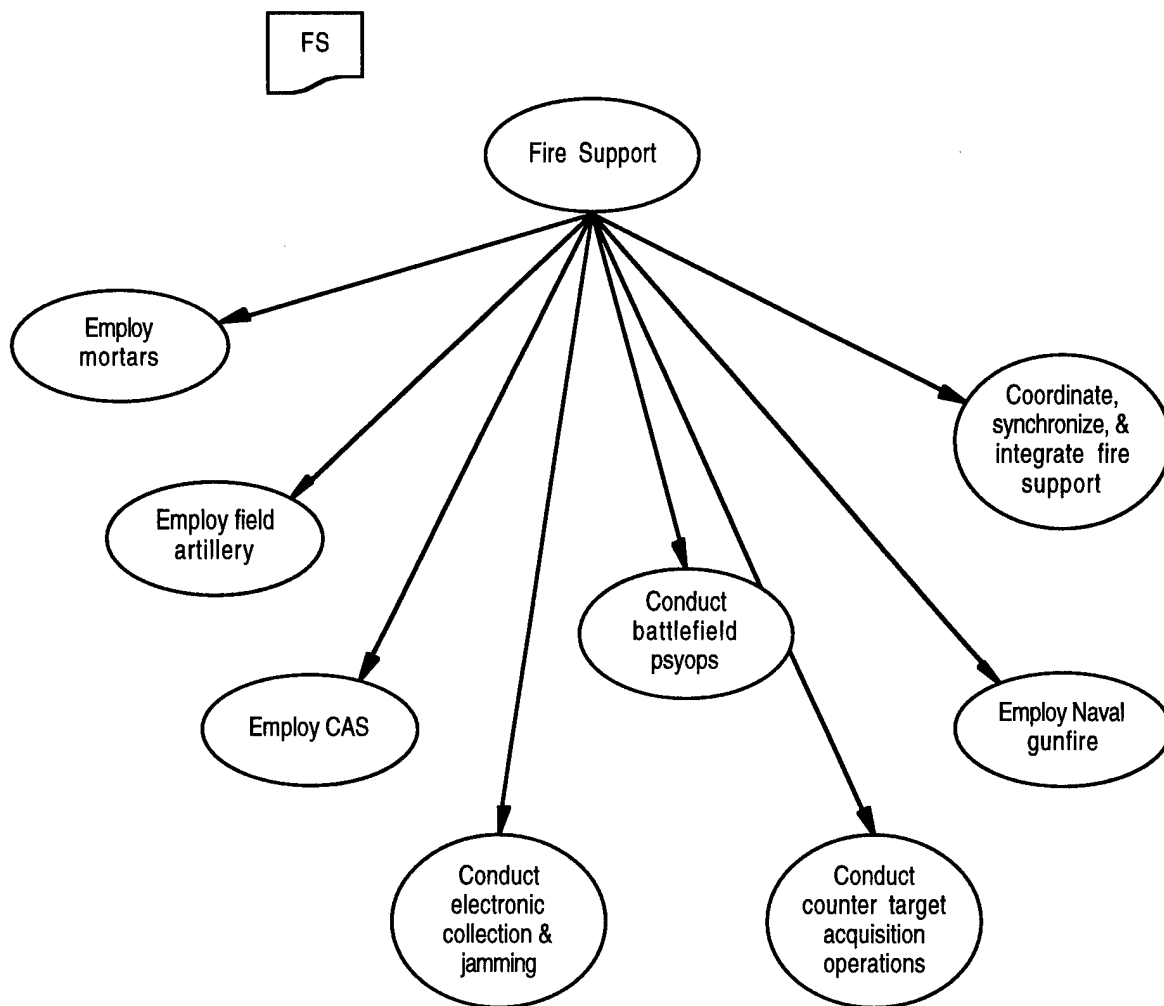
FM 71-3. Armored and Mechanized Infantry Brigade. Headquarters, Department of the Army. 1992

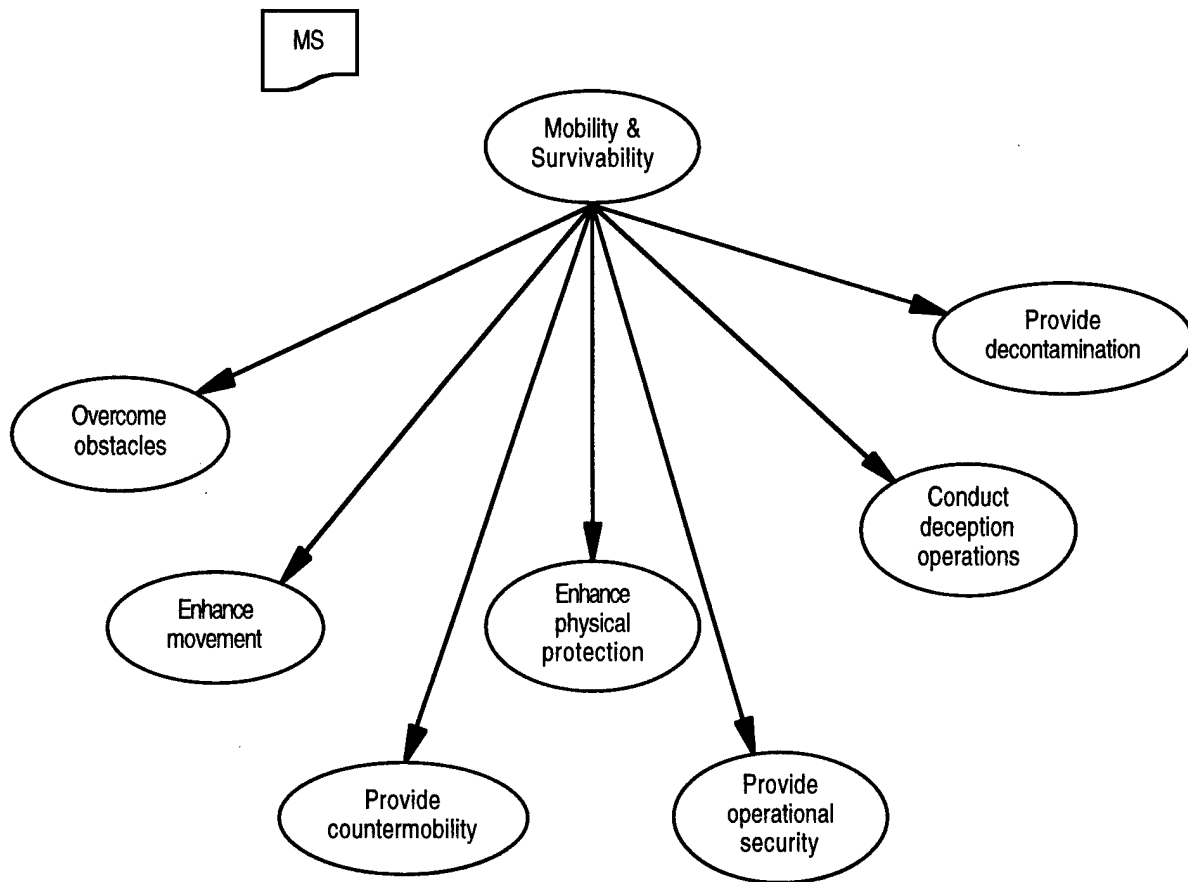
FM 71-123. Tactics and techniques for combined arms heavy forces: Armored Brigade, Battalion/Task Force, and Company Team. Headquarters, Department of the Army. 1992

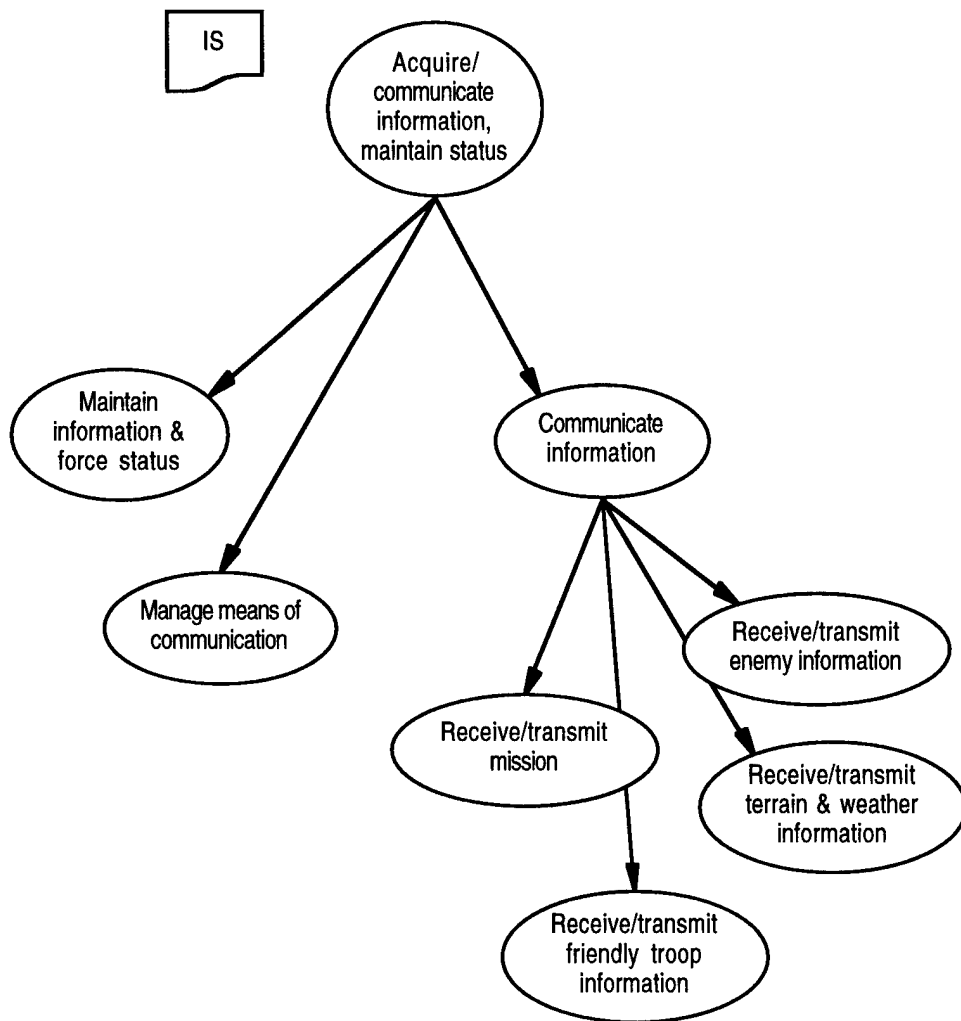
APPENDIX A: BRIGADE/BATTALION TASK DECOMPOSITION

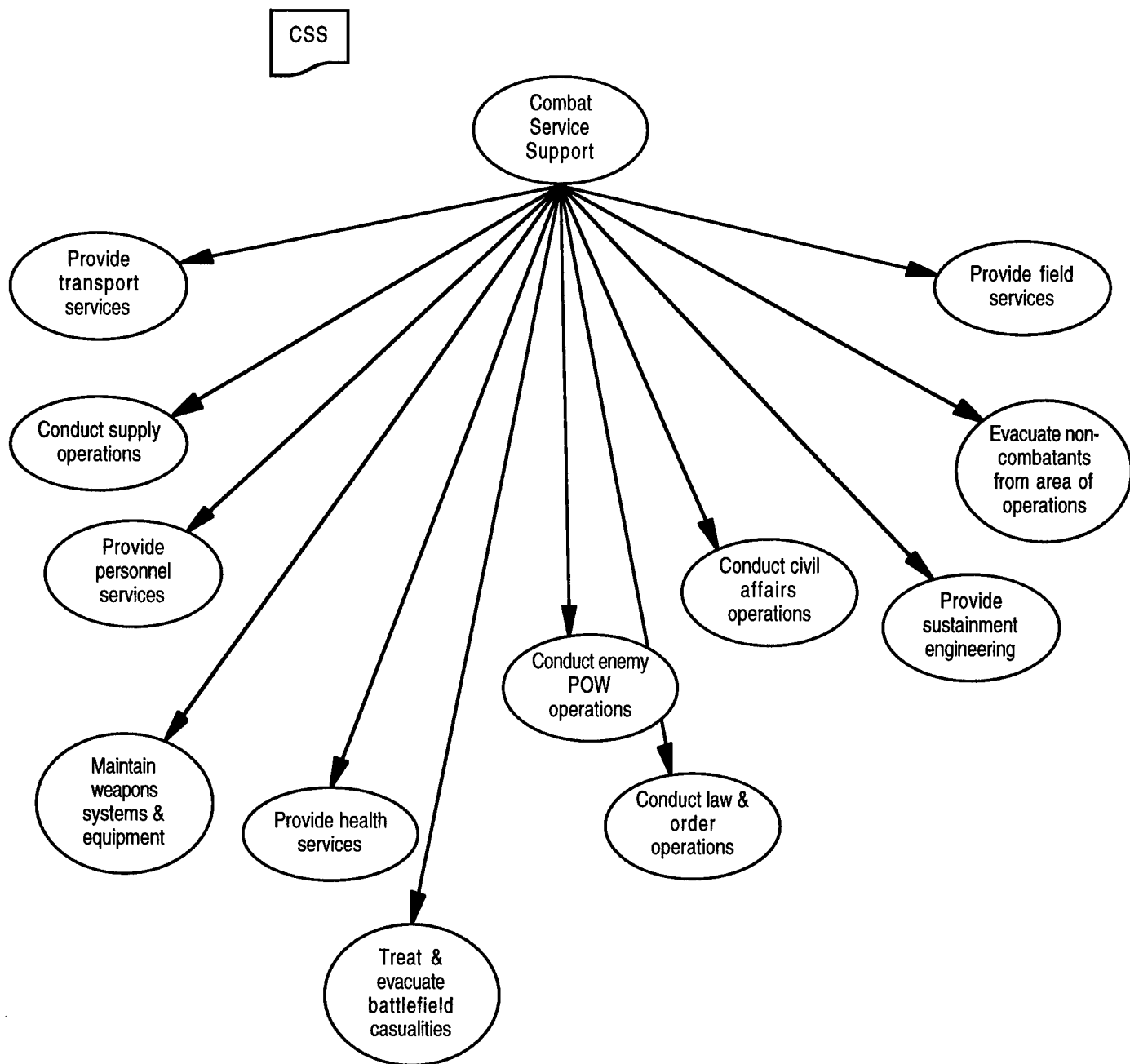
These charts show where reading and interpreting topographic maps fit into the broad and heterogeneous obligations of Brigade/ Battalion command. These charts are a composite of data from The Blueprint of the Battlefield, FM 71-3, FM 71-123, and interview materials from Lt. Col. Kenneth Bell, ret., a former Command & General Staff College instructor.

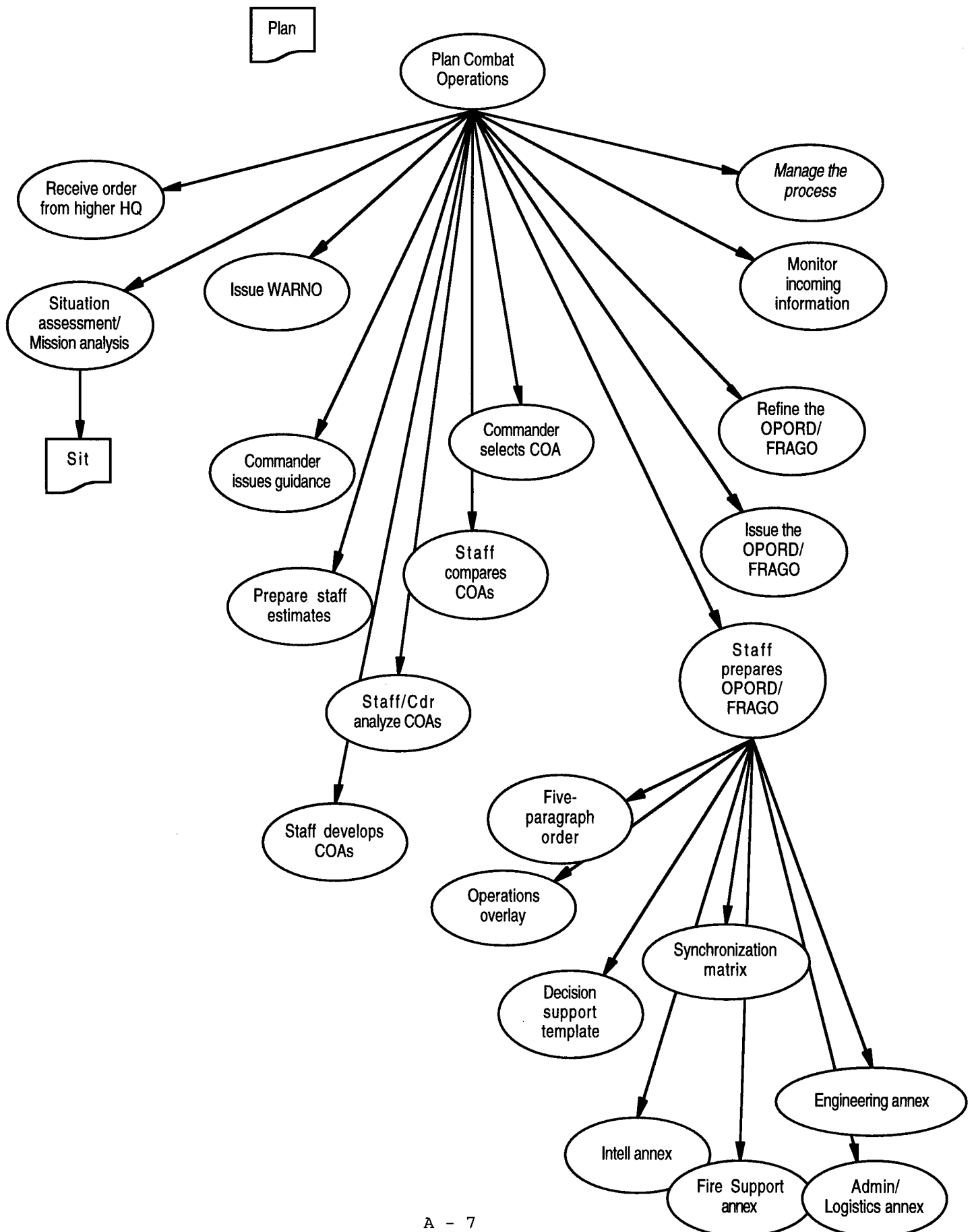


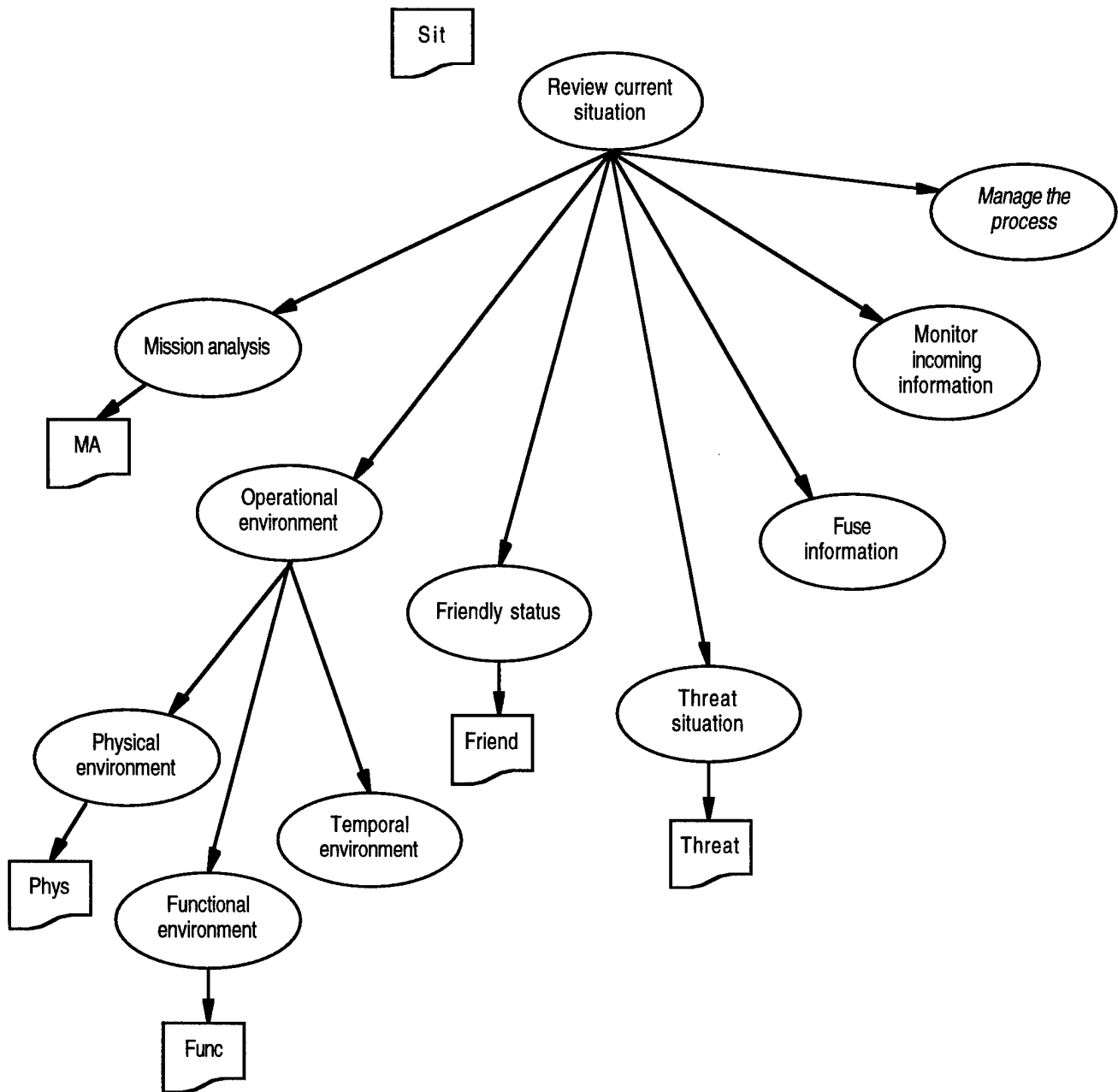


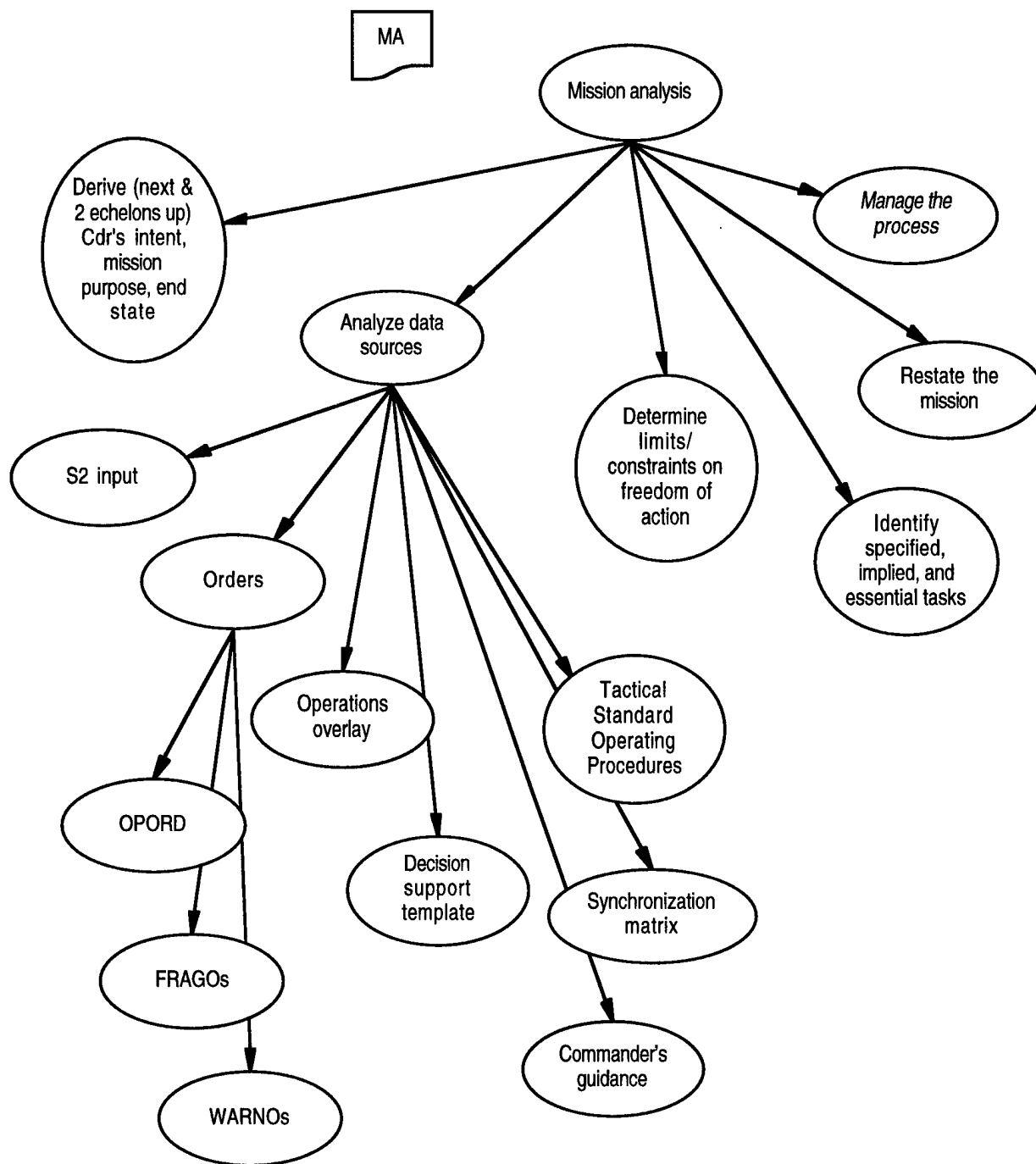


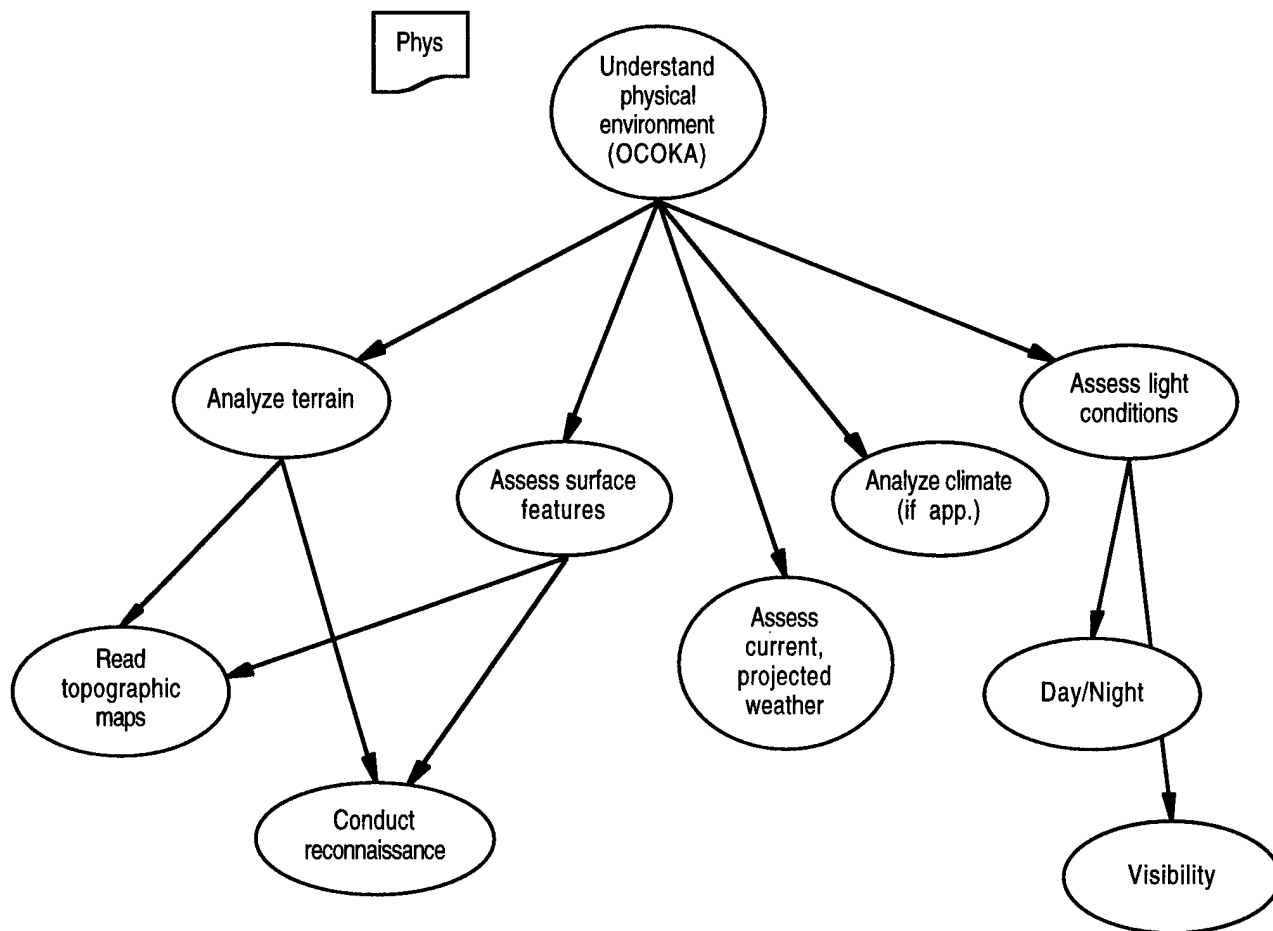


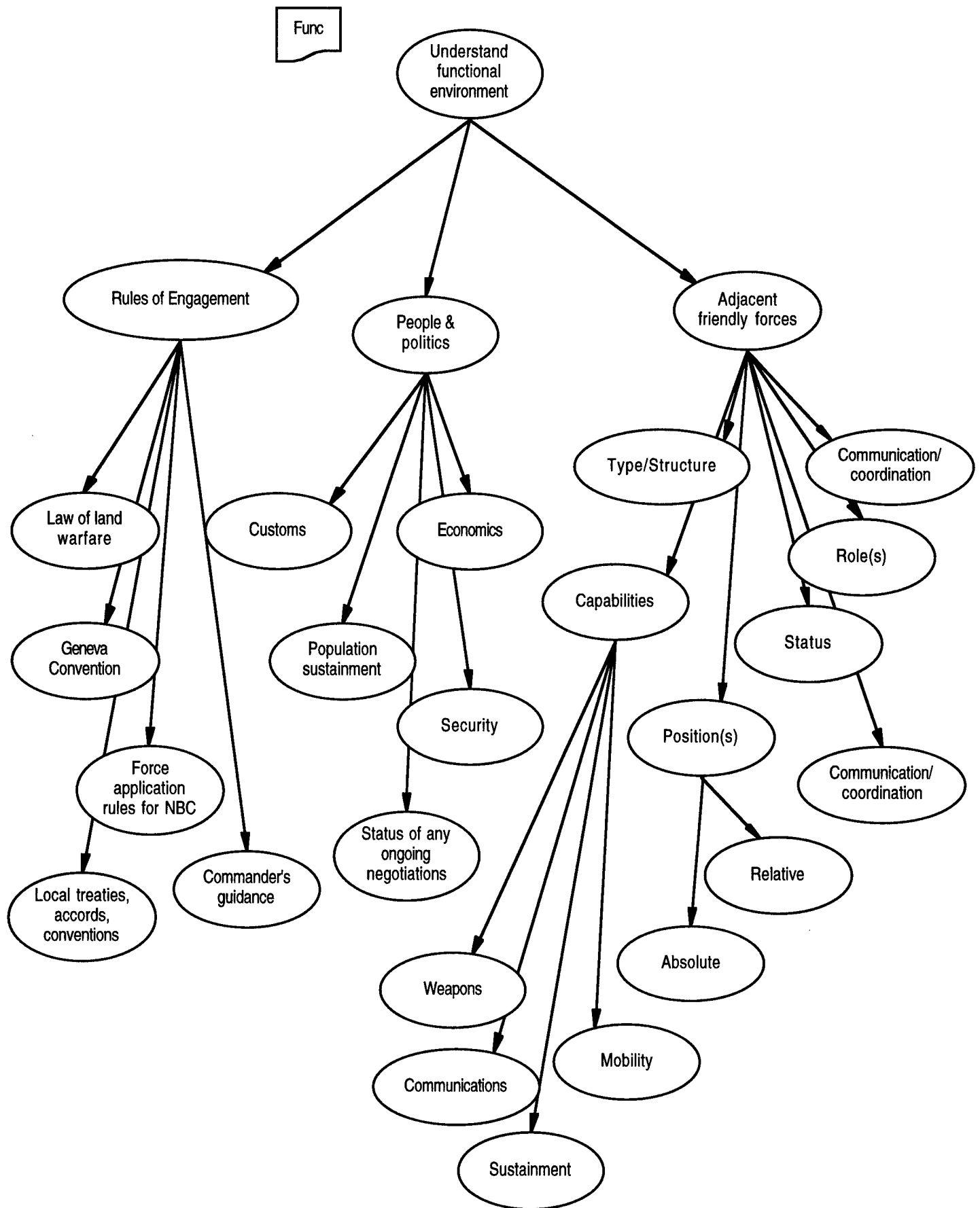












Friend

